

AI-Mediated Generative Relations: Contribution, Control, Capability, and Benefit

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

Once artificial intelligence enters an economy as something more than a tool, the natural question is who is exploiting whom. With capital, labor, the system, and users there are twelve directed pairs; adding the firm that builds the system and separating it from the firm that deploys one gives twenty. This paper argues that the question in that form cannot be answered, and that the reason is structural rather than evidentiary. A diagnosis of exploitation assigns four roles — who controls, who acts, whose capability falls, and who benefits through the fall — and a directed pair names at most two of them. The paper proves that topologies can agree on every quantity such a table is usually populated from — which node acted, whose capability changed, whose return rose — and still disagree about who, if anyone, exploited whom. It then proves an aggregation result explaining why the entity list had to be revised twice: merging a controller with the system it governs manufactures an action attribution that belonged to neither, and merging a provider with an adopter internalizes a benefit path and can convert a positive diagnosis into a negative one, so a merge is admissible for a question only when no edge of that question’s witness crosses or is internalized by the merge boundary. A third result concerns opacity. Where corpus membership, the counterfactual capability, and the benefit path are each unavailable, the observation set consistent with what is seen contains both positive and negative cases, so the audit is unresolved; and every observation that would resolve it is held by the party whose conduct is at issue. Disclosure is therefore a condition on the framework’s applicability rather than a policy preference, and reading an unresolved audit as a negative finding is an error with an identifiable beneficiary. A fourth result concerns speed: two parties with identical holdings and different expansion rates have different reachable sets at every finite horizon, so equalizing holdings does not equalize reachability. The paper also argues that the training corpus is a less diffuse object than the collective inheritance it is said to condense, since a corpus is finite, assembled, and bounded by decisions someone made; that the standing of the system itself is open, and that the framework must remain expressible under both answers. It classifies no system, provider, or firm, and makes no empirical claim. A closing section states, item by item, the evidence that any such claim would require and that this paper does not have.

1 Introduction

The question that opened this research programme was whether exploitation becomes more aggressive in an era of artificial intelligence. It was posed concretely. In the classical picture, a worker produces value and receives less than the value produced, and the difference is the object of the analysis. In the picture that replaced it, capital spends a small sum on machine-generated work and realizes a large return, and there is no wage from which a difference could be taken. The classical accounting has no place to attach.

The first response to this is usually to widen the cast. If the analysis cannot find its object between an employer and an employee, perhaps it can be found among capital, labor, the system,

and users — four entities, twelve directed relations, one of which may be the answer. The research discussion that generated this paper did exactly that, and then revised the cast twice: first by separating the firm that builds and governs the system from the capital that finances it, and then by separating the firm that deploys a system from the provider that supplies it. Each revision was made because a bundled entity was found to be doing two things at once. With five entities there are twenty directed relations, and the natural next step is to fill in a table.

This paper argues that the table cannot be filled in, and that the obstacle is not missing data. A diagnosis of exploitation, on the framework this programme uses, assigns four roles: a party that controls an arrangement, a party whose operation is the proximate cause of a change, a party whose effective generative capability is lower than it would be under an admissible alternative, and a party whose return is higher *through* that reduction. A directed pair names at most two of these. The paper proves that two topologies can agree on every quantity a pairwise table is normally populated from — which node acted, whose capability fell, whose return rose — while differing on whether anyone exploited anyone. The pairwise question is therefore not hard; it is underdetermined by its own evidence.

The paper’s second result explains the revisions. Aggregating two entities into one is a modeling decision with diagnostic consequences, and the consequences run in both directions. Merging a controller with the system it governs produces an aggregate that appears to act on a third party, an attribution neither constituent had. Merging a provider with an adopter internalizes a benefit path, and a diagnostic that requires the benefit to run through the capability reduction will report nothing where a path has become a loop. So a merge is admissible for a question exactly when no edge of that question’s witness crosses or is internalized by the merge boundary. The entity list was revised twice because two bundles failed this test, and the result says which questions each bundle was still adequate for.

The third result concerns the difficulty the discussion identified as decisive: that one cannot tell, from outside, whether a particular use of a particular work in training was an appropriation at all. This is treated here as structural rather than as a temporary deficiency of transparency. Where corpus membership, the counterfactual capability, and the benefit path are each unavailable, the set of topologies consistent with what is observed contains both positive and negative cases, so a three-valued audit returns its unresolved value; and every observation that would resolve it is held by the party whose conduct is in question. Two things follow. Disclosure is a condition on whether the framework applies at all, not a preference among remedies. And reading an unresolved audit as a negative finding is an error whose beneficiary can be named in advance.

A fourth result concerns what changes first. If a system lets one party traverse a possibility space faster than another, and positions in that space are exclusive once occupied, then two parties with identical holdings have different reachable sets at every finite horizon. Equalizing holdings does not equalize reachability. Rate is therefore a structural variable in its own right rather than a description of how quickly a distributional fact takes effect.

Alongside these the paper argues three things it does not formalize. The first concerns the paradox the discussion raised at the outset: if a system is built from the accumulated production of everyone, then everyone contributed, and a category that applies to everyone distinguishes nobody. The general form of this paradox is addressed in a companion project by separating causal contribution from standing. The paper adds an observation specific to the present case. A training corpus is not the collective inheritance of humanity; it is a finite assembled object whose boundary was drawn by identifiable decisions. The appearance of unlimited diffusion is an artifact of describing an artifact as if it were a heritage, and correcting the description converts a metaphysical puzzle into an institutional question about who assembled what, from where, and on what terms. The second is that the standing of the system itself is open, and that a framework which presupposes

either answer is defective; the paper is constructed so that both are expressible and states which of its results would require revision under each. The third is that the appropriate register for what the framework can currently deliver is diagnosis-conditional-on-disclosure, not diagnosis.

The paper classifies no system, provider, adopting firm, or episode. It establishes no prevalence. It does not claim that AI-mediated relations are exploitative, nor that they are not. Section 7 states, item by item, the evidence any such claim would require, and records that this paper has none of it. The portfolio entry from which the project derives anticipated exactly this deficit, and the appropriate response is to say so rather than to supply illustrations that would function as evidence without being evidence.

Section 2 reconstructs the question and shows where it goes wrong. Section 3 reviews the traditions and records what each supplies and forbids. Section 4 argues six theses with their strongest objections and replies. Section 5 declares the objects and proves four propositions. Section 6 replaces the pairwise table with a role-assignment table and exhibits five countermodels. Section 7 states the evidence gate. Section 8 records open questions, and Section 9 the limits and defeat conditions.

The philosophical and formal parts of this paper are *formally independent* and *justificatorily dependent*. No proposition cites a thesis, and every proposition survives the rejection of all six theses. What does not survive is the claim that the propositions matter.

2 The Question and Why It Is Malformed

This section performs the problem analysis. Its objective is to reconstruct the question as it was posed, to record the two revisions it underwent, and to locate the defect that no further revision of the same kind would repair.

2.1 The Initiating Comparison

The comparison that opened the inquiry sets a classical configuration against a current one. Classically, a worker's labor produces value, the worker receives compensation, and the analysis attends to the difference. Currently, a party spends a small sum on machine-generated work and realizes a large return. The classical difference is not small in the second case; it is undefined, because the party whose compensation would enter the comparison is not present in the transaction at all.

Two observations follow, and only the second is this paper's business. The first is that a value-theoretic accounting will not reach the configuration, which is a familiar result: exploitation theory has been reconstructed on property relations and on comparison against a feasible alternative arrangement without requiring a value substance [5, 21]. The second is that the parties whose position the comparison is about — those whose prior production entered the system, and those whose work the system now substitutes for — are not parties to any transaction under examination. They appear nowhere in the accounting, and no refinement of the accounting will make them appear.

That is why the inquiry widened to a cast of actors. The move is correct. What follows is an argument that the widened question was then posed in a form its own evidence cannot settle.

2.2 Four Entities, Then Five, Then Five Others

The first cast was capital, labor, the system, and users, with the question stated as which of the twelve directed relations among them are exploitative.

The cast was then revised twice, and the revisions are the most instructive part of the sequence. First, the firm that builds, trains, owns, and governs the system was separated from the capital that finances it, on the ground that financing and governing are different functions that had been bundled. Second, the cast was restated as the system, users, laborers, firms that adopt AI, and provider companies, separating the party that deploys a system in a workplace from the party that supplies it.

Each revision was made for the same reason: an entity in the previous list was found to be doing two things whose analysis differs. The paper takes the reason seriously and asks what governs such decisions. Section 5.3 answers: a merge is admissible for a question when no edge of that question's witness crosses or is internalized by the merge boundary, and both revised bundles fail that test for the questions at issue.

But the revisions do not repair the underlying difficulty, because they leave the question's form untouched. With five entities the table has twenty cells instead of twelve; a sixth entity would give thirty. The number of cells is not the problem.

2.3 Four Roles, Not Two Positions

The diagnostic this programme uses requires four things of an alleged relation between a benefiting party and an affected party: that the affected party has standing; that a declared alternative arrangement is admissible; that the affected party's effective generative capability is lower under the focal arrangement than under the alternative; and that the benefiting party's return is higher *through* that reduction rather than merely alongside it.

Satisfying these requires locating four roles. Someone governs the arrangement. Something operates and is the proximate cause of the change. Someone's capability falls. Someone's return rises through the fall. In the classical employment case these collapse: the employer governs, the process operates under its direction, the employee's position is affected, and the employer benefits. Collapse is why the pairwise form was ever adequate.

In the configurations at issue the roles come apart routinely. A system operates; a provider governs what it does; an adopting firm chooses to deploy it; a worker's capability is affected; and the return may rise at the provider, at the adopter, or at neither. A relation described as one between the system and the worker names the actor and the affected party and is silent on the two roles that determine whether the diagnosis holds at all. A relation described as one between capital and the worker names the beneficiary and the affected party and is silent on the mechanism.

This is the defect. A directed pair is a projection of a four-role assignment onto two coordinates, and different assignments project onto the same pair. Section 5.2 shows the projection loses exactly the information the diagnosis needs.

2.4 Requirements on an Adequate Account

Four requirements follow.

First, the unit of analysis must be a role assignment rather than a pair, and any tabulation must be indexed accordingly.

Second, entity individuation must be treated as a modeling decision with stated consequences, not as a description of the world that happens to be convenient.

Third, the account must state what is unresolved and why, and must be constructed so that an unresolved audit cannot be read as a negative finding.

Fourth, the account must not presuppose an answer to the standing of the system, since that question is open and the framework is meant to survive either resolution.

3 Traditions Bearing on AI-Mediated Relations

This section reviews the relevant literatures. Its objective is to record what each supplies, what constraint each imposes, and where each stops. No source in this corpus states the underdetermination result, the aggregation criterion, or the opacity result.

3.1 The Political Economy of AI Production

Crawford analyzes artificial intelligence as an extractive industry, tracing mineral extraction, low-wage labor, and data appropriation through the production of systems, and argues that the resulting infrastructure concentrates power [7]. This supplies the empirical picture the paper’s schematic topology abstracts from, and it imposes a caution the paper accepts: the extraction Crawford documents is heterogeneous, and treating it as one relation between two parties is precisely the error under examination here.

Srnicek analyzes platforms as intermediaries whose position derives from controlling the terms on which others interact [25]; Birch analyzes value appropriated through control rights rather than through production [3]; Khan argues that structural position can produce harm that price-based tests do not register [11]. Together these supply the form of a provider’s position. None distinguishes the provider from the adopter, which is the distinction the second revision of the entity list introduced.

Terranova analyzes free labor in digital economies [26], and Couldry and Mejias analyze data appropriation as a colonial structure [6]. These describe the pattern of distributed contribution meeting narrow capture that the training relation instantiates. Neither addresses the diagnostic question of whose capability falls and whose return rises through the fall.

3.2 Automation, Tasks, and Labor Demand

Acemoglu and Restrepo analyze automation through a task framework in which technology both displaces labor from tasks and reinstates it by creating new ones, so that the net effect on labor demand depends on which effect dominates [1].

This is the most disciplined available treatment of the substitution relation and it imposes the paper’s sharpest empirical constraint. Displacement is not assumed by the framework; it is one of two effects whose balance is an empirical question. A theory that treats every deployment as a capability reduction has assumed what this literature investigates, and the paper’s countermodels are written to keep the assumption out.

The task framework and the capability framework used here are not the same object. A task is a unit of production; effective generative capability is what a participant is positioned to develop. A deployment can reinstate tasks while contracting a participant’s developmental repertoire, and can displace tasks while leaving the repertoire intact. Neither result transfers to the other.

3.3 Opacity, Documentation, and Disclosure

Burrell distinguishes three forms of opacity in machine learning — intentional concealment, technical illiteracy, and opacity arising from the mismatch between the scale of learned models and human interpretive capacity — and argues that they call for different responses [4].

The distinction matters here. The opacity the paper’s third result concerns is the first kind, and it is the kind that documentation practices address rather than dissolve. Gebru and colleagues propose datasheets recording a dataset’s motivation, composition, collection process, and recommended uses [10]; Mitchell and colleagues propose model cards recording a model’s intended use, evaluation, and limitations [16]. Both are proposals for making determinable something that is otherwise held

privately, and both are cited here for that structural property rather than for any claim about adoption.

The European Union’s regulatory framework for AI supplies a verified record of staged obligations, with prohibited practices applying from February 2025, governance and general-purpose model obligations from August 2025, and obligations for systems used in employment and other listed high-risk areas from December 2027 [8]. The paper cites the record for its dates and draws no conclusion about the adequacy of any obligation.

Fricker’s analysis of epistemic injustice identifies wrongs done to someone in their capacity as a knower, including the case where the interpretive resources needed to articulate a harm are unavailable to the person harmed [9]. This is the closest normative treatment of the position the third result describes, in which a party cannot determine whether something happened to it because the determination depends on records held by another party.

3.4 Exploitation, Domination, Justice, and Standing

Roemer’s comparison against a feasible alternative arrangement supplies the diagnostic form [21]; Cohen establishes that the charge does not require a labor theory of value [5]. Marx’s account of the general intellect describes social knowledge becoming a direct force of production [14], which is the classical antecedent for a system that condenses accumulated collective production, and the surrounding value theory is declined here as it is in the companion projects [15]. Romer’s treatment of ideas as nonrival and partially excludable supplies the property that makes the training relation non-depleting [22].

Wertheimer establishes that mutually beneficial and consented transactions can be exploitative [28]; Sample locates the wrong in a failure to respect the value of the other party [23]; Vrousalis locates it in self-enriching instrumentalization of vulnerability [27]. Zwolinski’s constraint against inferring a transaction’s status from background structure applies with particular force to a setting in which almost everyone stands in some remote causal relation to the training corpus [31]. Wollner supplies the category for a wrong with no identifiable agent [29], and Young the form a forward-looking response takes [30].

The two normative frameworks the source discussion invoked are recorded with their standard content. Rawls’s difference principle permits inequalities only if they work to the advantage of the least advantaged [19], and Pettit’s republicanism locates unfreedom in subjection to another’s arbitrary power rather than in interference as such [18]. Both bear on the configuration and neither is applied to any case here.

Sen’s distinction between achieved functionings and the capability set supplies the affected object [24], with Robeyns recording the specification choices any operationalization requires [20]. Anderson’s relational criterion bears on why a question about who may determine another’s developmental conditions is not answered by a statement of holdings [2].

On the standing of systems, Long and colleagues argue that the welfare of AI systems is a question that can no longer responsibly be dismissed, while declining to assert that any current system is a moral patient [12]. The paper adopts exactly that posture and no more.

3.5 Causal Structure and Partial Identification

Pearl’s framework distinguishes causal structure from statistical association and supplies the vocabulary of mediation in which an intermediate node transmits an effect it does not originate [17]. The acting-node and beneficiary-node distinction on which the first result rests is a special case of that vocabulary, applied to a diagnostic predicate rather than to an outcome variable.

Table 1: Traditions bearing on AI-mediated generative relations, what each supplies, and the constraint each imposes on the paper’s claims.

Tradition	Supplies	Constraint imposed
Political economy of AI production	Extraction across minerals, labor, and data; concentration of infrastructure	The extraction is heterogeneous; treating it as one relation is the error under examination
Platform position and rent	The form of a provider’s advantage	No distinction between provider and adopter
Appropriated contribution	Distributed contribution meeting narrow capture	Silent on whose capability falls and whose return rises through it
Automation and tasks	Displacement and reinstatement as two effects whose balance is empirical	A theory assuming every deployment reduces capability has assumed what this literature investigates
Opacity	Three kinds, requiring different responses	Only concealment is addressed by disclosure; scale opacity is not
Documentation practices	Making corpus and model properties determinable	Cited for structure, not for any claim about adoption
Regulatory record	Verified staged dates of obligation	No conclusion about adequacy of any obligation
Epistemic injustice	Wrongs done in the capacity of a knower	Does not establish that opacity here is such a wrong
Exploitation theory	Comparison against an arrangement; consent is not exculpatory; a category for agentless wrongs	Remote causal contribution must not imply exploitation, or the category empties
Justice frameworks	Difference principle; non-domination	Neither is applied to any case here
Capability	Opportunity rather than achievement as the affected object	Specification choices are unavoidable
AI standing	The question is live and unresolved	No current system is asserted to be a moral patient
Causal structure	Mediation: a node transmits an effect it does not originate	Structure is assumed, not identified, in the constructions here
Partial identification	Report the consistent set, not a point	A set containing both answers supports neither

Manski’s framework for partial identification supplies the reporting discipline the third result requires [13]: where the data do not determine a quantity, report the set that is consistent with the data rather than a point.

4 Six Theses on Roles, Bundles, Corpora, Opacity, Standing, and Speed

This section argues the paper’s six theses. Its objective is to give each an argument, the strongest objection known to the author, and a reply. The section makes no use of the formal results and is not used by them.

4.1 The Unit of Diagnosis Is a Role Assignment

Thesis 4.1 (Roles, not pairs). The unit over which a diagnosis of generativity exploitation is defined is an assignment of four roles — controller, actor, affected participant, and benefiting party

— and not an ordered pair of entities. A tabulation indexed by ordered pairs cannot represent the diagnosis.

Argument. The diagnostic conditions quantify over four positions, and no two of them are guaranteed to coincide in the configurations at issue. A pair therefore fixes two coordinates and leaves two free, and the free coordinates carry the conditions that decide the case: whether the actor was governed, and whether the benefit ran through the reduction rather than alongside it.

The point is not that pairwise descriptions are imprecise. It is that they are projections, and Proposition 5.5 shows the projection is not injective on the quantities such tables are populated from. Two topologies can agree on which node acted, whose capability fell, and whose return rose, and disagree on whether anyone exploited anyone.

Strongest objection. This is a familiar point about mediation dressed up as a discovery. Everyone knows that a tool is not the party responsible for what is done with it, and the observation that one should trace causal chains rather than read off adjacency is a commonplace of causal analysis [17].

Reply. The commonplace is granted and the thesis is stronger. The claim is not that one should trace chains but that the pairwise data do not determine the chain, so tracing is not an improvement available within the pairwise framing. A tabulation of directed relations among five entities has twenty cells and the role assignments it is meant to summarize number in the hundreds; the table is not a coarse view of the answer but an index of the wrong space. This is a claim about representational adequacy, and it is proved rather than asserted.

4.2 Bundling Is the Dominant Failure Mode

Thesis 4.2 (Tested aggregation). Individuating entities is a modeling decision with diagnostic consequences, and an aggregation is admissible for a question only if the diagnosis of that question is invariant under it. Most errors in this area are bundling errors, and they are errors in both directions: a bundle can manufacture an attribution and can destroy one.

Argument. The evidence is the source sequence itself. The entity list was revised twice, each time because a single name was covering two functions with different diagnostics — financing against governing, then supplying against deploying. The revisions were made on intuition, and the intuition was right, and Proposition 5.8 states the criterion it was tracking.

The criterion’s two directions are both consequential. Merging a controller with the system it governs yields an aggregate that appears to act on third parties, an attribution that belonged to neither constituent: the controller did not act, and the system did not decide. Merging a provider with an adopter internalizes a benefit path, so a diagnostic requiring benefit to run *through* a capability reduction finds a loop where it needs a path, and reports nothing. The second direction is the more dangerous, because it produces false negatives that look like clean results.

Strongest objection. At a sufficient level of aggregation the bundles are correct. The same owners stand behind the financing, the provider, and often the adopter, and insisting on their separation is a formalism that obscures a unity everyone can see.

Reply. The objection may be true and does not license the merge. What it asserts is a claim about ownership, which is a claim about control; the criterion grants that the merge is admissible for control questions and denies that it is admissible for benefit-path questions, and these are different questions with different witnesses. The demand is not that entities be kept apart but that a merge be justified for the question at hand, which is a weaker and more defensible requirement than either separation or unity as a general policy.

4.3 A Corpus Is Not an Inheritance

Thesis 4.3 (Assembled boundary). The appearance that everyone contributed to a system, and hence that the category of exploitation applies to everyone or to no one, arises from describing an assembled artifact as if it were a collective inheritance. A training corpus is finite, assembled, and bounded by decisions someone made, and this is a different kind of object from the accumulated production of a civilization.

Argument. The paradox as posed runs: a system condenses the historical production of language, knowledge, and culture; almost everyone stands in some remote causal relation to that production; if everyone contributed then everyone is exploited, and a category holding of everyone distinguishes nothing.

The general form of this argument is answered elsewhere in the programme by separating causal contribution, which is historical and admits no threshold, from standing, which is a normative status tracking who stands to lose. That answer holds here and is not repeated.

What is added is specific to the object. The premise slides between two things. A civilization’s accumulated production is unbounded, unenumerable, and had no assembler. A training corpus is a set of items that were selected, obtained, processed, and used, by parties who made decisions about inclusion, and whose extent is in principle enumerable because it was in fact enumerated by whoever assembled it. The universality in the premise belongs to the first object and the appropriation in the conclusion concerns the second, and the argument works only by treating them as one.

Correcting this does not show that anyone was exploited. It changes the kind of question. Instead of asking whether a metaphysically diffuse contribution can ground a claim, one asks who assembled which items, from where, on what terms, and with what effect on the parties whose items were included. Those questions have answers, and Section 4.4 concerns why the answers are not available.

Strongest objection. The corpus is enumerable only in principle. In practice its composition is undisclosed, its scale defeats inspection, and much of it derives from material whose own provenance is untraceable. A boundary that cannot be examined is not usefully different from no boundary.

Reply. Undisclosed is not unbounded, and the difference is exactly the paper’s point. An unbounded contribution grounds no claim in principle; an undisclosed bounded one grounds a claim that cannot currently be evaluated. These call for different responses — conceptual revision in the first case, disclosure in the second — and collapsing them recommends the first where the second is needed.

4.4 Opacity Is Structural and Interest-Aligned

Thesis 4.4 (Interest-aligned opacity). The facts that would resolve a diagnosis in this setting are held by the parties whose conduct is at issue. This is a structural feature of the configuration rather than a passing deficiency of transparency, and it entails that an unresolved audit is the generic outcome and must not be read as a negative finding.

Argument. Three facts are required: whether a participant’s material entered the corpus, what the participant’s capability would have been under an admissible alternative arrangement, and whether the benefiting party’s return rose through the reduction rather than alongside it. The first is recorded, if anywhere, in the assembler’s records. The third is recorded, if anywhere, in the benefiting party’s accounts. The second is counterfactual and is unavailable to everyone, which is a separate difficulty addressed in a companion project.

So two of the three resolving facts are private to parties with an interest in the outcome, and Proposition 5.12 shows that without them the consistent set contains both positive and negative cases.

Two consequences follow. Disclosure requirements are not one policy option among others but a condition on whether the framework can be applied at all; documented corpus and model properties are proposals of exactly this kind [10, 16], and a regulatory record of staged obligations is a record of when certain determinations become possible [8]. And the specific error of treating an unresolved audit as a negative finding has an identifiable beneficiary, which is a reason to name the error rather than to rely on inferential good manners.

Fricker’s analysis describes the position of a party that cannot articulate what was done to it because the resources for the articulation are not available to it [9]. The paper notes the resemblance and does not claim that the configuration constitutes epistemic injustice in her sense, which would require an argument about the mechanism of the exclusion.

Strongest objection. Not all opacity is concealment. Burrell distinguishes opacity arising from the scale and character of learned models from opacity that anyone chose [4], and the second sort cannot be disclosed away because nobody holds the missing fact. Treating the whole difficulty as interest-aligned overstates it and invites remedies that will not work.

Reply. Accepted, and the thesis should be read as restricted. Corpus membership and benefit path are held facts; how a system’s behavior arises from its training is not, and no disclosure produces it. The restriction matters for what follows: the first two are the facts the diagnosis needs, and the third is not among its conditions. A framework that required an account of why a system behaves as it does would be defeated by scale opacity; this one is not, and that is a design consequence rather than an accident.

4.5 Standing Must Remain an Open Coordinate

Thesis 4.5 (Neutral construction). Whether the system itself has standing is open, and a framework that presupposes either answer is defective. The framework should be constructed so that both are expressible, and should state which of its claims change under each.

Argument. The question is live. There is serious argument that the welfare of AI systems can no longer be responsibly dismissed, offered without any assertion that a current system is a moral patient [12]. Constructing a framework that assumes a negative answer builds a contested premise into an apparatus meant to outlast the controversy; assuming a positive answer does the same in the other direction.

Neutrality here is cheap, and the paper should say so rather than claim a virtue. Every result in Section 5 is proved for a topology in which the system occupies the actor role and no standing role. Under a positive answer to the standing question, three things change and nothing else does. Relations directed at the system become diagnosable rather than undefined, so the role table acquires occupied cells it currently leaves empty. The aggregation criterion acquires a further constraint, since merging a system with its provider would then merge a standing-holder with a non-holder. And the system becomes a possible affected participant in the opacity result, with the peculiar feature that the records bearing on its position are held by its provider.

Strongest objection. Neutrality is evasion. The question of whether these systems have interests is substantive and consequential, and a framework that declines to answer it while claiming to analyze relations involving them has declined the hardest part of its subject.

Reply. The objection is right that the question is substantive and wrong that this paper is the place to settle it. What is offered is not silence but a conditional: the results hold under the

negative answer, and the enumerated three changes are what a positive answer would require. A reader who holds a view can therefore use the framework; a framework that had chosen would serve only readers who chose the same way.

4.6 Rate Is a Structural Variable

Thesis 4.6 (Foreclosure by rate). Where a party can traverse a possibility space faster than another and positions in that space are exclusive once occupied, a difference in rate forecloses without any difference in holdings. Rate is therefore a structural variable and not a description of how quickly a distributional fact takes effect.

Argument. A diagnosis compares an affected participant’s position under the focal arrangement against an admissible alternative. If the alternative requires time the participant does not have because another party will have occupied the relevant positions first, then the alternative is not available even though nothing in the participant’s holdings prevents it. Proposition 5.16 states this: identical holdings and different rates give different reachable sets at every finite horizon, and the difference does not close.

The significance for the present subject is that rate is the dimension along which a general-purpose system most obviously operates. Whatever else is contested about such systems, that they let some parties traverse certain spaces faster than others is not.

Strongest objection. This describes competition. Being first is how priority has always worked, and a framework that treats speed as a structural wrong will classify every race as a foreclosure.

Reply. Accepted as description and denied as inference, which is the reply the programme gives whenever a structural feature is mistaken for a verdict [31]. The thesis locates a variable and asserts nothing about permissibility. Its use is negative and specific: it defeats the inference from “holdings were equalized” to “the alternative was available,” which is an inference remedial proposals make routinely.

4.7 Dependence Structure Among the Theses

Rejecting Thesis 4.1 restores the pairwise table and with it the question the paper says cannot be answered. Rejecting Thesis 4.2 makes entity individuation a matter of convenience and removes the account of why the source’s revisions were corrections. Rejecting Thesis 4.3 returns the universal-contribution paradox in its AI form and with it the conclusion that the category applies to everyone or nobody. Rejecting Thesis 4.4 makes the evidentiary difficulty contingent and removes the argument that disclosure is a precondition rather than a preference. Rejecting Thesis 4.5 requires the framework to take a position on a question it has no means to settle. Rejecting Thesis 4.6 restores the inference from equalized holdings to available alternatives.

Theses 4.1 and 4.4 are load-bearing: the first is the paper’s reason for existing and the second is its reason for reaching no conclusions. Thesis 4.3 is the most contestable, since it rests on a distinction between an assembled artifact and an inheritance that a determined opponent can attack from either side.

5 Formal Results

This section declares the objects and proves four propositions. Its objective is to establish that pairwise data underdetermine the diagnosis, that aggregation has a stateable admissibility criterion, that opacity yields a generically unresolved audit whose resolving observations are privately held, and that a rate difference forecloses without a holdings difference. The section cites no thesis.

5.1 Declared Objects

Definition 5.1 (Typed topology). A typed topology is a tuple $T = (\mathcal{N}, K, A, \Gamma, \rho, b)$ where $\mathcal{N}$ is a finite set of nodes; $K \subseteq \mathcal{N} \times \mathcal{N}$ is a set of control edges, with $(x, y) \in K$ meaning that x governs y 's policy; $A \subseteq \mathcal{N} \times \mathcal{N}$ is a set of action edges, with $(x, y) \in A$ meaning that x 's operation is the proximate cause of a change in y 's conditions; $\Gamma : \mathcal{N} \rightarrow \mathbb{R}$ assigns each node its effective generative capability under the focal arrangement; $\rho : \mathcal{N} \rightarrow \mathbb{R}$ assigns each node its return; and b is a declared alternative arrangement inducing Γ^b and ρ^b .

Definition 5.2 (Diagnostic predicate). For nodes x, y with $x \neq y$, write $E_b(x, y) = 1$ when all of the following hold: y has standing; b is admissible; $\Gamma(y) < \Gamma^b(y)$; $\rho(x) > \rho^b(x)$; and there is a *benefit-through witness*, that is, a directed path w from the arrangement's point of intervention to x 's return which passes through the reduction at y , such that removing the reduction removes the increase. Write $E_b(x, y) = 0$ when some conjunct is determinately false, and $E_b(x, y) = ?$ when the available observations determine no value.

The predicate is the companion framework's, restated to the extent this paper uses it. Nothing here re-derives its motivation, and nothing here weakens it.

Definition 5.3 (Role assignment). A role assignment is a tuple $(c, a, y, x) \in \mathcal{N}^4$ with $y \neq x$, naming a controller, an actor, an affected participant, and a benefiting party. A pair (x, y) is the projection of (c, a, y, x) onto its last two coordinates.

Definition 5.4 (Marginal record). The marginal record of T is the triple $M(T) = (\mathbf{1}_A, \Gamma - \Gamma^b, \rho - \rho^b)$: which node acted on which, how each node's capability changed, and how each node's return changed.

The marginal record is what a pairwise tabulation is typically populated from: it says who did something, whose position worsened, and whose improved.

5.2 Pairwise Data Underdetermine the Diagnosis

Proposition 5.5 (Marginal underdetermination). *There exist typed topologies T_1, T_2 on the same node set with $M(T_1) = M(T_2)$ and a pair (x, y) such that $E_b(x, y) = 1$ in T_1 and $E_b(x, y) = 0$ in T_2 .*

Proof. Let $\mathcal{N} = \{c, s, l\}$. In both topologies let $A = \{(s, l)\}$, let $\Gamma(l) < \Gamma^b(l)$ with all other capabilities unchanged, and let $\rho(c) > \rho^b(c)$ with all other returns unchanged. Then $M(T_1) = M(T_2)$ by construction.

In T_1 let $K = \{(c, s)\}$ and let the point of intervention be c 's policy for s , with a witness path $c \rightarrow s \rightarrow l \rightarrow \rho(c)$: the reduction at l is what raises c 's return, and removing it removes the increase. Then $E_b(c, l) = 1$.

In T_2 let $K = \emptyset$, so s operates autonomously, and let c 's return rise through a channel disjoint from l — for instance an unrelated holding that appreciates over the same period. Then no witness path through the reduction exists, the last conjunct of Definition 5.2 fails determinately, and $E_b(c, l) = 0$. $\square$

Corollary 5.6 (The table indexes the wrong space). *On n nodes there are $n(n-1)$ ordered pairs and $n^3(n-1)$ role assignments satisfying $y \neq x$, of which $n(n-1)(n-2)(n-3)$ have four distinct occupants. For $n = 5$ these are 20, 500, and 120 respectively.*

The counting is trivial and its point is not. A tabulation of twenty directed relations is not an approximation to the answer; the objects it indexes are not the objects the diagnosis is defined on,

and Proposition 5.5 shows the map between them is not injective on the data that would populate the table. Adding entities makes this worse rather than better, since the assignment count grows as n^4 and the pair count as n^2 .

Two things the proposition does not show should be stated. It does not show that the diagnosis is never determinable; it shows that marginal data do not determine it, and control edges and witness paths are exactly the further data required. And it does not show that T_2 is innocent in every respect — only that the particular predicate is determinately false there, which is what a diagnostic framework is supposed to be able to say.

5.3 When an Aggregation Is Admissible

Definition 5.7 (Merge). For $u, v \in \mathcal{N}$ with $u \neq v$, the merge $T/\{u, v\}$ replaces them by a single node $\bar{u}$, redirects every edge with an endpoint in $\{u, v\}$ to $\bar{u}$, deletes edges with both endpoints in $\{u, v\}$, and sets $\rho(\bar{u}) = \rho(u) + \rho(v)$.

Proposition 5.8 (Aggregation criterion). *Fix a pair (x, y) with a witness path w in T , and let $u, v \notin \{y\}$.*

1. *If no edge of w has both endpoints in $\{u, v\}$, and $x \notin \{u, v\}$ or $\{u, v\} \subseteq \{x\} \cup (\mathcal{N} \setminus w)$, then $E_b(x, y)$ is unchanged in $T/\{u, v\}$.*
2. *If some edge of w has both endpoints in $\{u, v\}$, then w has no image in $T/\{u, v\}$, and if no alternative witness exists then $E_b(\bar{u}, y) \neq 1$.*
3. *If $(u, v) \in K$ and $(v, z) \in A$ for some z , then $(\bar{u}, z) \in A$ in $T/\{u, v\}$, although $(u, z) \notin A$ in T .*

Proof. (1) Under the stated condition every edge of w has an image in $T/\{u, v\}$ and the images form a directed path with the same endpoints, so the witness survives; the remaining conjuncts of Definition 5.2 concern y and the returns, and neither is altered by a merge disjoint from y whose return sum preserves the sign of the change at the benefiting party. (2) By Definition 5.7 edges internal to $\{u, v\}$ are deleted, so w is not a path in the merged topology; if no other witness exists the final conjunct fails and the predicate is not 1. (3) The merge redirects (v, z) to $(\bar{u}, z)$, which is an action edge incident to $\bar{u}$; since $(u, z) \notin A$, the aggregate carries an action attribution that u did not. $\square$

Corollary 5.9 (Question-relative admissibility). *An aggregation is admissible for a question exactly when no edge of that question’s witness is internalized by the merge and no role of the question’s assignment is split across the merged nodes. In particular, a merge of a financing party with a governing party is admissible for questions about control and inadmissible for questions whose witness runs between them; and a merge of a supplying party with a deploying party is inadmissible for any question whose witness runs from one to the other.*

Clause (2) is the one that produces silent errors. A merged provider and adopter present as a single firm that both trained a system and deployed it, and a benefit path that ran from deployment back to supply becomes internal and disappears. The resulting diagnosis is not merely coarse; it is negative, and it looks like a finding.

Clause (3) produces the opposite error and is the more visible of the two. Where a governing party and the system it governs are treated as one, the aggregate appears to act on those the system acts on. This is how a relation between a system and a worker comes to be asserted, and neither constituent supports it: the governing party did not act and the system did not decide.

5.4 Opacity Yields an Unresolved Audit With Privately Held Resolvers

Definition 5.10 (Observation and consistent set). An observation $\mathcal{O}$ is a set of facts about a topology. The consistent set $\mathcal{T}(\mathcal{O})$ is the set of typed topologies compatible with $\mathcal{O}$.

Assumption 5.11 (Unavailability). $\mathcal{O}$ determines neither (i) whether a given participant’s material entered the corpus from which the system was produced, nor (ii) the counterfactual capability Γ^b , nor (iii) the existence of a benefit-through witness.

Proposition 5.12 (Generic non-resolution). *Under Assumption 5.11, for the pair (x, y) at issue there exist $T_+, T_- \in \mathcal{T}(\mathcal{O})$ with $E_b(x, y) = 1$ in T_+ and $E_b(x, y) = 0$ in T_- . Hence the audit returns ?. Moreover any $\mathcal{O}' \supseteq \mathcal{O}$ with $E_b(x, y) \neq ?$ for all $T \in \mathcal{T}(\mathcal{O}')$ must determine at least one of (i)–(iii).*

Proof. The construction of Proposition 5.5 supplies T_+ and T_- agreeing on all marginals; by Assumption 5.11 the observation distinguishes neither their control structure nor their witness structure, so both lie in $\mathcal{T}(\mathcal{O})$ and the audit has no determinate value. For the second claim, suppose $\mathcal{O}'$ determines none of (i)–(iii). Then the same two topologies remain consistent with $\mathcal{O}'$, since each of (i)–(iii) is the only respect in which the relevant conjuncts differ, and the audit remains ?. $\square$

Corollary 5.13 (Disclosure as a decidability condition). *If facts (i) and (iii) are recorded only by the party assembling the corpus and the party whose return is at issue, then every $\mathcal{O}'$ resolving the audit includes information held by a party whose conduct the audit concerns.*

Three consequences are worth separating. First, an unresolved audit is the expected outcome rather than an anomaly, so a body of unresolved audits is not evidence that nothing occurred. Second, the inference from ? to 0 is invalid on the framework’s own three-valued semantics, and the party favored by that invalid inference is identifiable in advance, which is a reason to state the point explicitly rather than to treat it as a technicality. Third, a disclosure requirement changes the class of decidable propositions, which places it in a different category from remedies that alter outcomes; documentation practices for datasets and models are proposals of exactly this shape [10, 16].

Fact (ii) is not repaired by any disclosure, since no party holds a counterfactual. Its treatment is the subject of a companion project, which shows that such a repertoire admits bounds rather than a value, and the reporting discipline for a quantity with a consistent set rather than a point is standard [13].

5.5 Rate Forecloses Without a Holdings Difference

Definition 5.14 (Reachable set under a rate). Let V be a space of developmental positions with a cost function $\text{cost} : V \rightarrow \mathbb{R}_{>0}$. A participant with rate $r > 0$ has reachable set at horizon h

$$\mathcal{R}_r(h) = \{\tau \in V : \text{cost}(\tau) \leq r h\}.$$

Assumption 5.15 (Exclusive occupation). A position occupied at some time is unavailable to any other participant thereafter.

Proposition 5.16 (Rate foreclosure). *Let two participants have identical holdings and rates $r_A < r_B$. Then for every $h > 0$, $\mathcal{R}_{r_A}(h) \subsetneq \mathcal{R}_{r_B}(h)$ whenever cost takes a value in $(r_A h, r_B h]$. Under Assumption 5.15, the positions in $\mathcal{R}_{r_B}(h) \setminus \mathcal{R}_{r_A}(h)$ are unavailable to A at every later time, and the excluded set is non-decreasing in h and in r_B/r_A .*

Proof. Strict inclusion is immediate from the definition and the stated value condition. Under Assumption 5.15, B occupies those positions at times A cannot match, and occupation is permanent, so they are excluded from A 's attainable set thereafter. Monotonicity in h and in the ratio follows because the excluded set is $\{\tau : r_A h < \text{cost}(\tau) \leq r_B h\}$, whose defining interval widens in both parameters. $\square$

Corollary 5.17 (Holdings equalization is insufficient). *An alternative arrangement that equalizes holdings but not rates does not equalize reachable sets, so its admissibility as a comparison does not follow from its distributive properties.*

The result is elementary and its assumptions are strong. Exclusive occupation holds for priority, for certain standards positions, and for some market positions, and fails wherever a development can be undertaken independently by several parties. The proposition should be read as identifying a mechanism available in some regions of a possibility space rather than as a claim about possibility spaces in general.

6 The Role Table and Countermodels

This section replaces the pairwise tabulation with one indexed as the diagnosis requires, and exhibits five countermodels. Its objective is to show what a correctly indexed summary looks like and to mark the boundaries of the analysis by displaying configurations that fail it.

6.1 A Role-Indexed Summary

Table 2 lists configuration schemas by role assignment rather than by pair. Each row names who governs, what operates, whose capability is at issue, and where a return would have to rise for the diagnosis to apply; the final column records what would have to be shown. No row asserts that any configuration obtains, and the table is a specification of questions rather than a summary of findings.

Two features of the table are the point of it. The third and fourth rows differ only in where the return rises, and they are different diagnostic questions with different witnesses; a pairwise table has one cell for both. The last row is empty or occupied depending on the standing question, which is how Thesis 4.5's neutrality is implemented rather than asserted.

6.2 Countermodels

Countermodel 6.1 (Mediation without benefit). A system operates, a participant's capability falls, and no party's return rises through the fall. The final conjunct of Definition 5.2 fails determinately.

The configuration is harm and not exploitation, and it is the case a pairwise reading most reliably misdescribes, since the action edge and the capability change are both present and both visible.

Countermodel 6.2 (Benefit without mediation). A participant's capability falls for reasons unrelated to any deployment, and a party's return rises over the same period through an unrelated channel.

This is the second topology of Proposition 5.5, and it exists to show that co-occurrence of a reduction and an increase is not the predicate.

Table 2: Configuration schemas indexed by role assignment. Each row states what would have to be established, not what has been.

Controls	Acts	Affected	Benefits	What must be shown
Provider	System	Contributor to corpus	Provider	Corpus membership; a capability reduction against an admissible alternative; a witness from the reduction to the provider’s return
Provider	System	Adopting firm	Provider	Dependence on a governed condition; a reduction relative to an alternative provision; a witness through the dependence
Adopter	System	Worker	Adopter	Deployment as the intervention; task displacement distinguished from repertoire contraction; a witness from the contraction to the adopter’s return
Adopter	System	Worker	Provider	The same, with a witness crossing from adopter to provider; inadmissible under any merge of the two
Provider	System	User	Provider	Interaction traces as the acquired material; a reduction in the user’s own developmental position; a witness distinct from ordinary service revenue
Capital	Provider	Worker	Capital	Control of the provider’s policy; the chain from that control to the reduction; admissibility of merging capital with provider tested separately
—	System	Any	Nobody	Autonomous operation with no return increase anywhere: a candidate for harm without exploitation
Provider	System	System	Provider	Undefined unless the system has standing; the cell is empty under the negative answer and occupied under the positive one

Countermodel 6.3 (Reinstatement). A deployment displaces a participant from a set of tasks and creates others that the participant is positioned to take up, leaving the developmental repertoire undiminished or enlarged.

The countermodel encodes the constraint the automation literature imposes [1]: displacement and reinstatement are two effects and their balance is empirical. A framework that could not represent this configuration would have assumed its conclusion.

Countermodel 6.4 (Contribution without standing). A participant’s material is in the corpus and the participant has no developmental position that the system’s existence worsens.

This is the AI-specific form of the universal-contribution paradox’s dissolution. Corpus membership is a historical fact and standing is a normative status, and the countermodel exists to keep the first from being read as the second.

Countermodel 6.5 (Autonomous operation). A system operates with no control edge, a participant’s capability falls, and the system’s own returns rise.

The configuration is undefined under the negative answer to the standing question, since a node without standing has no return in the relevant sense, and becomes diagnosable under the positive one with the system in the benefiting role. It is included to make the dependence on the standing question concrete rather than rhetorical.

7 The Evidence Gate

This section states what evidence a claim in this area would require. Its objective is to convert the portfolio’s recorded readiness deficit into an itemized specification, so that the absence of the evidence is visible rather than implicit, and so that a subsequent empirical project has a target.

The paper makes none of the claims below and asserts none of the antecedents. Each gate states a claim, the evidence that would license it, and the reason the evidence is not substitutable.

Evidence Gate 7.1 (Corpus membership). *Claim:* a specific participant’s material entered the corpus from which a specific system was produced. *Required:* the assembler’s records of sources, acquisition, filtering, and retention, or an inference procedure whose error rate is characterized. *Not substitutable by:* demonstrations that a system reproduces material resembling the participant’s, which are consistent with several origins.

Evidence Gate 7.2 (Capability effect). *Claim:* a participant’s effective generative capability is lower than it would be under an admissible alternative arrangement. *Required:* a declared alternative, a specification of the repertoire, and either identification or characterized bounds. *Not substitutable by:* observed reductions in income, employment, or output, which are neither necessary nor sufficient for a repertoire contraction.

Evidence Gate 7.3 (Benefit path). *Claim:* a party’s return rose through the reduction rather than alongside it. *Required:* a witness path with the counterfactual property that removing the reduction removes the increase. *Not substitutable by:* correlation between a deployment and a return, or by the deploying party’s stated rationale.

Evidence Gate 7.4 (Concentration). *Claim:* control of the relevant conditions is concentrated to a degree that matters for the diagnosis. *Required:* a specification of which conditions are necessary for which continuations, and measurement over that specification. *Not substitutable by:* market share in adjacent product markets.

Evidence Gate 7.5 (Substitution). *Claim:* a deployment displaced participants without reinstating them. *Required:* the task-level evidence the automation literature specifies, plus a separate argument connecting task displacement to repertoire contraction [1]. *Not substitutable by:* either alone.

Evidence Gate 7.6 (Opacity). *Claim:* the relevant facts are withheld rather than unavailable in principle. *Required:* the distinction between concealment and scale opacity applied to the specific fact at issue [4]. *Not substitutable by:* the observation that an explanation was not provided.

Two remarks about the gate as a whole. It is demanding, and it is demanding in the same way for claims in either direction: a claim that no exploitation is occurring passes through Gates 7.1–7.3 exactly as a claim that it is. And it is not a counsel of despair, because Proposition 5.12 identifies which gates are opened by disclosure and which are not.

8 Open Questions

This section records what the project leaves unresolved, by register.

8.1 Conceptual and Formal

Open Problem 8.1 (Partial merges). Proposition 5.8 treats merges of whole nodes. Institutions overlap partially — shared ownership, interlocking control, joint ventures — and the criterion for a partial merge is unknown.

Open Problem 8.2 (Witness identification). The diagnostic requires a benefit-through witness, and nothing here says under what conditions such a path is identifiable from data rather than assumed. The causal apparatus for mediation exists [17]; its application to a predicate rather than an outcome has not been worked out.

Open Problem 8.3 (Cost structure of the possibility space). Proposition 5.16 assumes a cost function on developmental positions and exclusive occupation. Which regions of which spaces have these properties is unknown, and the proposition is empty where they fail.

Open Problem 8.4 (Aggregating unresolved audits). If most audits return ?, what a body of such audits supports collectively is undefined. The reporting discipline for individual quantities is available [13]; a discipline for populations of unresolved diagnoses is not.

8.2 Political-Economic

Whether the provider–adopter distinction is stable, or whether integration routinely internalizes the benefit paths that Corollary 5.9 identifies as diagnostically load-bearing, is an empirical question with direct consequences for whether the framework can be applied at all.

Whether the returns at issue accrue as rent on control rights [3], as platform position [25], or as ordinary product revenue determines which witness structures are even candidates, and the paper does not settle it.

Whether disclosure obligations of the kind now staged in at least one jurisdiction [8] in fact open Gates 7.1 and 7.3 is a question about implementation that the regulatory record does not answer.

8.3 Normative

Whether the position of a party that cannot determine what was done to it, because the determination depends on records held by another party, constitutes an epistemic injustice in Fricker’s sense [9] requires an argument about the mechanism of the exclusion that this paper does not supply.

Whether responsibility in configurations with no identifiable perpetrator takes the forward-looking structural form [30, 29] or admits liability depends on facts about anticipation that the framework deliberately does not require.

Whether the distributive and republican criteria the source discussion invoked [19, 18] deliver convergent or divergent verdicts on the same configuration is untested here and would be worth testing, since a configuration on which they diverge would be informative about both.

Whether a system’s standing, if established, would place its provider in a relation the framework already has vocabulary for [12, 27] is noted and not pursued.

8.4 Empirical

The evidence gate of Section 7 is the empirical programme. Nothing in it has been attempted here, and the design of a study that would pass even one gate is itself unresolved.

9 Limits and Revisable Research Program

This section states what the paper contributes, what would defeat it, and what it has not done.

9.1 Contributions

The conceptual contribution is the replacement of the pairwise question by a role-indexed one, together with the argument that a training corpus is an assembled artifact rather than an inheritance, which converts a metaphysical paradox into an institutional question.

The structural contribution is the aggregation criterion and its two directions: that merging a controller with a mediator manufactures an attribution, and that merging a supplier with a deployer destroys a witness and produces a negative result that resembles a finding.

The formal contribution is four results: marginal underdetermination with its counting corollary; the aggregation criterion; generic non-resolution under opacity with disclosure as a decidability condition; and rate foreclosure with the corollary that holdings equalization does not establish an alternative’s admissibility.

The methodological contribution is the evidence gate, which states what any claim in this area requires and applies symmetrically to claims in both directions.

9.2 Defeat Conditions

The conceptual contribution loses support if the four roles coincide in the configurations of interest, since the pairwise form would then be adequate and the paper would be solving a problem that does not arise.

The aggregation criterion loses support if diagnostic witnesses are robust to merges in practice — if, for instance, benefit paths in these settings are typically short and external, so that internalization is rare.

The opacity result loses support if the resolving facts turn out not to be privately held, which disclosure regimes could in principle bring about and which would be a welcome defeat.

The rate result loses support wherever exclusive occupation fails, which is wherever a development can be independently undertaken by several parties.

The project as a whole loses interest if the evidence gate is never passed, since a framework specifying conditions no investigation can meet is a description of an impasse rather than a tool.

9.3 Development Status

The portfolio records this project as an AI-specific comparative and empirical framework, and records that claims about training traces, capability effects, concentration, substitution, and opacity require current system-specific evidence and verified empirical sources.

The status of that deficit is as follows. The verified corpus stands at thirty-one entries and includes the political economy of AI production, the automation literature, the opacity and documentation literatures, and a verified regulatory record. The comparative framework is supplied by the role table and the countermodels. The empirical component is entirely absent: no system, provider, or firm is named, no case is examined, and no claim requiring any of the six gates is made. The paper is therefore a framework paper with an explicit specification of what it does not have, rather than the comparative and empirical framework the seed described.

Deferred questions are assigned. The diagnostic itself belongs to the capability framework. Transmission along chains belongs to the propagation project. Identification of repertoires belongs to the latent-generativity project. Closed configurations belong to the cycles project. The commons object belongs to the commons project. Standing control of expansion conditions belongs to the political-economic project, and induced dependence to the revocable-infrastructure project. Recovery and legal relations belong to projects not yet developed.

9.4 The Conservative Position

Stated as conservatively as the arguments permit: the question of who exploits whom among the parties to an AI-mediated economy cannot be answered in pairwise form, because the diagnosis quantifies over four roles and a pair fixes two; the entity list matters, and merging entities can both manufacture and destroy findings, so each merge must be justified for the question at hand; the facts that would settle the matter are largely held by the parties whose conduct is at issue, so most audits will be unresolved, and an unresolved audit is not a negative one; a system that lets one party move faster than another can foreclose without any difference in holdings; and a training corpus, whatever else it is, is a bounded object someone assembled, which makes the question about it institutional rather than metaphysical.

None of this establishes that any AI-mediated relation is exploitative, and none of it establishes that none is. The paper’s position is that the second claim is currently as unsupported as the first, and that saying so is more useful than choosing.

References

- [1] Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30, 2019.
- [2] Elizabeth S. Anderson. What is the point of equality? *Ethics*, 109(2):287–337, 1999.
- [3] Kean Birch. Technoscience rent: Toward a theory of rentiership for technoscientific capitalism. *Science, Technology, & Human Values*, 45(1):3–33, 2020.
- [4] Jenna Burrell. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 2016.
- [5] G. A. Cohen. The labor theory of value and the concept of exploitation. *Philosophy & Public Affairs*, 8(4):338–360, 1979.
- [6] Nick Couldry and Ulises A. Mejias. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Culture and Economic Life. Stanford University Press, Stanford, CA, 2019.
- [7] Kate Crawford. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, New Haven, CT, 2021. The Yale University Press catalogue record consulted lists the paperback, ISBN 9780300264630, and the eBook, ISBN 9780300252392.
- [8] European Commission. Regulatory framework for AI. Shaping Europe’s Digital Future, European Commission, 2026. Records that Regulation (EU) 2024/1689 entered into force on 1 August 2024 and became applicable on 2 August 2026, with prohibited-practice rules from 2 February 2025, governance and general-purpose model obligations from 2 August 2025, and obligations for systems used in employment and other listed high-risk areas from 2 December 2027. Consulted 16 August 2026.
- [9] Miranda Fricker. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, Oxford, 2007.
- [10] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021.

- [11] Lina M. Khan. Amazon’s antitrust paradox. *The Yale Law Journal*, 126:710–805, 2017.
- [12] 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, 2024.
- [13] Charles F. Manski. Nonparametric bounds on treatment effects. *The American Economic Review*, 80(2):319–323, May 1990. Papers and Proceedings.
- [14] Karl Marx. *Grundrisse: Foundations of the Critique of Political Economy*. Penguin Books, London, 1973. Translated with a foreword by Martin Nicolaus, in association with New Left Review; ISBN is that of the Penguin Classics reprint of this translation.
- [15] Karl Marx. *Capital: A Critique of Political Economy*, volume 1. Penguin Books, 1992. Translated by Ben Fowkes; original German edition published in 1867.
- [16] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 220–229. Association for Computing Machinery, 2019.
- [17] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2 edition, 2009.
- [18] Philip Pettit. *Republicanism: A Theory of Freedom and Government*. Oxford Political Theory. Oxford University Press, Oxford, 1997.
- [19] John Rawls. *A Theory of Justice*. Harvard University Press, Cambridge, MA, revised edition, 1999.
- [20] Ingrid Robeyns. The capability approach: A theoretical survey. *Journal of Human Development*, 6(1):93–117, 2005.
- [21] John E. Roemer. *A General Theory of Exploitation and Class*. Harvard University Press, Cambridge, MA, 1982.
- [22] Paul M. Romer. Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2):S71–S102, 1990.
- [23] Ruth J. Sample. *Exploitation: What It Is and Why It’s Wrong*. Rowman & Littlefield, Lanham, MD, 2003.
- [24] Amartya Sen. Capability and well-being. In Martha C. Nussbaum and Amartya Sen, editors, *The Quality of Life*, pages 30–53. Oxford University Press, Oxford, 1993.
- [25] Nick Srnicek. *Platform Capitalism*. Theory Redux. Polity Press, Cambridge, 2017.
- [26] Tiziana Terranova. Free labor: Producing culture for the digital economy. *Social Text*, 18(2 (63)):33–58, 2000.
- [27] Nicholas Vrousalis. Exploitation, vulnerability, and social domination. *Philosophy & Public Affairs*, 41(2):131–157, 2013.
- [28] Alan Wertheimer. *Exploitation*. Princeton University Press, Princeton, NJ, 1996.

- [29] Gabriel Wollner. Anonymous exploitation: Non-individual, non-agential and structural. *Review of Social Economy*, 77(2):143–162, 2019.
- [30] Iris Marion Young. Responsibility and global justice: A social connection model. *Social Philosophy and Policy*, 23(1):102–130, 2006.
- [31] Matt Zwolinski. Structural exploitation. *Social Philosophy and Policy*, 29(1):154–179, 2012.