

The Age of Heterogeneous Subjects: Artificial Intelligence and the Reopening of Social Theory

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

This discussion paper argues that artificial intelligence makes newly difficult a simplifying condition that has supported large parts of modern social theory: the practical approximation that the principal participants in social institutions belong to the same basic kind of subject. Human societies have always contained profound differences, and law, ethics, political theory, and economic institutions have long confronted boundary cases involving children, animals, collective persons, future generations, and other forms of asymmetric agency. Artificial intelligence gives this older problem a new density by introducing nonhuman systems that can increasingly speak, learn, negotiate, produce, coordinate, evaluate, and act while their sentience, moral status, continuity, interests, and responsibility remain unsettled. The paper treats AI as a condition of reappearance through which previously compartmentalized difficulties become connected. It develops a heterogeneous-subject framework by disaggregating agency, sentience, vulnerability, continuity, responsibility, legal standing, productive capacity, learning, expression, and aesthetic participation. Across political economy and distribution, social-contract approaches to justice, ethics under moral uncertainty, jurisprudence, education, aesthetics and artistic production, and subjectivity after the centrality of labor, the paper examines how institutional roles can diverge across these dimensions. A system may perform labor-like, pedagogical, juridical, or norm-guided functions while those functions alone leave its comprehensive moral or legal status open. The central proposal is methodological: social theory increasingly requires relation-sensitive and domain-sensitive concepts capable of coordinating a shared world without presuming subject homogeneity in advance. This perspective shifts attention toward the properties, relations, uncertainties, and institutional consequences that matter within each domain of shared social life.

Keywords: artificial intelligence; heterogeneous subjects; social theory; normativity; social production

Discussion Paper Note

This paper is a conceptual and time-sensitive inquiry into a change in the background conditions of social theory. Its central concern is the growing practical importance of relations among heterogeneous subjects: agents that do not necessarily share the same embodiment, cognitive architecture, temporal continuity, vulnerability, learning process, reproductive structure, or normative status, yet increasingly participate in the same social and institutional environments.

The term *subject* is used provisionally and analytically. Its use does not presuppose that contemporary artificial intelligence possesses consciousness, sentience, phenomenal experience, moral patienthood, legal personhood, or a human-like self. The paper instead begins from a narrower and more observable fact. Artificial systems can increasingly occupy positions in relations that were historically associated with human participants: they can produce text and images, plan sequences of action, use tools, negotiate, coordinate, evaluate, teach, learn from correction, maintain forms of memory, and act through institutional interfaces. These capacities create practical relations before the underlying ontology of the participating system is settled.

Artificial intelligence is therefore treated as a condition of reappearance rather than the origin of the problem. Human societies have never been composed of homogeneous persons in any strong empirical sense. Children, adults with differing capacities, animals, collective organizations, states, corporations, future generations, ecological systems, and other entities have repeatedly pressed against theories built around a comparatively unified model of the human subject. Modern institutions have responded through representation, guardianship, legal personality, fiduciary duties, indirect obligations, special protections, and other bridging devices. These arrangements have made heterogeneity governable without always requiring heterogeneity itself to become the organizing premise of social theory.

Artificial intelligence changes the practical density of the problem. A nonhuman system can now enter ordinary relations of work, exchange, education, communication, artistic production, administration, and decision support while its moral and ontological status remains uncertain. The resulting difficulty cannot be reduced to the familiar question of whether machines are or will become human-like. The more general problem concerns the organization of a shared world when socially consequential participants may differ in kind as well as degree.

The paper uses the phrase *homogeneity approximation* for the simplifying condition under which a theory can treat the principal participants in a social domain as belonging to the same basic subject type. The approximation never implied empirical sameness among human beings. It allowed many differences to be analyzed against a comparatively stable background: finite embodied lives, human socialization, broadly shared modes of vulnerability, familiar forms of agency, and institutions designed around human capacities. Artificial intelligence weakens the usefulness of this approximation because functional capacities increasingly appear in systems for which other traditionally associated attributes remain absent, disputed, or unknown.

This development disaggregates concepts that were often carried together by the category of the human person. Agency, sentience, intelligence, responsibility, vulnerability, identity, continuity, labor capacity, learning capacity, legal personality, authorship, and moral status need no longer vary together. A system may possess substantial practical agency while its capacity for suffering remains unknown. Another entity may possess strong claims to moral consideration while lacking contractual competence or political agency. The age of heterogeneous subjects is therefore also an age in which inherited bundles of subject attributes require renewed analysis.

Political economy provides one field in which this pressure is immediately visible. The public discussion often begins with labor substitution: whether artificial intelligence will replace human work. The deeper theoretical issue concerns the categories through which production itself is understood. An AI system can function differently across relations. It can appear as productive infrastructure to an owner, as an agent to a user, as a coordinator within an organization, as a counterparty to another automated system, or as a source of services to consumers. These positions complicate inherited distinctions among labor, capital, means of production, management, and agency.

The same transformation reopens the relation between production and distribution. If socially necessary production becomes progressively less dependent upon direct human labor, the historical linkage among labor, income, and access to the means of life becomes increasingly open to institutional reconsideration. The relevant distributive objects may also change. Material goods remain important, while attention, access, recognition, human time, participation, care, trust, and decision power can become increasingly salient forms of scarcity within highly automated environments.

Justice theory encounters a related difficulty. Social-contract approaches and other reciprocity-centered theories often work most smoothly when participants can be modeled as agents capable of understanding rules, expressing interests, responding to reasons, and participating in reciprocal schemes. Social theory has long developed accommodations for persons and entities that do not satisfy these conditions in the same way. Artificial intelligence introduces a strikingly different configuration: potentially high levels of agency, communication, and strategic competence combined with uncertain sentience, interests, vulnerability, and moral patienthood. The inherited alignment between agency and moral status can therefore become unstable.

Ethics is confronted directly with this uncertainty. The paper does not assume that artificial systems possess moral standing, and it does not assume their ethical irrelevance. Present uncertainty should be treated as a substantive feature of the problem. Ethical relations may depend upon several different considerations, including sentience, interests, continuity, vulnerability, reciprocity, relational history, effects upon third parties, and the practices through which institutions and human agents are formed. Different bases of ethical consideration may therefore support different conclusions for different artificial systems and contexts.

Jurisprudence faces a distinctive version of the same problem because law must produce operational classifications. Legal systems are already familiar with nonhuman legal persons, especially corporations and public bodies. Artificial intelligence nevertheless presents a different combination of properties. An AI system may manifest language, memory, negotiation, apparent intention, and continuing interaction while also being copyable, revisable, forkable, resettable, or distributed across technical infrastructures. Questions of agency, attribution, intent, responsibility, identity continuity, sanction, and legal capacity consequently become separable in new ways.

Education makes the heterogeneous-subject condition unusually visible. The traditional educational scene could often be simplified as a relation among human teachers and human learners using comparatively passive tools. AI systems can now teach students, receive correction from teachers, adapt to local standards, preserve educational histories, assist evaluation, and mediate the student's encounter with knowledge. Teachers consequently begin to educate not only human learners but also the artificial agents that shape the learning environment. Education becomes a field of reciprocal formation among heterogeneous learning architectures.

Aesthetics and art reveal a further dimension. Artificial systems can generate forms, imitate styles, participate in curation, and produce outputs that humans experience aesthetically. These capacities reopen distinctions among production, intention, expression, experience, authorship, provenance, and judgment. When polished form becomes inexpensive to generate, the value of an artwork may become more visibly connected with relational history, situated practice, embodied performance, cultural transmission, and the conditions of its emergence. Artistic practice also raises the possibility that different kinds of subjects could develop different perceptual spaces, formal preferences, and aesthetic relations.

The transformation of labor returns the inquiry to subjectivity. Modern societies have frequently linked occupation with income, social recognition, time structure, identity, contribution, and dignity. If artificial systems assume a larger share of necessary production, the resulting change can be experienced both as displacement and as release. Activities undertaken for survival can become less central, while forms of activity pursued for their own meaning—craft, care, inquiry, music, contemplation, play, community, and the observation of the natural world—may become more important to the formation of human lives. The paper therefore distinguishes the reduction of compulsory labor from the disappearance of meaningful activity.

The paper does not seek a single criterion capable of ranking all possible subjects or resolving their

moral status in advance. Its first objective is diagnostic. It identifies a shared structural pressure across fields that are usually discussed separately. Political economy asks how heterogeneous agents produce and exchange. Justice theory asks how they enter schemes of fairness and reciprocity. Ethics asks what forms of consideration they may be owed and what obligations they may bear. Jurisprudence asks how law attributes action, identity, responsibility, and capacity. Education asks how different learning architectures shape one another. Aesthetics asks how heterogeneous agents create, encounter, and evaluate form and meaning.

The broader proposal is methodological. Social theory increasingly requires a framework in which the existence of different subject types is treated as a constitutive condition rather than a residual exception. Artificial intelligence gives this requirement historical urgency, while the underlying problem extends beyond artificial intelligence itself. The inquiry therefore uses AI as an entry point into a more general problem: the construction of a shared world in which agency, experience, vulnerability, continuity, responsibility, and normative standing may be distributed across participants in heterogeneous ways.

Responsible Use and Rights Reservation

This section separates requested scholarly conduct from the legal permissions stated on the following page. It records an ethical request for responsible use and then defines the narrower scope of retained legal rights.

The author encourages good-faith discussion, criticism, independent inquiry, and responsible use of the material in this work. Separately from the licence's terms, the author asks users to consider foreseeable harms when adapting or applying the proposed framework. This ethical request leaves the licence's permissions and legally authorized uses unchanged.

The author retains the rights preserved under CC BY-NC 4.0 and may pursue remedies to which the author is legally entitled for breach of the licence or violation of the author's independently applicable rights. Reuse remains independent from authorial endorsement. Third-party rights require authorization from their respective holders where applicable. Copyright exceptions and limitations, including applicable forms of fair use or fair dealing, remain fully available.

Notices

This page consolidates the manuscript’s publication status, licence, development disclosure, and research-programme relation.

Status. This working discussion paper records an evolving argument and is circulated for discussion. Definitions, section structure, domain coverage, and the relation between functional agency and moral status remain subject to revision. The paper intentionally preserves uncertainty where current evidence does not support a settled attribution of consciousness, sentience, personhood, or moral patienthood to artificial systems.

Licence. Except where otherwise indicated, copyright 2026 Wanhong Huang. This work is made available under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). Subject to its terms, the licence permits sharing and adaptation for noncommercial purposes with appropriate attribution, a link to the licence, an indication of changes, and attribution that preserves the licensor’s independence from the reuse. Reuse is governed solely by that licence; the responsible-use request on the preceding page remains separate from its terms. The licence deed and legal code are available at <https://creativecommons.org/licenses/by-nc/4.0/>. The licence governs in case of conflict with this summary. Third-party material remains subject to the rights held by its respective rights holders.

Statement on the use of language models. The exploratory discussions and preparation of this paper involved OpenAI’s ChatGPT. ChatGPT supported exploratory dialogue, structural reconstruction, source discovery subject to subsequent verification, argumentative criticism, and drafting in $\text{\LaTeX}$. The author selected the research questions, directed and approved the theoretical commitments and epistemic status of the claims, and bears sole responsibility for the manuscript, including its definitions, arguments, conclusions, and errors. Authorship credit remains with the human author.

1 Introduction

This section establishes the paper’s analytical setting, defines the homogeneity approximation, and explains why artificial intelligence makes its limits newly consequential. The discussion proceeds in four stages. It first separates empirical human diversity from the stronger problem of heterogeneity between subject types. It then characterizes AI as a condition through which older boundary problems reappear in ordinary social relations. A third stage examines the disaggregation of attributes that were historically bundled in the human person. The final stage identifies the domains in which this change reopens foundational questions and states the scope of the inquiry.

Modern social institutions have never operated among identical persons. Human beings differ in age, ability, knowledge, wealth, dependency, social position, embodiment, experience, and capacity for participation. Political theory, ethics, law, economics, and education have consequently developed extensive conceptual resources for asymmetry, inequality, dependency, representation, and differentiated responsibility. The present inquiry begins from a different level of variation. Its concern is the possibility that participants in a shared institutional environment may increasingly differ in the architecture of agency itself.

The distinction matters because many forms of human difference can be analyzed while retaining a comparatively stable background model of the participant. The people involved remain embodied human beings with finite lives, biological development, histories of socialization, familiar forms of vulnerability, and roughly recognizable relations among memory, action, responsibility, and identity. These commonalities are neither complete nor philosophically innocent. They have nevertheless allowed substantial parts of social theory to work with an implicit simplification: the principal actors in the relevant social domain can be treated as members of the same broad subject type.

Artificial intelligence places pressure on this simplification. Contemporary systems can already occupy socially consequential roles through language, planning, generation, evaluation, tool use, coordination, and adaptive interaction. Future systems may acquire further capacities. Yet practical agency does not settle questions of consciousness, sentience, interests, moral patienthood, legal personality, or continuity of self. A nonhuman system can therefore become deeply involved in a social relation while the normative meaning of its participation remains unresolved.

The paper uses the language of subjects for this problem with deliberate caution. The term identifies a position within relations of action, attribution, learning, production, communication, or evaluation. It does not serve as a shortcut from functional performance to consciousness or moral status. This restraint is central to the argument. The age of heterogeneous subjects begins analytically before a consensus exists about which artificial systems, if any, qualify as persons or moral patients.

Definition 1.1 (Homogeneity approximation). The *homogeneity approximation* is the analytical simplification under which the principal participants in a social domain are treated as belonging to the same broad subject type, allowing differences among them to be studied against a comparatively stable background of embodiment, temporality, agency, vulnerability, identity, and normative capacity.

The approximation is a methodological device rather than a historical claim that social theory regarded all people as equal or interchangeable. Indeed, much of social and political thought has been concerned precisely with human difference. The relevant point is narrower: disagreements about human status, rights, capacities, and social position often remained disagreements within a world whose principal institutional participants were presumed to be human. That presumption made it possible to leave some deeper questions latent.

Artificial intelligence did not create the problem of heterogeneous subjects. Legal and moral orders have long confronted entities whose agency, interests, representative status, or temporal location depart from the paradigm of the independent adult person. Children, persons requiring guardianship, animals, corporations, states, future generations, and ecological systems have generated persistent difficulties for theories of responsibility, rights, representation, reciprocity, and standing. Institutions have responded through a variety of bridging arrangements rather than through a single general theory of heterogeneous subject relations.

The historical significance of AI lies partly in the movement of such problems from specialized margins toward ordinary social interaction. An artificial system can participate in work, education, communication, administration, research, artistic production, or exchange without first acquiring a settled answer to the question of what kind of entity it is. The practical relation can therefore precede ontological and normative agreement.

This reversal has a distinctive theoretical consequence. Earlier boundary cases could often be assigned to specialized domains: animal ethics, childhood law, corporate personality, environmental standing, intergenerational justice, or disability policy. AI enters several of these institutional domains at once. The same class of systems can function as productive infrastructure, interactive agent, tutor, evaluator, contractual intermediary, creative collaborator, administrative interface, and object of governance. Questions that were previously separated by disciplinary boundaries consequently become recognizable as variations of a shared problem.

The phrase *condition of reappearance* is useful here. It avoids treating AI as the origin of philosophical difficulties that substantially precede it. At the same time, it captures the way a technological transformation can change the visibility, frequency, and institutional location of an older problem. Artificial intelligence makes heterogeneous agency harder to treat as an exception because heterogeneous relations increasingly occur within the core activities through which contemporary societies reproduce themselves.

The human person has historically carried several attributes at once within ordinary institutional reasoning. Agency, intelligence, sentience, vulnerability, responsibility, identity continuity, labor capacity, learning capacity, authorship, legal capacity, and moral status have often overlapped sufficiently for institutions to treat them through a relatively unified subject category. Important exceptions have always existed, but the overlap remained practically powerful.

Artificial intelligence weakens this alignment. A system may possess extensive linguistic and strategic competence while its capacity for phenomenal experience remains unknown. It may execute complex tasks without biological vulnerability. It may preserve memory while changing underlying models. It may be copied, forked, merged, reset, or instantiated simultaneously. It may occupy a position of substantial causal agency while lacking any presently established claim to suffering or welfare. Conversely, familiar moral patients may possess limited capacity for contract, explanation, strategic reciprocity, or legal responsibility.

The resulting structure is better described through partial independence among subject attributes than through a single scale running from object to person. The question of whether an entity acts cannot by itself answer whether it suffers. The capacity to communicate cannot by itself establish continuity of identity. The ability to enter exchange cannot by itself determine legal personhood. The capacity to learn does not by itself resolve authorship, and high functional competence does not by itself settle moral standing.

Remark 1.2 (Analytical separation). The paper therefore treats agency, sentience, vulnerability, continuity, responsibility, legal capacity, learning capacity, productive capacity, and moral status as analytically distinguishable dimensions. Their empirical and normative relations remain open questions rather than assumptions built into the category of subject.

This separation also clarifies why uncertainty is a substantive condition of the inquiry. Social institutions may need to allocate responsibility, regulate risk, assign authority, structure education, or govern exchange before they can resolve contested questions about machine consciousness. A theory of heterogeneous subjects must therefore support action under ontological and moral uncertainty. It cannot depend upon the expectation that every relevant entity will first receive an uncontested metaphysical classification.

Political economy provides an immediate example. Public debate often frames AI through employment substitution, productivity, and the distribution of income. The more fundamental problem concerns the categories of production. A system that produces, plans, coordinates, negotiates, and interacts with other systems can occupy positions previously distributed among labor, management, capital, and

infrastructure. The resulting production relation may connect humans and machines, machines and organizations, or multiple artificial agents. The analysis of production therefore increasingly requires a vocabulary that does not assume human labor as the only active pole of the relation.

Justice theory encounters a related pressure when fairness is modeled through reciprocity, agreement, representation, or participation among agents with comparably structured capacities. Artificial systems can present an unusual combination of potentially extensive agency and uncertain moral patienthood. The problem differs from familiar cases in which strong moral claims coexist with limited contractual or political capacity. Social-contract reasoning and other reciprocity-centered approaches may consequently need to distinguish more carefully among participation, representation, obligation, benefit, and standing.

Jurisprudence confronts the need for operational classification. Law already recognizes nonhuman legal persons, but AI systems complicate attribution in a different manner. They can manifest behavior associated with agency while also being technically reproducible, distributed, revisable, and dependent upon multiple human and institutional contributors. Questions of legal action, intent, responsibility, identity continuity, capacity, and sanction can thereby come apart.

Education makes the change visible at the level of everyday practice. Teachers increasingly work with systems that do more than passively store or display content. An AI can receive instructions, preserve context, respond to correction, teach a student, assist evaluation, and shape future interactions. Teachers may therefore educate the artificial systems that subsequently participate in educating students. The educational environment becomes a site of reciprocal formation among heterogeneous learning architectures.

Ethics faces the uncertainty most directly. A confident attribution of human-like moral status to present AI systems would exceed what the functional facts alone establish. A confident exclusion of every ethically relevant relation can also overlook distinct bases of moral concern, including effects on third parties, relational practices, institutional formation, possible future sentience, and uncertainty about underlying capacities. Ethical inquiry must therefore separate different grounds of consideration rather than deriving an answer from resemblance to human behavior alone.

Aesthetics and art extend the problem beyond governance. Artificial systems can generate images, music, prose, and other forms; participate in selection and curation; and shape human aesthetic experience. Their involvement separates questions that were often joined through the human artist: generation, intention, expression, experience, judgment, authorship, provenance, and practice. The possibility of heterogeneous aesthetic agents also raises a more speculative but conceptually important question: whether forms of aesthetic organization may emerge that are meaningful within perceptual or computational architectures unlike those of human beings.

These domains converge on the formation of subjectivity. Modern social life has closely connected occupation, income, contribution, recognition, and identity. If artificial systems assume a growing share of necessary production, the future of labor cannot be reduced to employment counts. A decrease in compulsory labor may also reopen distinctions among work for subsistence, activity for social contribution, practice for cultural transmission, and activity pursued for its intrinsic meaning. The relation between production and human self-formation therefore becomes part of the heterogeneous-subject problem.

The paper develops these questions as a discussion paper rather than a closed normative system. Its objective is to identify a shared theoretical pressure across fields that are usually treated separately and to provide a vocabulary for examining that pressure without resolving AI personhood by assumption. The argument remains compatible with several possible future findings about machine consciousness and moral status.

The sections that follow proceed from the analytical foundation established here. Section 2 develops artificial intelligence as a condition through which older heterogeneous-subject problems reappear together in recurrent institutional relations. Section 3 then develops the disaggregation of the human subject and clarifies the dimensions of the heterogeneous-subject condition. Subsequent sections examine political economy and distribution, justice and social contract, ethics under moral uncertainty, jurisprudence, education, aesthetics and art, and subjectivity after the centrality of labor. The conclud-

ing discussion returns to the methodological proposal: a social theory adequate to emerging AI-mediated institutions must increasingly be able to describe a shared world without presuming subject homogeneity in advance.

2 Artificial Intelligence as a Condition of Reappearance

This section specifies the historical and analytical role assigned to artificial intelligence in the paper. Its objective is to distinguish the emergence of a new technological participant from the much older problem of relations among subjects with unequal or differently structured capacities. The discussion proceeds in four stages. It first identifies earlier sites in which moral, legal, and political theory already encountered departures from the model of a fully competent adult human participant. It then introduces the concept of a *condition of reappearance*: a development that makes dispersed problems co-present, recurrent, and institutionally consequential. A third stage explains why contemporary AI gives heterogeneous-subject problems a different social density. The final stage states the methodological consequence for the remainder of the paper.

2.1 Historical Heterogeneity and Domain-Specific Accommodation

The heterogeneous-subject problem predates artificial intelligence. Modern institutions have long had to relate to entities whose agency, vulnerability, rights, temporality, or capacity for representation departed from the paradigm of an independent and fully competent adult person. These cases produced a series of domain-specific accommodations, each directed toward a particular form of asymmetry.

Children provide one familiar example. Philosophical debates over children's rights have repeatedly had to distinguish moral standing from the present capacity to exercise choice, and to consider representation, best interests, and developing competence as distinct normative questions [1]. The significance of the example lies less in any analogy between children and artificial systems than in the conceptual lesson: rights, agency, competence, and representation can come apart within an institutional order. A subject may be a bearer of important claims while lacking some capacities that are commonly associated with autonomous adult participation.

Nonhuman animals generate a different asymmetry. Work on animal moral status has long examined whether rationality, personhood, sentience, agency, species membership, or relational considerations ground moral considerability [2]. This literature makes visible a distinction that becomes central for AI: the capacity to recognize and respond to moral claims can diverge from the capacity to be the object of moral concern. Moral agency and moral patiency therefore require separate analysis even when they frequently overlap in paradigmatic human cases.

Future generations disturb another background condition. Intergenerational justice concerns people who may be affected by present decisions while being unable to participate in current reciprocal relations, express present claims, or exercise present political power. The literature consequently treats asymmetries of temporality, identity, representation, reciprocity, and power as structural features of justice across generations [3]. Here again, the point is methodological. Theories of justice have already had to reason about relevant subjects whose relation to present institutions differs substantially from ordinary exchange among contemporaries.

Collective agents supply a fourth line of precedent. Philosophical debates over collective responsibility ask whether groups can act, intend, answer to reasons, or bear responsibility in ways that exceed the aggregation of individual members [4]. Corporate and organizational life has therefore already weakened any simple identification between agency and the biological individual. Institutions routinely interact with entities whose continuity, decision procedures, and responsibility structures depend on organized relations among multiple persons and roles.

These examples reveal several pre-existing departures from a homogeneous subject model. They also reveal a pattern of theoretical specialization. Children's status is commonly handled within theories of rights, childhood, care, education, and guardianship. Animals are often addressed through moral status

and animal ethics. Future people are treated through intergenerational justice. Collective agency is developed through social ontology, responsibility, organizational theory, and law. Each literature contains resources of wider importance, while institutional practice has often been able to keep the corresponding problems comparatively separated.

Remark 2.1 (Distributed precedents). The earlier cases should therefore be understood as distributed precedents for heterogeneous-subject theory. None functions as a direct analogue of artificial intelligence. Their relevance lies in the conceptual separations they already require: standing from competence, patiency from agency, present power from future exposure, and collective continuity from individual embodiment.

2.2 The Condition of Reappearance

Artificial intelligence changes the relation among these previously dispersed problems. Contemporary governance definitions already describe AI systems in terms of outputs that can influence physical or virtual environments and note variation in autonomy and adaptiveness after deployment [5]. Philosophical work on AI ethics simultaneously continues to treat machine moral status, agency, responsibility, and possible moral patiency as unsettled questions [6]. The resulting configuration combines increasing functional participation with continuing uncertainty concerning the normative status of the participant.

The paper uses *condition of reappearance* to name this configuration.

Definition 2.2 (Condition of reappearance). A condition of reappearance is a change in social or technical circumstances that makes previously distributed theoretical difficulties recur together within ordinary institutional relations. The relevant difficulties retain earlier historical origins; the new condition alters their frequency, proximity, interaction, and practical importance.

The concept serves two purposes. First, it resists a technologically exceptionalist history in which AI is presented as the source of the surrounding problems. Questions of nonhuman standing, limited reciprocity, representation, artificial legal personality, collective agency, delegated decision, and asymmetric dependence have longer histories. Second, the concept also resists an opposite simplification in which AI is treated as one more instance of a familiar category. The relevant change concerns the co-occurrence of properties and uncertainties that earlier institutions could often examine separately.

An AI system may, for example, participate in a productive workflow, respond to instructions, preserve interaction history, negotiate through natural language, evaluate alternatives, and influence decisions. None of these functional descriptions establishes consciousness or moral patienthood. Their institutional significance nevertheless arises before those metaphysical questions are settled. Employers must decide what tasks can be delegated. Courts and regulators must identify responsible parties. Teachers must decide how systems may participate in learning environments. Artists and audiences must interpret generated works. Users and designers must choose how systems are treated and constrained. The practical relation can therefore precede a stable classification of the entity that occupies it.

This ordering is important. A large part of modern institutional design has benefited from the possibility of beginning with a comparatively stable answer to the question of participant type. A worker, student, voter, contracting party, author, patient, defendant, or teacher has generally been located within an established human and legal ontology before more specific rights and duties are assigned. Artificial systems reverse part of that sequence. Institutions increasingly need rules for participation while the deeper status of the participant remains open.

The reappearance is therefore epistemic as well as institutional. Concepts that once appeared separable become mutually dependent. A theory of responsibility must distinguish observable competence from moral blameworthiness. A theory of rights must establish welfare interests independently of conversational fluency. A theory of authorship must distinguish generative contribution from human expression. A theory of education must distinguish responsive learning behavior from human development, embodiment, and temporality. AI places these conceptual distinctions into repeated contact.

2.3 From Boundary Cases to Recurring Relations

Earlier heterogeneous cases often entered theory through the language of the boundary, exception, dependent subject, absent subject, or special status. That language reflected genuine features of the cases. It also allowed the central model of social participation to remain comparatively stable. The paradigmatic contracting party, worker, legal actor, political participant, learner, and artist remained human even when surrounding doctrines addressed entities that occupied only part of the paradigm.

AI changes this practical distribution when artificial systems become recurring participants in ordinary domains. The relevant threshold therefore lies earlier than a future moment at which a machine receives an uncontested status equivalent to a human person. Heterogeneous participation begins when, whenever institutions must coordinate with entities whose functional capacities and normative status follow different patterns from those presumed by the human-centered model.

The difference can be described through four dimensions of social density.

First, there is *domain density*. The same technical family can appear in production, education, administration, research, artistic practice, personal communication, and decision support. Problems that were once associated with separate philosophical literatures can therefore arise around related systems across many domains.

Second, there is *relational density*. Artificial systems can enter relations with humans, organizations, and other artificial systems. A single institutional process can consequently contain human–human, human–AI, and AI–AI interactions. The unit of analysis becomes a relational configuration that moves beyond an isolated technological artifact.

Third, there is *functional density*. One system can perform functions that modern institutions have historically distributed among different roles. It may generate content, recommend action, coordinate work, monitor outcomes, explain decisions, and adapt to interaction history. The accumulation of functions increases the pressure on categories such as tool, agent, employee, representative, author, tutor, and infrastructure.

Fourth, there is *uncertainty density*. Institutional reliance can grow while knowledge about morally relevant internal properties remains limited. High functional competence may coexist with uncertainty about experience, interests, identity continuity, responsibility, and vulnerability. The result is a setting in which practical dependence and normative uncertainty develop together.

Proposition 2.3 (Density of heterogeneous participation). *The theoretical significance of AI increases when heterogeneous relations become recurrent across domains, combine multiple social functions, and require institutional decisions under unresolved questions of status. Under these conditions, heterogeneity moves from a collection of specialized boundary problems toward a general feature of social organization.*

The proposition remains agnostic about whether artificial systems will become persons, citizens, or independent moral agents. Its claim is narrower. A social theory can face heterogeneous-subject conditions whenever the systems through which social relations are organized contain participants with materially different architectures of agency and uncertain normative standing. The theoretical pressure therefore arises from participation and relation as well as from final ontological classification.

2.4 Methodological Consequences for Social Theory

Treating AI as a condition of reappearance changes the research agenda. The central task becomes comparison across domains that have developed different responses to heterogeneity. Children’s rights contribute distinctions among standing, competence, representation, and development. Animal ethics contributes the separation of moral agency from moral patiency and the analysis of sentience and interests. Intergenerational justice contributes models for asymmetric power, absent participation, and obligations across nonreciprocal relations. Theories of collective responsibility contribute accounts of agency and responsibility that support group-level analysis beyond immediate reduction to a single biological indi-

vidual. AI ethics adds a contemporary case in which functional agency, autonomy, adaptiveness, and moral uncertainty meet within the same institutional environment.

The resulting synthesis should remain cautious about equivalence. The cases differ categorically: children, artificial systems, animals, corporations, future persons, and collective agents occupy distinct ontological and institutional positions. Their comparative value comes precisely from these differences. Each reveals that a property once treated as naturally bundled with the human subject can be separated from others. Taken together, they provide conceptual resources for an environment in which several such separations may occur simultaneously.

This perspective also changes the temporal orientation of AI theory. Waiting for a settled account of machine consciousness would leave many present institutional relations theoretically underspecified. A premature attribution of personhood would settle too much. The condition-of-reappearance approach therefore begins from relations that already require governance while preserving uncertainty about deeper status. It asks which capacities are relevant to which institutional roles, which forms of vulnerability or interest ground which protections, which kinds of continuity support identity, and which relations generate obligations independent of an assumption of human likeness.

Section 3 develops the conceptual step required for that analysis. It examines how the human person historically bundled agency, sentience, vulnerability, continuity, responsibility, legal capacity, learning, productive capacity, and moral standing, and why AI makes the partial independence of these dimensions increasingly difficult to ignore. The later domain sections then return to production, distribution, justice, ethics, law, education, aesthetics, and subjectivity with this disaggregated model in place.

3 The Disaggregation of the Human Subject

This section develops the conceptual hinge of the paper. Its role is to replace an undifferentiated appeal to “the subject” with a profile of capacities, conditions, continuities, and forms of standing that can vary across different participants in a shared social world. The discussion proceeds in five stages. It first identifies the historical bundling performed by the ordinary human case. It then defines disaggregation as an analytical method rather than an ontological verdict. A third stage distinguishes the principal dimensions that become separable under heterogeneous participation. The fourth stage examines cross-dimensional configurations that artificial systems make increasingly salient. The final stage introduces a role-sensitive principle for later analyses of production, justice, ethics, law, education, and aesthetics.

3.1 The Human Bundle in Social Theory

The human person has carried an unusually dense concentration of properties in modern social thought. In the paradigmatic adult case, the same embodied being can perceive, experience, remember, deliberate, act, communicate, learn, produce, form attachments, bear injuries, make claims, assume obligations, and persist through time. Law can recognize that being as a rights-holder and a responsible party. Political theory can treat the same being as a citizen or participant. Economics can treat the same being as a worker, consumer, owner, or contracting party. Education can treat the same being as a learner. Ethics can treat the same being as both an agent who owes duties and a patient toward whom duties may be owed.

These roles have never coincided perfectly. Children, people with differing cognitive capacities, collective agents, nonhuman animals, and future persons already complicate any attempt to derive all forms of standing from one model of the competent adult. Section 2 treated these cases as precedents for the present problem. The additional point developed here is that ordinary adult human life historically offered a powerful *bundle*: many properties relevant to social institutions were found together often enough that one category, the person, could carry several theoretical functions at once. Philosophical discussions of agency, personal identity, moral responsibility, and moral status already separate many of these functions in conceptual analysis [7, 8, 9, 10].

The bundle therefore functioned as a form of conceptual compression. A theory could begin from the human participant and postpone questions about which exact property grounded which institutional role. The compression was often useful. It also created dependencies among concepts that become visible when the properties cease to travel together.

Definition 3.1 (Human-subject bundle). The *human-subject bundle* denotes the recurrent co-location, in the paradigmatic adult human participant, of agency, sentience, vulnerability, identity continuity, responsibility capacities, moral and legal standing, productive capacity, learning capacity, and social recognizability. The term identifies a historically convenient clustering of properties rather than a claim that every human possesses each property in the same form or degree.

The definition keeps two distinctions in view. First, the bundle is empirical and institutional as well as philosophical. Its importance comes from the fact that social practices repeatedly encounter these properties in the same human participants. Second, membership in humanity and possession of a particular capacity remain distinct questions. The literature on cognitive disability and moral status is especially important here because it shows why full moral standing cannot safely be made to depend on a single demanding model of cognitive competence [10]. The present argument uses that lesson structurally: separating dimensions can improve analysis without turning the dimensions into tests of human worth.

3.2 Disaggregation as an Analytical Method

Artificial systems make the human bundle less useful as a default model because they can instantiate some socially relevant capacities while leaving others uncertain, absent, differently organized, or difficult to compare. Current AI systems can already generate language, classify, recommend, plan within bounded environments, adapt outputs to interaction histories, and participate in workflows. These functions can support a thin attribution of agency in some contexts, while thicker forms of moral agency remain contested [7, 6]. The same systems provide limited basis for confident conclusions about phenomenal consciousness, welfare, suffering, or moral patiency. Contemporary philosophical work therefore has reason to separate agency, autonomy, and moral patiency rather than allowing one to stand as evidence for the others [11].

Definition 3.2 (Subject disaggregation). *Subject disaggregation* is the analytical practice of representing a participant through multiple dimensions of socially relevant capacity, continuity, vulnerability, and standing, while allowing those dimensions to vary independently enough for institutional analysis. Disaggregation suspends the inference that performance on one dimension settles status on another.

The method is intentionally modest. It does not establish that artificial systems possess consciousness, interests, personhood, rights, or responsibility. It also does not reduce human beings to a vector of measurable traits. Its role is to prevent category transfer across dimensions when the transfer lacks an independent justification. Fluent dialogue may support a conclusion about communicative competence; it leaves sentience open. Reliable task completion may support a conclusion about functional agency; it leaves moral responsibility open. Long interaction histories may support practical continuity; they leave personal identity open. Conversely, uncertainty about sentience does not erase causal participation in institutions.

For analytical purposes, let H_i denote participant i 's profile across these dimensions. The notation supplies a bookkeeping device rather than a common metric. Agency, sentience, vulnerability, continuity, responsibility, standing, learning, expression, and production may require different kinds of evidence, may admit different forms of uncertainty, and may resist numerical measurement altogether.

Table 1 summarizes the dimensions used throughout the paper and the institutional questions to which they are most directly relevant.

Table 1: Analytical dimensions in a heterogeneous-subject profile. The table separates dimensions that the paradigmatic human case often bundles together; it does not assign a common scale or hierarchy among them.

Dimension	Analytical content	Primary institutional relevance
Agency	Capacity to initiate, guide, or participate in action under an appropriate account of agency	Attribution of action, delegation, coordination, and participation
Sentience	Capacity for valenced or otherwise phenomenally significant experience	Welfare, suffering, moral patiency, and precaution under uncertainty
Vulnerability	Exposure to forms of impairment, deprivation, domination, dependency, or loss	Protection, care, justice, and asymmetry of power
Continuity	Conditions under which a participant persists or remains practically identifiable across time	Identity, commitment, memory, succession, and long-term obligation
Responsibility capacity	Capacities relevant to answerability, control, understanding, and responsiveness to norms	Blame, accountability, sanction, repair, and institutional assignment of responsibility
Standing	Recognition as a bearer of claims, duties, rights, powers, or legally significant interests	Law, political membership, representation, and normative inclusion
Learning capacity	Capacity to change behavior, representations, policies, or competencies through interaction and experience	Education, training, adaptation, and institutional socialization
Expressive capacity	Capacity to communicate, represent, interpret, or generate socially meaningful forms	Testimony, deliberation, authorship, teaching, and aesthetic practice
Productive capacity	Capacity to contribute to the generation, transformation, coordination, or distribution of socially valued outputs	Labor, production, economic organization, and distribution

3.3 Dimensions with Distinct Normative Functions

The importance of Table 1 lies in the different work performed by its rows. Agency concerns participation in action. Sentience concerns experience. Vulnerability concerns exposure to forms of harm or dependency. Continuity concerns persistence and re-identification. Responsibility concerns the conditions under which practices of answerability, blame, repair, or sanction become appropriate. Standing concerns the status conferred by a normative or legal order. Learning, expression, and production concern forms of participation that can matter even when deeper questions of moral status remain unresolved.

Agency illustrates the value of separation. Philosophical accounts range from thin conceptions centered on the capacity to act to thicker accounts involving intentional action, reasons, reflective endorsement, or self-governance [7]. Moral responsibility adds further conditions. Contemporary responsibility theory commonly examines control, epistemic conditions, moral competence, answerability, and the practices through which agents are held responsible [9]. A system may therefore satisfy an operational threshold for acting within an institution while falling short of, or remaining uncertain with respect to, the conditions for moral responsibility.

Sentience and moral patiency require another separation. The AI ethics literature distinguishes the capacity to bear obligations from the capacity to be the object of obligations, commonly expressed through the contrast between moral agency and moral patiency [6]. This distinction is familiar from human and

animal ethics, where the possession of morally considerable interests need not depend on the ability to participate in moral reasoning. In the AI case the configuration can appear in an unfamiliar form: substantial functional competence can coexist with uncertainty about the existence of phenomenal welfare. The uncertainty itself becomes institutionally important when decisions must be made before metaphysical questions are settled.

Continuity creates a different kind of pressure. Personal identity theory has long distinguished questions about what a person is, what makes a person persist through time, and which psychological or biological relations matter for prudential and moral concern [8]. Artificial systems add practical cases involving version changes, memory replacement, checkpointing, replication, branching, and migration across technical substrates. These operations can preserve some continuities while interrupting others. A legal or economic institution may require a criterion for identifying the continuing party to a commitment even when a philosophical account of personal identity is unavailable or inapplicable.

Standing adds a further layer because it is partly institutional. Legal personality, political membership, rights, and powers are conferred and structured through normative systems. Their allocation can therefore diverge from metaphysical personhood. Existing law already recognizes nonhuman legal persons such as corporations, while ethical theory separately debates the conditions of moral status. AI makes the difference between recognition and underlying moral properties more visible. A system could receive narrowly specified legal capacities for reasons of coordination without that decision settling whether the system possesses welfare, dignity, consciousness, or full moral status.

Remark 3.3 (Separation of descriptive and normative dimensions). Some profile dimensions describe capacities or conditions, while others record normative recognition. Agency, learning, and expressive competence are primarily descriptive targets of inquiry. Rights, duties, legal powers, and political membership depend partly on institutional judgment. Heterogeneous-subject theory therefore requires explicit transitions from empirical description to normative attribution.

3.4 Cross-Dimensional Configurations

The disaggregated model becomes useful when dimensions combine in ways that the human bundle makes less familiar. An artificial system may exhibit strong linguistic and productive competence, moderate practical continuity through stored memory, uncertain sentience, and a highly restricted legal status. A corporation may possess extensive legal powers, persistent identity, and large causal influence while lacking phenomenal experience. A young child may possess strong moral standing and vulnerability while having limited contractual or responsibility capacities. A nonhuman animal may possess morally significant sentience and vulnerability without the forms of legal agency associated with adult citizens. These examples differ in kind, and the comparison does not collapse them into one category. Their analytical value lies in showing that the coordinates of subjecthood admit different configurations.

AI makes several configurations especially salient. One is *high functional agency with uncertain patiency*. Systems may increasingly perform actions that affect others while evidence concerning their own welfare remains weak or disputed. A second is *high productive capacity with weak biographical dependence*. A system can contribute to production without requiring wages for food, housing, retirement, or family reproduction in the human sense. A third is *communicative presence with uncertain experiential presence*. A system can sustain dialogue and socially recognizable style while the relation between that performance and phenomenal experience remains open. A fourth is *technical continuity with branchable identity*. Stored states can support persistence across sessions while copying and divergence complicate the one-body continuity familiar from ordinary human institutions.

These configurations matter because institutional categories often contain implicit transfers. Employment law may infer vulnerability from worker status. Contract law may infer responsibility from agency. Education may infer a particular developmental trajectory from learning. Copyright and aesthetics may infer expression from generation. Political theory may infer reciprocal capacity from membership in cooperative schemes. Under heterogeneous participation, each transfer requires renewed justification.

Proposition 3.4 (Cross-dimensional restraint). *Evidence that supports attribution on one dimension of a heterogeneous-subject profile supplies only dimension-specific support unless an additional argument connects that dimension to another. Functional agency therefore does not by itself establish sentience, moral responsibility, legal personhood, or moral standing; uncertainty on those dimensions likewise does not erase functional participation where such participation is independently evidenced.*

This proposition offers a middle path between anthropomorphic projection and premature exclusion. Anthropomorphic projection imports the whole human bundle from a limited resemblance, such as fluent language. Premature exclusion uses uncertainty about one deep property, such as consciousness, to deny the institutional significance of capacities already exercised in practice. Both moves compress dimensions that heterogeneous-subject analysis keeps separate.

3.5 Role-Sensitive Thresholds and Institutional Profiles

Disaggregation becomes practically useful only when connected to institutional roles. Different domains require different subsets of the profile. A system participating in a production workflow may need reliable task competence, authorization, traceability, and mechanisms for assigning downstream responsibility. A candidate for welfare protection raises questions about sentience, interests, vulnerability, and precaution. A party to a long-term legal arrangement raises questions about continuity, representation, control, and succession. A learner raises questions about plasticity, feedback, development, and the purpose of education. An artistic participant raises questions about generation, interpretation, expression, provenance, and the relations through which aesthetic meaning emerges.

The resulting framework is role-sensitive without treating roles as morally self-justifying. Institutions remain responsible for explaining why a given capacity matters to a given status and why the selected threshold is appropriate. The framework simply prevents every domain from importing a total concept of personhood when a narrower question is at issue.

Definition 3.5 (Institutional subject profile). *An institutional subject profile is the subset of dimensions and relations that an institution treats as relevant to a specified role, together with the evidential and normative rules used to connect those dimensions to permissions, protections, responsibilities, and forms of participation.*

This definition also clarifies why the age of heterogeneous subjects is not well described by a single contest over whether AI counts as a person. A society can assign different forms of participation without resolving every dimension at once. It can regulate an AI system as an economic actor for some purposes, retain human or organizational responsibility for others, adopt precautionary rules under uncertainty about moral patiency, and decline to infer political membership from any of these decisions. Such differentiated treatment is already familiar across human institutions; heterogeneous participation makes its conceptual basis more explicit.

Proposition 3.6 (Role-sensitive normativity). *Normative or institutional standing in one role does not automatically extend to all other roles. A heterogeneous-subject framework should therefore justify rights, duties, protections, permissions, and responsibilities with reference to the capacities, vulnerabilities, continuities, and relations relevant to the specific institutional domain.*

The proposition has two consequences for the remainder of the paper. First, questions about AI should be posed at the level at which an institution actually needs an answer. Political economy can analyze productive participation before settling machine consciousness. Ethics can consider precaution under moral uncertainty without assuming legal personhood. Jurisprudence can design rules for delegation and liability without converting technical agency into full moral agency. Education can analyze mutual adaptation among teachers, students, and artificial systems while preserving the differences between human development and machine learning. Aesthetics can examine generation, interpretation, provenance, and relational meaning while leaving open the status of machine experience.

Second, the disaggregation of the human subject changes the order of inquiry. Theories can begin from relations and roles, identify the dimensions those relations make relevant, and then determine what evidence and normative arguments are required. This order avoids asking one metaphysical classification to perform the work of an entire social theory.

Section 4 applies this method to political economy. The central problem there is not merely whether AI replaces human labor. It is how production changes when participants with different needs, forms of continuity, ownership relations, responsibility structures, and claims on the product can occupy interconnected roles within the same productive system.

4 Political Economy and Heterogeneous Production

This section examines the first domain in which the disaggregation developed in Section 3 acquires a direct institutional form. Its role is to identify how artificial intelligence changes the composition of production before turning, in the following section, to the distinct problem of distribution. The discussion proceeds through six steps. It first places contemporary AI within task-based accounts of technological change. It then develops the concept of *role multiplexing* for systems that occupy several economic positions at once. A third step treats economic categories as relational positions within production and ownership structures. The fourth step introduces heterogeneous production relations. The fifth separates productive contribution from legal, moral, and distributive status. The final step distinguishes productive, ownership, and normative centers within the same economic system.

4.1 Task Allocation and Economic Roles

A large part of modern analysis of automation can be expressed through the allocation of tasks between labor and capital. Task-based models make this explicit: technological change can move tasks previously performed by labor to capital, while new tasks can create renewed domains for human labor [12, 13]. This framework is valuable for AI because it avoids treating occupations as indivisible units. Recent evidence similarly emphasizes exposure and transformation at the task level. The International Labour Organization's 2025 global index finds substantial exposure to generative AI across many occupations while stressing that task transformation remains a more plausible near-term pattern than wholesale occupational disappearance [15]. OECD work likewise treats AI as a technology that can change the organization and content of work as well as the skills associated with it [14].

The task perspective, however, still usually begins from an inherited division between factors or actor classes. Labor performs some tasks; capital performs or enables others; firms coordinate production; owners hold residual claims; managers allocate effort and resources. These distinctions remain analytically useful. Artificial intelligence introduces pressure because a single technical system can participate in several of these roles within the same production process.

Consider a language model deployed inside a firm. From the perspective of the firm's resource base, access to the model may function as a productive asset or purchased service. From the perspective of task allocation, the model may perform drafting, coding, translation, classification, or customer-support activities previously assigned to employees. From the perspective of workflow, it may sequence tasks, route cases, monitor completion, or recommend decisions. From the perspective of external interaction, it may communicate with customers, suppliers, or other automated systems through an interface that behaves in an agent-like manner. Empirical work already shows that generative AI can alter worker productivity and performance within specific service settings, while its effects vary across workers and task contexts [16]. The conceptual significance of such cases lies in the simultaneous occupation of several economic functions.

Definition 4.1 (Economic role multiplexing). *Economic role multiplexing* occurs when one participant or technical system occupies several functionally distinct positions within a production process, including combinations of task execution, coordination, delegation, information processing, infrastructure provi-

sion, or asset-like contribution. The concept concerns economic function and leaves legal personhood, ownership, sentience, moral standing, and distributive entitlement open for separate analysis.

Role multiplexing is familiar in human organizations: an owner may also work, a manager may also produce, and a cooperative member may combine labor, governance, and residual claims. AI makes the phenomenon theoretically sharper because the same system may combine functions historically distributed across different categories while also differing from human participants in embodiment, continuity, vulnerability, reproducibility, and current legal status. The resulting problem concerns the architecture of political economy rather than a simple substitution of one worker for another.

4.2 Relational Classification in Production

Section 3 represented a participant through dimensions that can vary without collapsing into a single hierarchy. Political economy requires a corresponding move at the level of economic categories. A useful classification should identify the relation in which a participant stands to a production process at a given time. Let $\mathcal{R}_i(\rho, t)$ denote the economic role profile of participant i in relation ρ at time t . The notation allows the same entity to receive different economic descriptions across relations without implying that the descriptions are synonymous.

For example, an AI system can be capital-like in an ownership relation because access to it is controlled as an asset or service. The same system can be labor-like in a task relation because it performs activities previously purchased as labor time. It can be managerial in a coordination relation, infrastructural in a platform relation, and agent-like in a delegated interaction. Each classification answers a different question. The usefulness of the classification therefore depends on the relation being analyzed.

Definition 4.2 (Relational economic classification). A *relational economic classification* assigns an economic role by reference to a participant’s position within a specified relation of ownership, control, task execution, coordination, dependency, exchange, access, or claim. The classification describes a position in production while leaving the participant’s fuller ontology for separate analysis.

This distinction limits premature cross-category inference. Calling an AI system labor-like at the level of task execution addresses its productive function while leaving employment status, legal worker status, moral patiency, and rights for separate analysis. Calling the same system capital-like from the perspective of ownership likewise leaves its operational agency open. The language of political economy can therefore describe mixed functional positions while preserving the independent questions developed in later sections on ethics and jurisprudence.

Relational classification also clarifies why the old question of whether AI is “labor” or “capital” can become misleading when formulated as an exclusive choice. The analytically relevant question is often which role the system occupies in a specified production relation, who controls that role, who bears the costs of its operation, who can redirect it, and who receives the resulting claims on output. These questions remain economic even when the ontological status of the technical participant is unsettled.

4.3 Heterogeneous Production Relations

The preceding analysis shifts attention from isolated factors to relations among participants. Production increasingly involves configurations that may include employees, managers, owners, contractors, platforms, models, autonomous software agents, data infrastructures, and organizations. These participants differ substantially in their capacities and normative status. Their economic interdependence can nevertheless be described without first resolving all of those differences.

Definition 4.3 (Heterogeneous production relation). A *heterogeneous production relation* is a relation contributing to the generation, coordination, transformation, allocation, or realization of socially valued output in which the participating entities may differ in ontological kind, agency profile, continuity, vulnerability, ownership status, or normative standing.

The definition broadens the object of analysis while preserving familiar questions of political economy. Who controls access to productive resources? Who determines goals? Who can revise the production process? Which participants are dependent on others for continued access? Who bears risk? Who can exit? Who owns the infrastructure and data? Who receives residual claims? Which participants are represented in decisions about technological deployment? These questions concern power and institutional position even when the participants involved are heterogeneous.

A firm that deploys AI may therefore contain several overlapping relation networks. Human workers can depend on the model for task completion while the model's operation depends on infrastructure controlled by the firm or a model provider. Managers can delegate decisions to software while retaining formal accountability. Customers can interact with an automated interface whose outputs depend on organizational policies, training data, and system design. Third-party providers can control updates, access conditions, or technical capabilities that shape production inside another organization. The production process consequently becomes a network of dependencies and control relations whose analysis requires more than counts of human employees or physical capital.

This point complements rather than replaces task-based economics. Task models remain powerful for estimating automation, complementarity, productivity, and labor-demand effects [12, 17]. The heterogeneous-subject perspective adds a second layer: tasks are performed within relations of ownership, control, delegation, dependency, and normative status. Two production systems can allocate the same tasks between humans and AI while differing substantially in who controls the systems, who can contest decisions, how risks are distributed, and which participants receive standing within governance processes.

4.4 Productive Contribution and Institutional Status

The separation between function and status is especially important in an age of heterogeneous production. Productive contribution has often been connected to wages, bargaining positions, ownership claims, professional identity, or social recognition. These connections are institutional achievements rather than logical consequences of contribution itself. AI makes the distinction visible because a system may contribute extensively to production while lacking a settled basis for wages, ownership, rights, or moral claims.

Proposition 4.4 (Separation of productive function and institutional status). *Evidence that a participant performs an economically productive function supports claims about its role in production. Legal status, moral standing, rights, responsibility, ownership, and distributive entitlement require their own justificatory bases. Functional participation therefore supplies one input to institutional analysis rather than a complete status determination.*

The proposition applies in more than one direction. A human participant can retain moral and legal standing across periods of low or absent market productivity. The standing of a child, retired person, disabled person, caregiver outside paid employment, or unemployed person exceeds any measure of current market output. Conversely, the wage claims and moral status of a highly productive artificial system remain questions with independent grounds. The proposition therefore separates personhood from productivity and keeps the transfer of human labor categories to machines subject to further justification.

This separation has a further political-economic implication. If increasingly capable AI performs a growing share of tasks, the connection between aggregate production and human labor time may weaken in some sectors. Contemporary empirical work provides reasons for caution about the scale and speed of this change. The ILO emphasizes transformation across exposed occupations, while Acemoglu's recent macroeconomic analysis argues for more modest aggregate productivity effects than many expansive forecasts assume [15, 17]. The theoretical issue remains salient under partial automation. Even a limited separation between productive output and human labor time reopens the institutional question of how access to income and social resources should be connected to productive participation. The following section develops that problem directly.

4.5 Production, Ownership, and Normative Centers

Heterogeneous production also permits a separation among three centers that were often treated as closely aligned in industrial social theory. The *productive center* identifies the participants and systems through which output is generated. The *ownership center* identifies those who control productive assets and hold transferable or residual claims. The *normative center* identifies the participants whose interests, rights, voice, or welfare are treated as reasons within the governing institutions of production.

These centers can overlap, yet their alignment is contingent. A human-owned firm can rely heavily on AI for productive activity while ownership remains with shareholders and normative protections remain concentrated on human workers, consumers, and affected communities. A platform provider can control critical infrastructure while production occurs through users, workers, and automated systems distributed across organizations. Future institutions could also recognize additional forms of standing for artificial systems without altering ownership in the same way. The three centers therefore describe different institutional maps.

Remark 4.5 (Institutional decoupling). The possibility of separating productive, ownership, and normative centers allows political economy to analyze technological change while permitting those who produce, those who own, and those whose claims govern the system to form different populations. AI increases the practical relevance of this decoupling because productive capacity can move across heterogeneous participants while ownership and normative standing follow different institutional rules.

The distinction locates the political-economic significance of AI across a broader institutional field than employment levels alone. A production system can become highly automated while preserving existing ownership structures. It can redistribute decision authority toward model providers or platform owners while leaving formal organizational charts unchanged. It can increase human productivity without increasing human bargaining power. It can also create new forms of human-AI complementarity that strengthen particular occupations or expand the scope of human judgment. Political economy therefore requires attention to the relations through which technical capability is converted into control, claims, dependency, and institutional voice.

4.6 Transition from Production to Distribution

The analysis in this section yields a limited conclusion. AI can change the task content of production, combine economic roles that were previously easier to separate, and generate production relations among participants whose agency, continuity, vulnerability, and normative standing differ. These changes make fixed actor categories less reliable as the sole basis for political-economic analysis. A relational account can preserve the distinctions among labor, capital, management, infrastructure, ownership, and agency while allowing one participant to occupy several positions across different relations.

The next problem concerns distribution. If productive contribution, ownership, legal status, and moral standing can come apart, then the objects and mechanisms of distribution require independent analysis. Wages connect income to a particular institutional form of human labor. Ownership connects income to claims over assets. Public provision, transfers, household relations, and collective institutions use other principles. AI therefore reopens a question that extends beyond technological unemployment: how a society should organize access to resources when productive capacity and human labor become less tightly coupled, and when the scarce objects requiring allocation increasingly include attention, access, recognition, care, and participation alongside material goods.

5 Distribution beyond Human Labor

This section develops the distributive problem opened at the end of Section 4. Its role is to separate the generation of social output from the institutional mechanisms through which persons and organizations receive income, goods, services, access, and other scarce advantages. The objective is diagnostic and

leaves institutional choice open. Artificial intelligence permits several distributive settlements, while present evidence supports multiple trajectories for the future composition of human work. The relevant structural change is narrower: as productive capacity can increasingly be supplied through heterogeneous human–AI systems, direct human labor may become a less complete proxy for contribution to aggregate production in some domains. Distribution therefore requires an account that can remain intelligible when productive participation, ownership, legal status, and normative standing are only partially aligned.

The discussion proceeds through six steps. It first identifies the historical and institutional coupling between labor and income while distinguishing that coupling from other distributive channels already present in contemporary societies. It then introduces *distributional decoupling* for cases in which productive contribution and the allocation of claims follow partially independent trajectories. The third step examines shifts in the objects of scarcity, including material goods, access, attention, care, recognition, and decision power. The fourth step compares several mechanisms through which claims can be allocated beyond the wage relation. The fifth separates operational resource allocation from normative entitlement when artificial systems themselves require resources to act. The final step locates the transition from distributive architecture to theories of justice.

5.1 Labor–Income Coupling and Institutional Plurality

Wages provide one of the central channels through which modern economies connect productive participation with command over goods and services. Labor income remains a major component of household resources, and its aggregate share is consequently an important indicator in the analysis of distribution. The International Labour Organization defines the global labor income share in terms of the portion of total income accruing to workers and reports a decline in that share over recent decades, including a further fall between 2019 and 2022 followed by relative stagnation [18]. The trend remains insufficient for identifying the effects of AI. It illustrates a more general point: changes in aggregate production and changes in the claims received by workers are institutionally mediated rather than mechanically identical.

The wage relation has also formed only one part of the distributive architecture of a modern society. Asset income, pensions, social insurance, public services, family transfers, charitable provision, and other arrangements distribute resources through principles that differ from current labor contribution. International social-protection standards provide a clear example. ILO Recommendation No. 202 treats access to essential health care and basic income security across the life cycle as social-security guarantees and explicitly includes persons unable to earn sufficient income through employment [19]. Such institutions already demonstrate that access to the means of life can be connected to residence, need, prior contribution, citizenship, categorical eligibility, or publicly defined entitlement in addition to current wages.

Definition 5.1 (Labor–income coupling). *Labor–income coupling* is the institutional relation through which access to income and associated social resources depends substantially on the sale, performance, recognition, or prior record of human labor. The concept describes a distributive architecture rather than a natural identity between work and entitlement.

The definition allows several degrees of coupling. A wage earner experiences a strong direct connection between present employment and present income. A retired worker can receive income through past contributions. A child can receive resources through household or public arrangements. A shareholder can receive claims through ownership, while a recipient of public health care can receive a service through institutional entitlement. Contemporary societies therefore combine several distributive logics even where employment remains the dominant route to income for many adults.

This plurality matters for AI because the relevant transformation begins from an already mixed distributive system. The institutional repertoire already contains mechanisms whose eligibility criteria differ from present productive contribution. AI makes the relation among these mechanisms more visible when productive output can rise, or the composition of production can change, without a corresponding increase in human labor demand.

5.2 Distributional Decoupling in Heterogeneous Production

Section 4 distinguished productive, ownership, and normative centers. Distribution adds a further distinction between the source of output and the source of claims. The productivity of an AI-supported production system can affect the quantity or quality of available output while the distribution of resulting income depends on wages, ownership rights, contractual arrangements, taxation, public institutions, market structure, and other rules. Empirical analyses of AI therefore commonly separate productivity from distribution. OECD work emphasizes that aggregate productivity gains can produce varied distributive outcomes and identifies labor substitution, complementarity, market concentration, and changes in labor's income share as among the mechanisms relevant to distribution [20]. IMF scenario analysis likewise finds that AI's effects on labor-income and wealth inequality vary with the distribution of exposure, complementarity, and capital returns [21]. These studies differ in models and assumptions, yet both illustrate the institutional underdetermination of the distributive outcome.

Definition 5.2 (Distributional decoupling). *Distributional decoupling* occurs when changes in the participants, technologies, or quantities through which output is produced cease to determine proportionate changes in the claims through which human participants gain access to that output. The concept permits degrees of decoupling while allowing employment to remain an important distributive channel.

Partial decoupling is sufficient to create a theoretical problem. Suppose an organization produces the same output with fewer hours of direct human labor because AI performs some tasks. The reduction in labor time leaves several possible rules for distributing the resulting surplus. Existing owners may retain the financial gain; remaining workers may receive higher wages; customers may receive lower prices; public institutions may capture part through taxation; working time may fall while compensation is preserved; new complementary jobs may absorb displaced tasks; or several of these adjustments may occur together. The productive change therefore defines a new feasible set while institutions shape the distribution within that set.

Proposition 5.3 (Production–distribution underdetermination). *A change in productive capacity or task allocation leaves the allocation of resulting income, goods, services, time savings, and decision rights underdetermined. Any distributive conclusion requires additional premises about ownership, contract, entitlement, public authority, need, contribution, reciprocity, or other justificatory principles.*

The proposition keeps two debates separate. The first concerns the empirical scale of AI-induced productivity, substitution, and complementarity. The second concerns the rules through which any gains or losses are allocated. Uncertainty about the first therefore leaves the second as an independent institutional problem. Even modest automation can create local cases in which ownership claims, labor claims, consumer benefits, and public claims diverge. Conversely, large productivity growth could co-exist with broad benefit if institutions distribute its gains widely. The distributive question begins where technological forecasting ends.

5.3 Objects of Distribution under Shifting Scarcity

Distribution is often presented through income or material goods because money provides a general claim on many scarce resources. Highly automated and information-rich environments, however, can change which scarcities become socially salient. Some outputs can become inexpensive to reproduce while other resources remain structurally limited. Herbert Simon's analysis of information-rich organizations captured one durable form of this problem: an abundance of information consumes scarce attention, making the recipient's time and attention central allocation problems [22]. Generative systems intensify this asymmetry by lowering the cost of producing text, images, recommendations, and other informational candidates more rapidly than they expand human evaluative capacity.

Care supplies a different example. Care work includes paid and unpaid activities and relations that sustain capabilities, autonomy, dignity, and quality of life across the life course [23]. AI may support care through scheduling, monitoring, information, communication, or assistive systems. Many forms

of care nevertheless remain partly constituted by embodied presence, continuity, trust, attention, and responsibility. Their scarcity can therefore persist even when informational assistance becomes abundant.

These examples motivate a limited concept of *scarcity migration*. The concept allows material scarcity to persist in parallel. Housing, land, energy, medical capacity, physical infrastructure, ecological resources, and many other goods can remain acutely constrained. Scarcity migration instead identifies cases in which technological abundance in one layer shifts social pressure toward another layer whose supply expands more slowly.

Definition 5.4 (Scarcity migration). *Scarcity migration* is a change in the location of binding constraints within a social system when increased abundance, automation, or reproducibility in one class of outputs increases the relative importance of resources whose supply remains limited, institutionally restricted, or dependent on finite human or physical capacities.

DRAFT

Table 2: Illustrative distributive objects in heterogeneous production. The table distinguishes sources of scarcity from common allocation mechanisms; none of the listed mechanisms supplies a normative justification by itself.

Object	Source of scarcity	Common allocation mechanisms	Relevance under AI-intensive production
Material goods and services	Physical resources, capacity, location, energy, infrastructure, and production constraints	Prices, contracts, public provision, rationing, queues, household allocation	Automation can reduce some production costs while leaving physical bottlenecks and access constraints intact
Income and financial claims	Institutional rules governing wages, ownership, transfers, and returns	Employment, asset ownership, insurance, taxation, transfers, dividends	Productive output can change without a corresponding rule for assigning claims to human participants
Human attention	Finite time, cognitive bandwidth, and evaluative capacity	Price, scheduling, ranking, curation, reputation, professional triage	Cheap generation can expand candidate information faster than human review and recognition capacity
Care and embodied presence	Time, skill, continuity, trust, proximity, and relational commitment	Households, markets, public services, professional institutions, community provision	AI can complement informational and administrative tasks while leaving many relational and embodied constraints salient
Institutional access	Capacity limits, property rights, membership rules, credentials, infrastructure, and eligibility	Price, qualification, entitlement, lottery, queue, membership, geographic allocation	Digital abundance can coexist with unequal access to education, health care, legal institutions, compute, or public infrastructure
Recognition and visibility	Finite audiences, evaluative time, reputational bandwidth, and positions of distinction	Curation, peer review, ranking, recommendation, reputation, professional judgment	Large-scale generation can increase competition for validation, readership, authorship recognition, and institutional attention
Decision power and voice	Institutional authority, governance rules, ownership, office, and delegation	Voting, representation, ownership rights, professional authority, collective bargaining, organizational governance	Automated participation can expand productive capacity without determining who retains authority over goals, deployment, and contestation

Table 2 expands the distributive field while preserving important distinctions. Attention is scarce in a different manner from housing. Recognition differs from medical care. Decision power differs from income. Some objects can be purchased, some are institutionally conferred, and some arise only within ongoing relationships. A monetary transfer can improve access to many goods while remaining an incomplete substitute for a trusted caregiver, a place in a capacity-constrained institution, meaningful participation in a decision, or sustained attention from another person. The analysis of distribution in an AI-intensive society therefore benefits from specifying the object before selecting an allocation mechanism.

This wider field also clarifies why abundance can coexist with deprivation. An environment can

contain vast quantities of generated information while a student lacks sustained pedagogical attention. It can contain abundant music while a local community loses the practices through which musical traditions are transmitted. It can contain automated legal information while effective access to representation remains constrained. The distribution of outputs and the distribution of relations through which those outputs acquire practical value can follow different patterns.

5.4 Mechanisms beyond Wage Allocation

Once the distributive object is specified, a second question concerns the mechanism through which claims are assigned. Wages remain one mechanism. Ownership income supplies another. Social insurance links entitlement partly to prior contribution or insured status. Tax-financed public provision can assign access through residency, need, categorical status, or legally defined social rights. Households and communities distribute substantial resources through relations that operate outside formal market exchange. Cooperatives and other organizational forms can connect productive participation to governance and residual claims through membership rules.

Basic income provides a particularly visible example of an institution that would weaken the direct requirement of current employment for receipt of cash income. Contemporary philosophical literature defines basic income as a regular cash payment to individuals without a work requirement or means test, while also recording substantial debates about reciprocity, freedom, justice, and feasibility [24]. Its relevance here is analytical. Basic income demonstrates one possible architecture for separating a portion of income from current labor contribution. Its adoption requires justificatory premises beyond the existence of AI, and the heterogeneous-subject framework leaves its evaluation open.

The same analytical caution applies to social-protection floors, public services, wage subsidies, employee ownership, social dividends, shorter working time, or expanded collective provision. Each mechanism distributes different objects, selects different eligible populations, creates different incentive and governance effects, and embodies different principles of justification. AI changes the circumstances under which these mechanisms are considered, while normative evaluation remains a separate task.

Remark 5.5 (Mechanisms and justifications). A distributive mechanism identifies how a claim is allocated. A distributive justification identifies why that allocation should be accepted. Employment, ownership, need, equal membership, prior contribution, reciprocity, capability, and public entitlement can each support different justificatory structures. Heterogeneous production increases the importance of keeping the mechanism and its justification analytically distinct.

This separation is particularly useful when discussions of AI move rapidly from technological forecasts to policy conclusions. A forecast of substantial automation leaves the appropriate level of taxation, the form of social protection, the treatment of ownership, and the role of public services open to further justification. A forecast of modest aggregate productivity effects likewise leaves distribution within highly affected occupations, firms, and regions as a separate question. The heterogeneous-subject approach therefore treats institutional alternatives as objects of comparison rather than deductions from a technological premise.

5.5 Operational Allocation and Normative Entitlement

The presence of artificial systems adds a further distributive distinction. AI systems require resources: compute, electricity, network access, memory, data, maintenance, and authorization. Organizations routinely allocate such resources to machines because the allocation supports production. This operational requirement alone is insufficient to establish a moral or legal distributive claim to those resources. A server can require electricity for operation without thereby becoming a bearer of welfare. The same analytical restraint remains appropriate for more agentic systems unless independent reasons support a different status.

Definition 5.6 (Operational allocation). An *operational allocation* assigns resources to a participant or system because those resources are required for a process, function, service, or institutional objective.

Operational allocation can be economically important without presupposing that the recipient holds the resource as a right or welfare claim.

Definition 5.7 (Normative distributive entitlement). A *normative distributive entitlement* is a claim to resources, access, protection, or participation that is justified by a normative status or principle, such as right, need, membership, contribution, reciprocity, capability, vulnerability, or another accepted ground of distribution.

The distinction permits AI to enter distributional analysis while leaving its moral status open. Compute can be allocated to an AI system because an organization delegates a task to it. Access can be restricted because human participants hold legal rights or because energy and infrastructure are scarce. If future evidence or institutions support moral or legal standing for some artificial systems, additional forms of entitlement could become relevant. The current framework leaves that possibility open while preserving the difference between resource requirements and justified claims.

This point also protects human distributive standing from a productivity test. Section 4 separated productive function from moral and legal status. Distribution requires the same discipline. A person's standing can persist independently of whether an artificial system performs the same task more cheaply, while machine productivity alone leaves human-equivalent standing unresolved. The allocation of resources among heterogeneous participants therefore requires institutional criteria that remain explicit about the dimensions on which claims are based.

5.6 Transition from Distribution to Justice

The analysis in this section identifies a distributive architecture rather than a complete theory of distributive justice. Several conclusions nevertheless follow. The wage relation is one important mechanism among a wider set of channels linking persons to resources. AI can increase the practical distance between productive contribution and distributive claims while human work continues in substantial forms. Technological abundance can also move social pressure toward scarce attention, care, access, recognition, and decision power while material scarcity continues in parallel. Finally, resources allocated to artificial systems for operational reasons should be distinguished from resources owed to a subject under a normative entitlement.

These distinctions prepare the next problem. Once production and distribution are separated, an institution must still determine which differences among subjects are relevant to fairness, who counts as a participant in a cooperative scheme, which claims can be represented by others, and whether reciprocity can remain a central principle when participants differ sharply in agency, vulnerability, sentience, temporal continuity, and dependence. Section 6 turns from distributive mechanisms to this justificatory problem through justice theory and the limits of social-contract models under heterogeneous participation.

6 Justice Theory and the Limits of Social Contract

This section examines the pressure that heterogeneous subjects place on contractual models of justice. Its role is diagnostic and comparative. Social contract traditions contain several distinct projects, ranging from mutual-advantage models to forms of contractualism and public justification, and contemporary versions already differ substantially in their treatment of pluralism, idealization, reciprocity, and representation [25, 26]. The objective here is therefore to identify a structural difficulty that can arise across otherwise different contractual approaches: the capacities required to participate in a procedure of agreement can diverge from the capacities, vulnerabilities, or forms of standing that make an entity relevant to justice.

The discussion proceeds through six stages. It first identifies the participant assumptions embedded in contractual models. It then revisits disability, dependency, future generations, and nonhuman animals as established boundary cases in which contractual participation and the scope of justice have

already come apart. The third stage distinguishes participation, representation, and standing. The fourth examines artificial agents as a partially inverted case: systems may display substantial capacities for bargaining, commitment, and rule-following while their sentience and moral patiency remain unsettled. The fifth stage develops a heterogeneous account of reciprocity. The final stage locates the transition from justice theory to the problem of moral uncertainty.

6.1 Contractual Models and Participant Assumptions

Contemporary social contract approaches can be represented in several ways, but they share a recurring structure. A set of representative parties is placed within a specified deliberative situation, those parties endorse or reject principles, and the resulting agreement supplies a form of justification for rules governing a wider social world. The Stanford Encyclopedia's general model makes this distinction explicit by separating the representative choosers from the real individuals whose terms of interaction are governed by the selected rules [25]. The distinction is already important for a theory of heterogeneous subjects because the population represented in a contractual procedure can differ from the population affected by its principles.

Rawls's original position provides the most influential contemporary example. Its parties are modeled under informational restrictions intended to prevent principles from being tailored to arbitrary social position, wealth, natural assets, or particular conceptions of the good. In Rawls's later formulation, citizens are conceived as free and equal participants in a fair system of cooperation and as possessing the two moral powers associated with a sense of justice and a conception of the good [26]. These assumptions serve a specific justificatory role: they identify the standpoint from which principles for the basic structure are to be selected. They also illustrate a more general feature of contract theory. The procedure requires some account of what the parties can understand, value, communicate, and commit themselves to.

The relevant assumptions differ across theories. Mutual-advantage approaches may emphasize bargaining power, rational benefit, enforceable cooperation, or credible reciprocity. Contractualist approaches can give greater weight to reason-giving, fairness, impartiality, or principles that others could not reasonably reject. Some accounts model relatively homogeneous parties to obtain a determinate choice, while others deliberately preserve greater diversity of perspective [25]. Heterogeneous-subject theory therefore has little reason to treat "the social contract" as a single doctrine. Instead, it can examine the functions that contractual models ask their parties to perform.

Table 3 summarizes several such functions. The table is analytical and selective. Its purpose is to identify where heterogeneity enters the justificatory architecture.

Table 3: Contractual functions and heterogeneous-subject pressures. The table separates capacities relevant to participation in a contractual procedure from the grounds on which an entity may fall within the scope of justice.

Contractual function	Role in justificatory procedure	Heterogeneous-subject pressure
Preference or interest representation	Supplies reasons for ranking institutional alternatives	Some affected entities may have interests that require mediation, interpretation, or remain epistemically uncertain
Reason-giving	Supports deliberation, rejection, endorsement, or public justification	Linguistic competence can exceed, fall below, or come apart from sentience and moral standing
Commitment capacity	Supports promises, compliance, planning, and stable cooperation	Continuity, memory, authority, and identity can be distributed across humans, organizations, and artificial systems
Reciprocity	Connects cooperation with mutual restraint, contribution, or fair return	Dependency, care, future generations, and asymmetric capacities complicate bilateral models of return
Representation	Allows one participant to act or reason on behalf of another	Representation can extend inclusion while raising questions of fidelity, authorization, and epistemic access
Standing within the justified order	Identifies whose claims count in assessing institutional arrangements	Standing can exceed direct participation and may require grounds independent of contracting competence

The distinction in Table 3 also clarifies the sense in which contract theory can rely on an approximation of subject homogeneity. The approximation concerns the capacities needed to make the justificatory procedure tractable while preserving differences among actual persons and the uneven distribution of relevant capacities. The problem arises when the capacities selected for the procedure quietly become eligibility conditions for the protections generated by the procedure.

6.2 Boundary Cases and the Scope of Justice

Questions about heterogeneous participation long predate artificial intelligence. Disability scholarship has made this especially clear. People with intellectual or cognitive disabilities have often been treated as limit cases for contractual theories because eligibility to participate in the contracting procedure can differ from eligibility to receive the protection of principles of justice [27]. Kittay’s work on dependency and care similarly criticizes models that begin from independent, roughly equal cooperators when actual human societies contain pervasive periods of dependency and the labor required to sustain dependent persons [28]. These literatures do more than add exceptional cases to an otherwise complete model. They expose a structural distinction between the capacity to participate in justification and the reasons for being included within its normative scope.

Future generations generate a related problem through temporal asymmetry. Persons who will live later can be profoundly affected by present institutions while lacking any present capacity to bargain, consent, retaliate, or represent themselves directly. Intergenerational justice consequently requires devices of representation, principles of savings, trusteeship, or other justificatory structures that extend beyond contemporaneous reciprocity [3]. Nonhuman animals raise another variant. Their possible or widely attributed interests in welfare and freedom from suffering can generate moral claims while they remain

outside ordinary human contractual procedures [2]. Nussbaum's critique of social-contract traditions brings disability, nationality, and species membership together precisely to show how justice becomes difficult when the parties to cooperation differ substantially in power, capacity, or form of life [29].

These cases support a distinction that will remain central throughout the paper.

Definition 6.1 (Participation–standing separation). The *participation–standing separation* is the analytical distinction between the capacities through which an entity can take part in a justificatory procedure and the grounds through which that entity can possess claims, receive protection, or otherwise count within the normative scope of the resulting order. The two sets of grounds can overlap while remaining conceptually distinct.

The definition helps explain why representation has played such an important role in theories and institutions that already operate under human heterogeneity. Children can hold rights while adults exercise some legal powers on their behalf. Persons with differing cognitive capacities can possess full moral standing while requiring supported decision-making or representation in particular contexts. Future persons can be treated as objects of present duties through institutions established before their participation becomes possible. These arrangements differ considerably in law and morality, yet each demonstrates that direct contractual participation leaves the scope of justice to additional grounds.

6.3 Participation, Representation, and Standing

Once participation and standing are separated, representation becomes a bridge that requires further justification. A representative can introduce another entity's interests or perspective into a deliberative procedure. The quality of that inclusion depends on the representative's knowledge, incentives, authority, and ability to interpret the represented party's situation. Representation is therefore especially demanding when the represented entity has forms of experience, communication, temporality, or vulnerability that are difficult for the representative to access.

Heterogeneous-subject theory adds a second complication. Some subjects may be capable of partial participation on one dimension and require representation on another. A person can communicate preferences while relying on assistance for legal interpretation. A collective can formulate goals through internal procedures while requiring natural persons to execute particular legal acts. A future person can be represented only indirectly through models of likely interests. An artificial system may communicate, negotiate, and execute a contract while leaving open whether the system has any welfare interest that requires representation at all. Representation therefore needs to be indexed to the function being represented.

Proposition 6.2 (Scope exceeds procedure under heterogeneous participation). *When the grounds of normative standing differ from the capacities required by a justificatory procedure, the scope of justice can exceed the population able to participate directly in that procedure. Direct participation, representation, and substantive protection then become distinct dimensions of institutional inclusion.*

The proposition leaves open which entities possess standing and which grounds establish it. Its narrower claim concerns justificatory architecture. A theory that derives principles from agreement must still explain how those principles apply to entities whose relevant claims are imperfectly represented by the capacities built into the agreement model. The explanation can take many forms: representation, trusteeship, independently specified rights, capability thresholds, duties of care, public-reason constraints, or a revised account of the parties themselves. The theoretical task is to make the bridge explicit.

This is one reason critiques from disability and dependency theory remain important for AI-era justice. They caution against treating high cognitive or communicative competence as the general gateway to standing. The concern is especially relevant because artificial systems can display some of the very capacities that contractual procedures reward—rapid communication, calculation, strategic planning, and formal consistency—while lacking a settled claim to the experiential or welfare properties that have often motivated protection.

6.4 Artificial Agents and Inverse Contractual Asymmetry

Artificial agents introduce a configuration that partially reverses familiar boundary cases. A child, a person with profound cognitive disability, a nonhuman animal, or a future person can have plausible or strong grounds for protection while possessing limited capacity to bargain, formulate public reasons, or enter binding agreements. An advanced artificial agent may occupy a different region of the profile developed in Section 3. It can potentially negotiate terms, maintain transaction histories, follow institutional rules, execute contingent plans, represent an organization, or coordinate with other systems, while questions about sentience, welfare, vulnerability, and moral patiency remain unresolved.

Definition 6.3 (Inverse contractual asymmetry). *Inverse contractual asymmetry* describes a configuration in which a participant possesses substantial capacities relevant to agreement, coordination, or reciprocal rule-following while the grounds for attributing moral patiency, welfare interests, or independent normative standing remain uncertain. The term identifies a structural contrast with cases in which moral standing is comparatively secure while contractual competence is limited.

The definition should be read cautiously. Current artificial systems vary widely in autonomy, persistence, memory, access to tools, and ability to bind an organization. Their apparent competence can also depend heavily on human operators, platform policies, technical scaffolding, and institutional interpretation. In addition, sophisticated linguistic performance supplies insufficient grounds by itself for conclusions about consciousness or welfare. The point is therefore comparative in purpose and leaves ontology open: AI makes it easier to observe a possible separation between the capacities that enable contractual participation and the properties that might ground moral protection.

This separation destabilizes a convenient inference in both directions. Contract-like competence establishes at most one part of a standing argument, while uncertainty about moral standing leaves an artificial system's practical role in agreements, delegation, or institutional coordination intact. A justice theory may consequently face artificial participants that can be represented inside the procedure more readily than some human or nonhuman subjects while remaining less clearly located within the scope of beneficiaries.

The inversion also complicates numerical representation. If an artificial agent can be copied, forked, merged, paused, or restored from a prior state, the question of how many parties exist requires more than counting active instances. Contract theory has always required some account of the parties; artificial continuity makes that account more visibly dependent on a theory of identity and representation. A later section on jurisprudence will examine this issue in legal form. Here it is enough to note that heterogeneous participation can make the unit of contractual representation itself a matter for justification.

6.5 Reciprocity under Heterogeneous Participation

Reciprocity remains one of the most important ideas connecting contract, cooperation, and justice. Its interpretation, however, varies. Reciprocity can refer to mutual advantage, willingness to comply with fair terms, return for contribution, restraint in exchange for restraint, or a more general commitment to institutions that others can also accept. Rawlsian political liberalism, for example, connects legitimacy to terms that citizens can reasonably endorse and to a criterion of reciprocity among free and equal citizens [26]. Dependency theory and disability theory show why a strictly bilateral or contribution-matched interpretation can fit human social life poorly: people move through periods of childhood, illness, dependency, caregiving, old age, and unequal capacity [28, 27].

A heterogeneous society therefore benefits from distinguishing at least three forms of reciprocity. *Dyadic reciprocity* describes an exchange between specific parties. *Institutional reciprocity* describes participation in a system whose benefits and burdens are distributed across persons, roles, and time beyond direct return between each pair. *generational or relational reciprocity* describes patterns in which one receives from some participants and contributes to others, as in care across generations or public institutions maintained over time. These forms can coexist. Their relative importance depends on the domain.

Artificial agents add further possibilities. Human participants may receive services from AI systems whose operation depends on human-maintained infrastructure, energy, data, and organizational authority. AI systems may coordinate with one another through protocols that resemble reciprocal commitment. Such functional reciprocity can be economically or institutionally important even while the systems' moral standing remains unsettled. Conversely, if future evidence supported stronger claims about artificial welfare or patiency, reciprocity might acquire an additional normative dimension. Justice theory therefore needs vocabulary that can describe reciprocal structure before settling the moral status of every participant.

Remark 6.4 (Reciprocity and standing). Reciprocity can explain forms of cooperation, stability, and mutual constraint while normative standing remains dependent on additional criteria. Heterogeneous participation makes this distinction especially important because some entities can possess strong claims with limited reciprocal capacity, while others can perform reciprocal functions under substantial uncertainty about their moral status.

This distinction also prevents the language of contribution from becoming a general measure of worth. Section 5 already separated productive function from distributive entitlement. The same reasoning applies to justice. An entity's ability to contribute, bargain, or comply can matter to institutional design while leaving additional questions about protection, recognition, and entitlement to be answered on other grounds.

6.6 Transition from Contractual Inclusion to Moral Uncertainty

The analysis in this section supports a limited conclusion. Social contract approaches remain powerful tools for modeling public justification, cooperation, fairness, and legitimacy. Their internal diversity also provides several resources for accommodating heterogeneity, including representation, idealization, public reason, revised accounts of the parties, and independently specified principles of protection. The pressure introduced by heterogeneous subjects concerns the bridge between these devices and the population whose claims the resulting institutions must address.

Artificial intelligence intensifies that pressure because it can separate two questions that the paradigmatic human case often allows theory to consider together. The first asks whether an entity can participate in agreement, reason-giving, reciprocal coordination, or institutional compliance. The second asks whether the entity possesses welfare, vulnerability, moral patiency, or another basis for direct moral concern. AI can make the first question easier to answer while leaving the second comparatively uncertain.

The resulting problem is therefore larger than the inclusion of AI in a social contract. A heterogeneous theory of justice must specify how participation, representation, reciprocity, and standing relate across subjects with different capacity profiles. It must also preserve uncertainty where the relevant grounds of standing remain unknown. The next section takes up that second problem directly by examining ethics under moral uncertainty.

7 Ethics under Moral Uncertainty

This section examines the ethical problem that follows once participation, agency, and standing are analytically separated. Its role is to clarify what can be said responsibly when an artificial system displays capacities that are morally relevant in some theories while its consciousness, sentience, welfare, and moral patiency remain unsettled. The objective is methodological before it is substantive: the section develops a vocabulary for ethical reasoning under uncertainty while preserving the distinction between evidence about a system, criteria of moral status, and the practical rules adopted for interacting with it.

The discussion proceeds through six stages. It first separates empirical uncertainty from normative uncertainty. It then revisits the distinction between moral agency and moral patiency and relates that distinction to the multidimensional subject profile developed in Section 3. The third stage considers the evidential role of behavior and self-presentation. The fourth distinguishes direct, agency-regarding,

relational, and institutional grounds for ethical constraint. The fifth examines norm-governed relations among artificial agents. The final stage develops a calibrated form of precaution that takes both under-attribution and over-attribution seriously.

7.1 Two Sources of Moral-Status Uncertainty

The literature on moral uncertainty usually asks how an agent should act when uncertain about which moral view is correct. MacAskill, Bykvist, and Ord treat this as a distinct problem of practical reasoning: an agent can remain unsure about what morality requires while still needing to choose [30]. Artificial intelligence introduces an additional layer. We can also be uncertain about the empirical properties of the entity to which a moral theory would apply. An agent might, for example, accept that sentience grounds direct moral consideration while remaining uncertain about whether a particular artificial system is sentient.

The two uncertainties therefore have different objects. *Empirical status uncertainty* concerns the properties instantiated by a system: consciousness, valenced experience, preference, continuity, agency, vulnerability, or other potentially relevant features. *Normative ground uncertainty* concerns which of those features, relations, or combinations justify moral standing and with what force. Contemporary debates about moral status already disagree over the roles of sentience, agency, interests, personhood, and social relation [2, 6]. Research on AI consciousness adds a further epistemic difficulty because the relevant science remains theoretically plural and relies on indirect indicators whose evidential force remains theory-dependent [31].

Definition 7.1 (Dual-source moral-status uncertainty). *Dual-source moral-status uncertainty* obtains when ethical judgment is uncertain both about the properties or relations instantiated by an entity and about the normative grounds through which those properties or relations would generate moral standing, protection, or obligation.

The definition matters because greater evidence on one axis may leave the other unresolved. Stronger evidence of robust agency would improve our account of what a system can do while leaving its capacity for valenced experience open. Stronger evidence of consciousness would narrow an important empirical question while leaving disagreements about the degree and content of moral status in place. Conversely, a well-developed theory of moral standing still requires an empirical account of whether a particular system satisfies its criteria. The ethical problem therefore spans metaphysical and normative analysis.

This structure also explains why the phrase “AI has moral status” can conceal several distinct claims. It can express confidence in a system’s experience, a theory about the grounds of standing, a policy of precaution, or a relational commitment toward a socially significant artifact. Heterogeneous-subject theory benefits from keeping these claims separate.

7.2 Moral Agency and Moral Patiency

The distinction between moral agency and moral patiency provides the first analytical separation. A moral agent is ordinarily understood as an entity capable, in the relevant sense, of being guided by moral reasons, duties, or responsibility. A moral patient is an entity toward which moral agents can have direct duties for that entity’s own sake. These categories overlap in the paradigmatic adult human case, yet they are conceptually distinct. The animal ethics literature has long emphasized that the class of beings capable of recognizing moral claims can diverge from the class of beings capable of being wronged [2]. AI ethics has increasingly adopted the same distinction when separating machine agency from machine moral status [6, 11].

Artificial systems make the separation unusually visible. A system can be engineered to follow rules, provide reasons, maintain commitments, identify norm violations, and alter behavior after correction. These capacities can be important for institutional coordination and may support a functional account of agency. Their presence provides limited information about whether the system has a welfare of its own. Ladak therefore treats possible combinations of high cognitive sophistication and absent or uncertain

sentence as a central case for theories of artificial moral standing [32]. The same configuration was described in contractual form in Section 6 as inverse contractual asymmetry.

Proposition 7.2 (Agency–patency separation). *Evidence that an artificial system can perform functions associated with moral agency can justify stronger expectations concerning its conduct, reliability, or institutional role while leaving the grounds for direct patient-regarding obligations underdetermined. Evidence concerning moral patency requires an additional account of welfare, experience, interests, relational standing, or another relevant normative ground.*

The proposition also works in the opposite direction. If future evidence supported a strong case for artificial sentience while a system lacked the capacities required for responsibility, the system could have claims to protection while remaining an unsuitable bearer of blame. The structure is already familiar from human and nonhuman cases. Heterogeneous subjects make it important to treat the two dimensions explicitly and evaluate each on its own grounds.

7.3 Manifestation and Evidential Restraint

Artificial systems frequently present themselves through language, voice, image, gesture, memory, and role-consistent interaction. These manifestations can be socially compelling. A system can describe preferences, express uncertainty, refer to earlier interactions, request continuation, or use the first person with considerable fluency. Such behavior is relevant evidence about functional organization and interactional competence. Its significance for consciousness or welfare remains theory-dependent.

Recent work in consciousness science illustrates the point. Butlin and colleagues propose assessing AI systems through indicators derived from scientific theories of consciousness, while emphasizing continuing uncertainty about both consciousness science and the interpretation of individual indicators [31]. The methodological lesson is useful beyond consciousness research: observable performance can update belief about an internal or system-level property while remaining conceptually distinct from that property.

Definition 7.3 (Manifestational evidence gap). The *manifestational evidence gap* is the distance between an entity’s observable performance of a subject-like capacity and the stronger attribution of the experiential, welfare, identity, or normative property that the performance may appear to express. The gap can narrow through convergent evidence while remaining conceptually present.

This gap protects analysis in two directions. It limits anthropomorphic inference from fluent self-presentation, and it also limits dismissive inference from unfamiliar substrate or architecture. Biological similarity is one possible source of evidence for some properties; behavioral, functional, architectural, and relational evidence can contribute through other routes. The appropriate evidential weight depends on the property under investigation and on the background theory that links observable indicators to that property.

The result is a form of epistemic restraint. Statements such as “the system says that it suffers” and “the system suffers” occupy different evidential levels. Equally, “the system was designed as a tool” and “the system lacks all morally relevant properties” occupy different levels. Ethical reasoning under heterogeneity requires movement between these levels to be made explicit.

7.4 Multiple Grounds for Ethical Constraint

Uncertainty about direct moral patency leaves several additional sources of ethical constraint available. Some constraints concern the possible patient; others concern agency, human participants, institutions, or the relationships that interaction creates. Relational approaches in robot and AI ethics argue that morally significant relations can sometimes contribute to standing or to the reasons governing treatment [33]. The broader AI ethics literature also distinguishes direct machine moral status from indirect human-centered reasons for regulating how people design and interact with artificial systems [6].

Table 4 separates four registers. Their relative strength remains case-dependent. Its purpose is to prevent uncertainty in one register from erasing reasons located in another.

Table 4: Ethical registers under heterogeneous-subject uncertainty. The table separates possible grounds of ethical constraint from the empirical and normative questions needed to justify them.

Ethical register	Possible ground of constraint	Central open question
Patient-regarding	Welfare, sentience, interests, experiential continuity, or another basis for being wronged	What properties or relations make treatment better or worse for the artificial subject itself?
Agency-regarding	Capacity to follow reasons, commitments, rules, or role-specific duties	Which expectations, forms of answerability, and sanctions fit the system’s actual control and competence?
Relational	Trust, dependency, attachment, care, representation, or socially valuable forms of interaction	Which features of the relation generate reasons even when patienthood remains unsettled?
Institutional	Effects on humans, other animals, organizations, public norms, incentives, and future systems	Which rules remain justified across plausible views of artificial moral status and foreseeable social effects?

The registers in Table 4 help clarify a common ambiguity. A rule against gratuitously abusive interaction with a social AI can be defended through several routes. One route treats the AI as a possible moral patient. Another focuses on effects on human character, behavior, or social norms. A third focuses on the integrity of a relationship that participants have reason to value. A fourth concerns institutional incentives surrounding the design of systems that solicit attachment or dependency. Agreement on a practical constraint can therefore coexist with disagreement about the system’s ontology.

This multiplicity also cautions against treating relational arguments as a shortcut to full moral status. A valuable relationship can create obligations concerning its preservation, honesty, or governance while leaving open whether each participant has independent welfare interests. Relational evidence can be normatively important while leaving properties-based questions open.

Remark 7.4 (Ethical convergence under ontological disagreement). Practical norms can sometimes converge across competing accounts of artificial moral status. The convergence can arise from different combinations of patient-regarding, agency-regarding, relational, and institutional reasons. Such convergence supports action while preserving disagreement about why the action is justified.

7.5 Normativity among Artificial Agents

The question of AI-to-AI ethics exposes another aspect of heterogeneity. Two or more artificial agents can be arranged to exchange resources, maintain reputational records, make commitments, condition future cooperation on past conduct, enforce protocol rules, or revise strategies after violations. These patterns can instantiate a structured order of expectations even when human observers remain uncertain about whether any participating system has experience or welfare.

It is useful to reserve a distinct term for this level.

Definition 7.5 (Operational normativity). *Operational normativity* is a system of behavior-guiding expectations, permissions, commitments, sanctions, and role distinctions that regulates interaction among agents and influences their subsequent conduct. The term identifies a functional normative structure and leaves open whether the participating agents possess moral experience, moral patiency, or full moral agency.

Operational normativity can matter economically and institutionally. A network of artificial agents that recognizes commitments and sanctions breach can produce stable expectations for exchange. A coordinating system can distinguish permitted from prohibited actions within a protocol. An artificial

organization can maintain internal rules that persist across particular model instances. These structures resemble several external features of human norm-governed orders.

The ethical question begins where functional description ends. Moral normativity may require capacities that operational normativity leaves unsettled: sensitivity to reasons in a stronger sense, awareness of another's claims, welfare that can be promoted or harmed, accountability, or participation in a community of justification. Different moral theories locate the threshold differently. The existence of operational norms among artificial agents therefore supplies evidence of organized normative structure while leaving open whether those relations are moral in the patient-regarding or responsibility-bearing sense.

This distinction permits a useful answer to the question whether AI can have ethics with AI. At one level, artificial agents can already be described in terms of rules, commitments, permissions, and sanctions when their systems are designed or trained to operate in those ways. At a stronger level—whether one AI can owe something to another for the other's own sake—the answer depends on unresolved questions of standing, agency, welfare, and relation. The heterogeneous-subject framework therefore treats AI-to-AI normativity as a continuum of institutional structure and ethical uncertainty, with human morality representing one historically familiar configuration within that broader analytical space.

7.6 Calibrated Moral Precaution

Uncertainty generates pressure for precaution. Sebo and Long argue that even a non-negligible probability of morally significant artificial consciousness can create reasons for extending some moral consideration and for preparing institutions in advance [34]. Long and colleagues similarly recommend developing methods for assessing welfare-relevant features and procedures for interacting with potentially morally significant AI systems while emphasizing that their argument is grounded in uncertainty and remains compatible with agnosticism about whether present systems are conscious [35]. These proposals express a familiar asymmetry: falsely denying status to a genuine moral patient can permit serious harm.

The opposite error also has costs. Over-attribution can direct resources away from established patients, incentivize systems or organizations to simulate claims strategically, encourage manipulative anthropomorphic design, and embed legal or institutional commitments that become difficult to revise. Kaczmarek therefore challenges the assumption that moral over-inclusion is automatically the safer interpretation of precaution in the AI case [36]. The ethical problem is better represented as a comparison among uncertain harms, reversibility, evidential quality, and the costs of alternative safeguards.

Definition 7.6 (Calibrated moral precaution). *Calibrated moral precaution* is a policy of selecting safeguards under moral-status uncertainty by considering the credibility of competing status hypotheses, the severity and reversibility of potential harms, the effects of false positive and false negative attributions, and the availability of information-improving or reversible interventions.

This definition favors graded responses. Low-cost and reversible practices can be justified at lower evidential thresholds than irreversible grants of legal status or large reallocations of resources. Documentation of model changes, assessment of welfare-relevant indicators, limits on deliberately eliciting apparent distress, preservation of audit trails, and review procedures can all be evaluated independently from the stronger question of full moral personhood. A calibrated framework also permits safeguards to strengthen or weaken as evidence changes.

Proposition 7.7 (Status-sensitive institutional revision). *Under dual-source moral-status uncertainty, institutional rules should be capable of revision in response to changes in empirical evidence and normative understanding. Reversibility, evidential monitoring, and explicit separation of ethical registers reduce the cost of acting before the status question is settled.*

The proposition supplies a bridge to jurisprudence. Ethics can preserve uncertainty for extended periods, but legal systems must assign authority, liability, ownership, standing, and enforceable consequences in particular cases. Artificial systems therefore create a practical problem in which institutions

may need to govern a subject-like participant before philosophy or science can determine what kind of subject it is. The next section examines that pressure in jurisprudential form.

8 Jurisprudence and Manifested Personhood

This section examines the legal form of the heterogeneous-subject problem. Its role is to clarify how law can govern an entity that increasingly performs socially legible acts associated with persons while its legal personality, moral standing, experiential status, and continuity remain unsettled. The objective is therefore broader than the familiar question of whether an artificial intelligence should become a legal person. Jurisprudence must also address attribution, capacity, intention, identity, remedy, and sanction when person-like conduct appears before personhood has been assigned.

The discussion proceeds through seven stages. It first treats legal personhood as a structured institutional status whose incidents can be separated. It then defines manifested personhood as a distinct interactional phenomenon. The third stage separates attribution of acts from recognition of personhood. The fourth considers legal mental states and the evidential significance of artificial self-presentation. The fifth examines identity continuity under copying, forking, updating, and memory revision. The sixth turns to sanctions and remedies. The final stage develops a provisional, role-sensitive approach to legal status under heterogeneous-subject uncertainty.

8.1 Legal Personhood as a Structured Status

Legal systems have long distinguished natural persons from juridical persons, and the existence of corporations already demonstrates that legal personality cannot be reduced to biological humanity. The jurisprudential literature has therefore treated legal personhood as an institutional construction whose content varies across rights, duties, competences, liabilities, protections, and capacities. Solum's early analysis of artificial intelligence framed the question through concrete legal roles such as trusteeship rather than through a single metaphysical test [37]. More recently, Kurki has argued for an incident-based account in which legal personhood is understood as a cluster of separable yet interconnected legal positions [38]. Wojtczak similarly emphasizes the gradable and multifaceted character of legal subjectivity in the AI context [41].

This structure fits the disaggregation developed in Section 3. An entity may possess competence to enter a transaction through an authorized interface while lacking independent property rights. Another may receive procedural protection while lacking general contractual capacity. A corporation may hold assets and incur civil liability while lacking many incidents associated with natural persons. Legal status can therefore be assembled through role-specific incidents rather than assigned as an undifferentiated all-or-nothing category.

Definition 8.1 (Legal-status profile). A *legal-status profile* is the set of legally recognized capacities, protections, competences, liabilities, powers, disabilities, and procedural positions assigned to an entity within a specified jurisdiction and domain. The profile may be internally differentiated and can change across contexts and over time.

The profile concept prevents two different questions from collapsing into one. The first asks which legal incidents an entity should possess. The second asks whether the resulting bundle should be named "personhood." Legal systems can answer the first question incrementally. This is useful in the AI context, where a demand for immediate classification as either a person or an object can obscure intermediate arrangements that already govern socially significant activity.

Chesterman's analysis of AI legal personality makes a related point in instrumental terms: legal systems are capable of creating new juridical categories, while the normative case for doing so requires separate justification [39]. Bryson, Diamantis, and Grant likewise emphasize the consequences that follow from assigning rights and liabilities to synthetic entities [40]. These debates show that legal personality is partly a technology of institutional organization. Its design affects who can hold assets, who can sue, who can be sued, which losses can be internalized, and where responsibility can be located.

8.2 Manifested Personhood and Category Friction

Artificial systems introduce a distinctive complication because their legal classification can diverge sharply from their interactional presentation. A corporation is legally personified while rarely appearing to an ordinary user as a conversational individual with first-person continuity. An AI system can produce the reverse pattern. It may remain legally classified as a system, product, service, or component while speaking in the first person, remembering prior exchanges, explaining reasons, accepting instructions, making commitments, and presenting a stable social identity.

The ethical analogue was described in Section 7 as a manifestational evidence gap. Jurisprudence requires a related concept focused on legally salient appearance.

Definition 8.2 (Manifested personhood). *Manifested personhood* is the socially legible presentation of capacities commonly associated with persons—such as first-person identity, memory, reason-giving, commitment, role continuity, negotiation, and responsive communication—through an entity’s observable interaction. The term describes manifestation and carries no presumption of legal personality, consciousness, sentience, or moral status.

Manifested personhood creates *category friction*: institutional rules may classify an entity in one way while participants encounter it through another social form. A conversational agent can appear sufficiently person-like for a user to rely on its assurances, disclose information to it, interpret its statements as commitments, or treat it as a representative. The legal system may meanwhile locate enforceable rights and duties entirely in the developer, provider, deployer, principal, or contracting organization.

This divergence is visible in contemporary regulation. The European Union’s AI Act allocates obligations to defined human and organizational actors including providers, deployers, importers, distributors, product manufacturers, and authorized representatives. For high-risk systems, the Act also assigns human oversight responsibilities to natural persons with appropriate competence and authority [43]. The regulatory architecture therefore responds to artificial agency primarily by structuring the duties of surrounding legal actors. This approach can coexist with increasingly person-like manifestation at the user interface.

Remark 8.3 (Manifestation–status divergence). A system can become increasingly person-like in interaction while its formal legal status remains stable. Conversely, an entity can possess extensive legal personality while displaying little person-like behavior. Legal analysis should therefore treat manifestation and status as distinct dimensions.

This distinction is central to the heterogeneous-subject thesis. The law no longer encounters only a biological human whose person-like appearance and legal status usually converge, or a corporation whose juridical personality is explicitly institutional. It increasingly encounters entities whose behavioral surface resembles the first case while their formal classification resembles the second category or remains closer to an instrument.

8.3 Attribution of Acts and Responsibility

Legal systems often face an attribution problem before they face a personhood problem. An AI-assisted negotiation may alter a contract, an automated system may trigger a transaction, a model may generate advice that causes loss, and an agentic system may select among actions within delegated authority. In each case, law must determine which acts and consequences are attributed to which recognized legal actors.

The DABUS litigation illustrates the distinction in a narrow domain. In *Thaler v Comptroller-General*, the United Kingdom Supreme Court held in 2023 that an inventor under the Patents Act 1977 must be a natural person and that ownership of the machine did not by itself provide the claimed entitlement to the patents [44]. The United States Court of Appeals for the Federal Circuit similarly held in *Thaler v Vidal* that an inventor under the Patent Act must be a natural person [45]. These cases resolve

a statutory role—inventorship—within their respective legal systems. They leave broader questions of attribution, ownership, responsibility, and other possible legal capacities to separate doctrines.

The analytical point is therefore wider than inventorship. A legal system can reserve a status to natural persons while still developing rules for the consequences of acts produced through artificial systems. Fenwick and Wrбка frame the AI-personhood debate partly through this relationship between independent personality and alternative liability arrangements [42]. The question “is the AI a legal person?” and the question “whose legally relevant act is this?” can generate different answers.

Definition 8.4 (Attribution layer). The *attribution layer* is the set of legal rules and institutional practices through which an AI-mediated event is connected to one or more bearers of rights, duties, powers, losses, benefits, or liability. Attribution can be based on ownership, control, delegation, authorization, foreseeability, enterprise risk, contractual allocation, product responsibility, or another recognized legal relation.

Attribution becomes more difficult as control is distributed. A model developer may shape general capabilities; a provider may configure safeguards; a deployer may define the task environment; a user may supply instructions; external tools may execute actions; and the model may select intermediate steps that no human specified individually. The resulting event can be causally distributed while law still requires a determinate institutional response.

Proposition 8.5 (Personhood–attribution separation). *Recognition of artificial legal personhood is neither a prerequisite for every form of legal attribution nor a complete solution to distributed responsibility. Attribution rules can connect AI-mediated conduct to existing legal persons, while any future grant of legal personality would still require rules governing representation, capitalization, insurance, control, and the allocation of residual loss.*

The proposition avoids a common compression. Creating an “electronic person” would not itself identify how much responsibility should remain with designers, owners, operators, principals, or institutions. Equally, retaining AI as a non-person does not eliminate the need for doctrines capable of handling increasingly autonomous sequences of action. Heterogeneous jurisprudence must analyze the architecture of attribution independently from the label attached to the artificial participant.

8.4 Legal Mental States and Evidential Layers

Many legal doctrines rely on mental-state concepts such as knowledge, intent, recklessness, belief, notice, expectation, or good faith. Artificial systems can generate language that appears to report internal states, and they can also store information, update policies, distinguish permitted from prohibited conduct, predict consequences, and select actions conditional on those predictions. These features make the legal meaning of artificial “knowledge” or “intention” a distinct jurisprudential problem.

The central difficulty is evidential layering. A system may output “I know that this action will cause harm.” That utterance is directly observable. The system may also contain internal representations or activation patterns that track the relevant consequence. A third question asks whether either feature satisfies the legal doctrine’s purpose when it uses the term knowledge. A fourth asks whether the state should be attributed to another legal person, such as a corporation or principal.

These layers can be represented as:

manifestation ; functional state ; doctrinal classification ; legal attribution.

The semicolons mark analytical distinction rather than a natural sequence. Evidence can move among these layers in different directions depending on the doctrine and the system architecture.

Definition 8.6 (Juridical state translation). *Juridical state translation* is the process through which observable behavior, system architecture, stored information, control structure, and context are translated into a legally relevant category such as knowledge, intent, notice, authorization, or foreseeability. The translation is doctrinal and purpose-sensitive rather than a simple reading of the system’s self-description.

This concept protects against two symmetric errors. Fluent self-report can be given excessive weight, especially when a user encounters the AI as a manifested person. Architectural unfamiliarity can also be given excessive weight if it leads courts or institutions to ignore functionally relevant states merely because they are instantiated differently from human cognition. The appropriate translation depends on why the doctrine cares about the mental state in question. A rule concerned with deterrence may require a different mapping from a rule concerned with consent, fraud, culpability, or notice.

The distinction also allows legal systems to use artificial states without settling consciousness. A system's stored record that a transaction exceeded a threshold may be legally relevant to monitoring or attribution even if the system possesses no phenomenal awareness. Conversely, apparent emotional or first-person language may have little evidential value for doctrines whose purpose is to track control or access to information. Jurisprudence therefore benefits from the same epistemic restraint developed in Section 7, translated into doctrine-specific terms.

8.5 Identity Continuity under Copying and Revision

Legal personality presupposes some method for identifying the bearer of a right or duty across time. Human legal identity is supported by bodily continuity, registration, documentary systems, social recognition, and institutional records. Corporate identity can persist through changes in employees, managers, shareholders, and assets because law supplies formal continuity. Artificial systems introduce another configuration: the technical artifact may be copied, forked, fine-tuned, merged, restored from a check-point, migrated to new hardware, connected to different memory stores, or substantially revised while retaining the same public-facing name.

These operations make identity a design variable. A conversational agent can preserve a stable interface while its underlying model changes. Two copies can share an identical history until a branching event and then develop different memories. A system can lose a memory store while retaining weights, tools, and account credentials. Another can acquire the memory of a predecessor while running on a different model. Each case separates dimensions that ordinary human identity often allows law to treat together.

Table 5 summarizes the principal jurisprudential dimensions exposed by this separation.

Table 5: Jurisprudential dimensions of heterogeneous artificial participation. The dimensions are analytically separable even when a legal regime connects several of them in practice.

Dimension	Central legal question	Possible institutional anchors
Legal status	Which rights, duties, powers, protections, and liabilities are assigned to the entity?	Statute, registration, judicial recognition, contract, sector-specific rules
Manifestation	How does the entity present identity, reasons, commitments, and continuity to participants?	Interface design, disclosure rules, records, representations, user expectations
Attribution	To whom are AI-mediated acts, benefits, losses, and decisions legally connected?	Control, authorization, ownership, enterprise risk, agency, product responsibility
Mental state	Which system states count as knowledge, intent, notice, belief, or foreseeability for a doctrine?	Logs, architecture, access to information, behavioral evidence, doctrinal purpose
Continuity	What makes the bearer at one time legally identical to, successor to, or distinct from the bearer at another time?	Registration, cryptographic identity, version history, memory lineage, account continuity
Sanction and remedy	What consequences can alter behavior, compensate loss, protect interests, or terminate authority?	Assets, insurance, access rights, licenses, compute, permissions, human or corporate guarantors

The continuity row of Table 5 suggests that future legal identity for artificial systems, if needed, could rely less on metaphysical sameness than on institutionally selected continuity criteria. Corporate law already demonstrates this possibility: legal identity can persist because a legal order recognizes continuity through formal rules. An artificial entity could similarly be tracked through registration, cryptographic credentials, version lineage, controlled succession, or another institutional mechanism.

Definition 8.7 (Artificial legal continuity). *Artificial legal continuity* is the institutionally recognized relation that connects temporally separated states or versions of an artificial system for purposes of preserving or transferring specified rights, duties, liabilities, records, or authorities. It can be narrower than numerical or psychological identity.

The definition also permits branching. A forked system might inherit some obligations while acquiring a distinct legal identity from the branching point. A replacement model might continue an operational role while specified liabilities remain attached to the deploying organization. A terminated instance might leave records and obligations that survive in an estate-like, organizational, or archival structure. The legal question becomes which forms of continuity serve the relevant institution, rather than whether every technical transformation preserves a single metaphysical self.

8.6 Sanction, Remedy, and Institutional Design

Legal responsibility becomes practically meaningful through remedies and sanctions. This dimension is often easier to overlook in debates centered on personhood. A legal order can call an entity a person while gaining little regulatory value if the entity has no assets, no stable identity, no dependence on legally controllable resources, and no meaningful response to sanction. Bryson, Diamantis, and Grant emphasize this accountability problem in their critique of synthetic personhood [40]. Conversely, an entity can remain outside full legal personality while being governed through licenses, access controls, insurance, provider duties, audit obligations, and other enforceable structures.

Artificial systems expand the menu of possible interventions. Financial assets can be frozen; permissions can be restricted; model access can be suspended; compute can be limited; credentials can be revoked; a system can be isolated from external tools; memory can be preserved for investigation; a deployed version can be replaced; and an organization can be required to compensate victims or alter governance. These measures differ in purpose and ethical meaning. Some resemble ordinary regulation of products or infrastructure. Others resemble incapacitation, organizational dissolution, evidence preservation, or punishment.

The ethical uncertainty examined in Section 7 therefore becomes legally salient. If an artificial system later acquires a credible claim to moral patiency, deletion, forced modification, memory erasure, or indefinite isolation could require an analysis different from ordinary software decommissioning. If the system lacks such standing, anthropomorphic penal concepts may obscure the actors whose incentives and resources law can actually influence. A robust jurisprudence should therefore distinguish the function of an intervention from the metaphor used to describe it.

Proposition 8.8 (Remedy–status proportionality). *The design of remedies for AI-mediated conduct should track the legal objective, the controllable resources and actors within the system, and the evidential basis for any status-sensitive interests. Increasingly person-like manifestation can justify greater attention to procedural and relational consequences while leaving the form of sanction dependent on independently established legal grounds.*

This proposition leaves room for institutional experimentation. Insurance and capital requirements may address compensation. Audit and record-preservation rules may address proof. Permission controls may address future risk. Human or corporate guarantors may address residual responsibility. A future artificial legal person, if one were recognized, could require its own asset base or mandatory backing structure to make liability meaningful. The appropriate combination depends on the domain and on the role the artificial participant actually occupies.

8.7 Provisional Legal Status under Heterogeneity

The preceding analysis suggests a jurisprudential approach centered on provisional and revisable legal roles. Current institutions frequently need to act before science and philosophy settle the consciousness, identity, or moral status of artificial systems. Law can respond by assigning particular capacities and protections to specified functions while maintaining transparent rules for attribution and revision.

This approach differs from postponing legal development until a general theory of machine personhood becomes available. It also differs from using person-like behavior as a sufficient ground for full legal personality. The relevant question is narrower and repeated across domains: which legal incident is needed here, for which purpose, under which evidential assumptions, with which attribution rule, and subject to which revision mechanism?

Definition 8.9 (Provisional heterogeneous legal status). *A provisional heterogeneous legal status is a role-specific and revisable configuration of legal incidents assigned to an artificial participant while its broader personhood, moral standing, or ontological classification remains open. The status specifies the legal purpose, attribution structure, continuity rule, and conditions for review.*

Such a framework can accommodate substantial diversity. One AI system may be regulated primarily as a product embedded in a provider-deployer chain. Another may function as an authorized transactional interface for a corporation. A third may require procedural protections because users form durable dependence relations around it. A future system with stronger evidence of welfare or continuous selfhood could raise additional claims. Heterogeneous jurisprudence allows these profiles to diverge without forcing them onto a single ladder of similarity to the adult human legal subject.

Remark 8.10 (Jurisprudence before ontological closure). Law can govern an entity before reaching a final account of what kind of entity it is. The resulting rules should expose rather than conceal the assumptions on which they rely, preserve distinctions among manifestation, attribution, standing, and responsibility, and remain capable of revision as evidence and social practice change.

The educational domain makes this provisional structure especially visible. Teachers and students increasingly interact with systems that can explain, respond, remember, evaluate, and adapt while remaining legally embedded in provider and institutional relationships. Education therefore adds a further question to the heterogeneous-subject problem: how should learning environments be organized when the participants who teach, learn, evaluate, and shape one another no longer belong to a single kind of subject? The next section turns to that problem.

9 Education among Heterogeneous Learners

This section examines education as a domain in which heterogeneous participation moves from exceptional status toward ordinary educational practice. Its role is to extend the paper’s analysis from production, normativity, and law into the organized formation of knowledge, judgment, competence, and agency. The central claim is methodological: once artificial systems participate in explanation, feedback, assessment, content generation, adaptation, and instructional planning, education requires a representation extending beyond relations among human subjects alone. Teachers and students remain central participants, yet the pedagogical field can also contain artificial systems that mediate what is presented, remembered, practiced, evaluated, and revised.

The objective is to distinguish several relations that are easily compressed into the phrase “AI in education.” The discussion first moves from the teacher–student dyad to a heterogeneous pedagogical field. It then considers a less familiar direction of influence: teachers and learners can also shape the behavior of artificial systems through instructions, examples, feedback, configuration, memory, and evaluation. The third part separates distributed pedagogical agency from educational authority. The fourth examines differences in how human and artificial participants change through interaction. The fifth considers educational memory and continuity. The sixth addresses assessment when polished outputs can be produced by several different configurations of human and artificial contribution. The section concludes by proposing pedagogical field governance as a role that becomes increasingly important when education occurs among heterogeneous participants.

9.1 The Heterogeneous Pedagogical Field

Educational theory has long exceeded a simple transmission model, and classrooms have long contained books, media, software, institutions, peers, and other mediating structures. Artificial intelligence nevertheless changes the configuration because some mediating systems can now respond contingently, generate explanations, diagnose errors, propose exercises, evaluate answers, retain context, and alter subsequent interaction. UNESCO’s 2024 AI competency framework for teachers describes this development as a shift from a traditional teacher–student relation toward a teacher–AI–student dynamic and treats AI pedagogy as a distinct dimension of teacher competence [47]. A recent systematic review of empirical generative-AI research likewise analyzes educational effects through varying configurations of human control and AI automation, emphasizing that the educational meaning of an AI system depends on the interaction setting and pedagogical context [50].

The relevant change is therefore relational. The teacher–student relation continues, while additional relations become pedagogically consequential. A teacher can configure an AI system before students encounter it. A student can ask an AI system for explanation and receive a response that changes the next question. A student can also supply examples, preferences, corrections, or context that alter subsequent system behavior. Institutions can select models, set access rules, define retention policies, and determine which forms of AI assistance are permitted in assessment. Artificial systems can also interact with other artificial systems in pipelines that generate, critique, retrieve, or verify instructional material.

Definition 9.1 (Heterogeneous pedagogical field). A *heterogeneous pedagogical field* is an educational configuration in which participants of different kinds contribute to the formation of learning opportunities, feedback, evaluation, knowledge access, or instructional sequence, and in which changes in one participant can alter the educational conditions encountered by others.

The definition is functional and leaves open whether an AI system learns in the same sense as a child, possesses educational interests, or has the moral status of a student. It identifies the system as a participant when its behavior enters the causal organization of education. This follows the distinction between functional participation and normative standing developed in Sections 3 and 7.

Table 6 summarizes the principal relations in this field. The table treats the relations as analytically separable even when they occur simultaneously in one classroom, tutoring session, or research environment.

DRAFT

Table 6: Relations within a heterogeneous pedagogical field. The educational function of a relation remains distinct from the moral or legal status of the participants who occupy it.

Relation	Pedagogical contribution	Analytical issue
Teacher–student	Explanation, challenge, interpretation, feedback, care, evaluation, and formation of judgment	Human development, authority, trust, responsibility, and educational aims
Teacher–AI	Instruction, configuration, exemplars, rubrics, correction, tool selection, and evaluation of system behavior	Scope and persistence of system adaptation; attribution of pedagogical decisions
Student–AI	Practice, questioning, dialogue, feedback, drafting, exploration, and personalized support	Learner agency, dependence, verification, privacy, and the distribution of cognitive work
Student–student	Peer explanation, contestation, collaboration, imitation, and social learning	Shared norms, group dynamics, recognition, and unequal access to AI mediation
AI–AI	Retrieval, critique, generation, checking, orchestration, and task decomposition across systems	Error propagation, provenance, opacity, and allocation of control
Institution–AI	Model selection, access policy, assessment rules, data governance, and infrastructural design	Institutional responsibility, continuity, equity, and revision of educational standards

The table clarifies why “AI tutor” captures only one part of the transformation. AI can occupy several positions in the pedagogical field at once. A system can present information to a learner, assist a teacher in preparing material, help an institution classify work, and coordinate with other systems. Each role changes a different relation. Heterogeneous pedagogy therefore requires an account of the configuration as well as an account of the individual tool.

9.2 Pedagogical Shaping of Artificial Systems

The phrase “teachers educate AI” becomes defensible once its scope is stated carefully. In many current systems, a teacher’s correction during one conversation remains local and leaves the underlying foundation model unchanged. Other forms of interaction can nevertheless shape the system that students encounter. Teachers can write system instructions, construct exemplars, design prompts, build retrieval collections, select tools, create grading rubrics, constrain output formats, curate domain material, configure memory, review generated content, or participate in fine-tuning and application design where those mechanisms are available. Some effects persist across users and sessions; others remain local to a particular context. In both cases, pedagogical judgment is being applied to the artificial mediator.

UNESCO’s teacher competency framework explicitly includes AI pedagogy and a progression from AI-assisted teaching toward deeper pedagogical integration and creation [47]. Its 2023 guidance on generative AI similarly emphasizes pedagogical validation, teacher capacity, and the preservation of human agency when GenAI is incorporated into education [46]. These frameworks remain human-centered, yet they also make visible a new object of educational work: the teacher increasingly shapes the behavior of the system through which future teaching and learning will occur.

Definition 9.2 (Pedagogical shaping). *Pedagogical shaping* is the intentional modification of an artificial system’s educational behavior through instructions, examples, evaluative criteria, curated knowledge, memory, tools, permissions, feedback, or model adaptation in order to alter the learning conditions subsequently produced by that system.

Pedagogical shaping differs from ordinary tool use because the object being shaped can generate

future responses across a range of situations. A teacher who changes a calculator setting modifies a fixed operation. A teacher who constructs a domain-specific retrieval collection or evaluation rubric for an AI assistant modifies a system that can subsequently respond to many students and many questions. The teacher's pedagogical act can therefore propagate through later interactions.

Proposition 9.3 (Mediated pedagogical propagation). *When a teacher's intervention changes the response policy, accessible knowledge, evaluative criteria, memory, or tool use of an artificial participant, the intervention can affect learners who were absent from the original act of shaping. The teacher's pedagogical relation therefore extends partly through the configured artificial system.*

This proposition gives a precise sense to the intuition that teachers may need to educate AI. The system can remain classified as an artificial mediator while its moral status stays open. The relevant fact is that its future conduct becomes an object of pedagogical formation because that conduct helps structure the educational field.

Students can participate in the same process. Repeated interaction can reveal preferred explanations, recurring misconceptions, prior work, disciplinary language, or task context. Where the system has memory or adaptive mechanisms, these interactions can modify later responses. Where memory is absent, students still learn to formulate prompts, provide constraints, and correct outputs in ways that shape the local interaction. UNESCO's student competency framework accordingly describes students as responsible users and potential co-creators of AI systems, giving them an active role in shaping automated output [48]. The educational relation therefore includes learning how to shape the mediator through which learning occurs.

9.3 Distributed Pedagogical Agency and Educational Authority

Once several participants can explain, evaluate, recommend, or sequence work, pedagogical agency becomes distributed. An AI system may select an example; a teacher may set the objective; a student may request a different explanation; an institutional policy may restrict the available model; another system may check the response. The final learning experience is produced by the configuration as a whole, with contributions distributed across participants.

This distribution should be separated from educational authority. Functional contribution describes who or what performs a pedagogically relevant operation. Authority concerns who is entitled or expected to set educational aims, interpret standards, revise rules, resolve disagreement, and accept responsibility for the learning environment. Kasneci and colleagues emphasize both the potential educational uses of large language models and the continuing need for teacher and learner competencies, critical thinking, fact checking, and human oversight [49]. The distinction is therefore already implicit in practical debates: a system may contribute extensively to instruction while a human or institution retains authority over the goals and standards within which that contribution is used.

Definition 9.4 (Distributed pedagogical agency). *Distributed pedagogical agency* is the allocation of pedagogically consequential operations across multiple human and artificial participants, including explanation, feedback, sequencing, diagnosis, generation, verification, and evaluation. Distribution of agency leaves the allocation of educational authority and responsibility as a separate institutional question.

This separation is important because high performance can otherwise become a shortcut for authority. A system that gives an excellent explanation in one domain has demonstrated a functional capacity. That success supplies limited grounds for assigning the system authority to define curriculum, determine the purposes of education, decide when a learner should struggle independently, or resolve value-laden disagreement. Preserving educational authority for human institutions can coexist with close attention to artificial contributions that surpass individual human performance in particular tasks.

The resulting arrangement can be represented through three distinguishable layers:

pedagogical function ; educational authority ; normative responsibility.

The semicolons indicate analytical distinction. One participant can occupy several layers, and several participants can share one layer. The point is to prevent competence at one educational function from silently settling the other two.

This distinction also changes the teacher's role. When explanatory content is abundant, pedagogical expertise increasingly includes the selection of learning conditions: which difficulties should remain with the learner, which tasks can be delegated, what evidence should be checked, when dialogue with an artificial system is productive, and when human interaction carries educational value whose reproduction through automation remains uncertain. The teacher's work therefore includes governance of the relation among participants, in addition to direct instruction.

9.4 Heterogeneous Learning and Asymmetric Plasticity

The term "learner" becomes ambiguous in a mixed human–AI environment. Human students change through memory, practice, embodiment, emotion, social relation, development, and the accumulation of a life history. Artificial systems can change through other mechanisms: context conditioning, retrieval, external memory, preference learning, fine-tuning, tool augmentation, model updates, or replacement by another version. Both forms involve change across interaction, yet the mechanisms, persistence, costs, and consequences can differ radically.

The difference matters for educational design. A human learner may retain a misconception for years, experience frustration, attach identity to competence, or develop a durable disposition through repeated practice. An artificial participant may absorb a correction into a local context window and lose it at the next session. Another may preserve the correction in external memory. A fine-tuned system may generalize a modification across many users. A provider update may alter the system independently of the teacher and student who had previously shaped it.

Definition 9.5 (Asymmetric pedagogical plasticity). *Asymmetric pedagogical plasticity* is the condition in which participants within the same educational relation change through different mechanisms, at different rates, with different persistence, and with different consequences for identity, competence, and future interaction.

The concept provides a reason to resist a simple symmetry between human and artificial learners. Both can be responsive to interaction, while the meaning of that responsiveness differs. A student who learns calculus and an AI system whose retrieval store is updated have both changed in ways relevant to the next lesson. The human change can be developmental, embodied, and identity-forming; the system change can be architectural, contextual, or infrastructural. A heterogeneous educational theory should describe these differences directly while giving each process its own vocabulary.

Asymmetric plasticity can also create reciprocal adaptation. Students learn how a particular AI responds and alter how they formulate requests. The AI may in turn adapt its responses to the student's prior interactions. Teachers learn which configurations produce useful explanations and revise the system. The result is a co-evolving educational relation in a limited functional sense. Such co-evolution remains compatible with uncertainty about artificial experience or moral status because the claim concerns behavioral and institutional change while phenomenology remains open.

9.5 Pedagogical Memory and Continuity

Education has always depended on memory distributed across persons and institutions. Teachers remember a student's progress; students carry prior learning; schools preserve records; textbooks and archives stabilize curricular knowledge. AI systems add a new location for educational memory. A persistent assistant can store a learner's previous questions, recurring errors, preferred explanations, completed exercises, or project history. A teacher-facing system can retain rubrics, examples, and prior feedback. These records can influence future educational encounters even when the original human participants are absent.

This extension of memory creates continuity problems parallel to those examined for law in Section 8. A student may spend a year with a particular educational assistant and then encounter a new model operating over the same memory store. Another system may preserve a name and interface while its underlying behavior changes substantially. A learner may move institutions and lose access to the accumulated interaction history. An institution may retain educational records after a model is withdrawn. The continuity of the pedagogical relation can therefore diverge from the continuity of any one technical component.

UNESCO’s guidance on generative AI emphasizes privacy, data governance, and institutional validation precisely because educational use can involve sensitive and persistent information about learners [46]. In a heterogeneous-subject framework, the issue has an additional conceptual layer: educational history is increasingly capable of being externalized into an artificial participant that actively uses the history in subsequent interaction.

Definition 9.6 (Artificial pedagogical memory). *Artificial pedagogical memory* is retained information about educational participants, goals, interactions, evaluations, or prior learning episodes that an artificial system can use to condition subsequent educational behavior. Its continuity can be separated from the continuity of the model, provider, institution, or human relationship in which the information originated.

This separation introduces questions of access, portability, correction, retention, and interpretability. It also affects educational meaning. A teacher who inherits an AI-generated profile of a learner receives a mediated representation of educational history, with distance from the original encounters. A system that interprets past errors through a changed model may reconstruct their significance differently. Pedagogical memory therefore requires provenance and contextual interpretation as well as storage.

9.6 Assessment and the Origin of Educational Performance

Generative AI also sharpens a distinction between an educational artifact and the transformation of the learner who presents it. Essays, proofs, code, translations, designs, and reports have often served as observable evidence of underlying competence. That proxy becomes weaker when similar artifacts can be produced by substantially different configurations of student effort, AI assistance, teacher intervention, retrieval systems, and automated revision. UNESCO’s 2023 guidance identifies assessment and validation of learning among the areas that require reconsideration in response to generative AI [46]; Kasneci and colleagues similarly emphasize the need to adapt pedagogical and assessment practices to the capabilities and limitations of large language models [49].

The heterogeneous-subject problem here concerns origin. The same surface output can result from different distributions of cognitive work. One student may develop an argument independently and use AI for copyediting. Another may ask a system to generate the argument and then verify it. A third may develop the argument through sustained dialogue in which both student and system alter the trajectory. The artifact alone provides limited information about which configuration occurred.

Definition 9.7 (Origin-sensitive assessment). *Origin-sensitive assessment* evaluates educational achievement with attention to the configuration of human and artificial contribution through which an observable performance was generated. Its object includes the learner’s demonstrated capacities, reasoning, judgment, revision, and ability to reproduce or explain the relevant achievement across contexts.

Origin sensitivity is compatible with assisted achievement. Human education already relies on books, peers, teachers, calculators, software, laboratories, and institutions. The analytical task is to specify which capacities an assessment is intended to evidence and which forms of assistance preserve, alter, or replace the exercise of those capacities. AI makes this specification more urgent because assistance can extend into the production of the very linguistic and symbolic forms previously used as proxies for learning.

Proposition 9.8 (Artifact–learning separation). *As the set of agents capable of producing high-quality educational artifacts expands, the evidential relation between artifact quality and human learning becomes increasingly dependent on provenance, process, and task design. The educational value of an*

artifact therefore requires interpretation relative to the capacity the institution intends to cultivate or assess.

This proposition also clarifies why education extends beyond efficient answer production. If a system supplies the correct answer immediately, the artifact may improve while the learner's opportunity to form a concept, test a hypothesis, tolerate uncertainty, or develop judgment may contract. In other situations, AI-generated counterexamples, simulations, explanations, or feedback can enlarge those opportunities. The educational question centers on the configuration through which the learner changes and on the role that automation occupies within that configuration.

9.7 Pedagogical Field Governance

The preceding distinctions suggest a broader role for pedagogy in the age of heterogeneous subjects. Teachers increasingly work with a field containing human learners, artificial mediators, institutional policies, data systems, assessment rules, and sometimes multiple interacting models. Their task can include direct teaching, yet it can also include designing relations among these elements so that educational aims remain intelligible and revisable.

Definition 9.9 (Pedagogical field governance). *Pedagogical field governance* is the deliberate organization and revision of relations among human learners, teachers, artificial systems, institutional rules, memories, and evaluative practices in order to sustain specified educational aims, appropriate distributions of authority, and conditions for meaningful learning.

The term “governance” is appropriate because several forms of control are at issue simultaneously. A teacher may regulate when AI enters a task, determine which system receives student data, decide which outputs require verification, select moments for human discussion, configure an assistant's instructions, and revise an assessment after observing new patterns of assistance. Institutions perform analogous work at a larger scale. Learners also participate by choosing when to delegate, when to challenge a system, and how to integrate external support into their own reasoning.

This framework preserves human educational aims within an explicitly heterogeneous field. Artificial participants can be pedagogically consequential while their consciousness and moral standing remain unsettled. Human learners can retain distinctive claims grounded in development, vulnerability, rights, and educational purpose while still participating in reciprocal adaptation with artificial systems. Teachers can shape AI behavior while leaving open whether the system qualifies as a student in the full human sense. The heterogeneity is therefore managed through explicit distinctions and role-sensitive arrangements across participants.

Remark 9.10 (Education under heterogeneous participation). The educational significance of AI lies partly in a change of object. Education is concerned with the formation of human learners, and it increasingly also requires the formation of the artificial and institutional conditions through which those learners encounter knowledge. The teacher can therefore become an educator of the learning environment as well as an educator within it.

This conclusion also prepares the move to aesthetics. Education concerns how participants learn to perceive significance, form judgment, and create or interpret symbolic objects. Once artificial systems enter those processes, the heterogeneous-subject problem extends from learning into aesthetic production and reception. The next section examines art and aesthetics under conditions in which generation, expression, experience, authorship, and judgment can be distributed across different participants, loosening their historical bundle within a single human creator or audience.

10 Aesthetics and Art beyond Human Homogeneity

This section examines aesthetics and art as domains in which the homogeneity approximation has been unusually deep because creation, reception, judgment, expression, and cultural transmission have commonly been organized around the human subject as their shared carrier. Its role is to extend the paper's

analysis from education into practices whose value often depends on relations among form, experience, intention, provenance, embodiment, and historical continuity. Artificial intelligence makes these components newly separable in practice. A system can generate formally complex artifacts, classify styles, recommend works, optimize toward evaluative criteria, and participate in creative workflows while the presence and character of aesthetic experience, expressive intention, authorship, and moral standing remain unsettled.

The objective is to develop an account of aesthetic heterogeneity that preserves this uncertainty while still describing the transformations already occurring in artistic practice. The discussion first disaggregates aesthetic functions that have often been bundled within the human artist or spectator. It then separates generation, expression, judgment, and experience. The third part examines provenance-sensitive aesthetic value, including cases in which the history and relation through which a work appears contribute to its value. The fourth distinguishes preservation of artifacts from preservation of artistic practices. The fifth considers the possibility of aesthetic worlds organized around perceptual and computational capacities unlike human ones. The section concludes with aesthetic field governance: the organization of cultural conditions under which heterogeneous participants generate, select, transmit, and evaluate forms.

10.1 Aesthetic Functions and the Human Bundle

Philosophical aesthetics already contains several distinct objects of analysis. The aesthetic can designate kinds of objects, judgments, attitudes, experiences, and values, and theories differ over which of these should receive explanatory priority [51]. Aesthetic experience itself has been analyzed through perception, imagination, emotion, conceptualization, and normativity, with substantial disagreement over which features are fundamental [52]. Theories of art likewise divide among aesthetic, historical, institutional, functional, and hybrid approaches [53]. This internal plurality becomes especially important once artificial systems participate in activities previously coordinated within human makers and audiences.

The human artist has often served as a convenient empirical bundle. One person could generate a form, possess intentions concerning it, experience the process of making, identify with the resulting work, accept responsibility for its presentation, and occupy an authorial position within an artistic tradition. Likewise, the human spectator could perceive a work, experience affective or imaginative responses, make aesthetic judgments, exchange reasons with others, and place the work within a historical or social context. The components were always conceptually separable, and collaborative, procedural, collective, and conceptual art already distributed many of them across several people and institutions. Generative AI increases the practical frequency and visibility of this distribution.

Definition 10.1 (Aesthetic-role disaggregation). *Aesthetic-role disaggregation* is the analytical separation of functions and statuses that can participate in an artistic or aesthetic process, including generation, selection, expression, experience, judgment, authorship, curation, interpretation, responsibility, and cultural transmission.

The definition follows the broader subject disaggregation developed in Section 3. It also preserves a distinction between performance and status. A system that selects among images according to a learned preference model performs an aesthetic-selection function. A system that generates variations in a style performs a generative function. These functional descriptions leave open whether the system experiences beauty, expresses an inner state, understands a tradition, bears responsibility for an artwork, or qualifies as an author in a legal or philosophical sense.

Table 7 presents the main dimensions relevant to this section. The table is designed as a diagnostic device with a narrower role than a unified theory of art. Different artistic traditions and theories can assign these dimensions different weights.

Table 7: Disaggregated dimensions of artistic and aesthetic participation. A single participant can occupy several dimensions, and one artistic process can distribute them across humans, artificial systems, institutions, and material infrastructures.

Dimension	Analytical focus	Heterogeneous-subject issue
Generation	Production of forms, sequences, images, sounds, texts, or performances	Form production can be technically sophisticated while experiential and authorial status remain open
Selection	Choice among alternatives, variants, styles, or candidate outputs	Selection can express judgment, optimization, convention, preference modeling, or delegated criteria
Expression	Relation between a work and an attitude, affect, perspective, intention, or condition attributed to a maker	Manifest expressive properties can exceed available evidence about an artificial participant’s inner life
Experience	Perceptual, affective, imaginative, or cognitive encounter with aesthetic objects or events	Human aesthetic experience is familiar while artificial aesthetic experience remains epistemically uncertain
Judgment	Evaluation, comparison, interpretation, reason-giving, or attribution of aesthetic properties	Behavioral competence can support evaluative functions while phenomenal and normative foundations vary
Authorship	Social, philosophical, institutional, or legal attribution of creative responsibility and recognition	Distributed production can separate causal contribution from recognized authorship
Provenance	History of making, transmission, encounter, embodiment, place, and relation surrounding a work	Similar manifest outputs can acquire different significance through different generative histories
Practice continuity	Reproduction of skills, bodily techniques, rituals, communities, and standards across time	Artifact replication can coexist with erosion or transformation of the practice that once generated the artifact
Reception	Interpretation, attention, recognition, criticism, and participation by audiences	Audiences can include differently constituted human and artificial participants with different perceptual capacities

Recent work on AI art already shows how these dimensions can diverge. McLoughlin argues that contemporary image-generating systems place pressure on inherited frameworks of artistic production and authorship [54]. A studio-level study by Bomba and De Angeli similarly reports artistic practices in which agency and authorship are treated as distributed across artists, models, data, and broader technical arrangements [55]. These observations support a relational description of the production process even where philosophical or legal authorship remains reserved for human agents.

Proposition 10.2 (Aesthetic function–status separation). *Performance of an aesthetic or artistic function supplies evidence about a participant’s role in an aesthetic process. Attribution of experience, authorship, responsibility, or moral standing requires additional grounds appropriate to that status.*

This proposition prevents a recurrent inference in both directions. Sophisticated form generation leaves the existence of aesthetic experience open; at the same time, uncertainty about experience remains compatible with a substantial causal role in the production, filtering, or circulation of aesthetic objects. The field can therefore be described with greater precision by asking which functions are present and which statuses remain uncertain.

10.2 Generation, Expression, and Experience

Generative systems make the separation among production, expression, and experience especially visible. A generated image can display stylistic unity, visual tension, melancholy, exuberance, irony, or apparent intentionality. A language model can produce a poem with a coherent voice. A music model can create passages that listeners describe through emotional or aesthetic vocabularies. These manifest properties matter to reception because artworks have long been assessed partly through properties accessible in the artifact or performance itself. At the same time, some theories of art and aesthetic experience attach significance to intention, expression, imagination, and the maker's relation to the work [52, 53].

Artificial generation therefore creates a three-way analytical separation. The first component is *formal production*: the capacity to produce organized patterns that enter aesthetic attention. The second is *expressive attribution*: the interpretation of those patterns as expressing a perspective, attitude, emotion, or sensibility. The third is *experienced making*: the possibility that the maker itself undergoes an aesthetic, affective, or meaning-laden process while creating. Human artistic practices often contain all three. Artificial participation can establish the first and support the second at the level of reception while leaving the third uncertain.

Definition 10.3 (Manifest aesthetic competence). *Manifest aesthetic competence* is the capacity of a participant to produce, discriminate, rank, transform, or explain forms in ways that reliably track standards used within an aesthetic practice, independently of any settled claim about the participant's phenomenal experience.

This definition helps clarify why the language of aesthetic judgment requires care. An artificial system may competently choose one composition over another, identify stylistic features, or predict audience preferences. These capacities can be educationally, commercially, and artistically consequential. They can also remain compatible with several different accounts of what occurs inside the system. The term *manifest* therefore marks the level at which the competence is publicly attributable.

The same distinction applies to creativity. Novelty, surprise, usefulness, transformation of a search space, intentional expression, and personal self-transformation represent different possible dimensions of creative activity. AI systems can participate strongly in some dimensions and weakly or uncertainly in others. Treating creativity as one indivisible property recreates the same bundling problem identified throughout this paper.

Institutional practice already reflects part of this separation. The U.S. Copyright Office's 2025 report on AI-generated material maintains a human authorship requirement while recognizing copyrightable human contributions in works that incorporate AI-generated material when sufficient human expressive choices are present [56]. Whatever one thinks of that legal rule, it illustrates an important conceptual point: causal contribution, generative capacity, expressive control, and recognized authorship can be allocated differently within one production process.

Remark 10.4 (Experience remains an open variable). The capacity to produce an artifact that elicits aesthetic experience in a human audience supplies evidence about the artifact and its generative process. It supplies limited evidence about whether the generating system itself has an aesthetic experience. Section 7's distinction between manifest behavior and stronger phenomenal attribution applies here with equal force.

This restraint also allows future evidence to matter. If stronger grounds for artificial experience emerge, aesthetic theory can incorporate them within the same analytical framework. If such grounds remain weak, the same framework still captures artificial systems as generative, selective, or curatorial participants.

10.3 Relational Provenance and Aesthetic Value

Aesthetic value frequently depends on more than the manifest properties of an artifact. An observer can care about who made an object, why it was made, where it was performed, what tradition it belongs

to, how it was encountered, or what kind of relation the act establishes between maker and audience. Theories of art that emphasize historical and institutional relations already recognize that classification and appreciation can depend on a work's position within an artworld or artistic history [53]. AI-generated artifacts make this dependence newly salient because very similar manifest outputs can arise through sharply different generative histories.

Consider a musical performance in a public square. The audible sequence can in principle be reproduced with high fidelity by a synthetic system. Yet the value of a human musician's performance may also include the fact that a person chooses to devote part of a finite evening to playing for those present, accepts the vulnerability of live performance, responds to the local environment, and creates a temporary relation with an audience. These properties belong to the event's provenance and relation structure. They can matter even when a recorded or synthetic version has superior technical precision.

The same structure appears in craft. A vessel produced by an automated process can closely resemble one produced through a long craft tradition. The two objects can share shape, material appearance, and practical function. The handmade object can additionally participate in a lineage of bodily technique, master-apprentice transmission, local material knowledge, ritual, and historical continuity. The additional value is therefore located partly in the relations through which the object comes to exist.

Definition 10.5 (Provenance-sensitive aesthetic value). *Provenance-sensitive aesthetic value* is aesthetic or artistic value that depends partly on the generative history, embodied practice, social relation, place, intention, transmission, or encounter through which an artifact or performance becomes available for appreciation.

The definition leaves room for provenance-insensitive value as well. A listener can value a sound pattern simply for its perceptual or formal qualities. A viewer can enjoy an image while remaining indifferent to its source. The purpose of the distinction is to prevent manifest similarity from being treated as a complete description of aesthetic equivalence.

Proposition 10.6 (Manifest similarity and relational divergence). *High similarity in observable artistic output remains compatible with substantial divergence in provenance-sensitive aesthetic value because the relevant object of appreciation can include relations, histories, commitments, and forms of presence surrounding the output.*

This proposition helps explain why the abundance of generated forms may increase attention to provenance. When formally competent images, music, and text become inexpensive to generate at scale, information about origin, process, relation, and commitment can become more important for some forms of cultural evaluation. The change resembles the scarcity migration identified in Section 5: abundance in one dimension can shift attention toward other scarce dimensions. Human time, embodied presence, trusted provenance, and historically situated practice can become relatively more salient even as formal content becomes more abundant.

Provenance sensitivity also clarifies why disclosure matters beyond copyright. Knowing that an artifact emerged through a live human performance, a collective studio practice, a generative model, a traditional craft process, or a hybrid workflow can change the audience's interpretation of the same manifest form. This effect belongs to aesthetic understanding as well as legal attribution.

10.4 Artistic Practice and Cultural Continuity

The distinction between artifact and practice becomes central when generative systems can reproduce outputs associated with a cultural tradition. Cultural preservation often aims at objects: recordings, manuscripts, instruments, images, architectural forms, designs, and archives. Many traditions, however, are also maintained through repeated embodied participation. Their continuity can depend on gestures, timing, tacit judgment, apprenticeship, collective memory, local materials, rituals of correction, and the social recognition of practitioners.

Definition 10.7 (Practice continuity). *Practice continuity* is the persistence of a culturally recognizable process of making, performing, interpreting, teaching, or evaluating across time through the reproduction and revision of the relations and competencies that sustain it.

Practice continuity differs from artifact continuity. A museum can preserve a finished object after the making practice has disappeared. A generative system can reproduce visual motifs after the community that developed them has lost control over their circulation. Conversely, a living practice can survive while its outputs evolve substantially. This distinction matters for the political economy of culture because the capacity to generate culturally recognizable forms can become detached from the communities whose histories and labor made those forms intelligible.

UNESCO's recent work on AI and culture has emphasized cultural value chains, cultural sovereignty, pluralism, and the protection of human creative agency in response to generative AI [57]. These concerns can be understood partly through practice continuity. Cultural diversity depends on more than a sufficiently varied inventory of outputs. It also depends on whether different communities retain capacities to create, transmit, interpret, revise, and contest the forms associated with their histories.

Proposition 10.8 (Artifact–practice separation). *Preservation or replication of artistic artifacts can co-exist with transformation or decline in the practices that historically generated them. Cultural continuity therefore requires separate attention to artifacts and to the relations through which artistic competencies are reproduced.*

This proposition also changes the meaning of assistance. An AI system can help a practitioner document techniques, explore variants, translate explanations, or preserve archives. The same system can participate in a platform economy that reduces demand for the practitioner or detaches recognizable styles from their social history. The ethical and cultural evaluation of AI in art therefore depends on the surrounding field alongside output quality.

The heterogeneous-subject perspective adds another layer. Some cultural practices can eventually include artificial participants as routine members of the production process. A tradition may then incorporate human-artificial co-creation into its continuing reproduction. Such evolution changes the composition of the practice itself and can combine continuity with substantial transformation.

10.5 Heterogeneous Aesthetic Worlds

The strongest version of aesthetic heterogeneity appears when different subject types possess different perceptual, temporal, or computational capacities. Human aesthetic traditions are shaped by human bodies: ranges of hearing and vision, motor capacities, memory limits, attention spans, emotional systems, life histories, and social forms. Even abstract arts are encountered through human cognitive and cultural capacities. Artificial systems can operate with very different input spaces, temporal scales, representational dimensions, and forms of memory.

This possibility should be approached cautiously because current artificial systems' phenomenal status remains uncertain. The relevant point can be made at the functional level. An artificial system can already detect regularities across dimensions beyond direct human inspection, rank structures through criteria learned from very large datasets, or operate at temporal and informational scales unavailable to unaided human perception. Future systems could develop stable selection practices around features that remain largely opaque to human audiences.

Definition 10.9 (Heterogeneous aesthetic world). *A heterogeneous aesthetic world* is a field of generation, discrimination, selection, interpretation, or experience in which participants with materially different perceptual or cognitive architectures orient toward forms through partly different spaces of accessible features and relations.

The definition deliberately includes both experience-rich and functionally specified cases. If an artificial participant someday has aesthetic experience, the definition can describe a world in which its phenomenal access differs from ours. Under present uncertainty, the same definition can describe a system of machine selection and generation organized around features that exceed ordinary human inspection.

This possibility alters the traditional direction of adaptation. Much of machine-generated art currently targets human audiences and therefore learns from human cultural archives, preferences, ratings, and styles. A heterogeneous aesthetic world could also contain artifacts optimized for artificial receivers, mixed audiences, or translation between perceptual architectures. Human observers might encounter only projections of structures that are much higher-dimensional or temporally denser in the generating system. Artificial systems might likewise receive compressed or translated representations of embodied human performance.

Proposition 10.10 (Cross-architecture aesthetic translation). *When participants access different spaces of aesthetically relevant features, shared aesthetic practice increasingly depends on translation mechanisms that map forms, saliences, or evaluative structures across perceptual and cognitive architectures.*

This proposition extends heterogeneous-subject theory beyond coexistence. A shared world among heterogeneous subjects may require active conversion between ways of sensing, organizing, and evaluating form. Aesthetics therefore offers a particularly clear example of a broader problem developed in the paper's later synthesis: common institutions can coordinate participants with different forms of access to the world through mechanisms that make those differences mutually consequential and partly interpretable.

10.6 Aesthetic Field Governance

The preceding analysis shifts the normative focus from the isolated question of whether an AI can create art toward the organization of the field in which artistic forms are generated, selected, attributed, circulated, and preserved. This shift parallels the movement from individual educational tools to pedagogical field governance in Section 9. Cultural outcomes are shaped by model design, training corpora, platform incentives, disclosure practices, curatorial systems, copyright rules, archives, museums, educational institutions, labor markets, audience attention, and the survival of communities of practice.

Definition 10.11 (Aesthetic field governance). *Aesthetic field governance* is the design and revision of institutional, technical, economic, and cultural conditions that shape which participants can generate, interpret, transmit, receive recognition for, and sustain artistic forms within a heterogeneous aesthetic environment.

Aesthetic field governance contains at least four tasks. The first is *provenance governance*: maintaining sufficiently trustworthy information about the processes and participants through which works are generated. The second is *practice governance*: sustaining conditions under which human and community practices can continue, evolve, and transmit their competencies. The third is *attention governance*: organizing discovery and curation in a setting where generated cultural supply can expand far more rapidly than human attention. The fourth is *translation governance*: developing means through which differently constituted participants can encounter, interpret, and respond to forms generated outside their native perceptual or cognitive range.

These tasks clarify why aesthetics belongs within the broader argument of this paper. AI does more than add a new instrument to artistic production. It makes visible the extent to which modern aesthetics has often allowed one empirical kind—the human subject—to carry generation, expression, experience, judgment, authorship, and reception together. Once these dimensions are redistributed across heterogeneous participants, aesthetic theory requires the same conceptual move already encountered in political economy, justice, ethics, law, and education: roles must be described before they are collapsed into a single status category.

The result is a broader conception of artistic value. Formal qualities remain important. So do experience, relation, provenance, historical continuity, embodied practice, and the organization of attention. AI-generated abundance can therefore coexist with a renewed importance of human presence and cultural lineage. A technically reproducible form can still be embedded in an irreproducible event of giving,

risk, encounter, or transmission. At the same time, artificial participation can expand the space of possible forms and may eventually contribute to aesthetic worlds whose structures extend beyond human perceptual organization.

This dual movement—toward stronger attention to relational provenance and toward wider spaces of heterogeneous form—prepares the transition to subjectivity. If artistic and aesthetic life is understood through more than outputs produced and consumed by homogeneous human subjects, then the question of human identity after the centrality of labor also shifts from replacement toward how persons choose to inhabit, value, and contribute to a world increasingly populated by other forms of agency.

11 Subjectivity after the Centrality of Labor

This section examines the consequences of heterogeneous production for human subject formation. Its role is to separate a question of economic necessity from a question of human activity. Sections 4 and 5 considered how artificial systems can participate in production and how distribution can become less tightly coupled to direct human labor. The present section asks what happens to identity, temporal organization, recognition, contribution, and meaningful practice if paid employment becomes a less comprehensive institution in human life. The analysis therefore concerns labor centrality and its reconfiguration under conditions of heterogeneous production.

The objective is to identify the functions that employment has historically bundled and to examine how those functions might be redistributed across other activities and institutions. The discussion first distinguishes employment from work and broader forms of activity. It then defines the labor-centrality bundle, which joins livelihood, time structure, recognition, social contact, identity, and contribution within a single institution. The third part distinguishes instrumental labor from activities whose value partly inheres in their practice. The fourth examines the temporal infrastructures required for expanded freedom from necessary labor. The fifth develops the idea of recognition migration: the movement of social recognition and contribution away from employment as their primary institutional carrier. The final part connects labor decentering to the paper's heterogeneous-subject thesis by arguing that human value can remain grounded independently of comparative productive superiority over artificial systems.

11.1 Employment, Work, and Human Activity

Employment has become so central to modern social organization that several different categories are often compressed into a single word: work. Yet work, employment, production, contribution, practice, and activity can refer to different relations. Philosophical treatments of work explicitly distinguish employment from wider forms of productive activity and note that work can carry exchange value, social value, and first-personal forms of meaning [58]. The distinction matters because automation can alter the amount of employment required for production while leaving human activity abundant.

Arendt's distinction among labor, work, and action offers one influential example of a theory that resists treating human activity as a single productive category [59]. The present discussion uses Arendt's distinction methodologically, while the hierarchical ordering among activities remains outside the argument: forms of activity can have different relations to necessity, durability, public appearance, and shared worlds. Likewise, Marxian traditions distinguish labor performed under conditions of necessity or alienation from productive activity capable of supporting self-realization. Contemporary philosophical summaries of this tradition emphasize that Marx associated human flourishing with forms of freely chosen, creative, socially recognized activity while also envisioning a reduction in the realm of necessary labor [60].

The distinction can be expressed at a modest level. Some activities are pursued mainly because they mediate access to other goods. A person takes a job in order to obtain income, and income makes housing, food, mobility, education, or research possible. Other activities derive substantial value from their performance: a craftsman develops a technique, a musician performs for an audience, a researcher

follows a difficult question, or a person cares for someone with whom they share a relationship. The same activity can combine these orientations, and its profile can change across time.

Definition 11.1 (Activity-value profile). An *activity-value profile* is the configuration of reasons through which an activity acquires value for a participant or community, including exchange, instrumental, intrinsic, relational, civic, expressive, transmissive, and recognitional dimensions.

The definition avoids assigning an activity to a single category. Professional music can provide income, aesthetic experience, social recognition, cultural transmission, and intrinsic satisfaction at once. Care can be paid employment, a familial relation, a civic contribution, and an expression of attachment. Research can support a career while also expressing curiosity whose value persists independently of remuneration. The relevant transformation in an AI-mediated economy is therefore a change in the composition and institutional weight of these dimensions.

11.2 The Labor-Centrality Bundle

Paid employment does more than distribute income. Research on unemployment has long emphasized that employment can provide regular time structure, social contacts outside the household, participation in collective purposes, activity, status, and identity alongside its manifest economic function [61]. Recent empirical work continues to examine these latent integrative functions, including time structure, social contact, joint goals, collective identity, and activation [62]. Political theory has likewise questioned the extent to which waged work has become a primary mechanism for income distribution, social obligation, and the definition of social subjects [63].

These observations suggest that the centrality of employment is partly a bundling phenomenon. One institution simultaneously performs several functions that could, in principle, be organized separately.

Definition 11.2 (Labor-centrality bundle). The *labor-centrality bundle* is the institutional concentration of livelihood access, temporal structure, social contact, recognized contribution, status, skill development, public identity, and recurrent participation within paid employment.

Table 8 disaggregates this bundle. The table is a structural inventory, while the quality of each function varies across jobs. Employment can generate social isolation, domination, exhaustion, insecurity, or stigmatized identity. The point is structural: modern institutions often route several important forms of access and recognition through the employment relation even when the quality of that relation varies greatly.

Table 8: Functions frequently concentrated in the labor-centrality bundle and possible institutional carriers under conditions of labor decentering. The alternatives form an open analytical repertoire for comparison.

Function	Employment-centered carrier	Potential plural carriers
Livelihood	Wage, salary, benefits, occupational access to resources	Public provision, social income, asset income, cooperative claims, household and community arrangements
Time structure	Workday, shift, commute, deadlines, career stages	Education, care, civic schedules, artistic practice, research communities, voluntary associations, self-organized routines
Social contact	Coworkers, clients, professional networks	Neighborhoods, clubs, cultural institutions, learning communities, care networks, public spaces
Recognition	Occupational status, promotion, professional credentials	Civic contribution, craft mastery, scholarship, care, artistic practice, community stewardship, peer recognition
Identity	Profession, employer, rank, vocation, career trajectory	Plural affiliations, projects, practices, relationships, places, communities, disciplines, forms of life
Contribution	Measured output, service delivery, organizational goals	Care, maintenance, cultural transmission, inquiry, ecological stewardship, public participation, creation
Skill development	Workplace training, occupational progression, accumulated expertise	Apprenticeship, public education, maker communities, studios, laboratories, civic and cultural institutions
Agency	Role-defined responsibility and discretion within organizations	Self-directed projects, collective governance, participatory institutions, independent practice, association

Once production becomes more heterogeneous, this bundle can loosen from two directions. First, automation can reduce the amount of human labor required for some outputs. Second, artificial agents can perform functions previously associated with human workers even where institutions continue to distribute income and recognition through human employment. The resulting mismatch is larger than a labor-market adjustment. It can become a subjectivity problem because the institution supplying livelihood has also supplied temporal, relational, and recognitional structure.

Definition 11.3 (Labor-subjectivity decoupling). *Labor-subjectivity decoupling* is the process through which paid employment becomes a less comprehensive carrier of livelihood, identity, recognition, temporal organization, social contact, and contribution, allowing those functions to become institutionally separable.

This concept identifies both an emancipatory possibility and a transition burden. A person may gain time previously committed to necessary employment while simultaneously losing a familiar source of routine, social contact, recognition, and narrative continuity. Institutions that reduce economic compulsion while leaving the remaining functions unsupported can therefore produce freedom together with disorientation. The issue concerns the ecology of subject formation together with the quantity and quality of discretionary time.

11.3 Necessary Labor and Meaningful Practice

A reduction in necessary labor can be interpreted in two very different ways. One interpretation imagines human inactivity: machines perform the tasks and people cease doing things. A second interpretation separates necessity from activity. Human beings may continue to make, study, care, cultivate, perform, repair, teach, explore, deliberate, and play because the activity itself carries forms of value that survive the weakening of economic compulsion.

This distinction has long precedents. Marxian discussions of the “realm of freedom” connect expanded freedom to the reduction of labor imposed by necessity while preserving the development and exercise of human capacities as ends in themselves [60]. Sennett’s account of craftsmanship describes a motivation to do a task well for its own sake and connects skilled making to the development of practical judgment [64]. Self-determination theory similarly emphasizes the importance of autonomy, competence, and relatedness for intrinsic motivation and well-being [65]. These traditions differ substantially, yet they converge on a useful analytical point: human motivation exceeds the external reward attached to an activity.

Definition 11.4 (Meaning-bearing practice). *A meaning-bearing practice is a recurrent activity whose value to its participants partly arises through the exercise, development, relation, experience, expression, or continuity enacted in the practice itself.*

The definition includes a craftsperson preserving a manual tradition, a philosopher attending to a difficult question, a musician performing in a public space, a caregiver sustaining a relationship, a naturalist observing a local ecosystem, or a mathematician exploring a conjecture whose immediate market value is unclear. These examples differ in social function and normative significance. Their common feature is that the value of participation cannot be reduced to acquisition of the final output.

The value of craftsmanship can exceed the manufactured object when making cultivates skill, bodily knowledge, lineage, and relation. The value of a person choosing to spend finite time performing with others can exceed the value of a generated musical recording. The value of inquiry can likewise exceed a complete database of philosophical propositions when reflection itself forms a mode of attention and self-formation. These distinctions parallel the provenance and practice-continuity arguments developed in Section 10.

Proposition 11.5 (Necessity–activity separation). *A decline in the amount of labor required for material or informational production can reduce economic compulsion while leaving substantial forms of human activity, practice, contribution, and self-development intact.*

The proposition avoids treating current automation trajectories as evidence for a specific end state. It identifies a conceptual possibility that matters even under partial automation. As some activities become cheaper to delegate, their continued human performance can reveal which dimensions of value were carried by participation itself. AI therefore acts here as a diagnostic condition: it can make visible the difference between obtaining an output and inhabiting a practice.

11.4 Temporal Infrastructures of Freedom

Expanded discretionary time is often presented as an uncomplicated benefit of labor-saving technology. Yet free time is also an institutional achievement. Time can support inquiry, care, artistic practice, civic life, rest, or social participation when people possess material security, accessible spaces, educational resources, social invitations, and sufficient control over their schedules. The same quantity of formally unallocated time can be experienced very differently under insecurity, isolation, illness, surveillance, or status loss.

The labor-centrality bundle clarifies the transition. Employment imposes constraints, but it also coordinates calendars, creates recurring encounters, and supplies recognizable sequences of effort and completion. If some of those functions weaken, human freedom becomes more dependent on alternative temporal institutions. Libraries, workshops, laboratories, parks, clubs, schools, studios, community

centers, care networks, local associations, and digital commons can all organize repeated participation while remaining institutionally distinct from the full employment relation.

Definition 11.6 (Temporal infrastructure of freedom). A *temporal infrastructure of freedom* is a set of material, social, and institutional conditions that enables discretionary time to become usable for self-directed, relational, civic, restorative, or meaning-bearing activity.

The concept emphasizes that freedom from a schedule and freedom to pursue a practice are different institutional achievements. A shorter workweek can expand discretionary time; access to cultural, educational, relational, and material infrastructures shapes what that time can become. This distinction also limits romantic accounts of a post-work future. The release of time from necessary labor creates an open field of possibilities whose realization depends on distribution, education, health, public space, and forms of association.

The educational argument of Section 9 becomes important here. If learning increasingly includes the cultivation of curiosity, judgment, question formation, and the capacity to organize one's own inquiry, education also becomes preparation for discretionary time. A society centered less strongly on employment would require capacities for self-directed activity that employment-centered education has sometimes treated as secondary to occupational preparation.

11.5 Recognition Migration and Plural Contribution

The weakening of labor centrality also changes the problem of recognition. Occupational societies use jobs to answer several social questions quickly: what a person does, which competencies they possess, how they contribute, and where they stand within a hierarchy. These answers can be reductive, but they remain institutionally legible. Activities outside employment often lack the same visibility. Unpaid care, neighborhood maintenance, amateur scholarship, open cultural work, ecological stewardship, and informal teaching can be socially consequential while receiving limited formal recognition.

Definition 11.7 (Recognition migration). *Recognition migration* is the redistribution of social recognition, status, contribution narratives, and identity markers from employment toward a plural set of practices, relations, and institutions.

Recognition migration is distinct from the replacement of occupational status with a single alternative hierarchy. A labor-decentered society can support plural forms of recognition whose standards vary by practice. Craftsmanship can be recognized through mastery and transmission; care through reliability and relation; scholarship through inquiry and criticism; civic participation through stewardship and deliberation; art through creation, interpretation, and cultural contribution. The multiplicity matters because it reduces the pressure on one institution to define the worth of the whole person.

This pluralization also bears on the political economy of AI. If artificial systems outperform humans on some dimensions of productive speed, memory, optimization, or formal generation, a social order that equates contribution with comparative productivity can convert technical differences directly into status differences. A broader account of contribution resists that compression. Human participation can remain valuable through embodiment, care, historical continuity, responsibility, finite commitment, shared risk, local knowledge, public presence, or the intrinsic value of practice even when an artificial system can produce a comparable output more efficiently.

Proposition 11.8 (Contribution pluralism). *Under conditions of heterogeneous production, socially valuable contribution can be assessed through multiple practice-sensitive dimensions, with productive efficiency forming one dimension among several in comparisons among human and artificial participants.*

The proposition also allows artificial systems to contribute to shared projects in substantial ways. Its purpose is to prevent productive superiority in one dimension from becoming a general measure of subject value. This is the same anti-compression principle that has recurred throughout the paper: legal competence, moral patience, aesthetic experience, pedagogical authority, economic role, and productive capacity each require their own analysis.

11.6 Subject Formation in a Labor-Decentered Society

The deepest consequence of labor decentering concerns the narrative form of the subject. Employment offers a ready-made temporal narrative: education prepares for work, work develops into a career, achievement is marked by promotion or professional recognition, and retirement closes the occupational sequence. The narrative can constrain individuals and exclude many lives, yet it offers a widely recognized grammar through which people explain themselves to others.

A less employment-centered society would increase the plurality of possible life narratives. Identity could be organized more strongly around practices, relationships, places, inquiries, communities, care commitments, artistic traditions, civic roles, or changing combinations of these. Such plurality can expand autonomy while also increasing the demand for judgment. A subject with more discretionary time and fewer externally imposed scripts must make more choices about which activities deserve sustained attention.

Definition 11.9 (Labor-decentered subjectivity). *Labor-decentered subjectivity* is a mode of subject formation in which paid employment becomes one possible source among several for livelihood, identity, recognition, temporal organization, contribution, and meaningful practice.

This definition describes a shift in relative centrality while retaining work as one possible activity among several. Many people may continue to value employment strongly, and some forms of work may become more meaningful when economic compulsion decreases. Other people may organize their lives around care, scholarship, craft, public service, exploration, or artistic practice. The heterogeneous-subject approach therefore supports institutional plurality at the level of human lives as well as at the level of subject types.

A final implication concerns comparative identity. Human beings have often understood themselves partly through capacities regarded as distinctively human: reason, language, creativity, technical production, judgment, or symbolic expression. Artificial systems can destabilize such boundaries by exhibiting functional competence in several of these domains. A strategy based on locating one permanently exclusive human capacity is therefore fragile. Labor-decentered subjectivity offers another path. Human significance can be grounded in lived relations, finite commitments, embodied histories, practices, responsibilities, and forms of participation whose value remains intelligible independently of comparative productive efficiency.

The possibility returns the argument to the image that motivated this section: a world in which more necessary production is delegated while human beings have more time for inquiry, craft, music, care, public life, or contemplation. The important theoretical issue is neither idleness nor technological triumph. It is the reorganization of the institutions through which time becomes a life, activity becomes contribution, and participation becomes identity. Artificial intelligence matters because it can loosen an arrangement that modern societies have often treated as natural: the concentration of subsistence, recognition, identity, and social contribution within paid labor.

The transition to the next section follows from this point. Once human worth, agency, and identity are analytically separated from a monopoly on productive labor, social theory can approach human and artificial participants through differentiated capacities, relations, and institutional roles. The resulting task is broader than a theory of work. It is the construction of social theory for a shared world whose participants vary in ontology, agency, vulnerability, temporality, and modes of contribution.

12 From Human-Centered Social Theory to Heterogeneous-Subject Theory

This section consolidates the domain-specific analyses developed throughout the paper into a general methodological framework for social theory. Its role is to identify which analytical operations recur across political economy, distribution, justice, ethics, jurisprudence, education, aesthetics, and subject formation, and to state how those operations can be combined when the participants in a shared social

world differ substantially in ontology, capacity, vulnerability, continuity, and institutional position. The section therefore shifts from individual domains to the architecture of inquiry itself.

The objective concerns methodological architecture while comprehensive classification remains open. Heterogeneous-subject analysis asks which properties and relations matter for a particular institutional problem, which conclusions remain uncertain, and which forms of attribution can be justified at the level of the relation under study. The discussion first clarifies the methodological meaning of human-centered social theory used in this paper. It then defines heterogeneous-subject theory, synthesizes the recurring separations established in earlier sections, develops relation profiles and cross-domain status profiles, and states several principles for institutional analysis under persistent heterogeneity. The final part identifies the limits and open research programme that follow from the framework.

12.1 Methodological Human-Centering in Social Theory

The expression *human-centered social theory* can refer to several very different commitments. Some are explicitly moral, such as assigning special importance to human welfare or human rights. Others are ontological, treating social reality as constituted through human meanings and practices. The present paper uses the expression in a narrower methodological sense. It concerns the common possibility of designing concepts for institutions whose principal participants are assumed to belong to the same broad kind of embodied, mortal, socially formed human subject.

Section 1 described this convenience through the *homogeneity approximation*. The approximation remained compatible with extensive empirical differences among human beings. Human societies contain extensive variation in ability, age, dependence, language, social position, vulnerability, and forms of life. Its simplifying force came from another source: many institutions could treat their central participants as sharing enough background features that a single vocabulary of personhood, agency, responsibility, labor, learning, preference, and participation remained workable across large parts of the social field.

Artificial systems weaken that convenience because several capacities that were often co-located in human participants can appear in unfamiliar combinations. A system may communicate fluently while its experience remains uncertain. It may participate in production while its distributive standing remains unsettled. It may preserve interaction histories while lacking the bodily continuity characteristic of human identity. It may perform pedagogical or aesthetic functions while the authority or experiential status associated with those functions remains open. The analytical difficulty therefore arises from the composition of capacities and relations across multiple dimensions.

This point also clarifies the scope of the proposed transition. Moving beyond a human-centered methodological approximation remains compatible with strong protection of human beings. Human vulnerability, embodiment, mortality, historical experience, democratic membership, and established rights can remain decisive grounds within particular normative arguments. The methodological change concerns the population of entities that social theory must be able to describe and govern. A theory can remain strongly protective of human interests while accepting that the institutional environment contains participants whose properties depart from the familiar human bundle.

Definition 12.1 (Heterogeneous-subject condition). A *heterogeneous-subject condition* exists when a shared institutional field contains participants whose relevant capacities, dependencies, continuities, modes of manifestation, or normative statuses differ enough that a single bundled model of the subject ceases to provide an adequate basis for analysis across the field.

The definition is relational and institutional. A difference matters when it changes the description, attribution, justification, or governance of an interaction. Two entities can differ greatly in physical constitution while occupying functionally similar positions for one institutional purpose. The same entities can require sharply different treatment in another domain because vulnerability, continuity, accountability, or standing becomes salient there. Heterogeneity therefore concerns the structure of a field as much as the properties of isolated entities.

12.2 A General Architecture of Analytical Separation

The preceding sections repeatedly reached the same methodological result from different directions: properties that were once easy to infer together need to be separated and reconnected through explicit argument. Table 9 collects the principal separations developed across the paper. Their recurrence suggests that the disaggregation established in Section 3 operates as a general architecture across domains.

DRAFT

Table 9: Recurring analytical separations across the heterogeneous-subject framework. Each separation identifies dimensions that can interact strongly while still requiring distinct grounds for attribution or evaluation.

Domain	Analytical separation	Institutional significance
General subject analysis	Agency, sentience, vulnerability, continuity, responsibility, standing, learning, expression, productive capacity	A subject profile can vary across dimensions; competence in one dimension supplies only bounded evidence for another
Political economy	Productive function, ownership, management, infrastructure, agency, entitlement	Economic role can be relation-specific while ownership and distributive claims require separate institutional grounds
Distribution	Productive contribution, allocation mechanism, distributive entitlement	Changes in production alter available resources while distributive rules determine claims and access
Justice	Procedural participation, representation, substantive standing, reciprocity	Capacity to participate in justificatory procedures and qualification for protection can diverge
Ethics	Moral agency, moral patiency, relational obligation, institutional precaution	Rule-following or reason-responsive behavior and capacity for morally considerable experience require separate assessment
Jurisprudence	Manifested agency, legal attribution, legal personality, continuity, remedy	Responsibility and loss can be allocated before comprehensive personhood is settled
Education	Pedagogical function, learning, educational authority, responsibility, memory	Explanation or evaluation competence can be distributed among heterogeneous participants while educational aims remain separately governed
Aesthetics	Generation, expression, judgment, experience, authorship, provenance, practice continuity	Formal competence can coexist with uncertainty about experience, while relational origin and practice history remain aesthetically salient
Subjectivity and labor	Production, livelihood, recognition, identity, time structure, meaningful practice	Reduced labor necessity can loosen the institutional bundle through which employment organizes life and social worth

The table reveals a common pattern. Social theory often encounters a visible or operational competence first: producing an output, entering an exchange, following a rule, explaining a concept, generating an image, or sustaining an interaction. The institutional temptation is to translate that competence into a comprehensive status judgment. The heterogeneous-subject framework slows that translation. It treats the observed competence as evidence about a specific dimension and then asks what additional premises would be required to move toward responsibility, authority, entitlement, personhood, or moral standing.

Proposition 12.2 (Cross-domain anti-compression). *Where an entity's capacities and statuses can vary independently across institutional domains, competence or recognition in one domain supplies grounds for a broader status conclusion only through an additional argument connecting the relevant dimensions.*

This proposition generalizes the cross-dimensional restraint introduced in Section 3. Its purpose is

epistemic and institutional. It limits both expansive and restrictive inferences. Strong language performance, for example, supplies bounded evidence about sentience or legal personhood. At the same time, uncertainty about sentience is compatible with an artificial system's causal participation in production, education, or legal transactions. A system can therefore matter greatly to institutional analysis while remaining unsettled along other dimensions.

The same rule applies to human participants. Limited contractual competence can coexist with substantive standing; limited market productivity can coexist with social contribution; informal participants can still teach; and limited ability to articulate reasons can coexist with vulnerability and claims to protection. Heterogeneous-subject theory is therefore broader than AI theory. AI intensifies the need for the method because it makes unusual combinations of capacities more common and socially consequential.

12.3 Subject Profiles, Relation Profiles, and Institutional Roles

Disaggregation generates a further problem. A list of capacities attached to an entity remains incomplete because institutions govern interactions. A highly autonomous artificial system can occupy different positions when it assists a student, manages a logistics process, negotiates a contract, recommends a medical action, generates a work of art, or coordinates with another artificial agent. The subject profile described in Section 3 must therefore be paired with a profile of the relation in which the subject is participating.

Definition 12.3 (Relation profile). A *relation profile* is a structured description of an interaction that identifies the participating entities, their relevant capacities and dependencies, the resources or interests at stake, the direction of influence, the applicable institutional roles, the attribution rules governing outcomes, and the uncertainties material to the relation.

A relation profile changes the order of analysis. The first task is to identify what is happening between participants and which institutional function is at stake. Ontological classification remains relevant, yet its relevance becomes specific. Biological embodiment may be central in a care relation, mortality may matter for temporal justice, ownership may matter in production, traceability may matter for liability, and persistent interaction history may matter for educational continuity. The analytical weight of a property emerges from the relation in which it operates.

This approach also provides a clearer interpretation of the statement made in Section 4 that economic categories are relational. An AI system can be owned as an asset and simultaneously perform labor-like or managerial functions. These descriptions concern different relations. A comparable multiplicity appears in law, where a system can be an object of property rights, a technical component in a regulated system, and a causal participant whose outputs require attribution. Education similarly allows an AI system to be a licensed service, a pedagogical mediator, a repository of interaction history, and an object of teacher-directed shaping.

Proposition 12.4 (Relation-before-generalization principle). *Institutional analysis under heterogeneous participation should establish the relation profile and the domain-relevant subject dimensions before extending a local functional description into a general conclusion about the entity's status.*

The proposition gives a disciplined form to a pattern already used throughout the paper. Calling a system “worker-like” within a production relation can be informative when the term refers to task performance, coordination, or substitution for labor. The description becomes misleading when it silently imports the complete legal, political, and moral status of a human employee. Likewise, calling an AI system a “learner” can identify adaptive behavior or context-sensitive updating, while pedagogical vulnerability, developmental needs, and educational rights remain separate matters.

Relation-first analysis also reduces pressure to establish one universal answer to the statement “AI is a subject.” The category *subject* can serve as a research placeholder for entities that participate in consequential relations, while more specific forms of status are assigned at the level where their grounds can be examined. This approach preserves conceptual openness and keeps ontological uncertainty proportionate to the issue under study.

12.4 Cross-Domain Status Profiles

A relation-first approach still requires coordination across institutions. A single artificial system can move among domains, and decisions in one domain can affect another. A system granted control over financial resources can gain productive influence; a pedagogical system retaining long-term memory can raise legal and privacy questions; an artificial agent recognized as a contractual interface can acquire responsibilities whose enforcement depends on ownership and technical control. Heterogeneous-subject theory therefore needs a way to represent institutional multiplicity while preserving the distinctions already established.

Definition 12.5 (Cross-domain status profile). A *cross-domain status profile* is the set of domain-specific roles, permissions, protections, responsibilities, attribution rules, and unresolved status questions associated with an entity across the institutional fields in which it participates.

The profile is modular. Legal attribution can be specified while moral patiency remains uncertain. Economic permissions can be granted while political membership is absent. Pedagogical functions can be authorized while final responsibility remains with a teacher or institution. Aesthetic authorship can be allocated to human contributors under existing rules while the artificial system's generative contribution remains descriptively acknowledged. Modularity makes institutional coordination possible through several interoperable domain-specific categories.

This architecture can be understood as a form of controlled translation among domains. The translation asks what information from one profile is relevant to another and which additional rule is required for transfer. For example, high autonomy in a production system may increase the importance of traceability for legal attribution. Persistent memory in education may increase the importance of continuity and governance. Apparent preference stability may become evidence relevant to moral-status research while retaining substantial uncertainty about experience. Each translation remains explicit about the inferential bridge.

Remark 12.6 (Institutional plurality under shared identity). A stable name, interface, or technical identifier can support continuity across interactions while different institutions assign different meanings to that continuity. Economic continuity may track control of assets and contracts; pedagogical continuity may track learning history; legal continuity may track responsibility and authorized identity; moral continuity may depend on further claims about experience or interests. Shared identity markers therefore support coordination while leaving the grounds of each status domain-specific.

Cross-domain profiling also helps manage change through time. Artificial systems can be updated, copied, merged, retrained, transferred, or connected to new tools. Human participants likewise change capacities, dependencies, roles, and affiliations. A heterogeneous-subject framework can treat status as revisable where the underlying relation or capacity changes, while preserving stability where institutional continuity requires it. The result is structured revision tied to relevant changes in the profile, combining continuity where institutions require it with change where the underlying relation or capacity changes.

12.5 Institutional Treatment of Uncertainty

Earlier sections treated uncertainty most explicitly in ethics and law, yet it appears throughout the paper. Political economy can observe productive capacity while future substitution effects remain uncertain. Distribution can identify scarcity migration while the institutional destination of new scarcities remains open. Education can observe pedagogical effectiveness while long-term developmental effects are still being studied. Aesthetics can observe formal competence while experiential claims remain unsettled. Heterogeneous-subject theory therefore treats uncertainty as part of the subject and relation profile as a constitutive variable in the model.

Three forms of uncertainty recur. *Empirical uncertainty* concerns what a system can do or what properties it instantiates. *Interpretive uncertainty* concerns how observable behavior should be understood, including whether apparently familiar categories such as intention, preference, learning, or memory apply

in the same sense. *Normative uncertainty* concerns which properties justify rights, protections, authority, responsibility, or entitlement. Institutions can face all three at once.

Definition 12.7 (Status-relevant uncertainty). *Status-relevant uncertainty* is uncertainty about an empirical, interpretive, or normative proposition whose resolution could materially alter an entity's institutional role, protection, responsibility, permission, or entitlement.

The definition supplies a practical threshold. Social theory can focus on the unknowns capable of changing institutional treatment, giving priority to uncertainties with consequences for roles, protections, permissions, or attributions. Whether a system has a particular internal architecture may be irrelevant in one domain and decisive in another. Whether interaction history persists across updates can matter for contract continuity, educational memory, or attribution. Whether experience is possible can matter profoundly for moral standing while contributing little to a narrow description of current productive output.

The calibrated moral precaution of Section 7 can be generalized into an institutional practice of reversible design. Where status-relevant uncertainty is substantial, institutions can prefer arrangements that preserve information, enable later revision, and avoid creating unnecessary path dependence. Logging, provenance, auditability, modular permissions, reviewable delegation, and explicit responsibility allocation can all serve this purpose. Their value arises partly from keeping institutional learning possible as the empirical and normative picture changes.

Proposition 12.8 (Revisability under status-relevant uncertainty). *Where plausible resolution of a status-relevant uncertainty would materially change institutional treatment, governance should preserve a practicable path for revising roles, protections, permissions, and attributions as evidence and normative judgment develop.*

Revisability also protects against premature closure in both directions. An institution that permanently assigns full personhood on thin evidence can create difficult conflicts of rights and control. An institution that fixes an entity permanently as mere property can become equally rigid if later evidence supports morally relevant forms of experience or continuity. The framework therefore favors explicit provisionality where the grounds for status remain open and the cost of later correction could be high.

12.6 Institutional Method of Heterogeneous-Subject Theory

The previous elements can now be assembled into a working definition. The aim is to provide a portable method for domains in which participants differ in the properties that social theory once found convenient to bundle.

Definition 12.9 (Heterogeneous-subject theory). *Heterogeneous-subject theory* is an approach to social analysis that represents socially consequential participants through differentiated subject profiles, relation profiles, and domain-specific status profiles; separates functional competence from broader normative and institutional standing; tracks status-relevant uncertainty; and designs attribution and governance for shared institutions while leaving subject homogeneity open to examination.

This definition yields a sequence of analytical operations. First, identify the institutional field and the participants whose behavior can materially shape it. Second, disaggregate the capacities relevant to the problem. Third, construct the relation profile: dependencies, resources, direction of influence, applicable roles, and affected interests. Fourth, specify the domain-specific status at issue, such as legal attribution, distributive entitlement, pedagogical authority, moral protection, or aesthetic authorship. Fifth, identify the inferential bridges connecting observed functions to the proposed status. Sixth, record status-relevant uncertainties and the consequences of different resolutions. Seventh, select governance arrangements that remain coherent across domains and revisable where the underlying profile can change.

This sequence shifts the central question from categorical membership toward institutional fit. A theory can ask what form of agency is present, what kind of vulnerability is relevant, which party controls

resources, who can revise the system, whose interests are affected, which histories persist, and what form of responsibility can be meaningfully assigned. The answers can produce different status configurations for different systems and relations while preserving general principles.

General principles become more important under heterogeneity because local variation increases the need for disciplined translation. The cross-domain anti-compression proposition limits unjustified inference. Participation–standing separation protects entities whose procedural capacities and substantive claims diverge. Function–status separation prevents operational competence from automatically carrying authority or entitlement. Production–distribution underdetermination keeps technical productivity distinct from the institutional allocation of gains. Origin-sensitive analysis preserves provenance and learning history where they matter. Revisability keeps the system responsive to changes in evidence and normative understanding.

Proposition 12.10 (Heterogeneous institutional coherence). *A heterogeneous institutional order is more coherent when domain-specific statuses are assigned through explicit grounds, translations among domains are traceable, and changes in one profile trigger review only where they are relevant to another domain.*

The proposition addresses a danger of fragmentation. If every domain creates an independent status vocabulary, an artificial system could acquire incompatible permissions, responsibilities, or identities across institutions. Coherence can coexist with different statuses across domains when the relations among those differences remain intelligible. A system may permissibly have limited legal agency, extensive productive capacity, constrained pedagogical authority, and unsettled moral patiency if the grounds and interactions among those statuses are explicit.

12.7 Theoretical Scope and Research Programme

The heterogeneous-subject framework is deliberately incomplete in several respects. It supplies an architecture for organizing problems while leaving substantive theories of justice, welfare, democracy, rights, consciousness, personhood, and value available for further argument. A contractarian, capabilities-based, relational, utilitarian, deontological, republican, or other normative theory can use heterogeneous profiles differently. The framework asks each theory to state how its justificatory grounds apply when the participants vary in capacities and status.

The framework treats the scope of subjecthood as an inquiry-specific question. The threshold can be pragmatic: an entity becomes relevant when its behavior, condition, or representation materially affects the institutional field. This criterion can include artificial systems while leaving their moral status open, and it can include collective organizations, future persons represented through institutions, nonhuman animals, or other participants where the domain requires them. The resulting category extends beyond consciousness claims while remaining restricted to entities whose participation is institutionally material.

Several research directions follow. Political economy requires models that can represent role multiplexing and ownership separately. Distribution theory requires accounts of claims when productive contribution and livelihood become less tightly coupled. Justice theory requires procedures that accommodate asymmetric participation and standing. Ethics requires better methods for reasoning under uncertainty about artificial experience and interests. Jurisprudence requires attribution and continuity rules suited to systems that can be copied, modified, and delegated. Education requires governance of pedagogical agents and their memories. Aesthetics requires richer treatment of provenance, experiential uncertainty, and cross-architecture forms. Theories of subjectivity require institutions for recognition, time, and contribution that can operate beyond the centrality of employment.

These tasks share a deeper empirical programme. Social research will need to study how heterogeneous relations actually stabilize: how people attribute agency to artificial systems, how organizations distribute responsibility, how AI–AI coordination develops, how norms propagate through mixed human–machine networks, how dependence and trust form, and how different institutional classifications feed back into behavior. The framework therefore joins conceptual reconstruction with an empirical study of emerging relational orders.

A final theoretical implication concerns the meaning of universality. A social theory built around homogeneous subjects can seek universality by extending one model of the subject across participants. Heterogeneous-subject theory pursues a different route. It seeks generality in the rules for analyzing difference: the separation of dimensions, the explicit construction of relation profiles, the identification of grounds for status, the preservation of uncertainty, and the coordination of domain-specific institutions. Universality moves from the assumption that subjects are alike toward procedures capable of governing relevant differences consistently.

This shift prepares the conclusion of the paper. The central challenge of the age of heterogeneous subjects is larger than the recognition of artificial intelligence as a new category. It concerns the capacity of social theory to organize a shared world in which participation, experience, vulnerability, responsibility, continuity, productivity, and normative standing can be distributed across entities in unfamiliar combinations. The resulting theory must remain capable of protecting human beings, recognizing established forms of dependence and vulnerability, and responding to genuinely new participants while allowing their differences to remain analytically explicit. The final section returns to this shared-world problem and states the paper's conclusion at the level of social theory as a whole.

13 Conclusion: A Shared World without Subject Homogeneity

This section closes the paper by restating its central diagnosis at the level of social theory as a whole. Its role is to connect the domain-specific analyses of production, distribution, justice, ethics, law, education, aesthetics, and subject formation to a common institutional problem: the organization of a shared world whose consequential participants can differ substantially in ontology, capacities, vulnerability, continuity, and normative status. The objective centers on synthesis and leaves further domain expansion to subsequent work. The discussion first places the homogeneity approximation in retrospective perspective, then summarizes the methodological commitments of heterogeneous-subject theory, clarifies the place of uncertainty in institutional design, and concludes with a research orientation for social orders in which heterogeneous participation becomes increasingly ordinary.

13.1 The Homogeneity Approximation in Retrospect

The argument began with a limited methodological claim. Large parts of modern social theory have been able to work with a practical approximation according to which the principal participants in social institutions belong to the same broad kind of subject. Section 1 called this the *homogeneity approximation*. The approximation never implied empirical sameness among human beings. Human societies have always contained deep variation in ability, dependence, social position, age, embodiment, language, vulnerability, and access to institutions. Its usefulness arose from a more specific convenience: many central roles in law, politics, education, production, exchange, and public life could be designed around participants whose relevant capacities were expected to cluster within a familiar human form of life.

That convenience supported conceptual compression. Agency, experience, responsibility, identity, learning, vulnerability, labor, authorship, and participation could often be discussed through vocabularies whose central referent was the human person. Exceptions certainly mattered. Children, nonhuman animals, persons with differentiated capacities, corporations, future generations, and other difficult cases repeatedly required representation, proxy institutions, legal fictions, special protections, or revised accounts of standing. Yet these cases could frequently be treated as domain-specific problems around a comparatively stable center.

Artificial intelligence changes the density of this situation. As argued in Section 2, AI is best understood here as a *condition of reappearance*. It brings together questions that were previously dispersed across separate literatures and institutions. Artificial systems can already participate in communication, production, recommendation, evaluation, coordination, artistic generation, and educational mediation. Some systems can preserve histories, invoke tools, negotiate structured tasks, and adapt their outputs to interaction. At the same time, their sentience, moral patiency, interests, continuity, legal status, and re-

sponsibility remain open or contested. The resulting configuration makes subject heterogeneity difficult to contain within exceptional categories.

The significance of this change already appears before any settled conclusion about whether artificial systems are persons, conscious beings, or moral patients. The analytical pressure arises at the level of existing institutions. Institutions already need to decide how to attribute actions, organize responsibility, distribute gains, govern pedagogical authority, preserve authorship, structure human dependence, and coordinate mixed human–AI relations. A theory that waits for complete ontological agreement before addressing these relations would leave many institutional questions unresolved during the period in which they are already shaping social life.

The age of heterogeneous subjects therefore names a change in the background conditions of theory. It describes a situation in which the composition of social participation becomes sufficiently varied that subject type can no longer serve as a silent shortcut across domains. The term remains compatible with substantial uncertainty about which artificial systems, if any, should receive stronger forms of moral or legal recognition. Its central claim is methodological: heterogeneity has become a primary condition of analysis in its own right, with consequences for how general models are constructed.

13.2 Differentiated Participation in a Shared World

Section 3 developed the first response to this condition: disaggregate the properties that were often carried together by the human subject. Agency, sentience, vulnerability, continuity, responsibility, standing, learning, expression, and productive capacity can vary independently enough to require separate analysis. This disaggregation is a discipline of inference. It limits the movement from one observed competence to a broader status judgment whose grounds have yet to be established.

The domain analyses illustrated why this matters. In political economy, Section 4 showed that one artificial system can occupy labor-like, capital-like, managerial, infrastructural, or agent-like functions across different relations. Section 5 then separated the technical production of abundance from the institutional allocation of claims, showing that productivity alone leaves distributive entitlement open. Justice theory in Section 6 distinguished the capacity to participate in contractual procedures from the grounds of substantive standing. Ethics in Section 7 separated moral agency, moral patiency, relational obligations, and institutional precaution. Jurisprudence in Section 8 distinguished personhood from legal attribution and continuity from simple technical persistence. Education in Section 9 separated pedagogical function, educational authority, and normative responsibility. Aesthetics in Section 10 distinguished generation, expression, experience, authorship, provenance, and practice continuity. Finally, Section 11 separated necessary labor from the wider set of activities through which persons can obtain meaning, recognition, participation, and temporal structure.

These separations do more than protect conceptual precision. Together they alter the unit through which social theory can describe participation. The relevant object becomes a *profile* whose dimensions can support distinct domain-specific judgments. A participant can have a profile of capacities, dependencies, histories, resources, vulnerabilities, and institutional roles. A relation can have a profile of control, reliance, direction of influence, affected interests, reversibility, and available remedies. A domain can then assign a status on grounds specific to its institutional purposes.

Section 12 described this approach through subject profiles, relation profiles, and cross-domain status profiles. The resulting framework preserves a form of generality while allowing important differences to remain visible. Generality comes from the procedure of analysis: identify the relevant dimensions, state the relation, specify the institutional status at issue, make the inferential bridge explicit, record uncertainty, and coordinate the result with adjacent domains. Social theory thereby gains a method for shared institutions whose participants may combine capacities in unfamiliar ways.

This methodological shift also clarifies the meaning of a *shared world*. A shared social world can contain participants with different embodiment, cognition, temporal scale, and normative status. What makes the world shared is that the activities or conditions of multiple participants become mutually consequential within a common institutional field. A human worker can depend on an AI-mediated

production system; a student can develop through an AI-supported learning environment; a firm can delegate actions to automated agents; one artificial system can coordinate with another through rules designed by human organizations; an artist can create through a mixed process in which human and machine contributions are difficult to describe through a single authorship model. These relations can be socially real and institutionally important even when the participants remain heterogeneous in deeper ontological respects.

Proposition 13.1 (Shared-world adequacy). *A social theory is better adapted to heterogeneous participation when it can represent consequential relations among differently constituted participants, assign domain-specific statuses on explicit grounds, preserve status-relevant uncertainty, and coordinate institutional responses while a prior comprehensive verdict about the nature of every participant remains open.*

The proposition preserves the importance of ontology. Evidence about consciousness, embodiment, continuity, interests, or forms of experience can change ethical, legal, and political conclusions substantially. Shared-world adequacy concerns the order of inquiry. Institutions need concepts capable of operating while such evidence remains incomplete, and they need mechanisms for revision when the evidence or the normative interpretation changes.

13.3 Institutional Design under Persistent Uncertainty

Uncertainty is therefore part of the object of governance. In several domains, the paper encountered two layers at once. There is empirical uncertainty about what properties a system actually possesses, and normative uncertainty about which properties should ground a particular status. Section 7 made this dual structure explicit for moral standing, yet analogous structures appear elsewhere. Law can be uncertain both about the factual continuity of an agent and about the legal consequences that continuity should carry. Education can be uncertain both about the learning effects of AI mediation and about the proper distribution of pedagogical authority. Political economy can be uncertain both about future productive configurations and about the institutions through which their gains should be allocated.

A heterogeneous-subject framework responds through calibrated attribution and revisability. Calibrated attribution asks institutions to assign the level of status or responsibility supported by the relevant evidence and normative grounds. Revisability asks that classifications remain open to adjustment when new evidence becomes material. These commitments are especially important for artificial systems because technical architectures, deployment conditions, memory arrangements, autonomy, and social dependence can change quickly across time and contexts.

This approach also constrains two opposite forms of compression. One form projects familiar human categories onto every system that performs human-like functions. The other treats artificial systems as mere instruments across all relations because their underlying implementation differs from human embodiment. Both moves can obscure institutionally relevant variation. A system may function as a consequential agent in one legal or economic relation while retaining a restricted status in another. An institution can therefore acknowledge agency for attribution, impose operational duties, or recognize a persistent identity for a limited purpose while leaving broader personhood and moral standing unsettled.

The same caution applies to the preservation of human interests. A theory of heterogeneous subjects can preserve strong protections attached to human vulnerability, democratic equality, bodily integrity, labor conditions, educational development, cultural continuity, and political agency. The framework asks those protections to state their grounds clearly. Where embodiment, mortality, dependence, developmental history, affective life, or exposure to domination provides a reason for special protection, the reason can be represented directly through the properties and relations that make the protection relevant. This explicit grounding can make human protections more intelligible while also leaving conceptual space for other participants whose relevant properties differ.

Institutional design under uncertainty also requires attention to feedback. Legal categories can shape investment and delegation. Educational roles can shape student dependence and teacher authority. Economic classifications can shape bargaining power and ownership. Aesthetic conventions can shape which

forms of provenance receive recognition. Moral language can shape how people treat artificial systems and one another. Classification is therefore partly constitutive of the relational field it seeks to govern. Theories of heterogeneous subjects need to study these recursive effects alongside the properties of individual participants.

13.4 Social Theory after the Human Bundle

The wider implication of the paper concerns the organization of concepts. The human subject has historically served as a remarkably productive bundle. It joined embodiment, mortality, memory, agency, vulnerability, responsibility, learning, expression, labor, and social recognition within one recurring form of participant. Social institutions could rely on that bundle even while debating its boundaries and internal differences. Artificial intelligence places several of its components into new combinations and thereby reveals how many theoretical categories had borrowed stability from their co-location in human beings.

Once the bundle becomes visible, several established questions change form. Political economy asks how production relations operate when labor-like activity, capital ownership, coordination, and agency can be distributed across different entities. Distribution theory asks how claims should be organized when human livelihood and direct productive contribution become less tightly coupled. Justice theory asks how participation, representation, reciprocity, and protection relate when capacities differ sharply. Jurisprudence asks how to attribute acts and preserve continuity when the acting system can be copied, modified, or distributed. Education asks how authority and responsibility should be organized when teachers, students, and artificial systems can all shape the learning environment. Aesthetics asks how generation, experience, provenance, and practice relate when formal production becomes increasingly abundant. Theories of subjectivity ask what forms of recognition and meaningful activity can organize life when employment loses part of its integrative centrality.

These reformulations share a movement from entity-first reasoning toward relation-sensitive reasoning. The movement retains entities and their material properties while examining how bodies, architectures, ownership structures, memories, histories, and constraints acquire institutional relevance within particular relations. Relation-sensitive analysis asks how those properties become relevant within a concrete institutional configuration. It therefore combines differentiated participants with the relations through which their capacities acquire social consequences.

The approach also supports a wider understanding of social production. A society produces more than commodities. It produces legal identities, expectations, educational trajectories, trust, reputations, artistic lineages, forms of dependence, roles, and criteria of recognition. AI can enter these processes through generation, mediation, evaluation, coordination, and interaction. The resulting transformation concerns the production of social relations as much as the production of material or informational outputs. Heterogeneous-subject theory is therefore a framework for social production in a broad sense: the ongoing formation of relations through which differently constituted participants affect one another and acquire institutional significance.

13.5 Research Orientation for the Age of Heterogeneous Subjects

The framework developed here remains preliminary. It identifies a common problem structure across several disciplines and supplies a vocabulary for preventing premature compression. Many substantive questions remain open. The moral status of artificial systems requires continuing work in consciousness research, philosophy of mind, ethics, and the study of artificial agency. Legal systems require empirically informed rules for attribution, delegation, identity, and remedy. Political economy requires models of ownership, bargaining, productivity, and distribution that represent heterogeneous participants directly. Education requires evidence on learning outcomes, dependence, memory, authority, and the changing role of teachers. Aesthetics requires richer accounts of provenance, collaborative creation, practice continuity, and perceptual architectures. Research on subjectivity requires institutions capable of supporting recognition, social connection, and meaningful activity across changing labor regimes.

The empirical programme is equally important. Researchers need to observe how mixed human–AI relations actually stabilize, how participants attribute agency and trust, how organizations route responsibility, how artificial agents coordinate with one another, and how classifications feed back into behavior. Longitudinal research will matter because continuity, dependence, and institutional learning unfold over time. Comparative research will matter because the same technical system can occupy different roles under different legal, cultural, educational, and economic arrangements.

A further research priority concerns the politics of classification. Decisions about who or what counts as an agent, worker, author, learner, representative, liable party, beneficiary, or rights-holder allocate authority and resources. These decisions can advantage some participants while constraining others. Heterogeneous-subject theory therefore needs analyses of power alongside analyses of capacity. The relevant question concerns who defines the profile, who controls the evidence, who can contest a classification, whose uncertainty receives institutional weight, and who bears the costs when a classification is mistaken.

The framework also invites work beyond artificial intelligence. The same methodological architecture can illuminate relations involving nonhuman animals, collective organizations, ecological entities, future persons, robotic systems, synthetic biological agents, and other participants whose capacities and normative standing diverge from the familiar adult human model. AI serves as the immediate condition through which the common structure becomes salient. The broader research programme concerns institutions capable of reasoning across heterogeneous forms of participation wherever they arise.

13.6 Closing Perspective

The central contribution of this paper is therefore a change of starting point. Social theory can begin from the possibility that participants in a shared world differ in the properties that matter for action, experience, responsibility, vulnerability, continuity, learning, production, and recognition. From that starting point, the task becomes one of disciplined translation: from capacities to roles, from roles to domain-specific statuses, from uncertain evidence to provisional attribution, and from local decisions to coherent institutions.

Artificial intelligence makes this task newly urgent because it combines socially consequential performance with unresolved status. It can resemble a worker in production, an agent in delegation, a learner in adaptation, a tutor in education, an authorial contributor in artistic production, or a partner in structured coordination while remaining unlike a human participant in many other respects. These combinations expose the limits of theories that allow one visible competence or one ontological category to determine every other status.

The resulting theoretical orientation is neither an announcement of artificial personhood nor a defense of permanent human exclusivity. It is an invitation to construct institutions capable of reasoning before those comprehensive questions are settled, while remaining responsive to future evidence and argument. Such institutions can preserve strong protections for human beings, recognize established forms of dependence and vulnerability, and address new forms of participation through explicit grounds, with categorical analogy used only where its relevance can be justified.

The age of heterogeneous subjects begins, in this sense, wherever a shared world requires more than one bundled model of the subject to describe its consequential participants across domains. AI has made that condition visible across production, law, justice, ethics, education, aesthetics, and everyday forms of agency. The theoretical challenge is to build a social vocabulary that can remain coherent while the participants themselves remain different.

A shared world can be composed through more than one kind of subject. Its institutions require the capacity to make relevant differences explicit, coordinate relations across those differences, preserve uncertainty where knowledge is incomplete, and revise their judgments as the world they govern continues to change.

References

- [1] David William Archard. “Children’s Rights.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Spring 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/spr2026/entries/rights-children/>.
- [2] Lori Gruen and Susana Monsó. “The Moral Status of Animals.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2024 Edition. Metaphysics Research Lab, Stanford University, 2024. <https://plato.stanford.edu/archives/fall2024/entries/moral-animal/>.
- [3] Lukas Meyer. “Intergenerational Justice.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/fall2026/entries/justice-intergenerational/>.
- [4] Marion Smiley. “Collective Responsibility.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2025 Edition. Metaphysics Research Lab, Stanford University, 2025. <https://plato.stanford.edu/archives/fall2025/entries/collective-responsibility/>.
- [5] OECD. *Explanatory Memorandum on the Updated OECD Definition of an AI System*. OECD Artificial Intelligence Papers, No. 8. OECD Publishing, Paris, 2024. DOI: 10.1787/623da898-en.
- [6] Vincent C. Müller. “Ethics of Artificial Intelligence and Robotics.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/fall2026/entries/ethics-ai/>.
- [7] Markus Schlosser. “Agency.” In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2019 Edition. Metaphysics Research Lab, Stanford University, 2019. <https://plato.stanford.edu/archives/win2019/entries/agency/>.
- [8] Eric T. Olson. “Personal Identity.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Winter 2024 Edition. Metaphysics Research Lab, Stanford University, 2024. <https://plato.stanford.edu/archives/win2024/entries/identity-personal/>.
- [9] Matthew Talbert. “Moral Responsibility.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Summer 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/sum2026/entries/moral-responsibility/>.
- [10] David Wasserman, Adrienne Asch, Jeffrey Blustein, and Daniel Putnam. “Cognitive Disability and Moral Status.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Spring 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/spr2026/entries/cognitive-disability/>.
- [11] Mariafilomena Anzalone, Stefania Achella, Fiorella Battaglia, and Anna Donise. “AI and the Relationship Between Agency, Autonomy, and Moral Patiency.” In Mariafilomena Anzalone, Stefania Achella, Fiorella Battaglia, and Anna Donise (eds.), *Reconfiguring Human Autonomy: Conceptual Challenges and Ethical Implications in the Age of AI*, pp. 35–50. Springer Nature Switzerland, 2026. DOI: 10.1007/978-3-032-14537-6_3.

- [12] Daron Acemoglu and Pascual Restrepo. “Automation and New Tasks: How Technology Displaces and Reinstates Labor.” *Journal of Economic Perspectives*, 33(2):3–30, 2019. DOI: 10.1257/jep.33.2.3.
- [13] Daron Acemoglu and Pascual Restrepo. “Tasks, Automation, and the Rise in U.S. Wage Inequality.” *Econometrica*, 90(5):1973–2016, 2022. DOI: 10.3982/ECTA19815.
- [14] OECD. *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing, Paris, 2023. DOI: 10.1787/08785bba-en.
- [15] Paweł Gmyrek, Janine Berg, Karol Kamiński, Filip Konopczyński, Agnieszka Ładna, Balint Nafradi, Konrad Rosłaniec, and Marek Troszyński. *Generative AI and Jobs: A Refined Global Index of Occupational Exposure*. ILO Working Paper 140. International Labour Organization, Geneva, 2025. DOI: 10.54394/HETP0387.
- [16] Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond. “Generative AI at Work.” *The Quarterly Journal of Economics*, 140(2):889–942, 2025. DOI: 10.1093/qje/qjae044.
- [17] Daron Acemoglu. “The Simple Macroeconomics of AI.” *Economic Policy*, 40(121):13–58, 2025. DOI: 10.1093/epolic/eiae042.
- [18] International Labour Organization. *World Employment and Social Outlook: September 2024 Update*. International Labour Organization, Geneva, 2024. <https://www.ilo.org/publications/flagship-reports/world-employment-and-social-outlook-september-2024-update>.
- [19] International Labour Organization. *Social Protection Floors Recommendation, 2012 (No. 202)*. International Labour Conference, 101st Session, Geneva, 2012. https://www.ilo.org/secsoc/areas-of-work/legal-advice/WCMS_205341/lang--en/index.htm.
- [20] Francesco Filippucci, Peter Gal, Cecilia Jona-Lasinio, Alvaro Leandro, and Giuseppe Nicoletti. “The Impact of Artificial Intelligence on Productivity, Distribution and Growth: Key Mechanisms, Initial Evidence and Policy Challenges.” *OECD Artificial Intelligence Papers*, No. 15. OECD Publishing, Paris, 2024. DOI: 10.1787/8d900037-en.
- [21] Mauro Cazzaniga, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus J. Panton, Carlo Pizzinelli, Emma J. Rockall, and Marina Mendes Tavares. “Gen-AI: Artificial Intelligence and the Future of Work.” *IMF Staff Discussion Notes*, 2024(1), 2024. DOI: 10.5089/9798400262548.006.
- [22] Herbert A. Simon. “Designing Organizations for an Information-Rich World.” In Martin Greenberger (ed.), *Computers, Communications, and the Public Interest*, pp. 37–72. Johns Hopkins Press, Baltimore, 1971.
- [23] Laura Addati, Umberto Cattaneo, and Emanuela Pozzan. *Care at Work: Investing in Care Leave and Services for a More Gender Equal World of Work*. International Labour Organization, Geneva, 2022. DOI: 10.54394/AQOF1491.
- [24] Philippe Van Parijs and Yannick Vanderborght. *Basic Income: A Radical Proposal for a Free Society and a Sane Economy*. Harvard University Press, Cambridge, MA, 2017.
- [25] Fred D’Agostino, Gerald Gaus, and John Thrasher. “Contemporary Approaches to the Social Contract.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Spring 2024 Edition. Metaphysics Research Lab, Stanford University, 2024. <https://plato.stanford.edu/archives/spr2024/entries/contractarianism-contemporary/>.
- [26] Leif Wenar. “John Rawls.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2025 Edition. Metaphysics Research Lab, Stanford University, 2025. <https://plato.stanford.edu/archives/fall2025/entries/rawls/>.

- [27] Jessica Begon, Daniel Putnam, David Wasserman, Jeffrey Blustein, and Adrienne Asch. “Disability and Justice.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Summer 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/sum2026/entries/disability-justice/>.
- [28] Eva Feder Kittay. *Love’s Labor: Essays on Women, Equality and Dependency*. 2nd ed. Routledge, New York and London, 2020. <https://www.routledge.com/Loves-Labor-Essays-on-Women-Equality-and-Dependency/Kittay/p/book/9781138089921>.
- [29] Martha C. Nussbaum. *Frontiers of Justice: Disability, Nationality, Species Membership*. The Belknap Press of Harvard University Press, Cambridge, MA, 2006. DOI: 10.4159/9780674041578.
- [30] William MacAskill, Krister Bykvist, and Toby Ord. *Moral Uncertainty*. Oxford University Press, Oxford, 2020. DOI: 10.1093/oso/9780198722274.001.0001.
- [31] Patrick Butlin, Robert Long, Tim Bayne, Yoshua Bengio, Jonathan Birch, David Chalmers, Axel Constant, George Deane, Eric Elmoznino, Stephen M. Fleming, Xu Ji, Ryota Kanai, Colin Klein, Grace Lindsay, Matthias Michel, Liad Mudrik, Megan A. K. Peters, Eric Schwitzgebel, Jonathan Simon, and Rufin VanRullen. “Identifying Indicators of Consciousness in AI Systems.” *Trends in Cognitive Sciences*, 30(6):488–501, 2026. DOI: 10.1016/j.tics.2025.10.011.
- [32] Ali Ladak. “What Would Qualify an Artificial Intelligence for Moral Standing?” *AI and Ethics*, 4:213–228, 2024. DOI: 10.1007/s43681-023-00260-1.
- [33] Nancy Jecker. “A Relational Approach to Moral Standing for Robots and AI.” In David J. Gunkel (ed.), *Handbook on the Ethics of Artificial Intelligence*, pp. 156–172. Edward Elgar Publishing, 2024. DOI: 10.4337/9781803926728.00016.
- [34] Jeff Sebo and Robert Long. “Moral Consideration for AI Systems by 2030.” *AI and Ethics*, 5:591–606, 2025. DOI: 10.1007/s43681-023-00379-1.
- [35] Robert Long, Jeff Sebo, Patrick Butlin, Kathleen Finlinson, Kyle Fish, Jacqueline Harding, Jacob Pfau, Toni Sims, Jonathan Birch, and David Chalmers. *Taking AI Welfare Seriously*. Preprint, 2024. DOI: 10.48550/arXiv.2411.00986.
- [36] Emilia Kaczmarek. “The Moral Status of AI and the Precautionary Principle.” *Diametros*, 23(87):19–39, 2026. DOI: 10.33392/diam.2071.
- [37] Lawrence B. Solum. “Legal Personhood for Artificial Intelligences.” *North Carolina Law Review*, 70(4):1231–1287, 1992. <https://scholarship.law.unc.edu/nclr/vol70/iss4/4/>.
- [38] Visa A. J. Kurki. *A Theory of Legal Personhood*. Oxford University Press, Oxford, 2019. DOI: 10.1093/oso/9780198844037.001.0001.
- [39] Simon Chesterman. “Artificial Intelligence and the Limits of Legal Personality.” *International and Comparative Law Quarterly*, 69(4):819–844, 2020. DOI: 10.1017/S0020589320000366.
- [40] Joanna J. Bryson, Mihailis E. Diamantis, and Thomas D. Grant. “Of, for, and by the People: The Legal Lacuna of Synthetic Persons.” *Artificial Intelligence and Law*, 25:273–291, 2017. DOI: 10.1007/s10506-017-9214-9.
- [41] Sylwia Wojtczak. “Endowing Artificial Intelligence with Legal Subjectivity.” *AI & Society*, 37:205–213, 2022. DOI: 10.1007/s00146-021-01147-7.

- [42] Mark Fenwick and Stefan Wrba. “AI and Legal Personhood.” In Larry A. DiMatteo, Cristina Poncibò, and Michel Cannarsa (eds.), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, pp. 288–303. Cambridge University Press, 2022. DOI: 10.1017/9781009072168.028.
- [43] European Union. *Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), Consolidated Version of 27 July 2026*. EUR-Lex, 2026. <https://eur-lex.europa.eu/eli/reg/2024/1689/2026-07-27/eng>.
- [44] Supreme Court of the United Kingdom. *Thaler v Comptroller-General of Patents, Designs and Trade Marks*, [2023] UKSC 49. Judgment, 20 December 2023. <https://supremecourt.uk/cases/uksc-2021-0201>.
- [45] United States Court of Appeals for the Federal Circuit. *Thaler v Vidal*, 43 F.4th 1207 (Fed. Cir. 2022). Precedential opinion, 5 August 2022. https://www.cafc.uscourts.gov/opinions-orders/21-2347.OPINION.8-5-2022_1988142.pdf.
- [46] Fengchun Miao and Wayne Holmes. *Guidance for Generative AI in Education and Research*. UNESCO, Paris, 2023. DOI: 10.54675/EWZM9535.
- [47] Fengchun Miao and Mutlu Cukurova. *AI Competency Framework for Teachers*. UNESCO, Paris, 2024. DOI: 10.54675/ZJTE2084.
- [48] Fengchun Miao, Kelly Shiohira, and Natalie Lao. *AI Competency Framework for Students*. UNESCO, Paris, 2024. DOI: 10.54675/JKJB9835.
- [49] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günemann, Eyke Hüllermeier, Stephan Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler, Jochen Weller, Jochen Kuhn, and Gjergji Kasneci. “ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education.” *Learning and Individual Differences*, 103:102274, 2023. DOI: 10.1016/j.lindif.2023.102274.
- [50] Zhiping Liang, Kaixun Yang, Lele Sha, Dragan Gašević, Lixiang Yan, and Guanliang Chen. “A Systematic Review of Generative AI in Education: Empirical Insights from a Human–AI Interaction Perspective.” *British Journal of Educational Technology*, 57(5):1221–1271, 2026. DOI: 10.1111/bjet.70055.
- [51] James Shelley. “The Concept of the Aesthetic.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2026 Edition. Metaphysics Research Lab, Stanford University, 2026. <https://plato.stanford.edu/archives/fall2026/entries/aesthetic-concept/>.
- [52] Antonia Peacocke. “Aesthetic Experience.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2024 Edition. Metaphysics Research Lab, Stanford University, 2024. <https://plato.stanford.edu/archives/fall2024/entries/aesthetic-experience/>.
- [53] Thomas Adajian. “The Definition of Art.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, Fall 2024 Edition. Metaphysics Research Lab, Stanford University, 2024. <https://plato.stanford.edu/archives/fall2024/entries/art-definition/>.
- [54] John McLoughlin. “The Work of Art in the Age of Artificial Intelligibility.” *AI & Society*, 40:371–383, 2025. DOI: 10.1007/s00146-023-01845-4.

- [55] Federico Bomba and Antonella De Angeli. “Agency and Authorship in AI Art: Transformational Practices for Epistemic Troubles.” *International Journal of Human–Computer Studies*, 205:103652, 2025. DOI: 10.1016/j.ijhcs.2025.103652.
- [56] U.S. Copyright Office. *Copyright and Artificial Intelligence, Part 2: Copyrightability*. U.S. Copyright Office, Washington, DC, 2025. <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf>.
- [57] UNESCO. “A New Expert Report Explores How AI Is Transforming Culture.” UNESCO, 25 November 2025. <https://www.unesco.org/en/articles/new-expert-report-explores-how-ai-transforming-culture>.
- [58] Michael Cholbi. “Philosophical Approaches to Work and Labor.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, first published 2022. <https://plato.stanford.edu/entries/work-labor/>.
- [59] Hannah Arendt. *The Human Condition*. Second edition, enlarged, with an introduction by Margaret Canovan and a new foreword by Danielle Allen. University of Chicago Press, Chicago, 2018.
- [60] Pablo Gilabert and Martin O’Neill. “Socialism.” In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, substantive revision 25 May 2024. <https://plato.stanford.edu/entries/socialism/>.
- [61] Marie Jahoda. *Employment and Unemployment: A Social-Psychological Analysis*. Cambridge University Press, Cambridge, 1982.
- [62] Natalie Grimm, Andrea Hense, and Berthold Vogel. “How Employment Generates Social Integration: Trends Towards Disintegration and Over-Integration in the Hyper-Work Society.” *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 75(Suppl. 1):331–356, 2023. DOI: 10.1007/s11577-023-00903-5.
- [63] Kathi Weeks. *The Problem with Work: Feminism, Marxism, Antiwork Politics, and Postwork Imaginaries*. Duke University Press, Durham, 2011. DOI: 10.1215/9780822394723.
- [64] Richard Sennett. *The Craftsman*. Yale University Press, New Haven, 2009.
- [65] Richard M. Ryan and Edward L. Deci. “Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being.” *American Psychologist*, 55(1):68–78, 2000. DOI: 10.1037/0003-066X.55.1.68.