

Symbolic Diagnosticity under Symbol Inflation

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

Social actors routinely infer effort, competence, attention, authorship, sincerity, scarcity, commitment, and institutional standing from symbolic form. This paper examines the changing *diagnosticity* of those forms under conditions of *symbol inflation*: a domain-specific expansion in the production, circulation, and strategic construction of socially consequential symbols that can outpace the interpretive and institutional capacities through which their underlying histories are distinguished. Symbol inflation is broader than generative artificial intelligence. Branding, templating, mass communication, institutional signaling, automated personalization, and other forms of organized symbolic production can also increase the supply and strategic optimization of symbolic surfaces; generative AI is a particularly strong contemporary accelerator because it widens the set of low-friction pathways capable of producing similar observable forms. The paper treats symbols as compression interfaces that permit coordination without direct access to the histories, relations, and practices behind them. As heterogeneous pathways converge on similar surfaces, diagnosticity can weaken for some inferred properties, remain stable for others, or migrate toward provenance, institutional assurance, trajectory, and relational context. The analysis develops this problem across communication, expertise, credentials, education, culture, branding, luxury, and institutional legitimacy. The term *symbolic crisis* is retained only for the narrower downstream condition in which consequential diagnostic mismatches persist while shared interpretive and institutional practices remain unsettled. Recent findings on AI-mediated communication, disclosure, trust, and authenticity provide one empirical window into this broader transition. The paper remains primarily diagnostic and leaves the design of response practices to a companion inquiry.

Keywords: symbolic diagnosticity; symbol inflation; social inference; generative AI; political economy

Discussion Paper Note

This paper examines the changing diagnosticity of symbolic forms under a condition described as *symbol inflation*. Its structural role is diagnostic. Symbols often function as socially economical interfaces through which recipients infer attention, competence, labor, authorship, sincerity, scarcity, commitment, status, or institutional standing without reconstructing the histories behind them. Symbol inflation occurs when the production, circulation, or strategic construction of such forms expands sufficiently that similar symbolic surfaces become compatible with increasingly heterogeneous histories. Diagnosticity can then weaken for selected inferences, remain stable for others, or migrate toward additional contextual evidence.

Symbol inflation is broader than generative AI. Commercial branding, advertising, templated communication, institutional signaling, reputational engineering, mass publishing, automated personalization, and other practices have long increased the volume or strategic intensity of symbolic production. Generative systems are especially consequential within this longer development because they lower several production frictions simultaneously and permit rapid variation and personalization at scale. The paper therefore treats AI as a major contemporary accelerator and empirical case rather than as the sole source of the underlying condition.

The immediate technological context is the rapid expansion of artificial intelligence-mediated communication. Hancock, Naaman, and Levy define AI-mediated communication as interpersonal communication in which an intelligent agent acts on behalf of a communicator by modifying, augmenting, or generating messages to accomplish communicative goals (Hancock et al. 2020). Generative systems have greatly expanded the range, quality, and accessibility of such mediation. An email, report, visual identity, apology, personal message, product narrative, or apparently expert explanation can now be produced through very different combinations of human intention, prompting, model generation, editing, review, and automated delegation.

This diversity of generative pathways creates an inverse-inference problem for recipients. A symbolic artifact can remain linguistically fluent, emotionally appropriate, or professionally polished while revealing less about how much human attention, understanding, effort, or commitment preceded it. Research on costly signaling provides one useful background: observable expenditures of time, effort, emotion, reputation, or sacrifice can contribute to perceived honesty because they make some signals harder to produce casually (Chaudhry and Wald 2022). Research on joint action similarly suggests that effort invested in informative coordination can influence trust and generosity (McEllin and Michael 2022). The present paper does not infer that effort is the essence of trust. Rather, these findings help explain why a change in production friction can disturb established social interpretations of symbolic form.

Recent empirical work already indicates that AI-mediated communication creates an unsettled trust environment. Sahebi and Formosa describe an epistemic-trust dilemma in which reduced trust toward AI-mediated communication can sometimes be warranted and sometimes generate epistemic injustice (Sahebi and Formosa 2025). Schilke and Reimann report across thirteen experiments that disclosure of AI use reduced trust in the disclosing actor in the studied settings, with perceived legitimacy helping explain the effect (Schilke and Reimann 2025). Sahebi, Formosa, and Bankins subsequently report lower perceived trustworthiness, authenticity, and knowledge uptake for AI-involved written communication in their experimental design, while also identifying a disclosure paradox (Sahebi, Formosa, and Bankins 2026). These findings support neither a universal anti-AI conclusion nor a single stable social norm. They instead show that familiar relations among authorship, transparency, authenticity, and trust are currently under negotiation.

The paper therefore distinguishes *symbol inflation* from symbolic worthlessness. Symbols remain indispensable for large-scale social coordination. Credentials, signatures, reports, brands, documents, rituals, images, and linguistic performances compress complex histories into portable forms. The difficulty is that the receiver often encounters the compressed surface without direct access to the generative

process. When many heterogeneous processes can produce a similar surface, the surface becomes a less determinate guide to its history. The central problem is located in this changing relation between representation and inference. The narrower term *symbolic crisis* is reserved for cases in which diagnostic mismatch becomes socially consequential while interpretive and institutional practices remain unsettled.

The analysis also distinguishes the crisis from a narrow problem of deception. A human can use generative AI responsibly, understand and endorse the resulting content, and still alter the evidential meaning of the artifact for a receiver. Conversely, a purely human-produced message can be deceptive, formulaic, or strategic. The relevant transformation concerns the structure of inference: what can reasonably be learned about a communicator, institution, or practice from the observable symbol alone?

A second distinction concerns communication itself. The paper resists treating communication as the mere presentation of symbols. Human communication can include mutual updating, responsiveness, vulnerability to reply, responsibility for commitments, and historical coupling among participants. AI-mediated communication can preserve some of these elements, transform others, and in some settings permit symbolic interaction with much less reciprocal human participation. The resulting difficulty is therefore partly ontological: what kind of event has occurred when a highly personalized message is generated and sent with little direct attention from the named sender?

A third focus concerns social value beyond interpersonal trust. Cultural and commercial systems often depend on symbolic scarcity, distinction, provenance, and recognized authorship. Symbol inflation can be intensified when the production cost of particular styles, narratives, images, and forms of personalization falls, or when organizations become more capable of strategically engineering symbolic associations at scale. This may displace scarcity toward other objects such as access, temporal depth, place, participation, craft history, or social recognition. The paper treats this as a political-economic problem because the capacity to create, certify, and control scarce relational conditions is unevenly distributed.

Existing provenance infrastructures illustrate one important institutional response while also revealing the limits of a purely technical solution. The C2PA specification defines provenance in terms of the history of an asset and its interactions with actors and other assets, and provides mechanisms for tamper-evident Content Credentials (Coalition for Content Provenance and Authenticity 2026). Such provenance can clarify how a digital artifact was produced or transformed. Yet an asset's technical provenance does not by itself establish whether a person understood, endorsed, cared about, or was meaningfully transformed by the communication. The diagnostic problem therefore extends beyond source verification.

The paper is intentionally separated from a companion inquiry into possible post-crisis practices. It does not assume that direct co-experience can replace symbols at scale; co-experience is costly in time, attention, exposure, and relational bandwidth. Nor does it assume that disclosure, provenance, or individual authenticity norms can solve an era-level transformation. The present task is to identify what has become unstable: the social compression performed by symbols, the inference from surface to generative history, the location of agency in mediated communication, and the historical habits through which people learn what symbolic forms deserve trust.

The intended contribution is a vocabulary for a transitional condition rather than a finished social theory. Symbolic diagnosticity is the organizing concept; symbol inflation names the broader condition under which established cue–history relations can become unstable. The paper also develops generative history, agency decomposition, symbolic compression, scarcity displacement, and temporal mismatch. It asks how these concepts interact across interpersonal communication, expertise, education, culture, luxury, credentials, and institutional legitimacy. The claim remains historically modest: contemporary generative systems appear to be changing the conditions under which symbolic surfaces can function as evidence, while human perception and institutions are likely to adapt more slowly and unevenly.

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. The framework developed here is diagnostic and should not be interpreted as a justification for deceptive communication, manipulation of trust, or strategic simulation of human relationships. The analysis of symbolic scarcity and credibility is intended to clarify changing social conditions rather than to provide techniques for exploiting them. 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 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.

DRAFT

Notices

Status. This working draft records an early stage of a conceptual inquiry into changing symbolic diagnosticity. The term *symbol inflation* identifies a context-dependent expansion and strategic intensification of symbolic production rather than a universal historical period. The narrower term *symbolic crisis* is retained for cases in which consequential diagnostic mismatches and unsettled interpretive practices coincide. The paper does not assume that all symbolic signals lose credibility, that AI-generated communication is necessarily deceptive, or that human-authored artifacts are inherently more valuable. Empirical claims are treated as domain-specific and revisable.

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, and an indication of changes. The licence deed and legal code are available at <https://creativecommons.org/licenses/by-nc/4.0/>.

Statement on the use of language models. The exploratory discussions and preparation of this paper involved OpenAI’s ChatGPT. ChatGPT supported exploratory dialogue, source discovery followed by website verification, conceptual reconstruction, argumentative criticism, and LaTeX drafting. The author selected the research problem, directed and approved theoretical commitments, and bears sole responsibility for the manuscript, including its definitions, claims, interpretations, and errors. Authorship credit remains with the human author.

1 Introduction

This section establishes the structural problem of the paper. Its objective is to specify *symbolic diagnosticity* under conditions of *symbol inflation*. The discussion first reconstructs the social function of symbols as compression interfaces through which recipients infer histories that they do not directly observe. It then identifies several mechanisms through which the production and strategic circulation of symbolic surfaces can expand, with generative AI treated as one especially powerful contemporary accelerator. The section next distinguishes changing diagnosticity from a universal decline of trust or symbolic value. Finally, it reserves the term *symbolic crisis* for the narrower downstream condition in which consequential diagnostic mismatches persist while shared interpretive and institutional practices remain unsettled.

Symbol inflation does not require artificial intelligence. Firms and institutions have long invested in brands, credentials, rankings, visual identities, professionalized communication, advertising, public relations, ritual forms, and other symbolic constructions that shape recognition, expectation, status, or demand. Digital publishing, platform distribution, templating, automation, and data-driven personalization further increase the quantity and strategic adaptability of such forms. The present argument therefore concerns a broader political economy of symbolic production. Generative AI matters because it can accelerate several of these processes at once while widening the range of actors able to produce polished and differentiated symbolic surfaces.

Artificial intelligence-mediated communication predates the current generation of large generative models. Predictive text, automated suggestions, profile construction, and algorithmic language assistance already complicated the relation between communicator and message. Hancock, Naaman, and Levy therefore defined AI-mediated communication broadly as interpersonal communication in which an intelligent agent modifies, augments, or generates messages on behalf of a communicator (Hancock et al. 2020). What has changed more recently is the scale and quality of accessible generation. A user can now produce fluent correspondence, reports, images, presentations, summaries, personal messages,

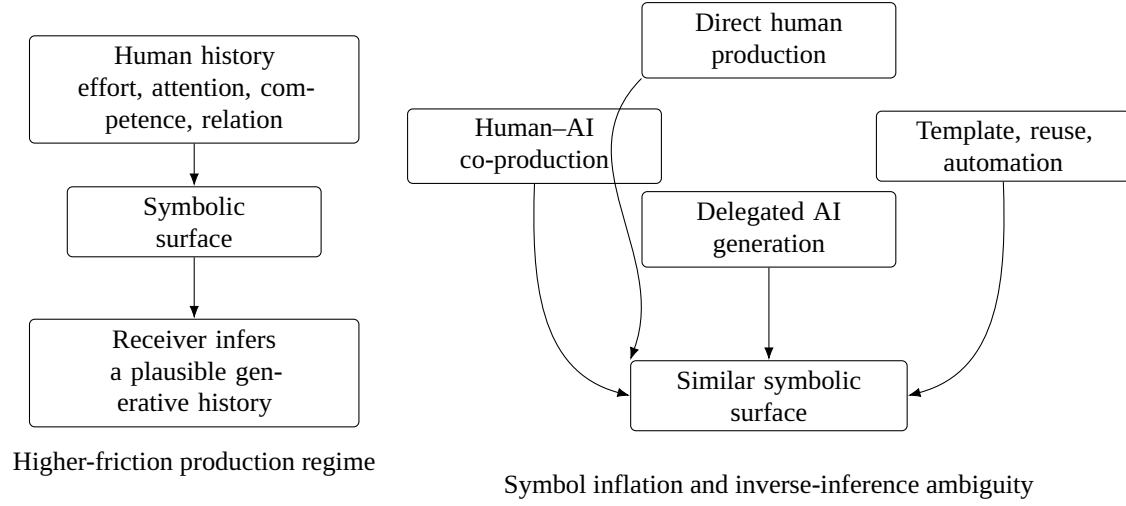


Figure 1: Conceptual contrast between a higher-friction symbolic regime and a condition in which heterogeneous generative pathways can converge on similar symbolic surfaces. The diagram does not imply that earlier symbolic inference was reliable in every case or that current symbols lack evidential value. It isolates the inverse-inference problem examined in this paper.

and stylistically differentiated texts with comparatively little direct production effort. The relevant transformation is thus neither the first appearance of mediation nor the first possibility of deception. It is a change in the set of generative pathways that can lead to a similar symbolic surface.

The distinction in Figure 1 can be expressed without assuming that a symbolic artifact uniquely reveals its origin. Let S denote an observable symbolic artifact and let H denote a generative history that can include human attention, prior competence, institutional process, model assistance, editing, delegation, and relational context. The receiver encounters S and forms some inference about H . The present problem concerns a widening set of histories consistent with a similar surface. The relevant set is defined in Equation 1.

$$\mathcal{H}(S) = \{H_1, H_2, \dots, H_n\}. \quad (1)$$

As $\mathcal{H}(S)$ becomes more heterogeneous, the artifact can remain useful as content while becoming less determinate as evidence about the process that produced it. This is an inverse-inference problem rather than a claim that the content is false.

Definition 1.1 (Symbol inflation). *Symbol inflation* is a domain-specific condition in which the production, circulation, or strategic construction of socially consequential symbolic forms expands faster than the interpretive and institutional capacities used to distinguish the histories, properties, or commitments those forms are taken to indicate. Symbol inflation does not imply that every symbol loses value. Diagnosticity can weaken, remain stable, or migrate toward additional cues and contexts.

The term deliberately extends beyond a monetary analogy. The relevant process is an increase in symbolic supply or strategic symbolic intensity together with a change in the evidential relations through which symbols acquire social force. A market can contain more branding without every brand becoming worthless; a professional field can contain more polished performance without every performance becoming uninformative. The analytical question is which target inferences become harder to sustain from the surface alone.

Definition 1.2 (Symbolic crisis). A *symbolic crisis* is a narrower condition in which symbol inflation or another change in symbolic production produces consequential diagnostic mismatches while inherited interpretive conventions and institutional practices remain insufficiently recalibrated. Crisis therefore

names a possible phase or domain-specific outcome of changing diagnosticity rather than the paper's primary object.

These definitions contain a historical component. A polished letter, carefully composed report, technically accomplished image, personalized response, or specialized explanation has never guaranteed sincerity, competence, or direct human attention. Forgery, ghostwriting, formulaic politeness, bureaucratic language, delegated correspondence, advertising, and strategic self-presentation long predate generative AI. The proposed crisis therefore cannot rest on a nostalgic contrast between a truthful human past and an artificial present. The relevant change is comparative. Some forms of symbolic production now require less direct human participation than recipients' inherited interpretive habits may continue to presume.

This temporal mismatch matters because symbolic forms perform a practical compression function. Large societies cannot require direct co-experience before every act of reliance. Credentials compress educational histories; signatures compress authorization; brands compress accumulated reputational expectations; professional prose can signal familiarity with a field; ritual forms compress institutional continuity; personal messages can signal attention; and aesthetic objects can carry assumptions about craft, scarcity, and authorship. These compressions are imperfect, but their imperfection does not make them socially unnecessary. They reduce the cost of coordination by allowing recipients to act on portable representations of histories they did not personally witness.

Diagnostic pressure increases when the compression interface remains visible while the relation between interface and compressed history becomes harder to infer. A sophisticated explanation can be produced by a domain expert, by a novice working carefully with a model, by an automated agent, or by a person who understands only part of the resulting text. A warm email can reflect sustained attention, minimal prompting, institutional automation, or some combination of these. The symbolic surface can therefore continue to accomplish an immediate communicative task while carrying less information about who participated, how much they understood, what they endorse, and what relational exposure occurred in producing it.

Research on costly signaling provides one partial background for this problem. Observable expenditures of time, effort, emotion, reputation, or sacrifice can contribute to perceived honesty because recipients interpret some costs as difficult to fake, verifiable, or self-sacrificing (Chaudhry and Wald 2022). Experimental work on joint action similarly suggests that effort invested in informative sensorimotor coordination can increase trust and generosity toward a co-actor (McEllin and Michael 2022). These findings do not establish that trust should be allocated according to effort. High effort can produce falsehood, while low-effort AI assistance can support accurate and responsible communication. Their relevance is narrower: social inference often uses cues about the process of communication, not merely the propositional content of a message.

Generative AI changes the availability and interpretation of some of those process cues. This can produce a difficult empirical pattern. Sahebi and Formosa argue that AI-mediated communication creates an epistemic dilemma because reduced trust can sometimes be justified and sometimes produce an unjustified credibility deficit (Sahebi and Formosa 2025). Schilke and Reimann report that disclosure of AI use reduced trust across the tasks studied in thirteen experiments, with legitimacy perceptions mediating the effect (Schilke and Reimann 2025). Sahebi, Formosa, and Bankins find lower perceived trustworthiness, authenticity, and knowledge uptake for AI-involved written communication in their preregistered experiment, while also reporting that participants valued disclosure even as disclosure itself carried a penalty (Sahebi, Formosa, and Bankins 2026). In marketing communication, Kirk and Givi likewise report lower positive responses to AI-authored emotional messages in their experimental settings, with authenticity playing an explanatory role (Kirk and Givi 2025).

These findings are important precisely because they do not yield a simple rule. They do not establish that AI mediation should always lower trust, that disclosure always improves communication, or that recipients possess stable norms for interpreting AI assistance. They indicate an unsettled transition in which the social meaning of authorship, assistance, authenticity, and legitimacy is being renegotiated.

Where these mismatches become consequential and socially unsettled, the narrower symbolic-crisis diagnosis becomes relevant. The underlying problem is collective rather than an individual moral instruction. A person cannot resolve it merely by deciding to trust different signals tomorrow. Interpretive habits are learned through institutions, technologies, repeated practices, educational systems, markets, and accumulated social experience.

Proposition 1.3 (Temporal mismatch). *A symbolic crisis becomes especially salient when the production conditions of symbolic artifacts change faster than the social practices through which those artifacts are interpreted. The resulting mismatch can preserve inherited responses to symbolic form after the evidential relations that supported those responses have materially changed.*

This proposition helps explain why contemporary recipients can experience a well-formed message in two incompatible ways. The message can still produce the ordinary phenomenology of attention, expertise, or warmth, while knowledge of AI involvement can abruptly reclassify that same artifact as less authentic or less trustworthy. The issue is neither located wholly in the text nor wholly in the recipient. It arises from a historically learned relation between symbolic form and assumptions about its generation.

The problem also reaches beyond trust. Cultural and economic value can depend on scarcity, distinctiveness, authorship, provenance, and recognized participation. Generative systems can lower the production cost of particular visual styles, narrative forms, personalized messages, and apparently expert performances. This does not abolish scarcity. It can displace scarcity toward other objects: access to specific communities, temporal depth, embodied participation, place-specific experience, craft history, recognized provenance, or durable relational trajectories. Such displacement creates a political-economic question because control over those scarce conditions is unevenly distributed. Luxury, branding, cultural distinction, and professional authority can all be reorganized when symbolic exclusivity becomes easier to imitate.

Nor is the crisis exhausted by content authenticity. Technical provenance systems can help recipients understand how a digital artifact was created and modified. The C2PA technical specification, for example, defines provenance in terms of the history of an asset and its interaction with actors and other assets, and provides mechanisms for tamper-evident Content Credentials (Coalition for Content Provenance and Authenticity 2026). Such infrastructure may become increasingly important. Yet a verified production history does not by itself answer whether the named human understood the message, endorsed its commitments, experienced the relation it describes, or remains available for reciprocal response. Technical provenance and social meaning therefore intersect without becoming equivalent.

This point leads to a broader reconsideration of communication. If communication is defined only as the successful delivery of symbols, then an automated system that produces semantically appropriate outputs can satisfy much of the definition. Human communicative practice, however, can also include reciprocal updating, vulnerability to reply, attribution of responsibility, historical continuity, and the possibility that participants are changed by interaction. AI mediation can preserve some of these dimensions and weaken or redistribute others. The relevant question is consequently not whether a symbol is “really human,” but what kind of relational event the symbol participates in.

Working Claim 1.4 (Diagnostic displacement). Under generative abundance, the evidential center of gravity can shift away from the isolated symbolic surface toward additional information about generative history, agency, temporal continuity, relation, and responsibility. This claim is diagnostic rather than prescriptive: it identifies a pressure on existing trust practices without specifying the institutions that should replace them.

The distinction between diagnosis and prescription is essential for the structure of this project. Direct co-experience can generate information that an isolated symbol cannot provide, yet co-experience is expensive in time, attention, emotional exposure, and coordination. A large society cannot replace symbolic mediation with deep personal interaction. Similarly, individual trajectory records, provenance

mechanisms, disclosure conventions, and new institutional credentials may strengthen particular forms of inference without resolving the underlying historical transition by themselves. These possible responses belong to a companion inquiry. The present paper first asks what has become unstable.

The following sections therefore develop the diagnosis across several domains. Section 2 reconstructs symbolic mediation as social compression. Section 3 examines generative abundance and changes in production friction. Section 4 develops the concept of symbolic diagnosticity. Sections 5–8 examine communication, delegated agency, trust, disclosure, and authenticity. Sections 9–13 extend the analysis to expertise, education, institutional credentials, culture, luxury, and the political economy of symbolic abundance. Sections 14–16 address provenance, temporal mismatch, and historical adaptation. The final sections clarify the boundaries of the symbolic-crisis concept and identify open problems for empirical and philosophical research.

Research Obligation 1.5 (Historical discipline). Claims about symbolic crisis should compare specific symbolic practices across identifiable production regimes rather than treating “before AI” and “after AI” as homogeneous periods. The analysis should remain open to cases in which symbolic credibility is preserved, strengthened, or reorganized rather than generally diminished.

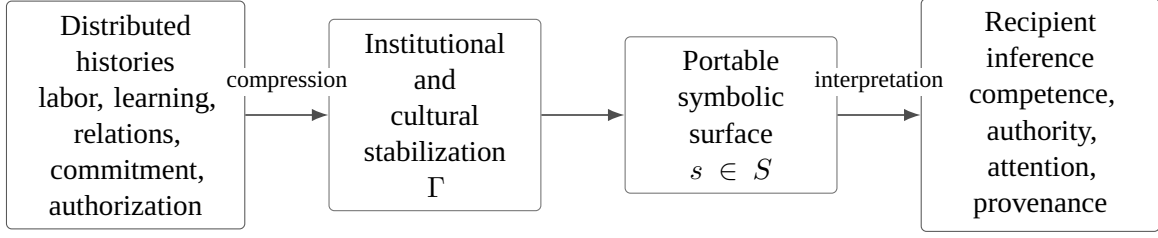
This obligation keeps the concept from becoming a totalizing description of the present. The symbolic crisis is proposed as a way to investigate a changing relation among representation, production, and social inference. Its value will depend on whether later sections can identify where that relation has changed, where it has remained robust, and which social domains exhibit genuinely new forms of ambiguity rather than familiar problems under a new technological label.

2 Symbolic Mediation and Social Compression

This section establishes the social function that makes the symbolic crisis consequential. Its objective is to explain how symbolic forms permit coordination among actors who cannot directly inspect one another’s histories, capacities, intentions, or relations. The discussion proceeds in six steps. It first situates symbolic mediation within distributed social life. It then introduces *social compression* as an analytical term for the reduction of extended histories into portable symbolic surfaces. It next distinguishes evidential, coordinative, and performative functions of symbols. The section then examines the institutional stabilization of symbolic inference, the lossiness and ambiguity of compression, and the relation between symbolic mediation and limited relational bandwidth. The final part clarifies why the symbolic crisis concerns a transformation of compression conditions rather than the disappearance of mediation itself.

Human social life routinely requires action under conditions of inaccessible history. A recipient of a diploma normally does not observe the years of study, examination, supervision, institutional selection, and administrative verification condensed into the credential. A customer does not directly witness the full production history of a brand. A recipient of an email does not usually observe the attention, drafting, revision, or delegation that preceded the message. A signature does not display the entire institutional process through which the signer acquired authority. In each case, social action proceeds through a limited observable surface that permits inferences about a much larger and only partially accessible history.

This problem resembles, without being reducible to, the problem of uncertainty in signaling. Spence’s account of job-market signaling showed how observable features can carry information when relevant qualities are not directly visible to the receiver (Spence 1973). Luhmann’s analysis of trust approached the issue at a broader social scale by treating trust as a mechanism through which actors cope with complexity that exceeds immediate calculation or observation (Luhmann 1979). The present paper joins these intuitions to a more general claim: symbolic forms frequently function as socially stabilized interfaces through which large, distributed, and temporally extended histories become usable in comparatively low-bandwidth encounters.



The symbolic surface does not reproduce the history. It makes selected aspects of that history socially portable under conventions that can remain stable, contested, or historically revised.

Figure 2: Social compression as a mediated relation between distributed histories and portable symbolic surfaces.

Definition 2.1 (Social compression). Let H denote a set of socially relevant histories, relations, practices, or institutional processes, and let S denote a set of observable symbolic forms. *Social compression* refers to a historically and institutionally mediated relation $C_\Gamma : H \rightarrow S$ through which a comparatively small symbolic surface becomes usable as a portable indicator, authorization, memory, or coordination device for a larger social history. The parameter Γ represents the conventions, institutions, interpretive habits, and enforcement structures that stabilize the relation.

The term *compression* is analytical rather than computational. It does not imply that social meaning can be reduced to an information-theoretic code or that a symbolic object stores a complete representation of its history. Its purpose is to capture an asymmetry of scale: a recipient may inspect a short document, title, seal, score, logo, signature, or sentence while relying on a much larger set of practices that the symbol does not reproduce. Figure 2 summarizes this relation.

2.1 Symbolic Mediation in Distributed Social Life

The importance of symbolic mediation increases with social distance, scale, and specialization. Direct acquaintance can provide rich information about a person’s conduct, responsiveness, memory, and practical competence. Large societies cannot rely on this form of knowledge for most interactions. A patient cannot apprentice under every physician before accepting treatment. A firm cannot personally reconstruct every educational history represented in an applicant’s record. A citizen cannot directly observe every administrative act carried out under a public office. Social coordination therefore depends on artifacts and classifications that make selected relations portable across distance.

The portability of a symbol depends on shared interpretive infrastructure. A diploma is useful because recipients recognize an institution, an educational category, and a set of procedural expectations. A legal signature matters because a surrounding institutional order assigns consequences to signing. A professional title has meaning because communities distinguish the title from neighboring categories and connect it to expected forms of qualification or responsibility. Even a personal message relies on learned conventions through which particular linguistic forms can be interpreted as attention, apology, affection, expertise, or commitment.

Goffman’s analysis of self-presentation emphasizes that social interaction is structured through performances, settings, expressive equipment, and audience interpretation (Goffman 1959). The present framework extends this insight toward the infrastructural side of symbolic mediation. A social performance becomes scalable when recognizable forms allow recipients to orient themselves without reconstructing the complete interactional history behind those forms. The resulting portability supports coordination while also creating room for strategic presentation, misrecognition, and contestation.

Proposition 2.2 (Scalability through symbolic compression). *As the direct observation of relevant social histories becomes more costly, the coordination value of institutionally stabilized symbolic compression tends to increase, provided that recipients retain sufficient grounds for interpreting the symbolic surface as connected to the relevant history or authority.*

This proposition does not imply that compressed signals are fully trusted. Many social environments combine provisional symbolic reliance with audits, recommendations, guarantees, sanctions, or repeated interaction. The point is narrower: symbolic mediation reduces the amount of direct relational access required before coordination can begin.

2.2 Compression of Inaccessible Histories

Social compression can be represented as a many-to-one relation. Several distinct histories may produce similar symbolic surfaces. This many-to-one relation is represented in Equation 2.

$$C_{\Gamma}(h_1) = C_{\Gamma}(h_2) = s, \quad h_1 \neq h_2. \quad (2)$$

The recipient therefore faces an inverse problem. Observation of s does not uniquely recover the history that produced it. Instead, the recipient infers a set of plausible histories conditional on the conventions and institutional environment in which the symbol appears.

This many-to-one structure is present even before generative AI. Two graduates holding the same degree can possess different competencies. Two signatures can express different levels of understanding or commitment. Two polished reports can result from different divisions of labor. Two apologies can have similar wording while differing in sincerity, reflection, or consequence. Symbolic compression has therefore always been lossy.

The social usefulness of compression does not require exact reconstruction. A credential can be useful when it narrows uncertainty enough for a subsequent interaction. A brand can coordinate expectations without revealing every production decision. A professional title can establish a starting presumption that remains open to later revision. The relation can therefore be informative even when the inverse mapping remains incomplete.

Remark 2.3 (Compression and equivalence). Social compression does not establish equivalence between a symbolic surface and the history associated with it. The symbol supports a socially organized inference, authorization, or act. Its practical role can remain substantial even when the underlying histories are heterogeneous.

This distinction matters for the symbolic-crisis argument. The paper does not depend on a nostalgic premise that symbols once transparently revealed reality. Historical symbolic orders contained forgery, strategic display, counterfeiting, credential inflation, ritualized language, and institutional exclusion. The relevant baseline is therefore probabilistic and institutional, not transparent. A symbolic crisis can emerge when previously workable inferential relations become less stable, less discriminating, or differently distributed across contexts.

2.3 Evidential, Coordinative, and Performative Functions

A symbolic form can participate in social life through several functions at once. First, it can have an *evidential function*: a recipient treats the surface as evidence concerning competence, provenance, attention, status, or intention. Second, it can have a *coordinative function*: the surface provides a common reference that allows actors to organize action without reconstructing every relevant history. Third, it can have a *performative function*: the symbol participates in creating a social relation or institutional consequence.

Austin’s analysis of performative utterances remains important for the third function because some expressions do more than describe a pre-existing state; under appropriate conventions, their utterance can constitute an act (Austin 1962). A promise, declaration, authorization, or contractual formula can acquire force through the conditions surrounding its use. The symbolic surface in such cases is connected to an institutional practice in which saying, signing, or recording contributes to the event itself.

These functions should remain analytically separate. A signature may validly authorize an act even when it provides weak evidence of the signer’s understanding. A polished report may coordinate organizational action even when its authorship is distributed. A credential can retain legal or institutional status while its evidential value for a particular competence changes. The symbolic crisis can therefore affect one function while leaving another comparatively stable.

Definition 2.4 (Compression profile). For a symbolic form s , its *compression profile* is the context-dependent configuration defined in Equation 3.

$$\Pi(s) = \langle E_s, C_s, P_s, A_s \rangle, \quad (3)$$

In Equation 2, E_s denotes evidential function, C_s coordinative function, P_s performative or constitutive function, and A_s the surrounding accountability or enforcement structure. The components need not move together when production conditions change.

The compression profile prevents the analysis from treating credibility as the sole social function of symbols. A decline in evidential diagnosticity does not automatically eliminate institutional usability. Conversely, a legally valid symbol can remain socially distrusted. This separation will become important later when the paper examines credentials, AI-mediated communication, provenance systems, and institutional legitimacy.

2.4 Institutional Stabilization of Symbolic Inference

A symbolic surface becomes useful through more than its visible form. Institutions stabilize the relation between the surface and the histories or capacities that recipients infer from it. Examinations, accreditation, professional regulation, contractual enforcement, archives, identity systems, reputational sanctions, and organizational procedures all help sustain particular compression relations.

Spence’s signaling model illustrates one mechanism through which observable characteristics can become informative when different underlying types face different signaling conditions (Spence 1973). The symbolic-crisis framework takes a broader view. A signal can acquire social credibility through heterogeneous mechanisms: production cost, difficulty of imitation, third-party certification, legal consequence, repeated exposure, reputational risk, technical provenance, or embeddedness within an institution that can be held accountable. Costly signaling is therefore one component of a larger ecology of stabilization.

The same surface can consequently change meaning when its institutional environment changes. A title issued by an unknown entity, a badge without enforceable criteria, or a signature detached from accountable identity may preserve visual form while losing much of the relation that previously made it informative. Conversely, relatively simple symbols can become highly consequential when institutions reliably connect them to authorization, liability, or verification.

Working Claim 2.5 (Institutional support of compression). The practical reliability of a symbolic compression relation depends jointly on the observable form and on the institutions, conventions, histories, and accountability structures that support interpretation of that form. Changes in any of these components can alter the inferences recipients reasonably draw from an otherwise unchanged surface.

This claim provides one reason that the symbolic crisis should be treated as institutional as well as technological. Generative systems alter production possibilities, yet the resulting social consequences

depend on how workplaces, universities, professions, platforms, families, legal systems, and cultural industries revise the conventions surrounding symbolic artifacts.

2.5 Lossiness, Ambiguity, and Strategic Presentation

Every compression relation omits information. The omission can be socially productive because it allows an actor to proceed without processing the full history. It can also create ambiguity because different histories remain compatible with the same surface. This dual character is structural rather than accidental.

The receiver therefore operates with a posterior interpretation rather than a recovered history. This posterior interpretation is represented schematically in Equation 4.

$$P(h \mid s, \Gamma) \quad (4)$$

represents the receiver’s context-sensitive inference concerning histories h after observing symbolic surface s under interpretive environment Γ . The expression is conceptual; it does not imply that recipients perform explicit Bayesian calculation. It emphasizes that symbolic meaning depends on priors, conventions, institutions, and contextual knowledge as well as on the surface itself.

Strategic presentation becomes possible because actors can shape s while recipients care about properties located in h . Goffman’s account of impression management provides a classical sociological vocabulary for this gap between presented self and inferred social reality (Goffman 1959). Contemporary generative systems expand the technical capacity to shape the surface, but the underlying problem of presentation and inference is older.

The important historical question is therefore comparative. Which kinds of surface previously required a narrower set of generative histories? Which forms were difficult, expensive, institutionally constrained, or socially risky to reproduce? Which inferential shortcuts depended on these production conditions? Which of those conditions are now changing? The next section will address these questions through the concept of production friction.

2.6 Relational Bandwidth and the Necessity of Compression

The persistence of symbolic mediation follows from a basic resource constraint. Direct co-experience requires time, attention, exposure, and repeated interaction. These resources are limited. A person cannot build deep relational histories with every professional, institution, seller, researcher, official, collaborator, or stranger whose actions matter to daily life.

Let B_i denote the relational bandwidth available to actor i , understood as the limited capacity for direct observation, repeated interaction, and co-experience. Let D_i denote the number and diversity of social dependencies relevant to that actor. As D_i exceeds what can be directly verified within B_i , mediated forms of trust and coordination acquire increasing practical importance. Symbolic compression can therefore be understood as one response to the gap between social interdependence and finite relational bandwidth.

This resource perspective also clarifies why a response to symbolic crisis cannot consist simply in replacing symbolic mediation with direct experience. Such a replacement would impose very high verification costs and would remain infeasible across many large-scale institutions. The diagnostic problem is therefore to understand how compression relations change when symbolic surfaces become easier to generate while relational bandwidth remains scarce.

Proposition 2.6 (Compression under relational scarcity). *When socially relevant dependencies exceed available relational bandwidth, some form of mediated compression remains necessary for scalable coordination. A decline in the diagnostic value of particular symbolic surfaces therefore creates pressure for changes in mediation, certification, provenance, or trust practice rather than a general return to direct verification.*

This proposition prepares the larger argument of the paper. The symbolic crisis is consequential because symbols solve a real coordination problem. Their compression function cannot simply be discarded. At the same time, the relation between surface and history is historically contingent. When new technologies alter the production conditions of surfaces, the inherited interpretive infrastructure can lag behind.

2.7 Compression Conditions before Generative Abundance

The analysis does not require an idealized pre-AI period. Symbolic systems have always contained ambiguity, convention, power, strategic presentation, and unequal access to certification. The relevant historical contrast concerns production conditions and inferential habits, not purity.

Many symbolic practices nevertheless developed under environments in which particular forms were coupled to nontrivial frictions. A carefully written letter generally required more direct linguistic production than a templated acknowledgment. A polished visual artifact required access to relevant skills or production resources. A detailed report often implied some combination of research, drafting, editing, institutional support, or delegated labor. A personalized message often carried at least a weak association with attention because personalization itself consumed time. Costly-signaling research provides one empirical reason such frictions can matter: observable expenditure of time, effort, emotion, reputation, or sacrifice can affect perceived honesty (Chaudhry and Wald 2022).

These associations were never universal, yet they could become embedded in ordinary interpretation. Recipients learned to treat particular surfaces as provisional traces of particular histories. Institutions likewise developed procedures that assumed relatively stable relations among production cost, expertise, authorship, and symbolic form.

Generative abundance changes this environment by widening the set of histories capable of producing similar surfaces. The next section therefore turns from the general architecture of social compression to the changing production friction of symbolic artifacts. The central issue is neither that compression has failed everywhere nor that symbolic surfaces have become meaningless. It is that the mapping between history and surface can become increasingly many-to-one while inherited habits of inference remain calibrated to earlier production conditions.

3 Production Friction under Symbol Inflation

This section specifies one major production-side mechanism of symbol inflation: the reduction and redistribution of friction involved in reaching socially recognizable symbolic surfaces. The objective is narrower than explaining symbol inflation as a whole. Symbolic production can intensify through branding, institutional signaling, mass communication, platform distribution, automation, and other organizational processes as well as through generative AI. The present section focuses on generative systems because they can reduce selected temporal, cognitive, technical, and coordination frictions while widening the range of pathways capable of producing similar observable forms. The discussion defines production friction as a multidimensional property of a pathway, treats *generative abundance* as one mechanism within the broader inflationary condition, examines the multiplication of generative histories behind similar surfaces, and traces the migration of costs toward verification, attention, selection, and responsibility. The final part identifies the transition from production friction to diagnostic ambiguity that motivates the following section.

The concept of production friction should be separated from a nostalgic contrast between arduous human production and effortless machine production. Human symbolic production has long involved templates, secretaries, editors, translators, bureaucratic routines, software, search engines, stock imagery, and delegated labor. Generative AI enters an already mediated production ecology. Its significance lies in the scale, generality, speed, and accessibility with which some forms of symbolic production can be

Table 1: Illustrative dimensions of production friction and their possible transformation under generative assistance.

Dimension	Higher-friction production condition	Possible transformation under generative assistance
Temporal expenditure	Drafting, formatting, revision, and adaptation consume substantial elapsed time.	First-pass production, reformulation, and variation can occur at substantially higher speed.
Expertise requirement	Surface quality is strongly coupled to linguistic, visual, technical, or disciplinary skill.	Some surface-level competence becomes accessible through model-mediated generation and editing.
Attention requirement	Personalization and stylistic adaptation require sustained human attention to each artifact.	Reusable context and automated generation permit large numbers of individually adapted outputs.
Coordination burden	Production requires several people, specialist roles, or sequential handoffs.	Some functions can be integrated into a single human-model workflow, while new review roles may arise.
Resource requirement	Specialized software, production infrastructure, or professional services restrict access.	General-purpose interfaces lower some access barriers while retaining computational and platform dependencies.
Institutional constraint	Authorization, accountability, certification, or legal responsibility constrains who may validly produce the artifact.	Formal constraints can remain largely intact even when the visible surface becomes easier to generate.

reorganized. The historical question is therefore comparative: for a given class of symbolic surfaces, which combinations of time, expertise, attention, coordination, and institutional access were previously common, and which combinations become feasible after generative systems enter the production process?

Definition 3.1 (Production friction). Let p denote a generative pathway capable of producing symbolic surface s . The *production-friction profile* of p relative to s is the context-dependent vector defined in Equation 5.

$$\Phi(p, s) = \langle T, E, A, C, R, I \rangle, \quad (5)$$

In Equation 5, T denotes temporal expenditure, E domain or expressive expertise, A sustained attentional input, C coordination burden, R material or computational resource requirements, and I institutional or consequential constraints attached to production. The components can move independently and need not admit a common scalar measure.

Production friction is thus broader than monetary cost. A message can be inexpensive in money while demanding substantial attention. A report can be generated quickly while still requiring institutional authorization, factual verification, and later responsibility. A visual artifact may require little domain expertise at the point of generation while depending on expensive computational infrastructure supplied elsewhere. The vector form prevents the analysis from treating all forms of friction as interchangeable.

Table 1 makes two points that will remain important throughout the paper. First, generative systems can lower particular frictions without lowering all of them. Second, the frictions most directly connected to visible symbolic form can decline while accountability, authorization, verification, or liability remain comparatively resistant to automation. A polished artifact can therefore become easier to produce without becoming easier to warrant.

3.1 Generative Abundance within Symbol Inflation

For a symbolic surface s , let $\mathcal{P}(s)$ denote the set of generative pathways capable of producing a surface that recipients would classify as belonging to the same practically relevant form. A pathway can

include direct human composition, collaborative production, template use, organizational delegation, model-assisted drafting, model-generated first passes, or more autonomous agentic production. The symbolic-crisis argument concerns changes in this pathway set.

For a context-dependent friction ceiling τ , the corresponding definition is given in Equation 6.

$$\mathcal{P}_\tau(s) = \{p \in \mathcal{P}(s) : \Phi(p, s) \preceq \tau\}, \quad (6)$$

In Equation 6, $\preceq$ indicates that the relevant dimensions remain within an acceptable friction profile for the actor or institution. The notation is schematic because the dimensions need not be commensurable. It provides a way to express the central historical change: generative technologies can make pathways feasible that were previously too slow, too skill-intensive, too coordination-heavy, or too expensive for routine use.

Definition 3.2 (Generative abundance). For a class of symbolic surfaces $\mathcal{S}$, *generative abundance* denotes a condition in which the number and diversity of feasible generative pathways increase substantially across relevant actors while one or more important components of production friction decline. The concept concerns expanded production possibility rather than a claim of zero cost, universal access, or unlimited quality.

This definition makes generative abundance a production-side mechanism of symbol inflation rather than a synonym for the broader condition. Mass printing, photography, recorded media, office software, and the internet each increased the availability of symbolic objects. Generative abundance concerns a further change in the relation between a desired symbolic form and the pathway needed to produce it. The actor can increasingly request, vary, personalize, transform, or regenerate the form rather than merely copy an existing artifact.

Experimental evidence provides concrete support for this friction-reduction claim in bounded domains. In a preregistered experiment involving professional writing tasks, access to ChatGPT reduced average completion time by approximately 40 percent while average output quality increased by 18 percent (Noy and Zhang 2023). A later field study of 5,172 customer-support agents found that access to a generative AI assistant increased issues resolved per hour by 15 percent on average, with heterogeneous effects across workers and task types (Brynjolfsson et al. 2025). These studies do not establish a universal productivity law. They demonstrate that, in identifiable forms of language-intensive work, generative systems can alter the time and expertise relations through which usable symbolic outputs are produced.

Proposition 3.3 (Expansion of feasible pathways). *When generative assistance lowers one or more binding components of production friction for a class of symbolic surfaces, the set of actors and generative histories capable of reaching those surfaces can expand even if formal authorization, factual validity, or downstream responsibility remains unchanged.*

The proposition is deliberately asymmetric. Easier production of the surface does not imply easier production of the underlying competence, evidence, institutional authority, or relational history that recipients may associate with the surface. This asymmetry is the central bridge from falling production friction to changing symbolic diagnosticity.

3.2 Surface Quality and Historical Friction

Many symbolic inferences rely implicitly on a relationship between visible quality and historical friction. A fluent technical explanation can suggest expertise because producing such explanations has often required familiarity with the domain. A carefully personalized message can suggest attention because personalization has historically consumed time. A polished presentation can suggest organizational resources because coherent visual production has required relevant skills, software, or specialist labor. These relations have always been probabilistic and institutionally variable.

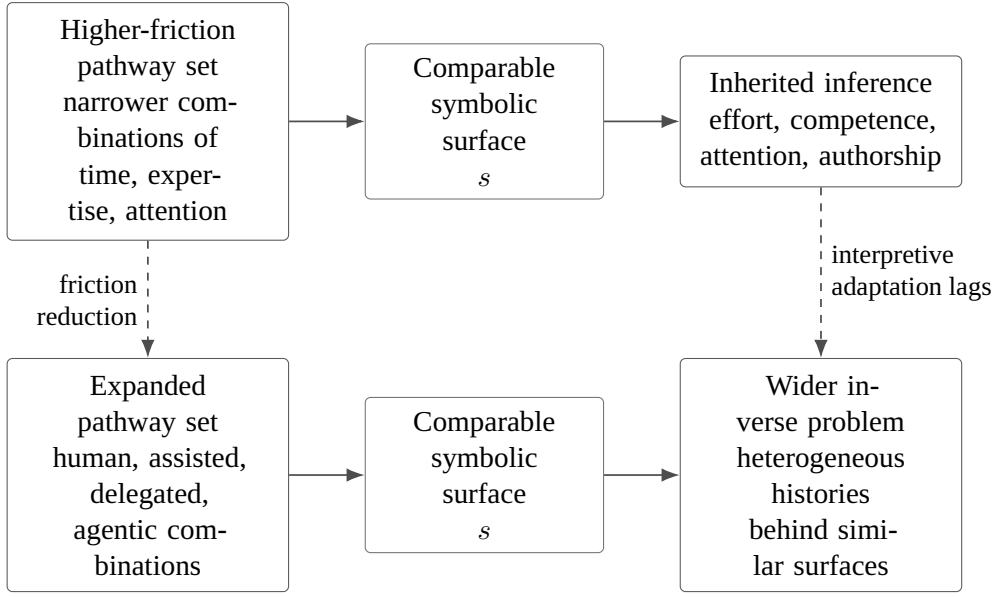


Figure 3: Generative abundance expands feasible production pathways while inherited surface-to-history inferences can persist.

Generative systems weaken some of these relations by changing the production function for surface quality. Let $q(s)$ denote an observable or recipient-perceived quality dimension of surface s , and let ϕ_j denote a particular component of the production-friction vector. Under one production regime, recipients may learn an empirical association.

The corresponding expression is given in Equation 7.

$$q(s) \uparrow \implies \mathbb{E}[\phi_j | s] \uparrow. \quad (7)$$

After generative assistance lowers the friction required to reach comparable $q(s)$, the same surface-level observation supports a wider range of plausible histories. The inferential problem changes even if the recipient’s perceptual response to the surface remains initially similar.

Figure 3 depicts the resulting asymmetry. The visible surface may remain familiar while the set of generative histories behind it becomes more heterogeneous. Recipients therefore face an expanded inverse problem. The crisis emerges when interpretive habits remain calibrated to a narrower historical pathway set.

Working Claim 3.4 (Surface–friction decoupling). For some symbolic forms, generative abundance weakens previously useful associations between observable polish, personalization, fluency, or stylistic coherence and the amount or type of human time, expertise, and attention directly invested in producing the surface. The degree of decoupling is domain-specific and can coexist with persistent institutional or consequential constraints.

Surface–friction decoupling should not be interpreted as evidence that generative outputs lack value. It identifies a change in what the surface alone can reasonably support as an inference about its production history. A highly useful report can be produced quickly. A thoughtful professional can use automation. A low-effort artifact can still be correct. The diagnostic point concerns inference, not merit.

3.3 Pathway Multiplication and Generative Heterogeneity

The symbolic crisis becomes more intelligible when generative abundance is treated as a multiplication of pathways rather than a binary division between human and AI production. Consider a polished organizational memo. Its generative history may include direct human drafting; human drafting followed

by model editing; model drafting followed by expert revision; several employees coordinating through a model; an autonomous agent composing from organizational data; or a hybrid workflow in which authorship, factual checking, approval, and delivery are distributed across humans and systems.

These pathways differ in ways that matter socially. One pathway may involve deep domain expertise and little direct writing. Another may involve extensive writing skill and weak factual verification. A third may involve minimal human attention but strong institutional oversight. The visible memo can remain similar across these histories.

Let $H(s)$ denote the set of socially relevant generative histories compatible with observed surface s . The effect of pathway multiplication on the size of the relevant history set is represented in Equation 8.

$$|H(s)| \uparrow \quad (8)$$

This increase can occur for at least some practically important classes of surfaces. The increase concerns relevant heterogeneity, not merely the mathematical existence of infinitely many microscopic causal histories. More distinctions become socially plausible behind the same ordinary symbolic appearance.

Proposition 3.5 (Generative heterogeneity). *Generative abundance tends to increase the heterogeneity of socially plausible production histories behind comparable symbolic surfaces. As this heterogeneity increases, recipients require additional context when they wish to infer authorship, attention, competence, endorsement, or responsibility from the surface itself.*

This proposition is central to the paper’s diagnosis because it relocates the problem from ontology to inference. The important issue is not whether a message is ontologically “human” or “AI.” The issue is which parts of the generative history remain relevant to the receiver and how much of that history the symbolic surface continues to disclose.

3.4 Friction Migration and Residual Scarcity

Generative abundance can lower production friction while creating or intensifying other costs. A first draft may become cheap while verification becomes more important. Personalization may become cheap while recipient attention remains scarce. Many plausible designs may become cheap while selection and institutional approval remain costly. An organization may accelerate document production while increasing the burden of checking factual accuracy, policy compliance, provenance, and responsibility.

It is therefore useful to speak of *friction migration*. Let Φ_{prod} denote the frictions directly associated with producing a symbolic surface and Φ_{post} denote frictions of selection, validation, integration, accountability, and reception. Generative assistance can produce the pattern stated in Equation 9.

$$\Delta\Phi_{\text{prod}} < 0, \quad \Delta\Phi_{\text{post}} \geq 0 \quad (9)$$

This pattern can hold for some dimensions and institutional settings. The inequalities are illustrative rather than universal.

Definition 3.6 (Friction migration). *Friction migration* denotes a redistribution of effort, expertise, attention, or institutional burden across stages of symbolic production and use after a production technology lowers costs at one stage. The concept directs attention to changed bottlenecks rather than total disappearance of cost.

This migration is especially relevant to symbol inflation because recipients experience the output side of expanding symbolic production while many verification costs remain hidden. A sender can generate ten variants of a message quickly; the recipient still has finite time to determine which messages deserve attention. A research group can produce more candidate analyses; evaluation remains constrained by expertise and review capacity. A brand can generate large numbers of culturally adapted expressions; the social meaning of those expressions still depends on reception, history, and institutional conduct.

The persistence of residual scarcity means that generative abundance is compatible with increasingly scarce attention, judgment, accountability, and relational bandwidth. The relevant political-economic transformation therefore concerns the relative location of scarcity within a symbolic ecology.

Remark 3.7 (Abundance and scarcity). Generative abundance in symbolic production can coexist with severe scarcity in attention, verification, legitimacy, responsibility, and direct relational experience. The term *abundance* should therefore remain indexed to the production of candidate symbolic forms rather than generalized to the entire social process in which those forms acquire meaning.

3.5 Uneven Abundance across Domains

Production-friction reduction is heterogeneous. Some tasks have clear linguistic structure, abundant training examples, and outcomes that can be evaluated rapidly. Other tasks depend on tacit knowledge, embodied access, rare events, confidential information, difficult factual verification, or institutional standing that a generative model cannot independently supply. Even within one occupation, generative assistance can affect workers differently. The customer-support field study reported by Brynjolfsson, Li, and Raymond found larger gains among less experienced and lower-skilled workers, while effects varied across problem types (Brynjolfsson et al. 2025). This heterogeneity cautions against describing generative abundance as a uniform compression of all expertise.

A symbolic surface can also remain scarce because production is constrained by legitimate authority rather than expressive difficulty. A court order, medical prescription, licensed certification, or binding organizational decision derives part of its status from institutional relations that remain consequential even if the wording can be generated instantly. Section 2 described this distinction through the compression profile. Generative abundance can strongly affect the evidential appearance of competence while leaving performative authorization comparatively stable.

The same unevenness applies across languages, cultural contexts, modalities, and access conditions. Model quality, platform cost, computational infrastructure, data access, and organizational permission shape which actors experience lower friction. Generative abundance can therefore redistribute symbolic productive capacity while reproducing or creating other forms of inequality. The present paper does not attempt to measure these distributions comprehensively. It treats unevenness as a boundary condition on any broad historical claim.

Research Obligation 3.8 (Domain-indexed abundance). Empirical claims about generative abundance should identify the symbolic domain, task class, relevant friction dimensions, actor population, and institutional setting. Evidence that production friction falls in professional writing or customer support does not by itself establish equivalent changes in scientific discovery, legal authority, artistic practice, interpersonal care, or other domains.

This obligation is methodological as well as conceptual. The symbolic-crisis framework becomes less informative if every increase in AI capability is treated as equivalent evidence of social transformation. The relevant question concerns which historically meaningful inference relations are altered in a given domain.

3.6 Personalization, Scale, and the Attention Asymmetry

One of the most consequential dimensions of generative abundance is the combination of personalization and scale. Historically, highly individualized symbolic production often implied some allocation of human attention because adaptation had to be performed separately for each recipient. Generative systems can reduce the marginal effort required to create messages that incorporate a recipient’s name, history, preferences, language, organizational role, or recent interaction.

This change creates an attention asymmetry. Sender-side production can scale faster than recipient-side interpretation. If one actor can cheaply produce n tailored messages while each recipient must

still allocate time to interpret, verify, or emotionally process the message, then the pressure on recipient attention can be summarized by the ratio in Equation 10.

$$\frac{\text{symbolic outputs generated}}{\text{recipient attention available}} \quad (10)$$

can rise sharply. The relevant scarcity migrates toward reception.

This asymmetry matters because personalization itself has often been treated as evidence of attention. Under generative abundance, individualized surface features can remain psychologically salient even when the marginal attentional investment of the sender has fallen. The relationship between personalization and attention therefore becomes a particularly clear instance of surface–friction decoupling.

The point should again be framed probabilistically. A generatively assisted personalized message can reflect genuine care, careful approval, or deep prior knowledge. Generative assistance can also allow a person with limited expressive skill to communicate intentions more effectively. The symbolic crisis arises because the same surface property supports a wider range of generative histories. Recipients consequently need more context when attention itself is the property they wish to infer.

3.7 The Inverse Problem under Expanding Production

The previous section described social compression as a many-to-one relation from distributed histories to portable symbolic surfaces. Generative abundance intensifies this structure by increasing the number of low-friction routes into the same region of symbolic space. The forward production problem becomes easier for many actors, as represented in Equation 11.

$$\text{history or intention} \mapsto s. \quad (11)$$

The receiver’s inverse problem can become harder, as represented in Equation 12.

$$s \mapsto \{h_1, h_2, \dots, h_k\}. \quad (12)$$

This asymmetry gives the diagnostic problem its distinctive form. The technology that helps producers reach desirable symbolic surfaces does not automatically provide recipients with equivalent information about how those surfaces were generated. Production capacity and interpretive capacity can therefore develop at different rates.

Proposition 3.9 (Production–inference asymmetry). *As generative abundance expands the feasible pathway set for symbolic production, the forward task of producing a socially recognizable surface can become easier while the inverse task of inferring its relevant generative history becomes more ambiguous. A symbolic crisis becomes more plausible when interpretive conventions adapt more slowly than the production regime and the resulting mismatch becomes consequential.*

This proposition leads directly to the next section. Production friction matters because historical friction supplied part of the background through which symbolic surfaces became diagnostically useful. Once similar surfaces can be reached through more heterogeneous and lower-friction pathways, the central question becomes the changing *diagnosticity* of the surface itself. The following section therefore examines which properties recipients can still infer from symbolic form, how diagnosticity varies across domains, and why generative abundance need not produce either universal distrust or universal equivalence.

4 Diagnosticity of Symbolic Surfaces

The previous section treated falling production friction and expanding feasible pathways as one mechanism of symbol inflation. This section examines the corresponding receiver-side problem. Its structural role is to specify what is meant when a symbolic surface becomes a weaker guide to the history behind it. The central concept is *diagnosticity*: the degree to which an observable symbolic feature supports discrimination among competing hypotheses about an unobserved property of its producer, production process, or institutional setting. Diagnosticity is therefore always relative to a target inference. A polished report may remain highly informative about the report’s explicit content while becoming less informative about the amount of unaided human labor that produced it. A personalized message may still communicate a proposition accurately while providing weaker evidence of the sender’s direct attention. The discussion develops this distinction through six steps: target-relative diagnosticity, posterior ambiguity, asymmetric erosion across properties, compensating context, institutional redistribution of diagnosticity, and the limits of global distrust.

The term *diagnosticity* has an established history in social judgment research, where cues differ in the extent to which they discriminate among underlying traits or hypotheses (Skowronski and Carlston 1987). The present paper uses the term more broadly and explicitly in relation to symbolic production. The source-monitoring literature likewise emphasizes that judgments about origin are inferential rather than direct: people attribute remembered or observed information to sources by evaluating available characteristics, contextual information, and prior beliefs (Johnson et al. 1993). These literatures do not by themselves establish a generative-AI-specific crisis. They provide conceptual support for a more modest starting point: recipients ordinarily infer hidden origins and attributes from imperfect cues, and the usefulness of a cue depends on how strongly it differentiates among plausible underlying histories.

4.1 Target-Relative Diagnosticity

Let s denote an observed symbolic surface and let q denote a target property that a recipient wishes to infer. The target might concern authorship, competence, attention, sincerity, effort, institutional authorization, scarcity, or some other latent feature. Let $\mathcal{H}_q = \{h_1, \dots, h_k\}$ denote the set of plausible histories relevant to that property. We can then use *diagnosticity* as a deliberately general label for the degree to which observing s changes discrimination among the members of $\mathcal{H}_q$.

Definition 4.1 (Target-relative symbolic diagnosticity). For a symbolic surface s and target property q , *symbolic diagnosticity* denotes the extent to which s , together with available context, supports discrimination among plausible generative histories relevant to q . Diagnosticity is indexed to the target inference, the comparison set, and the recipient’s information environment.

This definition prevents several overextensions. A symbol does not possess one universal level of credibility. The same surface can be highly diagnostic for one property and weakly diagnostic for another. Nor is diagnosticity equivalent to truth. A sentence can be propositionally accurate while revealing little about who composed it. Conversely, a surface can provide strong evidence about origin while containing a false claim. The paper therefore separates at least three judgments: whether a proposition is accurate, whether a surface is attributable to a particular generative history, and whether that history warrants a social inference such as competence or sincerity.

A simple notation makes the distinction visible. Target-relative diagnosticity is denoted in Equation 13.

$$D_q(s \mid C) \tag{13}$$

represent the diagnosticity of surface s for target q under contextual information C . A possible asymmetry between content and effort diagnosticity is stated in Equation 14.

$$D_{\text{content}}(s \mid C) \gg D_{\text{effort}}(s \mid C), \tag{14}$$

without any contradiction. Symbol inflation matters because the mechanisms that produce it can alter these relations unevenly. It may reduce the diagnosticity of fluency for unaided language competence while leaving the diagnosticity of a verified institutional signature for authorization largely intact. It may reduce the diagnosticity of polished prose for time investment while increasing the importance of other cues such as revision history, accountability arrangements, or external anchoring.

Proposition 4.2 (Property-specific diagnostic erosion). *Generative abundance need not reduce the diagnosticity of symbolic surfaces uniformly. It can weaken selected inferences by increasing the number of plausible generative histories compatible with a surface while leaving other inferences comparatively stable or shifting their evidential basis toward additional context.*

4.2 Posterior Ambiguity and Surface Equivalence

The forward production relation considered in Section 3 maps heterogeneous histories into observable symbolic surfaces. The inverse problem asks what a recipient can infer from the surface. Suppose two generative histories h_a and h_b differ substantially in human attention, expertise, or delegated agency, while both can generate surfaces inside the same recognizable region $\mathcal{S}$. As the overlap of feasible surface regions increases, visual or linguistic inspection alone may provide less leverage for distinguishing those histories.

The relevant change can be described as *posterior ambiguity*. Let $P(h_i \mid s, C)$ denote the recipient’s posterior belief that history h_i generated surface s under context C . Generative abundance can increase ambiguity when several histories receive non-trivial posterior weight after the surface is observed. The argument does not require equal likelihood among histories. It requires only that a previously useful surface property becomes compatible with a wider range of plausible origins.

Definition 4.3 (Surface equivalence under heterogeneous histories). Two generative histories are *surface-equivalent with respect to a recipient* when the observable features available to that recipient do not reliably discriminate between them for the target inference at hand.

Surface equivalence is relational rather than intrinsic. A recipient with access only to the final text may treat two histories as equivalent, while an editor with access to a version archive, model-use disclosure, and correspondence may distinguish them easily. Diagnosticity therefore depends on the architecture of observation as well as on the artifact itself.

This helps clarify why attempts to locate the symbolic crisis entirely inside artifacts are likely to be incomplete. The same artifact can occupy different epistemic environments. A report accompanied by auditable data, a named responsible officer, and a stable institutional review process is embedded differently from an anonymous report circulating without provenance. Likewise, an email whose sender is accountable for its commitments differs from a message whose apparent personalization has no stable relation to a human principal. The surface can be similar while the surrounding inferential environment differs substantially.

Figure 4 illustrates the basic asymmetry. Generative heterogeneity does not erase semantic content. It changes the evidential relation between the surface and selected features of its history. Symbol inflation can therefore produce a redistribution of inferential load: properties that were once inferred cheaply from surface cues may increasingly require supplementary evidence.

4.3 Diagnosticity Profiles across Symbolic Domains

Because diagnosticity is target-relative, the symbolic crisis should be analyzed through profiles rather than global declarations. Table 2 gives an illustrative set of target inferences. The table is conceptual rather than empirical. Its purpose is to specify what future research would need to measure separately.

The distinction between direct attention and propositional content is especially important for communication. A recipient may correctly understand what a message says while remaining uncertain about

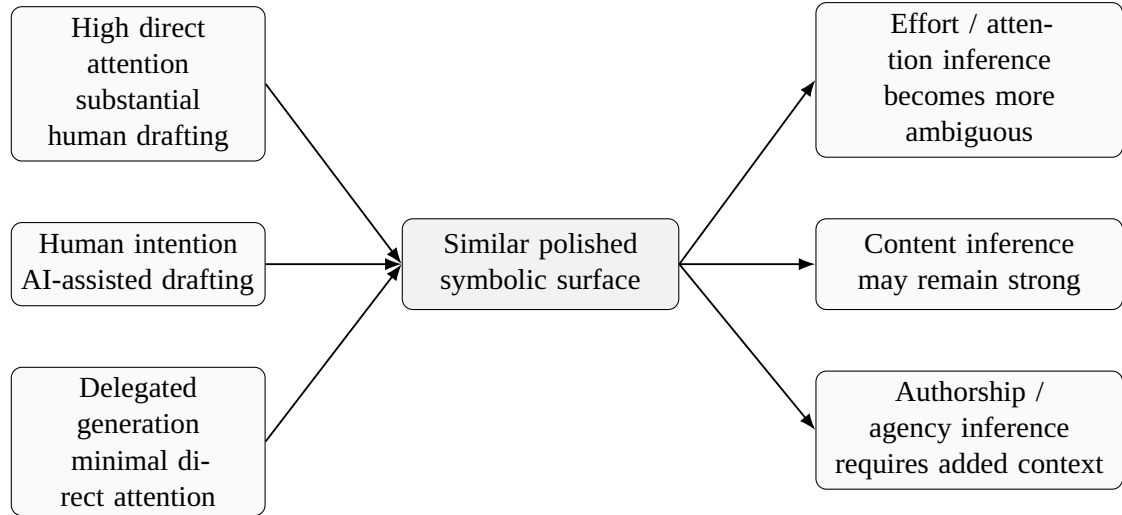


Figure 4: Target-relative diagnosticity under generative heterogeneity. Similar symbolic surfaces can remain useful for some inferences while providing weaker discrimination among histories relevant to effort, attention, authorship, or delegated agency.

how much attention the sender personally allocated to composing it. Similarly, a manager may receive a technically competent report without being able to infer how much of the analysis was produced, checked, or merely presented by the nominal author. These are different epistemic questions.

Existing research on AI-mediated communication illustrates this plurality rather than supporting a single social reaction. Studies report penalties in perceived authenticity or trust when AI authorship or assistance is disclosed in some communicative and marketing contexts (Schilke and Reimann 2025; Sahabi, Formosa, and Bankins 2026; Kirk and Givi 2025). Yet AI-mediated systems can also improve task performance or lower production costs, which means that decreased diagnosticity for one latent property does not imply decreased instrumental usefulness of the output (Noy and Zhang 2023; Brynjolfsson et al. 2025). The symbolic crisis is therefore compatible with simultaneous productivity gains and interpretive instability.

4.4 Costly Cues and Changing Evidential Weight

Historical production friction often contributed to diagnosticity because some symbolic performances were harder to produce without possessing particular resources or commitments. Signaling theory formalizes one version of this intuition: signals can become informative when their costs differ across underlying types (Spence 1973). Communication research similarly suggests that costly signaling can influence perceived honesty under some conditions (Chaudhry and Wald 2022), and sensorimotor communication research indicates that effort and signal utility can contribute to trust and generosity (McEllin and Michael 2022).

The present argument does not infer that costly production is inherently more authentic. High cost can be wasteful, exclusionary, performative, or strategically engineered. The narrower point is that a change in the cost structure can alter the evidential meaning of a cue. If elaborate personalization once required substantial direct attention and can later be produced through low-friction delegation, then the same visible personalization supports a different inference about attention unless supplementary information restores discrimination.

This can be described as an *evidential reweighting* problem. Recipients need not abandon the cue. They may reduce its weight relative to other information. Let w_j denote the weight assigned to cue c_j in an informal inference process. A transition in production technology can motivate the reweighting

Table 2: Illustrative diagnosticity profile of symbolic surfaces under generative abundance.

Target inference	Historical surface cue	Potential change under generative abundance
Direct attention	personalization, detailed response, stylistic care	Similar cues become feasible with less direct attention; contextual confirmation carries greater weight.
Unaided competence	fluency, technical vocabulary, polished explanation	Surface performance can be partially decoupled from unaided production capacity.
Authorship	stylistic continuity, first-person voice, distinctive phrasing	Multiple forms of assistance and delegation make authorship more compositional.
Effort	length, polish, formatting complexity, bespoke presentation	Observable production quality can require less time or coordination than inherited heuristics assume.
Institutional authorization	signatures, credentials, formal templates	Diagnosticity may remain comparatively strong where authorization is externally anchored and enforceable.
Sincerity	warmth, vulnerability, individualized language	Surface cues become compatible with wider ranges of felt commitment and delegated expression.
Propositional content	explicit claims, data statements, instructions	Content can remain directly inspectable even when origin-related inferences become more ambiguous.

summarized in Equation 15.

$$w_{\text{polish}} \downarrow, \quad w_{\text{external anchoring}} \uparrow, \quad w_{\text{accountability}} \uparrow, \quad (15)$$

This reweighting can occur for selected target properties. These arrows summarize a conceptual redistribution of evidential weight; they are not proposed as an empirical law.

Working Claim 4.4 (Diagnostic reweighting). A symbolic crisis can proceed through reweighting rather than abandonment. Recipients may continue to use inherited symbolic forms while changing which latent properties those forms are allowed to support without supplementary evidence.

This claim helps explain why adaptation may be slow. Symbolic interpretation is embedded in habits, professional routines, aesthetic conventions, institutional procedures, and learned expectations. Production regimes can change faster than these interpretive practices. A polished document can therefore continue to produce an intuitive impression of effort or competence even when recipients explicitly know that the relevant surface can now be generated through several lower-friction pathways.

4.5 Compensating Context and Institutional Redistribution

Declining surface diagnosticity does not imply that societies must reconstruct every generative history directly. Such a requirement would destroy the compression advantage that made symbolic mediation valuable in the first place. Instead, diagnosticity can migrate from isolated surfaces toward larger evidential assemblages. These may include provenance records, review procedures, institutional liability, accountable principals, external references, persistent identifiers, or repeated interactions. The present paper treats these as diagnostic redistributions rather than completed solutions; their design and political economy remain open questions.

A useful distinction is between *intrinsic surface cues* and *relationally anchored cues*. Intrinsic surface cues are properties visible in the artifact itself: polish, stylistic coherence, personalization, formal vocabulary, visual sophistication. Relationally anchored cues depend on links beyond the artifact: a responsible institution, a verifiable publication history, a contractual obligation, a prior interaction, or a technical provenance chain. Generative abundance can reduce the relative diagnostic advantage of some intrinsic cues while increasing demand for relational anchoring.

Content provenance standards such as C2PA illustrate one technical attempt to preserve selected origin information through authenticated metadata (Coalition for Content Provenance and Authenticity 2026). Such systems can support claims about content history or credentialed actors, but they do not settle the broader social inferences at issue here. Knowing which tools participated in production does not by itself reveal sincerity, competence, attention, responsibility, or the quality of the underlying relationship. Provenance is therefore one source of context rather than a universal repair of symbolic diagnosticity.

Institutional redistribution also raises questions of power. If recipients increasingly rely on platforms, credential issuers, employers, publishers, or identity infrastructures to interpret symbolic origins, diagnostic uncertainty can increase the authority of intermediaries. The symbolic crisis may therefore shift rather than simply reduce dependence. A recipient who trusts an institutional verification layer does not escape mediated trust; the locus of trust moves from the artifact to the infrastructure that certifies aspects of its history.

Remark 4.5 (Redistribution of trust). A decline in the diagnosticity of isolated symbolic surfaces can increase reliance on institutional, relational, or infrastructural context. This process redistributes trust and verification burdens rather than eliminating mediation.

4.6 The Limits of Global Distrust

A diagnosis of symbolic crisis can easily overreach into the claim that nothing can be trusted once generative systems become widespread. That conclusion does not follow. First, many symbolic systems retain external enforcement, institutional anchoring, or material constraints. A legally valid authorization, authenticated transaction, or verified measurement may remain useful even if the surrounding presentation is machine-generated. Second, recipients can learn new cue weights. Third, domains differ substantially in the relation between surfaces and latent properties. Fourth, human symbolic practices have always included delegation, editing, impersonation, convention, and strategic presentation. Generative AI changes the scale and distribution of these possibilities rather than creating mediated inference from nothing.

The more precise diagnosis is therefore selective. Symbolic surfaces become especially problematic when recipients continue to infer latent properties from cues whose production conditions have changed substantially. The crisis is strongest where three conditions coincide: the latent property matters socially, the surface was historically used as a shortcut for that property, and generative abundance expands the set of plausible histories behind the cue faster than interpretive practices adapt.

Definition 4.6 (Diagnostic mismatch). A *diagnostic mismatch* occurs when a recipient or institution continues to assign substantial evidential weight to a symbolic cue for target property q even though changes in production conditions have materially widened the set of plausible generative histories compatible with that cue.

This definition gives the symbolic-crisis thesis a more discriminating form. It does not predict universal distrust. It predicts a period in which inherited mappings between surfaces and latent properties become unevenly miscalibrated. Some mappings will remain robust. Some will weaken. Some will be replaced by new contextual architectures. Others may persist culturally after their evidential basis has changed.

4.7 Diagnosticity and the Transition to Communication

The analysis so far has treated symbolic surfaces as inferential objects. Communication introduces a further difficulty because many communicative practices involve more than the transmission of propositional content. A message can carry commitments, expose a speaker to response, establish responsibility, invite reciprocal adjustment, or alter the relation between participants. If generative systems can reproduce the symbolic surface of these practices while changing how participation is distributed among humans and technical agents, then diagnosticity alone is insufficient to describe the resulting problem.

The next section therefore moves from inference to communication. Its question is not whether an AI-generated sentence can contain meaningful information; clearly it can. The question is which dimensions of communication depend on participation, attribution, reciprocal updating, commitment, and exposure to consequences, and how these dimensions change when symbolic expression is delegated. This transition is necessary because the symbolic crisis concerns both what recipients can infer from symbols and what kinds of social relation those symbols actually instantiate.

5 Communication beyond Symbol Transmission

The preceding sections treated symbolic surfaces as compression interfaces whose diagnostic relation to their generative histories can weaken under conditions of generative abundance. This section shifts from the epistemics of inference to the structure of communication itself. Its role is to distinguish the successful transmission of symbolic content from the broader social processes that can accompany communication: attribution, attention, mutual updating, commitment, exposure to response, and responsibility for consequences. The distinction matters because generative systems can preserve or improve the first dimension while redistributing the others across human and technical participants. The resulting problem is therefore wider than uncertainty about authorship. Similar messages can instantiate different communicative relations even when their propositional content is identical.

Research on AI-mediated communication already treats computational systems as participants in the mediation, modification, or generation of messages exchanged among humans (Hancock et al. 2020). Speech-act theory provides a complementary reminder that utterances can function as actions within social settings rather than merely as containers of propositions (Austin 1962). Interactional approaches likewise emphasize that communication participates in the presentation, interpretation, and maintenance of social relations (Goffman 1959). These traditions differ substantially in method and scope. The present argument draws a limited conclusion from them: communication has several analytically separable dimensions, and symbolic transmission alone does not exhaust the social event.

5.1 Semantic Carriage and Relational Participation

Let m denote a symbolic message delivered from an apparent sender A to a recipient B . A minimal transmission model asks whether the propositional or directive content of m reaches B with adequate fidelity. For many practical purposes, this remains an important criterion. Instructions can be clear, facts can be accurate, and requests can be unambiguous even when their wording was heavily assisted or generated by an AI system.

The symbolic-crisis problem appears when successful semantic carriage is treated as sufficient evidence that the surrounding social relation has remained unchanged. Consider three messages with nearly identical wording. In the first, the sender reads the recipient’s earlier message, drafts a response, revises it, and sends it personally. In the second, the sender instructs a model to draft a response, reads the result carefully, edits it, and adopts it. In the third, an agent receives the message, infers the sender’s likely intention from stored preferences, generates a response, and sends it without contemporaneous review. The recipient can receive the same proposition in all three cases. The distribution of attention, interpretation, authorship, revision, and responsibility differs substantially.

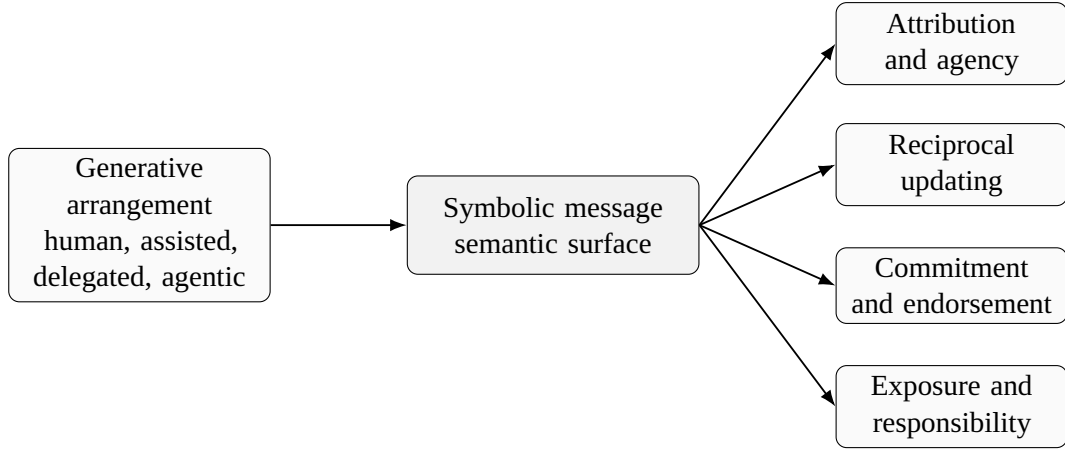


Figure 5: Symbol transmission and the wider communicative participation profile. Similar symbolic surfaces can preserve semantic content while differing in attribution, reciprocal updating, commitment, exposure, and responsibility.

Definition 5.1 (Communicative participation profile). For a communicative event e , the *communicative participation profile* is the context-dependent vector defined in Equation 16.

$$\Pi(e) = \langle S, A, U, K, E, R \rangle, \quad (16)$$

In Equation 16, S denotes semantic contribution, A attribution and identifiable agency, U reciprocal updating, K commitment formation, E exposure to response and consequence, and R responsibility for the communicative act. The dimensions can be distributed across several human and technical participants and need not be equally present in every communicative practice.

The vector does not define one ideal form of communication. Broadcast notices, emergency alerts, automated receipts, contractual negotiations, intimate conversations, and organizational reports legitimately occupy different regions of the profile. Its purpose is diagnostic. It makes visible the possibility that semantic performance can remain high while other dimensions of participation change.

Proposition 5.2 (Semantic-participatory divergence). *Under mediated or delegated symbolic production, preservation of semantic adequacy does not by itself establish preservation of the communicative participation profile. Equivalent or near-equivalent symbolic content can be generated through arrangements that distribute attention, updating, commitment, exposure, and responsibility differently.*

This proposition prevents two symmetrical errors. One error treats AI mediation as communicative failure whenever the human principal did not compose each word. The other treats accurate content as evidence that mediation is socially irrelevant. Both positions collapse a multidimensional event into a single criterion.

Figure 5 places symbolic transmission inside a larger communicative structure. The figure should not be read as requiring every dimension in every message. It shows why changes in the production arrangement can matter even when the message itself remains intelligible.

5.2 Attribution and Communicative Agency

Communication ordinarily involves some account of who is speaking, on whose behalf, and with what authority. The apparent first-person voice of a message can conceal several distinct agency structures. A human may have composed the message directly. A model may have generated wording that the human reviewed and adopted. An employee may have sent language approved by an organization. An

autonomous agent may have acted under standing instructions. These arrangements can all support meaningful communication, yet they allocate agency differently.

The attribution problem therefore extends beyond the binary distinction between human and AI authorship. A more useful analysis separates at least four questions: who initiated the communicative act, who selected or generated the content, who endorsed the resulting message, and who can be held responsible for commitments created by it. These roles can coincide in one person, or they can be distributed.

The relevant role assignments are collected in Equation 17.

$$\mathcal{A}(e) = \{a_{\text{init}}, a_{\text{gen}}, a_{\text{endorse}}, a_{\text{resp}}\} \quad (17)$$

denote the set of role assignments in event e . The notation is schematic; a role can be shared among several participants. The relevant point is that a symbol can present a unified voice while the underlying agency structure is compositional.

Definition 5.3 (Agency compression). *Agency compression* occurs when a symbolic surface presents a communicative act through a unified or simplified apparent voice while initiation, generation, endorsement, and responsibility are distributed across multiple participants or systems.

Agency compression long predates generative AI. Institutions speak through officers, lawyers draft statements for clients, speechwriters prepare addresses, and organizations use standard templates. Generative systems broaden the range and granularity of tasks that can be delegated while preserving the appearance of a single communicative voice. The diagnostic change therefore concerns degree, accessibility, and opacity rather than the invention of mediation itself.

5.3 Reciprocal Updating and Exposure to Response

A second dimension concerns whether the parties are mutually exposed to the interaction. In many socially significant exchanges, communication changes subsequent states of the participants. A recipient replies, challenges, misunderstands, requests clarification, or introduces new information. The sender then updates. The communication is temporally extended and recursively coupled.

A simple representation of the relation is given in Equation 18.

$$A_{t+1} = f(A_t, B_t, m_t, C_t), \quad B_{t+1} = g(B_t, A_t, m_t, C_t), \quad (18)$$

In Equation 17, C_t denotes the surrounding context. The equations do not claim that interpersonal communication follows a deterministic law. They represent a minimal relational point: the state of each participant can become partly conditioned by the other through repeated exchange.

Generative mediation can preserve this reciprocal structure when the human principal remains attentive to the evolving interaction. It can also thin it when the principal delegates interpretation and response so extensively that the other party is interacting mainly with a proxy optimized to represent the principal’s preferences. The recipient may still receive competent responses, yet the principal’s own beliefs, affect, priorities, or understanding can remain comparatively unchanged.

This motivates a distinction between *symbolic responsiveness* and *participant updating*. A system can generate an appropriate reply to the latest message without producing a corresponding update in the human principal. In some domains this is desirable. Customer-service systems, scheduling agents, and routine administrative communication are often designed precisely to prevent every interaction from consuming direct human attention. In other domains, recipients may reasonably care whether the apparent interlocutor has actually encountered the information being exchanged.

Working Claim 5.4 (Reciprocity gap). A *reciprocity gap* can arise when a communication system produces context-sensitive symbolic responses while the human principal undergoes little corresponding

updating, exposure, or revision. The significance of the gap depends on the practice and should be evaluated separately from semantic quality.

The claim is intentionally conditional. It does not treat direct human updating as universally required. It identifies a possible mismatch between the surface appearance of reciprocal engagement and the distribution of actual updating among participants.

5.4 Commitment, Endorsement, and Consequence

Communication can also create or modify commitments. Promises, approvals, refusals, authorizations, acknowledgments, and assurances often have consequences because recipients treat them as attributable to an agent or institution. Austin’s account of performative utterances is relevant here: some utterances function partly through the social conditions under which saying something counts as doing something (Austin 1962). The present argument extends this concern into mediated production without assuming that every message is performative in the technical sense.

When generative systems participate in communication, the important distinction is often between generation and endorsement. A model can generate the words “I approve this proposal,” yet the communicative force of the sentence depends on whether an authorized actor actually adopts it. Similarly, an agent may be given authority to commit within a bounded domain, such as scheduling a meeting or approving a routine purchase. In that case the communicative force derives partly from the delegation architecture rather than from contemporaneous human composition.

The distinction is summarized in Equation 19.

$$\text{surface production,} \quad \text{endorsement,} \quad \text{authorization,} \quad \text{consequence bearing.} \quad (19)$$

These dimensions often coincided more closely in low-automation interpersonal communication. Generative and agentic systems make their separation more visible and potentially more consequential.

Proposition 5.5 (Commitment-location principle). *The social force of a mediated message depends less on who physically produced its wording than on where endorsement, authorization, and consequence-bearing responsibility are located. Generative systems can therefore alter communication substantially without eliminating human or institutional commitment, provided those relations remain identifiable and operative.*

This proposition is important for the diagnosis of symbolic crisis because it shows that the problem cannot be reduced to authorship detection. A perfectly reliable classifier of AI-generated text would reveal little about whether a human principal endorsed the message, whether the institution stands behind it, or whether the sender will bear consequences for non-performance. The location of commitment is a relational and institutional question.

5.5 Attention and the Experience of Recognition

Another dimension concerns attention. Many communicative practices carry an implicit claim that the recipient has been attended to. A detailed reply can be experienced as recognition because the recipient assumes that someone read, interpreted, and responded to their specific situation. Generative systems can reproduce the symbolic form of such attention at greatly reduced marginal cost.

The resulting ambiguity should be described carefully. The recipient’s experience is not automatically false merely because assistance was used. A human can use a model to articulate a genuinely attentive response more clearly. An organization can use AI to make previously inaccessible levels of personalization feasible. Conversely, a highly individualized message can be generated from stored data with almost no contemporaneous human attention. The same surface property therefore spans a wider range of attentional histories.

Table 3: Illustrative communicative dimensions under increasing mediation.

Dimension	What the recipient may care about	Potential effect of generative mediation
Semantic carriage	Whether content, instructions, or claims are accurately conveyed	Can remain strong or improve even when human drafting decreases.
Attribution	Who is speaking and on whose behalf	Can become distributed across principal, model, agent, employee, and institution.
Reciprocal updating	Whether the other party actually encounters and updates from the exchange	Symbolic responsiveness can exceed direct human updating.
Commitment	Whether the message creates or expresses an endorsed obligation	Depends on adoption, authorization, and delegation architecture more than on wording alone.
Attention	Whether the recipient’s particular situation received direct consideration	Personalized form can be produced across increasingly heterogeneous attentional histories.
Responsibility	Who bears consequences for error, promise, or misuse	Can remain human or institutional even when production is automated, but may become less visible.

Let $H_A(e)$ denote the amount and type of direct human attention allocated to event e , and let $P_A(e)$ denote the level of personalized responsiveness observable in its symbolic surface. Generative abundance can weaken the inherited association, as stated in Equation 20.

$$P_A(e) \uparrow \implies H_A(e) \uparrow. \quad (20)$$

The implication becomes less reliable even when personalization remains useful to the recipient.

This is a communication-specific instance of the diagnosticity problem developed in Section 4. Its significance is relational because perceived attention can influence whether recipients experience themselves as heard, recognized, or taken seriously. The problem is therefore not the use of polished language. It is the increasing uncertainty surrounding the participation that the polished language appears to represent.

5.6 Communicative Thickness and Domain Variation

The preceding dimensions can be summarized through the concept of *communicative thickness*. The term is used here descriptively rather than normatively.

Definition 5.6 (Communicative thickness). *Communicative thickness* denotes the extent to which a communicative practice couples symbolic exchange to identifiable agency, reciprocal updating, commitment, exposure to response, and responsibility across time. Thickness is multidimensional and context-dependent; greater thickness is not assumed to be superior in every domain.

Routine machine-generated notices can be intentionally thin. Their value lies in reliable information transfer rather than reciprocal relation. A legal negotiation, therapeutic conversation, diplomatic exchange, intimate apology, or research collaboration may place more weight on sustained updating, accountability, and relational exposure. The same level of automation can therefore have very different significance across domains.

Table 3 shows why a single authenticity judgment is often too coarse. A mediated message can score differently across semantic, relational, and institutional dimensions. The appropriate analysis therefore depends on the practice and the inference at stake.

5.7 Participation Gaps under Symbolic Abundance

The symbolic crisis becomes especially visible when a symbolic surface continues to evoke expectations formed under a thicker participation regime while the actual participation profile has changed. We can call this a *participation gap*. Let $\Pi^{\text{exp}}(e)$ denote the participation profile that recipients reasonably expect from a communicative form in context, and let $\Pi^{\text{act}}(e)$ denote the profile instantiated by the actual production and delegation arrangement. The gap is then represented schematically in Equation 21.

$$\Delta_{\Pi}(e) = \Pi^{\text{exp}}(e) - \Pi^{\text{act}}(e), \quad (21)$$

In Equation 18, subtraction is purely conceptual because the dimensions need not be commensurable.

Definition 5.7 (Participation gap). A *participation gap* arises when the symbolic form of a communicative act supports expectations of attention, updating, endorsement, exposure, or responsibility that are substantially different from the participation structure actually instantiated by the communication process.

The definition avoids equating automation with deception. A gap can be small when institutional conventions clearly establish that a message is automated, when delegated authority is stable, or when recipients care mainly about semantic reliability. It can become more consequential when the communicative form invites reliance on interpersonal attention, personal endorsement, or reciprocal exposure that the underlying arrangement does not provide.

Existing work on AI-mediated communication and disclosure already suggests that recipients respond to perceived mediation in heterogeneous ways (Hancock et al. 2020; Schilke and Reimann 2025; Sahebi, Formosa, and Bankins 2026). The present framework interprets this heterogeneity as theoretically expected. Different communicative settings place different weights on semantic quality, personal participation, authenticity, and institutional accountability. A general theory should therefore predict variation rather than one universal AI penalty or one universal acceptance effect.

5.8 Communication as Relational Change

The deepest implication of this section is that communication can be understood partly through the changes it makes possible in relations among participants. A communicative event can transmit information while also altering expectations, commitments, mutual knowledge, obligations, and future behavior. The social significance of an exchange therefore extends across time.

Let $R_t(A, B)$ denote a relational state between participants at time t . A communicative event e_t is represented in Equation 22.

$$R_{t+1}(A, B) = \mathcal{F}(R_t(A, B), e_t, \Pi(e_t), C_t). \quad (22)$$

The function $\mathcal{F}$ remains unspecified because the paper does not assume a universal model of relational change. The notation expresses a narrower claim: the consequences of communication depend partly on the participation structure through which the symbolic event is generated and received.

This matters for the symbolic crisis because generative systems increase the possibility that the same surface event is embedded in substantially different relational processes. The recipient sees the symbolic message. The underlying relation may involve direct encounter, assisted articulation, institutional delegation, autonomous agentic execution, or some hybrid arrangement. Symbolic abundance therefore increases the importance of distinguishing the message from the relational process in which the message participates.

Working Claim 5.8 (Relational non-equivalence). Surface-equivalent messages need not be relationally equivalent. When communicative participation profiles differ, identical or near-identical symbolic

content can create different conditions for attribution, updating, commitment, exposure, and future interaction.

This claim prepares the next analytical step. Once communication is treated as a distributed social process, the binary category of authorship becomes insufficient. The relevant problem becomes the architecture of delegated agency: which communicative functions are assigned to humans, models, agents, and institutions; how authority travels across those assignments; and where endorsement and responsibility remain located. The next section therefore turns from the dimensions of communication to the structure of mediated agency itself.

6 Delegated Agency and Mediated Expression

Section 5 separated symbolic transmission from the wider participation structure of communication. The present section develops the next analytical layer: the architecture of delegation through which communicative agency is distributed among human principals, institutions, models, and autonomous or semi-autonomous agents. Its role is diagnostic. It identifies which communicative functions can be delegated, how several forms of mediation differ, and why a unified symbolic voice can conceal a heterogeneous agency structure. The section proceeds through four distinctions: delegation of execution and delegation of judgment; generation and endorsement; authority and responsibility; and episodic assistance and persistent agentic representation. These distinctions provide the conceptual basis for the analysis of trust in Section 7.

AI-mediated communication was already defined before the current generation of large language models as communication in which an artificial agent modifies, augments, or generates messages in pursuit of communicative goals (Hancock et al. 2020). Generative and agentic systems broaden this phenomenon by allowing increasingly large parts of the communicative process to be assigned to technical systems. Recent work on large-language-model use likewise distinguishes the offloading of well-specified execution from the delegation of substantive judgment, emphasizing that these practices carry different implications for epistemic control and responsibility (Mühlhoff 2026). The resulting landscape is better represented as a family of delegation architectures than as a binary distinction between human and AI communication.

6.1 Functional Decomposition of Communicative Agency

A communicative act that appears unitary at the symbolic surface can be decomposed into the functions collected in Equation 23.

$$\mathcal{F}_e = \{f_G, f_C, f_I, f_D, f_E, f_A, f_T, f_M, f_R\} \quad (23)$$

denote a schematic function set for communicative event e . Here f_G denotes goal formation; f_C context acquisition; f_I interpretation of the interlocutor and situation; f_D drafting or symbolic generation; f_E evaluation and revision; f_A authorization; f_T transmission; f_M monitoring of subsequent interaction; and f_R repair or response when the exchange develops unexpectedly. The list is deliberately functional rather than ontological. A human, organization, model, or agent can perform several functions, and one function can be distributed across several participants.

This decomposition clarifies why the phrase “AI-generated message” is often too coarse. Consider four cases. In a spelling assistant, the human performs almost every communicative function while the system modifies linguistic form. In a drafting assistant, the human sets the goal and context while the model proposes language that is later evaluated and authorized. In an interpretive assistant, the system also summarizes the recipient’s request, infers priorities, and proposes a substantive answer. In an autonomous service agent, the system may interpret incoming messages, select responses, transmit them, monitor replies, and continue the interaction under standing instructions. All four involve AI mediation, yet the agency architecture differs sharply.

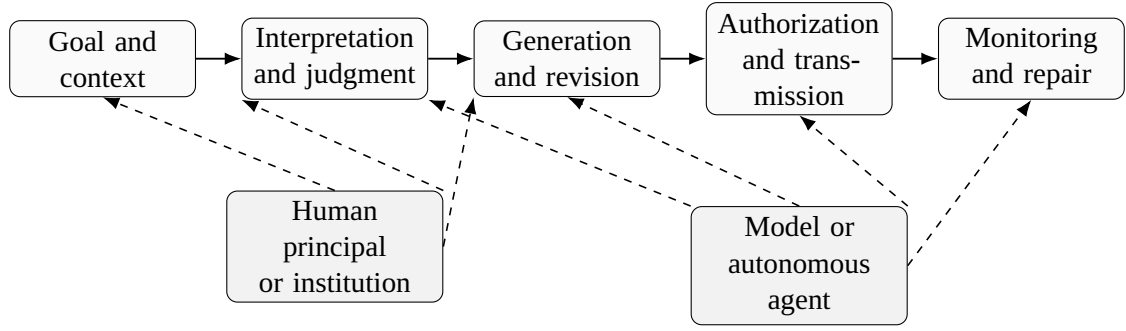


Figure 6: Communicative agency as a delegation architecture. The same outward message can result from different assignments of goal formation, interpretation, generation, authorization, transmission, monitoring, and repair. Dashed links indicate illustrative allocations rather than a required workflow.

Definition 6.1 (Delegation architecture). A *delegation architecture* for communicative event e is the assignment defined in Equation 24.

$$\mathcal{D}(e) : \mathcal{F}_e \mapsto \mathcal{P}_e, \quad (24)$$

In Equation 23, $\mathcal{F}_e$ is the set of communicative functions and $\mathcal{P}_e$ is the set of participating human, institutional, and technical actors. The architecture records which participants perform, supervise, authorize, or remain accountable for each function.

The concept extends the communicative participation profile developed in Section 5. The participation profile describes dimensions of the social event; the delegation architecture describes how the functions producing that event are distributed.

Figure 6 represents communication as a sequence of functions that can be distributed differently across human and technical participants. The analytical value lies in revealing that mediation can move upstream. A system may begin as a surface-level language assistant and later acquire roles in interpretation, prioritization, timing, or interaction management. The visible message can remain stylistically continuous while the locus of agency changes.

6.2 Execution, Interpretation, and Judgment

Delegation becomes socially more consequential as it moves from bounded execution toward interpretation and judgment. A system asked to correct grammar receives a comparatively specified task. A system asked to “reply appropriately,” “understand what the customer really wants,” or “handle this relationship according to my preferences” must determine relevance, infer context, and weigh competing considerations. Mühlhoff describes this difference through a spectrum from execution to delegated judgment and connects deeper judgment delegation to weaker epistemic control (Mühlhoff 2026).

For communication, the three analytical levels are ordered in Equation 25.

$$\text{symbolic execution} \subset \text{situational interpretation} \subset \text{communicative judgment}. \quad (25)$$

Symbolic execution concerns form within an already specified communicative act. Situational interpretation concerns the meaning of the recipient’s message, context, and likely expectations. Communicative judgment concerns what should be said, what should remain unsaid, which commitments are appropriate, how much risk can be accepted, and when the interaction should be escalated to another actor.

Definition 6.2 (Delegation depth). *Delegation depth* denotes the extent to which a communicative system performs upstream functions that shape the interpretation, judgment, authorization, and continuation of an interaction rather than only the downstream realization of already specified symbolic content.

Delegation depth is independent of message length or linguistic sophistication. A long report can involve shallow delegation when the human supplies the substantive structure and uses a model for editing. A two-sentence response can involve deep delegation when an agent independently decides what the recipient means, whether a concession should be offered, and whether the sender should be bound by it.

Proposition 6.3 (Functional-depth principle). *The social significance of AI mediation depends more directly on the functions delegated than on the proportion of surface text produced by the system. Delegation of interpretation, judgment, authorization, or relational repair can alter the communicative relation substantially even when the visible symbolic contribution is small.*

This principle also limits the usefulness of authorship percentages. A statement such as “80 percent AI-written” can describe token production while remaining silent about who determined the communicative goal, evaluated the recipient’s situation, approved commitments, or retained the ability to revise the interaction.

6.3 Generation, Adoption, and Endorsement

The decomposition of agency also separates symbolic generation from adoption. A human can endorse wording they did not personally compose, just as institutions routinely authorize statements drafted by lawyers, staff, or communications teams. Generative AI expands the range of possible drafters, yet the social force of many messages continues to depend on whether an authorized principal adopts the content.

Let $g(m)$ denote the generative source of message m , and let $a(m)$ denote the actor or institution that adopts the message for communicative purposes. In direct composition, $g(m)$ and $a(m)$ may coincide. Under assisted drafting, they can differ. The distinction becomes especially important when the message contains promises, evaluations, instructions, apologies, or representations of institutional policy.

Definition 6.4 (Mediated expression). *A mediated expression is a symbolic act whose wording, structure, interpretation, or timing is materially shaped by an intermediary while its social standing depends partly on adoption, authorization, or attribution to another principal.*

Mediated expression is historically ordinary. The present change lies in the scale, speed, accessibility, and granularity with which mediation can occur. Generative systems make it feasible to mediate large numbers of individually tailored exchanges, including exchanges in which recipients reasonably infer direct attention or contemporaneous judgment from the apparent speaker.

This distinction yields an important diagnostic implication. Detection of machine generation cannot by itself establish absence of human endorsement. Conversely, human endorsement cannot establish direct human attention to every inference or phrase contained in the message. The relationship among generation, adoption, attention, and responsibility therefore needs to be represented explicitly rather than compressed into a single authorship label.

6.4 Authorization and Standing Commitments

Delegation also changes the temporal structure of authority. A human can authorize a specific draft after reading it. An institution can authorize a class of routine messages through templates or policies. An autonomous agent can receive standing authority to respond within a domain without contemporaneous approval of each output. These arrangements differ in the temporal distance between authorization and symbolic act.

Let $\tau_A(e)$ denote the interval between the principal’s relevant authorization and communicative event e . A short interval can involve message-specific review. A long interval can involve standing instructions, general preferences, or policy constraints. Temporal distance alone does not determine legitimacy or responsibility. It changes what the recipient can infer from the message about contemporaneous human awareness.

Definition 6.5 (Standing communicative delegation). *Standing communicative delegation* occurs when a principal authorizes a technical or organizational delegate to conduct a class of future communicative acts under previously specified goals, constraints, or preferences without message-by-message approval.

Standing delegation is especially important for agentic systems. It allows communication to continue during periods in which the apparent principal is absent from the interaction. The recipient can therefore encounter a responsive symbolic persona that remains operational across time while the human principal’s own state remains unchanged. This possibility intensifies the participation gap identified in Section 5: responsiveness at the symbolic interface can become increasingly independent of contemporaneous participation by the represented person.

6.5 Epistemic Control and Delegation Opacity

Delegation has an epistemic dimension in addition to an operational one. A principal can authorize a system while possessing different levels of understanding of how the system interprets context, generates conclusions, or selects responses. The AI-responsibility literature has long examined cases in which learning systems create difficulties for conventional responsibility attribution because human operators cannot fully predict particular machine actions (Matthias 2004). Later work has distinguished several forms of responsibility gap and emphasized technical, organizational, legal, and ethical sources rather than treating responsibility as one undifferentiated problem (Santoni de Sio and Mecacci 2021). Recent work on agentic AI continues the disagreement: some accounts argue that delegation preserves principal answerability for foreseeable classes of outcome even when the principal lacks control over particular acts (Kong 2026). The present paper does not resolve that normative debate. It draws a diagnostic conclusion: communicative delegation can separate formal authorization from detailed epistemic access to the act being authorized.

We can represent this separation through two variables. Let $A_u(e)$ denote the scope of authorization granted for event e and let $K_u(e)$ denote the principal’s epistemic access to the reasoning, contextual interpretation, and likely consequences of the event. Several arrangements are then possible. High authorization can coexist with high epistemic access, as in careful message-specific review. High authorization can also coexist with low epistemic access, as in broad standing delegation to an agent operating across many interactions.

Definition 6.6 (Delegation opacity). *Delegation opacity* denotes the degree to which recipients, affected parties, or even principals lack access to the functional allocation, interpretive process, authorization path, or responsibility structure through which a mediated communicative act was produced.

Delegation opacity differs from technical opacity. A system can be technically interpretable while the recipient still lacks knowledge of who authorized its use, which parts of the interaction were reviewed, or whether the apparent principal will honor the generated commitment. Conversely, an opaque model can participate in a communicative process whose institutional responsibility structure is clear. The symbolic crisis therefore concerns socio-technical opacity as much as model opacity.

6.6 Agency Topology and the Unified Voice

The preceding distinctions can be integrated through the concept of an *agency topology*. The term refers to the relational arrangement among the actors and functions that jointly produce a communicative event.

Definition 6.7 (Agency topology). For communicative event e , the *agency topology* is the structured relation defined in Equation 26.

$$\mathcal{T}_A(e) = (\mathcal{P}_e, \mathcal{F}_e, \mathcal{D}(e), \mathcal{L}_e), \quad (26)$$

In Equation 24, $\mathcal{P}_e$ denotes participating actors, $\mathcal{F}_e$ communicative functions, $\mathcal{D}(e)$ their delegation assignment, and $\mathcal{L}_e$ the authorization, supervision, and responsibility links among participants.

Table 4: Illustrative delegation architectures in mediated communication.

Architecture	Principal retains	Technical system contributes	Primary diagnostic ambiguity
Surface assistance	Goal, interpretation, substantive judgment, authorization	Editing, translation, formatting, stylistic realization	Linguistic form becomes weak evidence of unaided expressive skill.
Draft delegation	Goal, context, evaluation, authorization	Substantive drafting and alternatives	Surface authorship becomes distributed while endorsement may remain message-specific.
Interpretive delegation	Goal and final authorization	Context interpretation, prioritization, drafting, recommendation	Recipient cannot infer how much of the apparent understanding occurred in the principal.
Standing agentic delegation	High-level objectives and constraints	Interpretation, judgment, generation, transmission, monitoring, routine repair	Symbolic responsiveness can persist without contemporaneous principal attention.
Institutional multi-agent delegation	Organizational policy and supervisory structure	Distributed triage, drafting, escalation, and transaction execution	Unified organizational voice can conceal a complex responsibility and authorization topology.

The topology can be simple or highly distributed. Its significance arises because the symbolic interface often compresses this topology into one apparent speaker. First-person language, a personal email address, a company account, or a familiar conversational style can present a continuous voice while underlying functions are distributed across humans and systems.

Working Claim 6.8 (Unified-voice compression). Generative and agentic mediation can preserve a coherent symbolic persona across communicative events while substantially changing the underlying agency topology. The continuity of voice therefore becomes increasingly weak evidence of continuity in attention, interpretation, judgment, or contemporaneous participation.

This claim extends the diagnosticity analysis from Section 4. The relevant ambiguity concerns more than the origin of words. It concerns which agentic relations the words represent.

Table 4 illustrates why mediation should be classified by functional distribution rather than by a single human/AI label. The categories are heuristic and can overlap. Their purpose is to reveal different inferential problems created by different locations of delegated agency.

6.7 Delegation, Responsibility, and Recipient Reliance

Responsibility becomes especially salient when recipients act in reliance on a mediated message. A system may communicate a price, promise a service, accept a deadline, provide professional guidance, or express a relational commitment. The recipient then needs more than information about token provenance. They may need to know whether the message is authorized, whether the represented principal stands behind it, and where repair is available if the message causes harm or conflict.

The responsibility-gap literature shows that responsibility in AI systems is already contested and multidimensional (Matthias 2004; Santoni de Sio and Mecacci 2021). Meaningful-human-control approaches likewise distinguish responsiveness to relevant human reasons from traceability to human agents capable of understanding their moral role (Santoni de Sio and Hoven 2018). These debates were developed primarily for autonomous systems and decision contexts rather than ordinary communication.

Their relevance here is structural: delegation can widen the distance between the symbolic act and the person or institution expected to answer for it.

For the symbolic-crisis diagnosis, the crucial issue is therefore *responsibility visibility*. A recipient can receive a highly competent message while remaining uncertain about whether the apparent principal has seen it, whether its commitments are institutionally binding, and who can revise or repair the interaction. Generative abundance can increase the frequency of such cases because the marginal cost of maintaining responsive symbolic interfaces falls dramatically.

Proposition 6.9 (Delegation-legibility principle). *As delegation depth increases, the social reliability of mediated communication depends increasingly on the legibility of authorization, endorsement, escalation, and responsibility relations. Symbolic fluency alone supplies progressively less information about these relations.*

This proposition remains diagnostic. It does not imply that every delegated communication requires exhaustive disclosure of its full technical workflow. Different domains will require different levels of agency legibility. The point is that the relevant social object extends beyond the generated text.

6.8 Delegated Agency within the Symbolic Crisis

Delegated agency intensifies the symbolic crisis through three interacting mechanisms. First, it increases *surface continuity*: institutions and individuals can maintain stable voices across increasingly automated workflows. Second, it increases *generative heterogeneity*: the same apparent voice can correspond to direct human composition, assisted expression, reviewed delegation, or autonomous execution. Third, it increases *relational uncertainty*: recipients can have less information about who interpreted them, who updated, who endorsed the result, and who remains exposed to later consequences.

These mechanisms can be summarized schematically in Equation 27.

$$\begin{aligned} \text{deeper delegation} \implies & \text{greater divergence among agency topologies} \\ & \text{compatible with one symbolic surface.} \end{aligned} \tag{27}$$

The implication is probabilistic and context-sensitive. Strong institutional conventions can narrow the divergence. Explicit authorization rules can make delegated communication highly reliable. Some domains can also tolerate extensive automation because semantic accuracy dominates other communicative dimensions. The crisis appears where inherited symbolic cues continue to support expectations about personal or institutional participation that the underlying delegation architecture no longer reliably supplies.

Working Claim 6.10 (Delegated-agency diagnosis). The symbolic crisis in AI-mediated communication concerns a widening gap between the apparent unity of symbolic voice and the possible plurality of agency topologies that can generate, authorize, and sustain that voice. The resulting uncertainty concerns participation, judgment, commitment, and responsibility in addition to authorship.

This diagnosis leads directly to trust. Trust judgments historically use symbolic form as evidence of attention, competence, commitment, and accountable agency. Delegated communication changes the inferential basis of those judgments while preserving many of the surfaces through which trust has conventionally been formed. Section 7 therefore examines how trust changes when symbolic fluency, relational participation, and responsibility become increasingly separable.

7 Trust under AI-Mediated Communication

Section 6 showed that a unified symbolic voice can be generated through substantially different agency topologies. The present section develops the corresponding trust problem. Its role is to distinguish the

objects, dimensions, and inferential bases of trust when communication is mediated by generative or agentic systems. The section does not assume that AI mediation produces a uniform decline in trust. It examines a more selective transformation: recipients may continue to rely on the informational or operational performance of a message while becoming less certain about the attention, commitment, endorsement, competence, or responsibility represented by its symbolic surface. The relevant question is therefore not simply whether a recipient “trusts AI.” It is which relation is trusted, for which purpose, on the basis of which evidence, and with what exposure if that trust proves misplaced.

Trust has long been treated as a means of acting under conditions in which complete verification is unavailable. Luhmann emphasizes its role in reducing social complexity by allowing actors to proceed without reconstructing every relevant contingency (Luhmann 1979). The symbolic-crisis argument adds a specific contemporary complication. Many of the surfaces through which trust has been compressed—fluency, personalization, responsiveness, polished expertise, apparent attentiveness, and coherent voice—can now be produced through a wider range of generative histories. As Sections 2–6 argued, the social significance of this change depends on which inference the recipient is trying to make. Trust therefore needs to be disaggregated before the consequences of AI mediation can be understood.

7.1 Trust Targets within Mediated Communication

Consider a communicative event e in which an apparent principal P is represented through a technical agent A , operating under a delegation relation D , within an institutional setting I , and producing symbolic content m . A recipient can form separate trust judgments about each element. For present purposes, the trust-target profile is defined in Equation 28.

$$\mathcal{T}(e) = \langle T_P, T_A, T_D, T_I, T_m \rangle, \quad (28)$$

In Equation 28, T_P denotes trust in the principal, T_A trust in the technical agent or mediation system, T_D trust in the delegation relation, T_I trust in the institution that authorizes or constrains the interaction, and T_m epistemic or practical reliance on the message content. The notation is schematic rather than additive. High trust in one target can coexist with low trust in another.

Definition 7.1 (Trust-target profile). *A trust-target profile is the differentiated pattern of reliance, expectation, and vulnerability directed toward the principal, technical agent, delegation relation, institution, and communicative content involved in a mediated interaction.*

The profile immediately clarifies several cases that would otherwise be described ambiguously as “trusting” or “distrusting AI.” A recipient can trust that an automated scheduling agent will select an available meeting slot while remaining uncertain whether the represented person has read the accompanying message. A customer can trust an organization’s chatbot to state a refund policy accurately while remaining unsure whether a generated concession binds the organization. A colleague can accept the factual content of an AI-assisted report while reducing the weight placed on the report as evidence of the author’s unaided competence. These are different trust relations, even when they occur around the same symbolic artifact.

Proposition 7.2 (Target-specific trust). *AI-mediated communication should be analyzed through differentiated trust targets. Changes in trust caused by mediation can be concentrated in one target or function while leaving others comparatively stable.*

Figure 7 represents trust as a profile rather than a scalar. The distinction is especially important under deep delegation because recipients may possess strong evidence about the message content and weak evidence about contemporaneous principal participation, or strong institutional guarantees and limited understanding of the technical system that generated the response.

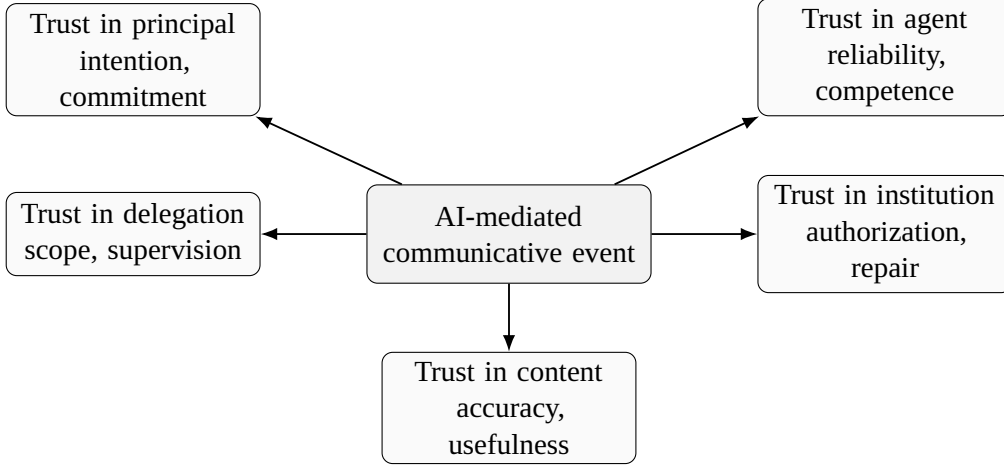


Figure 7: Trust targets in AI-mediated communication. One symbolic event can support distinct judgments about the represented principal, technical agent, delegation relation, institutional environment, and message content.

7.2 Epistemic, Operational, and Relational Trust

Trust targets can also be evaluated along different functions. Three distinctions are particularly useful for the present diagnosis. *Epistemic trust* concerns whether information is accepted as credible enough to inform belief or judgment. Sahebi and Formosa identify generative AI as creating a distinctive problem for epistemic trust in mediated environments because recipients can face difficulty determining the relation between communicative content and its source (Sahebi and Formosa 2025). *Operational trust* concerns whether an agent or institution can be relied upon to perform a bounded task adequately. *Relational trust* concerns expectations about attention, commitment, reciprocity, and accountability across an ongoing relation.

The distinctions can diverge. A technically mediated service can be operationally reliable while relationally thin. A message can be epistemically credible while carrying weak evidence that the apparent sender personally attended to the recipient. Conversely, a close interpersonal relation can support strong relational trust even when the participants recognize that neither party is an expert on the factual issue being discussed.

Definition 7.3 (Trust-function profile). For a trust target z , the *trust-function profile* is the context-sensitive relation defined in Equation 29.

$$\Theta(z) = \langle T^E(z), T^O(z), T^R(z), T^N(z) \rangle, \quad (29)$$

In Equation 29, T^E denotes epistemic trust, T^O operational trust, T^R relational trust, and T^N normative or legitimacy-oriented trust concerning authorization, fairness, and responsibility.

The fourth dimension matters because recipients often need more than reliability. A system can produce accurate outputs through a process that recipients regard as unauthorized, manipulative, or illegitimate. Schilke and Reimann’s experiments are relevant here: across multiple task contexts, disclosure of AI use reduced trust in the AI user, and perceived legitimacy mediated part of that effect (Schilke and Reimann 2025). The result should be interpreted carefully. It demonstrates a context-sensitive relation among disclosure, legitimacy judgments, and trust; it does not establish a universal rule that transparency or AI use always lowers trust.

Working Claim 7.4 (Trust-function divergence). The symbolic crisis can produce divergence among epistemic, operational, relational, and normative trust. A communication system can improve one trust function while weakening the evidential basis of another.

This divergence explains why aggregate trust measures can obscure the structure of the problem. “I trust this system” can mean that its factual answers are usually accurate, that it executes tasks reliably, that its principal stands behind its commitments, or that its use is institutionally legitimate. These judgments need not move together.

7.3 Trust Transfer across Delegated Agency

Delegated communication creates pathways through which trust can be transferred, borrowed, or displaced. A recipient may initially trust principal P and therefore accept outputs produced by agent A under delegation D . Alternatively, a recipient may trust institution I because it guarantees repair, making the reliability of the specific technical agent less important. In other cases, repeated positive interaction with the agent can produce operational trust that exceeds trust in the represented principal.

Let T_P^0 denote prior trust in the principal before a mediated interaction. The posterior judgment is represented in Equation 30.

$$T_P^1 = F(T_P^0, E_m, E_D, E_I, E_A, C), \quad (30)$$

In Equation 30, E_m denotes evidence from the message, E_D evidence about the delegation arrangement, E_I institutional evidence, E_A evidence about agent performance, and C the interaction context. The function F is intentionally unspecified. The point is that the symbolic surface is only one evidential channel among several.

Definition 7.5 (Trust transfer). *Trust transfer* occurs when confidence in one component of a mediated relation contributes to reliance on another component whose qualities have not been independently established to the same degree.

Trust transfer is socially ordinary. Institutional brands, professional credentials, recommendations, and contractual guarantees routinely allow trust to move across relationships. AI-mediated communication changes the topology through which such transfer occurs. A familiar personal voice can lend the principal’s accumulated trust to an agent. An institution can lend legitimacy to automated outputs. Conversely, a conspicuous system failure can feed back into trust in the principal or organization even when the principal did not generate the particular content.

This makes trust a relationally distributed object. The relevant failure mode is not always false information. It can be *misallocated trust*: the recipient relies on one actor because symbolic presentation causes confidence generated elsewhere to be attributed to that actor.

Proposition 7.6 (Trust-transfer asymmetry). *A mediated symbolic surface can transmit trust across an agency topology more readily than it reveals the topology itself. Recipients can therefore rely on a principal, agent, or institution through trust borrowed from another component of the relation while possessing limited information about the underlying allocation of agency.*

7.4 Effort, Attention, and Commitment Cues

Several trust practices historically rely on costly or effort-bearing signals. Chaudhry and Wald review evidence that costly communication can increase perceived honesty because bearing a cost can function as evidence that the communicator has something at stake (Chaudhry and Wald 2022). McEllin and Michael experimentally show that effort invested in informative movement adaptations during joint action can increase trust and generosity toward a co-actor (McEllin and Michael 2022). These findings do not imply that effort is necessary for trust or that expensive communication is inherently more honest. They establish a narrower point: effort can function as one cue within trust formation.

Generative systems complicate this cue because the relation between symbolic polish and human effort becomes less stable. A long, personalized, emotionally calibrated message can result from substantial personal attention, shallow prompting, delegated drafting followed by review, or standing autonomous

Table 5: Illustrative trust targets and ambiguities in AI-mediated communication.

Trust target	Central recipient question	Potentially informative evidence	Characteristic ambiguity under mediation
Principal	Will the represented person or organization stand behind this interaction?	Prior relationship, explicit endorsement, later conduct, repair behavior	Fluent first-person expression can persist without contemporaneous principal attention.
Technical agent	Will the system execute the bounded communicative task reliably?	Performance history, error rates, domain constraints, monitoring	Reliability in one task can be generalized beyond the conditions that support it.
Delegation relation	Is the system acting within an authorized and supervised scope?	Delegation rules, escalation paths, review practices, authorization limits	Recipients can see the output while lacking visibility into the delegated scope.
Institution	Are commitments enforceable and failures repairable?	Policies, accountability mechanisms, institutional history, remedies	Organizational voice can obscure internal responsibility allocation.
Message content	Is the proposition accurate or useful enough to rely upon?	Independent verification, source quality, consistency, evidence	High-quality content can be mistaken for evidence of personal expertise or participation.

representation. The surface can therefore retain the appearance of costly attention while the relevant cost has migrated elsewhere in the production process.

For trust analysis, the important variable is not raw labor time. A recipient may care about whether the sender attended to their situation, accepted exposure to the consequences of the message, or invested scarce judgment in the interaction. Generative assistance can reduce linguistic production cost while leaving these other investments intact. It can also allow a communicative surface to simulate their conventional signs without the corresponding participation.

Definition 7.7 (Commitment-cue ambiguity). *Commitment-cue ambiguity* arises when a symbolic feature historically associated with attention, effort, or relational investment remains perceptually available while its correlation with the underlying commitment has become substantially less stable.

This concept links trust to the diagnosticity analysis in Section 4. The problem is not that every warm or carefully written message has become meaningless. The problem is that the recipient needs more contextual evidence before treating warmth or polish as evidence of the sender’s own attention or commitment.

Table 5 shows why the trust problem cannot be reduced to authorship disclosure. Different trust targets require different evidence, and one disclosure statement rarely resolves all of them.

7.5 The AI Penalty and the Heterogeneity of Recipient Response

Recent empirical work provides evidence that AI mediation can alter trust judgments, while also showing why a single universal effect should be avoided. Sahebi, Formosa, and Bankins report a preregistered experiment in which AI-assisted and fully AI-generated communication was perceived as less trustworthy and less authentic than human-authored communication, with lower willingness to use the content for knowledge uptake; the study also identifies a disclosure paradox in which respondents endorse disclosure norms while disclosed AI use attracts a penalty (Sahebi, Formosa, and Bankins 2026). Schilke and

Reimann similarly find across thirteen experiments that AI disclosure can reduce trust in the user, with legitimacy perceptions contributing to the effect (Schilke and Reimann 2025). Kirk and Givi document related authenticity and consumer-response effects in marketing communication when AI authorship is salient (Kirk and Givi 2025).

These findings are important evidence of an unsettled social transition. Their domain should remain explicit. They concern particular tasks, samples, disclosure conditions, and institutional expectations. They do not establish that recipients always distrust AI-mediated communication, that AI assistance necessarily reduces objective quality, or that the observed penalties will remain stable as social norms evolve. Indeed, the productivity findings discussed in Section 3 show that generative systems can improve task performance in several settings. The symbolic-crisis hypothesis is therefore compatible with simultaneous gains in performance and losses in some forms of trust.

Working Claim 7.8 (Performance–trust divergence). Under generative abundance, improvements in symbolic or task performance can coexist with weaker recipient confidence in authorship, authenticity, legitimacy, or relational participation. Performance and trust should therefore be measured as separate outcome families.

This divergence is central to the historical transition. Organizations and individuals have strong incentives to use systems that lower production cost or improve output quality. Recipients can simultaneously experience uncertainty about what the resulting surface reveals about the people or institutions represented through it. The social pressure toward AI use and the recipient penalty for disclosed AI involvement can therefore coexist rather than cancel each other.

7.6 Operational Trust and Relational Thinness

AI-mediated communication can also create a form of trust that is highly functional while relationally thin. A recipient may learn that an agent responds quickly, remembers preferences, follows policy, and completes transactions reliably. These regularities can support operational reliance. Yet they need not imply that the represented person has become more attentive, knowledgeable, or committed to the relationship.

We can represent this divergence through two trajectories. Let T_t^O denote accumulated operational trust and T_t^R relational trust in the represented principal. Repeated successful automation can produce the divergence stated in Equation 31.

$$T_{t+1}^O > T_t^O \quad (31)$$

Under this pattern, T_{t+1}^R can remain approximately stable or become more uncertain. The notation is illustrative. It shows that trust accumulation can occur in different locations of the agency topology.

Definition 7.9 (Relational thinness). *Relational thinness* denotes a condition in which a communicative arrangement supports reliable symbolic exchange or task completion while providing comparatively little reciprocal updating, personal exposure, or evidence of commitment by the represented human principal.

Relational thinness is not inherently defective. Many communications are appropriately transactional. Automated reminders, status notifications, routine customer-service exchanges, and scheduling interactions can benefit from reliable low-cost mediation. The problem appears when the symbolic interface continues to invite interpretations associated with thicker relational participation. In such cases, operational success can conceal a participation gap.

Proposition 7.10 (Operational-relational separation). *Repeated success of an AI-mediated interface can generate stable operational trust without establishing equivalent relational trust in the represented principal. The distinction becomes increasingly important as agentic systems sustain long-running interactions on behalf of people or institutions.*

7.7 Recipient Vulnerability and Trust Exposure

Trust becomes socially consequential because recipients act under vulnerability. They disclose information, accept advice, make purchases, rely on promises, allocate attention, alter plans, or form expectations about future conduct. AI mediation can change the location of this vulnerability. If the recipient believes they are interacting with a principal who is personally attentive, they may reveal information or make commitments they would not make to an automated delegate. If they believe an institutional message is binding, they may act before verifying whether the agent had authority to create the commitment.

This suggests distinguishing trust level from *trust exposure*. Let $V(e)$ denote the potential consequence to the recipient if the relied-upon expectation in event e is not met. Two interactions with similar reported trust can involve very different exposure. A casual recommendation and a contractual commitment therefore require different degrees of agency legibility and repair capacity.

Definition 7.11 (Trust exposure). *Trust exposure* is the magnitude and character of the recipient’s practical, epistemic, relational, or institutional vulnerability created by acting on a trust judgment within a mediated interaction.

The concept shifts analysis from attitudes alone toward the consequences of reliance. A low-stakes interaction can tolerate substantial ambiguity. High-exposure domains create stronger reasons to clarify authorization, review, and responsibility because the cost of misplaced trust is larger.

Working Claim 7.12 (Exposure-sensitive legibility). The socially relevant degree of agency legibility should increase with the recipient’s trust exposure. The same delegation opacity can be tolerable in a low-stakes exchange and problematic where the recipient relies on the message for consequential commitments or decisions.

The claim remains diagnostic because the appropriate institutional standard depends on domain, law, professional norms, and the distribution of risk. It nevertheless identifies a structural relation between trust, opacity, and consequence that a simple disclosure/no-disclosure distinction misses.

7.8 Trust Recalibration within the Symbolic Crisis

The analysis of this section suggests that the symbolic crisis does not imply the disappearance of trust. Trust is likely to be reallocated across targets and functions as production practices change. Recipients can reduce the evidential weight assigned to polished symbolic form, increase reliance on institutional guarantees, distinguish content reliability from relational participation, or demand clearer information about delegation in high-exposure contexts. These possibilities belong partly to the companion paper on response directions. The diagnostic point here is that such recalibration becomes necessary because inherited trust cues no longer map as stably onto the generative histories behind them.

Proposition 7.13 (Trust-recalibration pressure). *When symbolic surfaces remain socially familiar while the agency topologies behind them become more heterogeneous, recipients face increasing pressure to recalibrate which cues support epistemic, operational, relational, and normative trust.*

The pressure can be gradual and uneven. Some communities may normalize extensive delegation rapidly. Others may retain strong expectations of direct human participation. Institutional rules can stabilize certain inferences even when production technology changes. Trust practices therefore need not converge on one universal equilibrium.

The next analytical problem concerns disclosure. If recipients need information about mediation in order to recalibrate trust, disclosure appears attractive. Yet recent evidence shows that disclosure can itself reduce perceived trust and legitimacy (Schilke and Reimann 2025; Sahebi, Formosa, and Banks 2026). Disclosure can also reveal only one layer of a more complex agency topology. Section 8 therefore examines disclosure, authenticity, and legitimacy as interdependent problems rather than treating transparency as an automatic resolution of the symbolic crisis.

8 Disclosure, Authenticity, and Legitimacy

Section 7 showed that trust in AI-mediated communication is distributed across several targets and functions. The present section examines a response that appears immediately attractive once this differentiation is recognized: disclosure. If a recipient cannot infer the generative history of a symbolic surface reliably, an explicit statement that AI was used seems capable of restoring the missing information. This section argues that disclosure is important while remaining structurally incomplete. Its role is to distinguish disclosure from *agency legibility*, authenticity from unaided human production, and transparency from legitimacy. The discussion proceeds through four related problems: the informational thinness of binary disclosure, the relation between disclosed mediation and authenticity judgments, the institutional legitimacy of delegated communication, and the conditions under which greater transparency can coexist with lower trust.

Organizational research commonly treats transparency as a multidimensional property involving disclosure, clarity, and accuracy rather than mere information release (Schnackenberg and Tomlinson 2016). The symbolic-crisis framework adds a generative dimension to this insight. A disclosure can be literally true while leaving the relevant agency topology almost entirely unspecified. “AI assisted this message” does not reveal whether the system corrected grammar, generated the argument, inferred the recipient’s emotional state, selected the communicative goal, or transmitted the message without contemporaneous human review. The analytical problem is therefore not a simple opposition between disclosed and undisclosed AI use. It is the relation between the information disclosed and the recipient inference that must be made.

8.1 Disclosure Content and Agency Legibility

The agency topology of communicative event e is represented in Equation 32.

$$A(e) = \langle G, I, J, X, U, M, R \rangle, \quad (32)$$

In Equation 32, the components denote goal formation, interpretation, judgment, expression generation, authorization, monitoring, and responsibility, following Section 6. A disclosure $d(e)$ communicates some subset or summary of this topology to the recipient. Binary disclosure provides the coarse representation in Equation 33.

$$d_B(e) \in \{\text{AI used}, \text{AI not used}\}. \quad (33)$$

A richer disclosure can specify which functions were delegated, who reviewed the output, whether the principal endorses the content, and what recourse exists if the communication fails.

Definition 8.1 (Agency legibility). *Agency legibility* is the degree to which a recipient can identify the practically relevant distribution of intention, interpretation, generation, authorization, monitoring, and responsibility behind a mediated communicative event.

Agency legibility is therefore relational rather than exhaustive. A recipient need not receive a complete technical log of model calls, prompts, intermediate drafts, or internal organizational workflow. The relevant question is whether the information available is sufficient for the trust judgment at stake. A scheduling message can require little more than confirmation that the agent is authorized to book within specified constraints. A consequential professional recommendation can require clearer information about human review, evidential responsibility, and the scope of delegated judgment.

Proposition 8.2 (Disclosure–legibility distinction). *The presence of AI disclosure and the legibility of communicative agency are analytically distinct. A disclosure can be accurate yet insufficient for the recipient to infer who exercised judgment, who endorsed the resulting message, and who remains responsible for its consequences.*

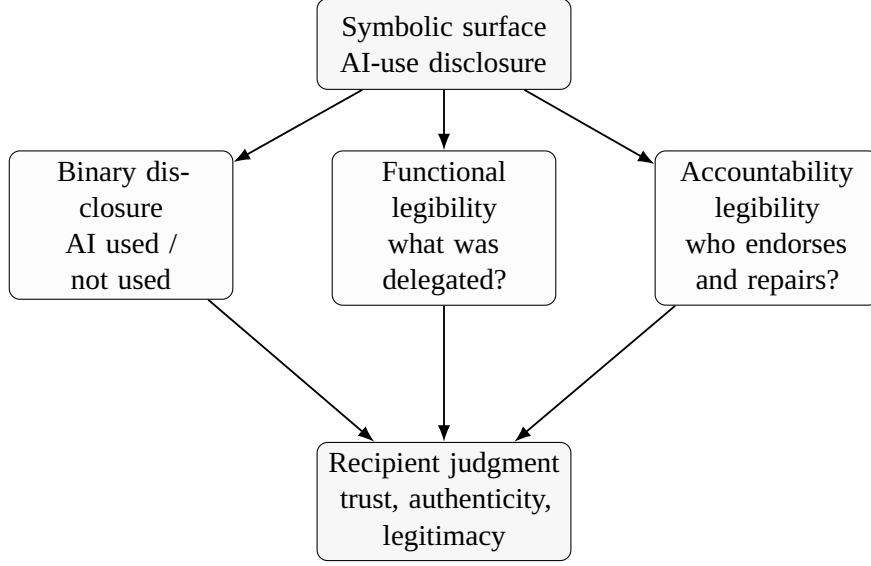


Figure 8: Disclosure can reveal AI involvement without making the relevant agency topology legible. Functional and accountability information can therefore matter independently of a binary AI-use label.

Binary labels can therefore become socially overloaded. “AI-generated” can collapse a wide range of arrangements into one category, while “human-authored” can conceal substantial technical assistance. Each label can support fewer inferences than recipients attribute to it.

Figure 8 presents disclosure as one input into recipient judgment rather than as a complete solution. The figure also clarifies why more information can change trust in either direction. Disclosure can reduce uncertainty while simultaneously revealing that an interaction contains less human participation than the recipient expected.

8.2 The Disclosure Sufficiency Gap

The difference between available disclosure and decision-relevant agency information can be represented as a *disclosure sufficiency gap*. Let $L^*(e, q)$ denote the level of agency legibility required for a recipient question q , and let $L_d(e)$ denote the legibility supplied by the disclosure practice in event e . The corresponding definition is given in Equation 34.

$$\Delta_L(e, q) = L^*(e, q) - L_d(e). \quad (34)$$

The expression is conceptual rather than metrically fixed. A positive gap means that the recipient is asked to make a judgment whose relevant agency information is not sufficiently represented by the disclosure.

Definition 8.3 (Disclosure sufficiency gap). A *disclosure sufficiency gap* occurs when the information disclosed about mediation is too coarse, ambiguous, or incomplete for the recipient to evaluate the trust, authorization, authenticity, or responsibility question that the interaction makes practically relevant.

The gap can arise in several ways. A disclosure can identify the presence of AI while omitting the degree of human review. It can name the model while omitting whether the principal supplied the substantive judgment. It can identify an automated agent while leaving its authority to make commitments unclear. It can disclose generation without identifying who will repair an error. These are not failures of honesty in the narrow sense. They are failures of correspondence between the granularity of disclosure and the granularity of the recipient’s decision.

Working Claim 8.4 (Exposure-sensitive disclosure). The informational sufficiency required of disclosure tends to increase with the recipient’s trust exposure and with the depth of delegated judgment. Low-

exposure interactions can remain workable under coarse disclosure, while high-exposure interactions create stronger reasons for functional and accountability legibility.

This claim extends the exposure-sensitive legibility principle developed in Section 7. It does not imply that every interaction should carry a long technical notice. Excessive disclosure can itself burden recipients and reduce communicative usability. The design problem concerns fit between the information supplied and the reliance invited.

8.3 Authenticity beyond Unaided Production

The symbolic crisis also complicates authenticity. A simple human/AI binary encourages the inference that human-originated expression is authentic while generated expression is artificial. That binary is too coarse for mediated communication. A person can use substantial assistance while genuinely endorsing the final message, and a person can manually type words that strategically simulate attention or feeling they do not possess. Authenticity therefore cannot be identified with production technique alone.

For present purposes, authenticity can be treated as an attribution concerning the relation between a symbolic expression and the agency, commitments, or identity it purports to manifest. Let P_e be the participation actually present in event e and P_e^{imp} the participation reasonably implied by the surface and context. The symbolic-crisis problem becomes especially salient under the participation mismatch stated in Equation 35.

$$P_e^{\text{imp}} \gg P_e. \quad (35)$$

The notation does not claim that participation can be placed on one universal scale. It indicates a mismatch between represented and actual participation.

Definition 8.5 (Participatory authenticity). *Participatory authenticity* is the degree of fit between the human or institutional participation reasonably attributed to a communicative expression and the participation, endorsement, and responsibility that actually sustain it.

This definition leaves room for authentic mediated expression. A researcher can use AI to improve grammar while retaining substantive judgment and responsibility. A multilingual speaker can rely on translation support while fully endorsing the communicated meaning. An organization can authorize an automated notification whose institutional force is genuine even though no employee composed the particular sentence. Conversely, highly human-produced expression can remain inauthentic when it strategically misrepresents commitment or identity.

Proposition 8.6 (Production–authenticity separation). *The degree of technical mediation and the authenticity of a communicative event should be evaluated separately. Mediation changes the evidence available for authenticity judgments, but authenticity depends on the fit among attributed participation, actual participation, endorsement, and responsibility rather than on unaided production alone.*

Recent empirical work nevertheless shows that recipients often use AI involvement as an authenticity cue. Sahebi, Formosa, and Bankins report lower perceived authenticity and trustworthiness for AI-assisted and AI-generated communication in their experimental setting (Sahebi, Formosa, and Bankins 2026). Kirk and Givi similarly identify authenticity-related penalties in AI-authored marketing communication (Kirk and Givi 2025). These findings are compatible with the present definition because recipient judgments can respond to perceived participation gaps even when the actual normative status of a particular mediated communication remains context-dependent.

8.4 Legitimacy of Mediated Communicative Arrangements

Authenticity concerns the relation between expression and attributed participation. Legitimacy concerns a broader question: whether the communicative arrangement is regarded as appropriate, acceptable, or

Table 6: Distinct evaluative questions around disclosure, authenticity, and legitimacy.

Dimension	Central question	Characteristic evidence	Failure mode under symbolic abundance
Disclosure accuracy	Is the stated information about mediation true?	AI-use statement, process record, technical metadata	A true label can remain too coarse for the relevant inference.
Agency legibility	Who interpreted, judged, generated, authorized, and remains responsible?	Functional disclosure, review history, authorization scope, repair path	Unified symbolic voice obscures distributed agency.
Authenticity	Does the expression correspond to the participation or commitment it appears to manifest?	Endorsement, relational context, later conduct, participation evidence	Surface warmth or personal voice can overstate actual human participation.
Legitimacy	Is this form of delegation appropriate and acceptable in this context?	Domain norms, institutional authorization, recipient expectations, accountability	Efficient mediation can remain normatively contested or socially unfamiliar.
Trust	What reliance is warranted, toward which target, and at what exposure?	Performance history, guarantees, relationship, disclosure, independent evidence	One disclosure cue is generalized across several distinct trust functions.

entitled to structure the interaction. Organizational legitimacy research distinguishes pragmatic, moral, and cognitive forms of legitimacy, emphasizing that acceptance can depend on self-interest, normative approval, and comprehensibility or taken-for-grantedness (Suchman 1995). AI-mediated communication can affect all three.

A recipient can find an automated arrangement pragmatically useful because it is fast and accurate, morally questionable because it conceals consequential delegation, and cognitively unfamiliar because norms for agentic communication remain unsettled. Conversely, a highly routinized automated interaction can become cognitively legitimate even when little relational participation is expected. Legitimacy therefore cannot be inferred directly from whether AI is present.

Definition 8.7 (Mediation legitimacy). *Mediation legitimacy* is the context-dependent acceptability of an arrangement through which communicative functions are delegated, represented, authorized, and made consequential for recipients.

Mediation legitimacy depends partly on role expectations. The communicative significance of delegation differs across scheduling, customer support, scholarship, intimate exchange, professional advice, and institutional commitment. A form of assistance that is unremarkable in one domain can violate participation expectations in another. This domain dependence helps explain why disclosure effects can vary even when the underlying technology is similar.

Table 6 prevents several categories from collapsing into one another. Accurate disclosure can coexist with low authenticity judgment. Authentic mediated expression can occur within an institutionally illegitimate delegation arrangement. Legitimate automation can support operational trust while remaining relationally thin. The dimensions interact without being interchangeable.

8.5 The Transparency Dilemma

The empirical finding that disclosure can reduce trust should therefore not be treated as evidence that transparency is socially harmful in general. Schilke and Reimann report that AI disclosure can reduce

trust in the person using AI, partly through perceived legitimacy (Schilke and Reimann 2025). Sahebi, Formosa, and Bankins likewise identify a disclosure paradox in which respondents favor disclosure norms while disclosed AI involvement attracts a trust and authenticity penalty (Sahebi, Formosa, and Bankins 2026). These results can be interpreted as a transparency dilemma: information that improves epistemic accuracy about the production process can simultaneously expose a mismatch with inherited expectations about participation.

Definition 8.8 (Transparency dilemma). *A transparency dilemma arises when disclosure improves the recipient's knowledge of a mediated communicative arrangement while the newly revealed arrangement receives a lower trust, authenticity, or legitimacy evaluation than the undisclosed surface would have received.*

There is no logical contradiction in this pattern. The recipient can become better informed and less trusting at the same time. Indeed, if undisclosed communication encouraged an inference of direct human participation that was inaccurate, reduced trust after disclosure can reflect successful recalibration rather than a failure of transparency. This point matters because an exclusive focus on preserving trust can create incentives to retain ambiguity.

Proposition 8.9 (Epistemic value without trust preservation). *The value of disclosure cannot be assessed solely by whether it preserves or increases trust. Disclosure can have epistemic value by enabling more accurate trust allocation even when the recipient's reported trust decreases.*

At the same time, transparency research cautions that useful transparency requires more than maximal information release. Clarity and accuracy matter alongside disclosure (Schnackenberg and Tomlinson 2016). In AI-mediated communication, this implies that meaningful disclosure should help recipients understand the relevant agency arrangement rather than merely announce that a technology was present.

8.6 Authenticity Norms under Historical Transition

The symbolic-crisis thesis treats present authenticity judgments as historically situated rather than permanent. Current recipients have learned expectations under communicative regimes in which fluent personalized expression was more tightly coupled to direct human production and attention. Generative abundance loosens that coupling before social norms have fully adapted. The AI penalty documented in recent experiments can therefore be interpreted partly as evidence of a transitional normative environment rather than as a timeless response to mediated expression.

This interpretation remains open to several futures. Direct human participation could retain strong normative value in some domains. Other domains could normalize extensive delegation while developing new conventions for authorization and endorsement. Institutional forms could stabilize specific agency arrangements, making them cognitively familiar and easier to evaluate. The diagnostic task is to track how authenticity and legitimacy norms move as production practices change rather than to predict one universal endpoint.

Working Claim 8.10 (Normative lag). *When generative practices change faster than shared expectations about legitimate participation, authenticity and legitimacy judgments can display a temporal lag: recipients continue to interpret symbolic surfaces through norms formed under earlier production regimes while new agency topologies become increasingly common.*

Normative lag is one mechanism within the broader temporal mismatch developed later in the paper. It helps explain why organizations can face simultaneous pressure to adopt AI for efficiency and to conceal, minimize, or carefully frame its use in contexts where disclosure carries a penalty. The resulting incentive structure is ethically significant because it can reward opacity precisely when agency legitimacy becomes more important.

8.7 Disclosure within the Symbolic Crisis

Disclosure therefore occupies an intermediate position in the symbolic crisis. It remains necessary in many contexts because recipients cannot recalibrate trust without information about mediation. Yet disclosure does not reconstruct the full generative history behind a symbol, determine whether the resulting expression is authentic, or settle whether the delegation arrangement is legitimate. Those evaluations require contextual knowledge about participation, endorsement, responsibility, institutional norms, and recipient exposure.

Proposition 8.11 (Disclosure insufficiency). *Disclosure can improve the epistemic conditions of mediated communication without resolving the symbolic crisis. Its contribution depends on whether it makes the agency relations relevant to the recipient’s judgment sufficiently legible.*

The proposition preserves the diagnostic character of this paper. The companion inquiry can later examine what forms of agency legibility, trajectory records, relational practice, or institutional redesign might respond to the insufficiency. The present paper needs only the narrower conclusion: a social order cannot restore the previous diagnostic force of symbols simply by attaching an “AI used” label to increasingly heterogeneous generative histories.

Section 9 turns from communication to expertise. Generative systems can produce sophisticated explanation, analysis, and disciplinary language at low marginal cost. The resulting problem resembles the one developed here but concerns a different inference: what does an impressive epistemic performance reveal about the competence, understanding, and intellectual history of the person who presents it?

9 Expertise and Epistemic Performance

Sections 4–8 examined how generative abundance changes the evidential meaning of symbolic surfaces in communication. The present section extends that diagnosis to expertise. Its role is to distinguish an observable epistemic performance from the broader capacities that make an actor expert across varying tasks, changing conditions, and repeated encounters with error. The distinction matters because generative systems can produce sophisticated explanations, analyses, summaries, code, designs, and professional language at a marginal cost that is increasingly weakly coupled to the human user’s prior disciplinary formation. The resulting problem is neither that AI-assisted performance is necessarily inferior nor that unaided human production is intrinsically authoritative. It is that the evidential relation between a polished performance and the expertise attributed to the person presenting it becomes less stable.

The problem of identifying expertise predates generative AI. Goldman frames expert selection as an epistemic problem for non-experts who must judge testimony without possessing the same domain competence as the people they evaluate (Goldman 2001). Collins and Evans likewise distinguish several forms of expertise and emphasize that fluency in the language of a specialization does not exhaust practical contribution within that specialization (H. Collins and Evans 2007). Generative systems intensify this older problem because interactional fluency and task-level performance can increasingly be supplied, amplified, or partially substituted by technical systems. The resulting symbolic crisis concerns what a present performance still licenses an observer to infer about the performer.

9.1 Observable Performance and Latent Expertise

Let $p(i, q, t)$ denote the observable performance of actor i on task q at time t . A performance can be evaluated along dimensions such as correctness, coherence, explanatory quality, speed, relevance, or professional form. Let E_i denote the broader expertise profile attributed to the actor. The central infer-

ential distinction is stated in Equation 36.

$$p(i, q, t) \not\equiv E_i. \quad (36)$$

The expression marks a category distinction rather than a claim that performance is irrelevant. Performance is evidence of expertise, but a single performance is a task-bound observation generated under particular tools, prompts, institutional supports, and opportunities for correction.

Definition 9.1 (Epistemic performance). An *epistemic performance* is an observable production or judgment through which an actor or socio-technical arrangement answers, explains, analyzes, classifies, predicts, designs, or otherwise handles an epistemically relevant task.

Definition 9.2 (Expertise profile). An *expertise profile* is a temporally extended and task-distributed configuration of capacities that can include domain knowledge, problem formulation, judgment, transfer across cases, error recognition, boundary recognition, justification, repair, learning, and responsibility for consequences.

The definition is intentionally plural. Expertise can be highly specialized, institutionally distributed, and dependent on instruments or collaborators. It also leaves open how much tacit knowledge, embodied practice, or social accreditation should be required in a particular domain. Collins and Evans’s distinction between contributory and interactional expertise is relevant here because an actor can participate fluently in specialist discourse without possessing the same practical capacity to contribute to all activities of the specialist community (H. Collins and Evans 2007). Generative systems add another layer: fluent specialist discourse can now be produced through a human–model arrangement even when the human participant has limited interactional or contributory expertise of their own.

Proposition 9.3 (Performance–expertise separation). *The quality of an epistemic performance and the breadth or durability of the performer’s expertise should be evaluated separately. High task-level performance can provide evidence of expertise without uniquely identifying the human capacities, technical assistance, or institutional supports that generated it.*

Figure 9 extends the inverse-inference problem developed earlier in the paper. The observer sees the output but may not know which generative history produced it. The relevant ambiguity is therefore not only authorship. It concerns what kind of expertise, if any, the symbolic performance reveals about the person whose name, role, or identity accompanies it.

9.2 Performance–Expertise Decoupling

Generative abundance can weaken the coupling between baseline human competence and observable task performance. This effect is not uniform across tasks or workers. Brynjolfsson, Li, and Raymond find that generative AI assistance in customer support produced larger productivity and quality gains for less experienced and lower-skilled workers, while gains for the most experienced and highest-skilled workers were smaller and included a small decline in quality on some measures (Brynjolfsson et al. 2025). In that setting, AI assistance partially compressed performance differences associated with prior experience.

Dell’Acqua and colleagues report a complementary pattern among management consultants. For tasks within the capabilities of the studied AI system, AI-assisted participants completed more tasks, worked faster, and produced higher-quality outputs; for a task selected outside the system’s frontier, AI-assisted participants were less likely to reach the correct solution (Dell’Acqua et al. 2026). The finding is important for the symbolic-crisis argument because it demonstrates that visible performance can improve substantially without producing a uniform increase in underlying task reliability across a knowledge workflow.

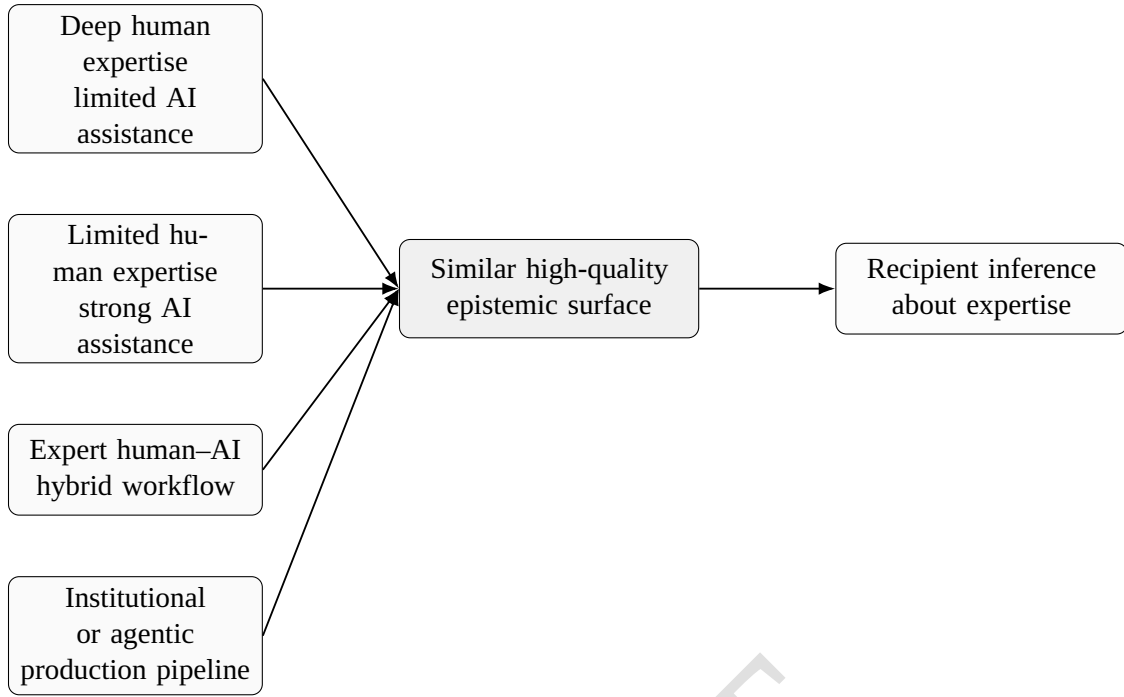


Figure 9: Expertise inference under generative heterogeneity. Similar epistemic surfaces can emerge from substantially different combinations of human expertise, AI assistance, and institutional support.

Definition 9.4 (Performance–expertise decoupling). *Performance–expertise decoupling* is the weakening of the statistical or practical relation between an actor’s baseline domain competence and the quality of a particular observable performance because technical assistance, institutional support, or delegated cognition contributes materially to that performance.

Decoupling is not equivalent to de-skilling. AI assistance can support learning, extend access to expertise, and improve real performance. Brynjolfsson, Li, and Raymond report evidence of faster movement down the experience curve for newer workers in their setting (Brynjolfsson et al. 2025). The diagnostic claim is narrower: when assistance alters the mapping from baseline skill to visible output, observers need more caution before using the output as a direct measure of the individual’s independent competence.

Working Claim 9.5 (Performance compression). Where generative assistance disproportionately improves the performance of actors with lower baseline skill, observable output differences can narrow faster than underlying expertise differences. The symbolic surface then becomes a weaker discriminator of baseline expertise even when aggregate performance improves.

This is a socially ambivalent development. Performance compression can lower barriers, improve service quality, and make high-quality work available to more people. At the same time, institutions that continue to infer expertise from the same surface cues can experience a diagnostic mismatch. A polished memo, technically sophisticated explanation, or fluent disciplinary discussion can remain useful as an artifact while revealing less about the person who presents it.

9.3 The Jagged Frontier of Assisted Performance

The distinction between performance and expertise becomes especially important when AI capabilities are uneven across tasks. Dell’Acqua and colleagues describe a *jagged technological frontier*: tasks that appear similar in human difficulty can fall on opposite sides of the boundary at which AI assistance im-

proves or worsens performance (Dell’Acqua et al. 2026). This finding complicates any simple inference from successful AI use to broad competence.

Suppose an actor performs strongly on tasks $q_1, \dots, q_k$ with assistance. If the underlying capability frontier is irregular, that success does not by itself show that the actor can recognize when a neighboring task q_{k+1} lies outside the region in which the same workflow remains reliable. The capacity to recognize this boundary is itself epistemically significant.

Definition 9.6 (Boundary-recognition competence). *Boundary-recognition competence* is the capacity to identify when one’s own knowledge, an assisting system, or a human–AI workflow is approaching conditions under which previously reliable methods cease to warrant the same confidence.

Boundary recognition is one component of durable expertise because epistemic performance includes decisions about when to defer, verify, seek another method, or refuse an apparent solution. A generative system can supply fluent continuation even where the surrounding human lacks the competence to identify the relevant failure mode. In such cases, the central risk is not that the output looks polished. It is that the recipient and the human user can both over-attribute reliability to a performance whose production process exceeds the user’s epistemic control.

Mühlhoff’s distinction among forms of LLM delegation is relevant here because delegation can vary in how much judgment and epistemic control remains with the human user (Mühlhoff 2026). The symbolic-crisis problem therefore concerns not only whether assistance occurred, but whether the human participant retained enough domain understanding to evaluate the generated performance at the level appropriate to the decision.

Proposition 9.7 (Boundary-recognition requirement). *The diagnostic value of successful AI-assisted performance depends partly on whether the actor can recognize, test, and respond to the boundaries of the assisting system’s competence. Repeated success within one capability region does not establish reliable judgment outside that region.*

9.4 Verification Burden and Epistemic Control

Generative abundance reduces the cost of producing candidate explanations, reports, analyses, and solutions. It does not necessarily reduce the cost of validating them by the same amount. In some domains, verification remains easier than generation; in others, correct evaluation requires expertise comparable to that needed for original production. This asymmetry creates a verification bottleneck.

Let $C_G(q)$ denote the cost of generating a plausible response to task q , and $C_V(q)$ the cost of validating that response to an appropriate standard. Generative abundance can produce the cost asymmetry stated in Equation 37.

$$C_G(q) \ll C_V(q). \quad (37)$$

The relation is task-dependent. The analytical point is that a rapid fall in generation cost can shift scarce resources toward checking, comparing, tracing, and deciding whether the output should be trusted.

Definition 9.8 (Verification burden migration). *Verification burden migration* is the relocation of epistemic cost from producing an apparently adequate symbolic output toward evaluating its correctness, scope, provenance, boundary conditions, and consequences.

This migration matters for expertise attribution. An actor who can obtain a sophisticated answer may still lack the competence required to validate that answer. Conversely, an expert may use AI extensively while retaining strong epistemic control because the expert can recognize errors, challenge assumptions, reformulate the problem, and determine when additional evidence is required. The amount of generated text therefore reveals little by itself about the distribution of expertise within the workflow.

Working Claim 9.9 (Epistemic-control distinction). Technical assistance can increase epistemic performance while leaving the user’s epistemic control unchanged, increasing it through learning and comparison, or reducing it through dependence on outputs the user cannot adequately evaluate. These possibilities should be distinguished empirically rather than inferred from AI use alone.

9.5 Individual Expertise and Hybrid System Competence

The symbolic crisis also requires a distinction between expertise attributed to a person and competence attributed to a socio-technical arrangement. Consider a recurring workflow in which a human, an AI system, organizational databases, review procedures, and specialist colleagues jointly produce high-quality decisions. The arrangement may be highly competent even if no individual possesses all of the capabilities manifested in the final output.

Define a hybrid epistemic arrangement in Equation 38.

$$H = \langle P, A, I, R \rangle, \quad (38)$$

In Equation 36, P is the human participant or principal, A the technical agent, I the informational and institutional infrastructure, and R the review and repair process. Let $K(H, q)$ denote the competence of the arrangement on task q . The inference evaluated in Equation 39 requires additional evidence.

$$K(H, q) \text{ is high} \implies E_P \text{ is equally high} \quad (39)$$

is generally unwarranted without additional evidence.

Definition 9.10 (Hybrid epistemic competence). *Hybrid epistemic competence* is the task-sensitive capacity of a human–technical–institutional arrangement to produce, evaluate, and repair epistemic outputs at an appropriate standard.

The distinction protects two kinds of judgment. It prevents the under-recognition of genuinely effective hybrid systems merely because their performance is not unaided human work. It also prevents the competence of the system from being automatically credited to the individual who presents its output. In professional and scholarly settings, both judgments can matter simultaneously: an institution can reasonably value hybrid performance while still needing to know what the individual can judge, explain, or take responsibility for when the technical system fails.

9.6 Expertise as a Social Signal

Expertise is not only a productive capacity. It is also a social attribution used to allocate authority, employment, responsibility, access, and trust. Goldman emphasizes that laypeople often must choose among experts using evidence available from outside the expert domain itself (Goldman 2001). Credentials, reputation, argumentative quality, track record, institutional affiliation, and visible performance all participate in that judgment. Section 2 described such signs as forms of social compression.

Generative abundance weakens some of the traditional shortcuts. An articulate answer can be less diagnostic of unaided knowledge. A polished report can be less diagnostic of writing or analytical skill. Rapid production can be less diagnostic of accumulated experience. The relevant implication is not that these surfaces lose all value. Their inferential targets must be recalibrated.

Definition 9.11 (Expertise-attribution mismatch). An *expertise-attribution mismatch* occurs when observers assign a symbolic performance more evidential weight for the performer’s durable expertise than the current production conditions warrant.

Table 7 shows why a single category of “AI-assisted expertise” is too coarse. The same artifact can carry strong evidence about one property and weak evidence about another. A correct answer remains

Table 7: Dimensions of epistemic performance and expertise under generative abundance.

Dimension	Observable surface	Underlying capacity of interest	Diagnostic pressure under generative abundance
Answer quality	Correct, coherent, relevant response	Domain knowledge and judgment	High-quality output can arise from several human–AI competence distributions.
Explanatory fluency	Specialist vocabulary and structured exposition	Conceptual understanding and ability to defend claims	Language production becomes less tightly coupled to the speaker’s independent understanding.
Problem handling	Rapid solution to a bounded task	Transfer, reformulation, and boundary recognition	Success inside an AI capability region can coexist with weak recognition of nearby failure conditions.
Error repair	Plausible correction after feedback	Diagnosis of failure and selection of corrective method	Generated repair can conceal whether the human identified the underlying error.
Professional form	Polished report, code, design, or presentation	Durable production skill and domain practice	Surface polish becomes less diagnostic of production history and baseline skill.
Responsible reliance	Confident recommendation or decision	Epistemic control, uncertainty judgment, and accountability	Confidence can be generated more cheaply than warranted capacity to validate consequences.

evidence that a competent answer was produced. It may be much weaker evidence that the named human could reproduce, defend, transfer, or appropriately limit that answer without the same assistance.

9.7 Epistemic Performance within the Symbolic Crisis

The expertise problem therefore follows the general structure of the symbolic crisis. Under earlier production conditions, sophisticated symbolic performance was often costly enough that observers could use it as a relatively informative proxy for training, attention, domain fluency, and practical competence. The proxy was always imperfect. Generative systems widen the set of histories capable of producing the same surface and thereby increase inverse-inference ambiguity.

The result should not be described as the disappearance of expertise. Expertise remains necessary for many forms of problem selection, validation, boundary recognition, responsibility, and repair. Nor does the analysis imply that AI-assisted work is epistemically suspect by default. The diagnosis concerns the social meaning of observable performance.

Proposition 9.12 (Expertise-diagnostics shift). *Under generative abundance, observable epistemic performance can remain valuable while becoming less diagnostic of the human performer’s durable expertise. Institutions that continue to allocate authority or trust from unchanged surface cues therefore face increasing pressure to distinguish artifact quality, hybrid system competence, and individual expertise.*

This shift creates a bridge to the next two sections. Section 10 examines symbolic achievement in education, where assignments, essays, examinations, and other performances have historically served

both pedagogical and diagnostic functions. Section 11 then examines credentials as institutional compression devices whose social value depends partly on the reliability of the processes and histories they summarize. The present paper remains diagnostic: later response-oriented work can consider whether trajectories, longitudinal records, co-experience, or new verification infrastructures offer scalable ways to rebuild credibility when isolated epistemic performances cease to carry the same evidential weight.

10 Education and Symbolic Achievement

The preceding section separated epistemic performance from expertise. The present section extends that distinction into education, where symbolic artifacts do more than display task performance: they are routinely used to infer learning, mastery, effort, intellectual formation, and readiness for subsequent study or responsibility. Its role is to identify how generative mediation changes those inferences. The section proceeds in four steps. It first distinguishes an *achievement surface* from the learner formation attributed to it. It then examines the separation between assisted performance and durable learning, using recent empirical studies to show that generative systems can either support or inhibit learning depending on pedagogical design. It next connects this separation to educational measurement, where validity depends on the interpretation made from a performance rather than on the artifact alone. Finally, it develops the diagnostic problem created when institutions retain inherited achievement symbols while the production conditions of those symbols become increasingly heterogeneous.

Educational assessment has always involved inference. A written solution, essay, oral presentation, laboratory report, examination score, or project is observable; the learning that educators wish to attribute to the student is only partially observable through that performance. Messick’s account of construct validity treats the interpretation of performances as an evidential problem and emphasizes that performance assessments require attention to the meaning and consequences attached to the resulting judgments (Messick 1995). Kane’s argument-based account similarly locates validity in the interpretations and uses built from scores, with the inferential chain from response to broader claim requiring explicit support (Kane 2013). Generative systems intensify this longstanding measurement problem by altering the process through which educational performances can be produced.

Definition 10.1 (Symbolic achievement). A *symbolic achievement* is an observable educational performance, artifact, score, or representation that is used to support an inference about a learner’s knowledge, skill, judgment, formation, or readiness.

The definition concerns evidential function rather than intrinsic merit. An essay can be intellectually valuable even when it is a weak measure of unaided writing ability. A collaboratively produced presentation can demonstrate coordination even when it reveals little about each member’s independent competence. A project completed with generative assistance can be educationally productive while leaving the specific contribution of the learner underdetermined. The symbolic crisis appears when these distinctions remain implicit while the artifact continues to carry inherited evidential weight.

Table 8 shows why the educational problem extends beyond academic integrity. Integrity concerns remain important, but the more general issue is inferential. A symbolic achievement can be legitimately produced under permitted assistance and still become a different kind of evidence from the same surface produced under another mediation regime. The educational question is therefore partly a question of *what the artifact is meant to evidence*.

10.1 Achievement Surfaces and Learner Formation

Let $A_{i,q,t}$ denote learner i ’s observable achievement on task q at time t . Let L_i denote a broader learner-formation profile that can include retained knowledge, transfer, problem formulation, error recognition,

Table 8: Educational achievement surfaces and the inferences commonly attached to them. The table separates the observable artifact from the broader educational property that an institution may attempt to infer.

Achievement surface	Possible educational inference	Generative-era ambiguity
Essay or report	disciplinary understanding, writing, synthesis, judgment	drafting, structuring, explanation, or revision may be distributed across learner and system
Problem set	procedural knowledge, conceptual understanding, error correction	correct solutions may be produced with different levels of learner comprehension
Presentation	understanding, organization, communication, preparation	slides, scripts, examples, and anticipated questions may be extensively mediated
Project or code artifact	problem solving, design, implementation, integration	successful artifact may depend on delegated generation, debugging, or architectural choices
Examination score	performance under specified assessment conditions	interpretability depends on whether assistance conditions correspond to the intended construct
Portfolio	development, range, continuity, reflective capacity	polished outputs may conceal heterogeneous levels and forms of participation across time

independent reconstruction, disciplinary judgment, and capacity to continue learning. Let $M_{i,q,t}$ represent the mediation configuration through which the achievement is produced. Then an educational inference has the general form stated in Equation 40.

$$A_{i,q,t}, M_{i,q,t}, C_{q,t} \rightsquigarrow \hat{L}_i, \quad (40)$$

In Equation 40, $C_{q,t}$ represents the assessment conditions and interpretive context. The notation emphasizes that the same observed achievement can warrant different conclusions when mediation and task conditions differ.

Definition 10.2 (Achievement–formation decoupling). *Achievement–formation decoupling* is a weakening or redistribution of the relation between an observable educational performance and the learner capacities that institutions conventionally infer from that performance.

Generative systems can contribute to such decoupling because they can supply language, intermediate reasoning, examples, revisions, hints, code, formatting, or complete candidate solutions. Yet the direction of the effect remains open. Assistance can substitute for learner activity, scaffold it, expose errors, accelerate feedback, or create new opportunities for practice. The diagnostic problem therefore concerns the instability of the mapping rather than a presumption that assistance reduces learning.

Proposition 10.3 (Achievement-surface ambiguity). *As educational production pathways become more heterogeneous, similarity in symbolic achievement provides less information by itself about similarity in learner formation. Additional information about task design, mediation, and subsequent performance becomes increasingly relevant to educational interpretation.*

This proposition extends the diagnosticity argument from Section 4. An educator observing two equally strong essays may reasonably judge the essays themselves to be similar in quality while remaining uncertain about whether the students developed similar capacities through producing them. The ambiguity concerns the inference from artifact to learner rather than the value of the artifact as text.

Figure 10 identifies the central ambiguity. The figure does not rank the three production pathways. A well-designed AI scaffold may generate more learning than an unaided task, while an assistance pattern

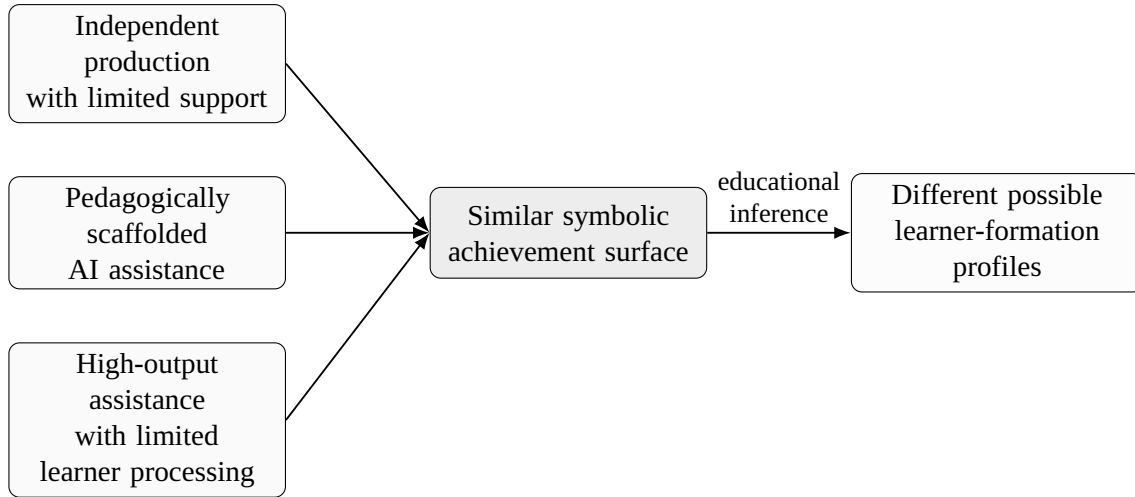


Figure 10: Educational inference under heterogeneous production pathways. Similar symbolic achievements can emerge from learning processes that differ substantially in learner participation, scaffolding, and durable formation.

that supplies answers too readily may support immediate performance with weaker durable formation. The evidential meaning of the surface therefore depends on the learning process and on what later capacities are retained.

10.2 Assisted Performance and Durable Learning

Recent evidence illustrates why assisted performance and learning should be kept analytically separate. Bastani and colleagues conducted a large field experiment in high-school mathematics in which students used two forms of GPT-4 support. Access to a general-purpose interface improved performance on practice problems, yet students in that condition later performed worse than controls when AI access was removed. A second tutor, designed with pedagogical safeguards and teacher-authored hints, substantially improved practice performance while largely mitigating the later learning loss (Bastani et al. 2025). The result is especially important for the present argument because the visible achievement during assisted practice and the subsequent unaided performance moved differently depending on the mediation design.

A different randomized trial by Kestin and colleagues found that an AI tutor designed around research-based pedagogical principles produced larger short-term learning gains than an in-class active-learning condition in an undergraduate physics setting, with students also reporting greater engagement and motivation (Kestin et al. 2025). These studies examine different populations, subjects, designs, and outcome windows, so they should not be placed into a simple competition over whether AI is educationally beneficial. Together they show something more useful for the symbolic-crisis thesis: the educational significance of generative assistance depends heavily on the interaction design and on the outcome being measured.

Definition 10.4 (Assisted-performance gap). The *assisted-performance gap* is the difference between what a learner can produce under a given mediation configuration and what the learner can subsequently understand, reconstruct, transfer, critique, or perform under a different configuration.

An assisted-performance gap can be positive, small, or negative depending on how the learning process is organized. A learner may perform better with assistance and also learn more. A learner may perform better while retaining little of the relevant capacity. A learner may initially perform more slowly under a demanding scaffold while developing stronger transfer or error-recognition capacity. The symbolic achievement observed during the activity therefore cannot be treated as a transparent measure of the educational trajectory.

Working Claim 10.5 (Mediation-sensitive learning interpretation). The educational interpretation of an AI-assisted performance should depend on the relation between assistance design and the target of learning. Performance gains supply evidence about task accomplishment; claims about learner formation require additional evidence concerning retention, transfer, reconstruction, or later judgment.

This distinction matters because many educational institutions historically used the same activity for two purposes: the activity was intended both to produce learning and to demonstrate learning. Generative assistance can separate these functions. An AI tutor may be highly effective for learning while making the resulting practice artifact weak evidence of independent mastery. Conversely, a tightly restricted assessment may improve evidential control while providing less pedagogical value. The tension is structural rather than a temporary enforcement problem.

10.3 Assessment Constructs under Changing Production Conditions

Educational measurement already provides a vocabulary for this problem. Kane’s validation framework emphasizes that score meaning depends on a chain of inferences from observed responses to claims about broader performance and use (Kane 2013). Generative mediation changes some of the assumptions inside that chain. If an essay score was previously interpreted as evidence of an individual’s capacity to research, structure, formulate, and express an argument under a given set of conditions, extensive delegation of some of those functions changes the construct represented by the resulting score unless the intended construct explicitly includes assisted performance.

Definition 10.6 (Assessment-construct drift). *Assessment-construct drift* occurs when the production conditions of an assessment change sufficiently that an established score or artifact increasingly represents a different configuration of capacities from the one that its users continue to attribute to it.

Construct drift can arise without deception. A course may openly permit generative tools while continuing to use a traditional essay rubric. The rubric may still assess coherence, evidence, structure, and argument quality accurately at the artifact level. Yet the educational interpretation of that score can change if the institution continues to treat it as evidence of capacities that have been materially delegated. The problem concerns the gap between *artifact scoring* and *person-level inference*.

Messick’s treatment of performance-assessment validity is relevant because it links score interpretation to both evidential support and the consequences of assessment use (Messick 1995). Under generative abundance, a familiar assessment format can preserve surface continuity while its inferential assumptions change. This creates a form of institutional temporal mismatch: the task and grading vocabulary remain recognizable even as the generative processes through which students reach those performances diversify.

Proposition 10.7 (Assessment-inference instability). *Where assistance conditions materially affect the relation between observed performance and the target educational construct, stable grading criteria can coexist with unstable person-level interpretations. Assessment continuity at the surface therefore does not guarantee continuity in what the resulting achievement symbol evidences.*

This proposition helps distinguish the symbolic-crisis perspective from a narrow cheating framework. Unauthorized substitution is one issue. A broader issue remains even under authorized use: institutions must decide which capabilities are educational targets, which can be delegated, and which forms of mediated performance count as evidence of learner development. Those decisions become increasingly visible as generative mediation expands.

10.4 Source Detection and the Limits of Achievement Verification

One possible response to inferential instability is to identify whether AI contributed to an educational artifact. Yet source identification supplies only part of the information needed. Fleckenstein and colleagues

found that both novice and experienced teachers had substantial difficulty identifying AI-generated essays among student-written texts and were overconfident in their judgments (Fleckenstein et al. 2024). This result reinforces the source-monitoring problem discussed in Section 4: symbolic form can become weak evidence of its production history.

Even perfect source detection would leave a deeper educational ambiguity. Knowing that an AI system contributed to a text does not specify what the learner understood, which parts were delegated, whether the system provided hints or solutions, how extensively the learner evaluated the output, or whether the interaction strengthened later competence. The relevant educational variable is therefore richer than binary source attribution.

Let the mediation profile of an educational performance be written in Equation 41.

$$M^{\text{edu}} = (g, h, e, r, v), \quad (41)$$

In Equation 41, g represents generation supplied by the system, h pedagogical hinting or scaffolding, e learner evaluation of intermediate outputs, r revision or reconstruction by the learner, and v later verification through transfer or independent performance. The vector is illustrative rather than a standardized metric. Its purpose is to show why the category “AI used” compresses educationally different processes.

Definition 10.8 (Educational mediation profile). An *educational mediation profile* is a structured description of how technical assistance participates in generation, scaffolding, evaluation, revision, and subsequent verification during an educational task.

The concept parallels the agency topology developed in Section 6. In education, however, the normative and epistemic focus shifts toward formation. Two learners can submit equally strong artifacts with the same declared AI tool while engaging in very different cognitive and pedagogical processes. Source disclosure improves legibility, yet the educational meaning of the achievement still depends on the topology of participation.

10.5 Symbolic Achievement under Symbol Inflation

The educational dimension of the symbolic crisis can now be stated more precisely. Modern education relies heavily on symbolic achievements because direct observation of learner formation is expensive. Essays, examinations, presentations, grades, and projects compress extended learning histories into portable representations. Their value partly depends on a socially stabilized mapping from performance conditions to educational inference.

Generative abundance places pressure on this mapping in three ways. First, it broadens the set of production pathways capable of reaching similar achievement surfaces. Second, it makes the relation between assisted task performance and durable learning more dependent on mediation design. Third, it increases the difficulty of inferring participation and understanding from the artifact alone. These changes do not imply the end of educational assessment. They create pressure to separate more carefully the quality of an artifact, the performance of a learner under specified conditions, and the longer trajectory of learner formation.

Working Claim 10.9 (Symbolic-achievement instability). Under generative abundance, the same educational symbol can increasingly compress heterogeneous learning processes. The resulting crisis lies in the persistence of inherited person-level inferences after the relation between symbolic achievement and learner formation has become more conditional on mediation, pedagogy, and verification context.

The diagnosis remains compatible with extensive beneficial use of generative systems. Kestin and colleagues demonstrate that carefully designed AI tutoring can support substantial learning gains in a specific educational setting (Kestin et al. 2025); Bastani and colleagues show that different forms of generative assistance can produce sharply different relations between assisted performance and subsequent

unaided learning (Bastani et al. 2025). The symbolic-crisis framework therefore gives no general reason to prefer unaided production. It instead directs attention to the inferential architecture through which educational institutions translate performances into claims about persons.

This architecture also provides the bridge to the next section. Educational institutions rarely stop at interpreting one assignment. They aggregate performances into grades, degrees, rankings, certificates, and reputational signals that travel beyond the original learning context. The next section therefore turns from symbolic achievement to *credentials and institutional compression*, where local uncertainties about performance become embedded in portable institutional representations of competence.

11 Credentials and Institutional Compression

The preceding section examined the instability of educational achievement surfaces. The present section moves one institutional layer upward. Educational systems rarely ask employers, professions, public agencies, or other organizations to inspect the full history of a learner's formation. They issue credentials that compress that history into portable institutional signs: degrees, licences, certificates, professional titles, grades, honours, and affiliations. These signs coordinate decisions across social distance because downstream actors can often rely on the issuing institution instead of reconstructing every assessment event themselves. The role of this section is to identify how generative abundance affects that compression mechanism. It proceeds in five steps. First, it distinguishes a credential surface from the upstream processes it summarizes. Second, it separates institutional issuance from the underlying capabilities that recipients may infer. Third, it develops a credential-dependence paradox in which weaker diagnosticity of isolated outputs can increase reliance on credentials even while the meaning of those credentials becomes less stable. Fourth, it examines how institutional prestige and verification capacity may acquire greater relative importance under symbolic abundance. Finally, it shows how ambiguity can be inherited through institutional compression rather than removed by it.

The starting point is familiar in both economics and sociology. Spence's signaling model treats education as a signal used under conditions in which employers cannot directly observe productivity (Spence 1973). Arrow's filtering account similarly emphasizes the possibility that higher education communicates information about ability rather than functioning only through the production of human capital (Arrow 1973). Collins's analysis of credentialism places greater weight on educational expansion, status competition, occupational closure, and the social power attached to formally recognized qualifications (R. Collins 2019). These perspectives differ substantially in causal emphasis, yet each treats credentials as more than a descriptive record of classroom participation. Credentials mediate allocation because they allow downstream institutions to act on compressed information about persons, training histories, selection processes, or institutional standing.

Definition 11.1 (Credential compression). *Credential compression* is the institutional process through which a distributed history of selection, participation, assessment, completion, and authorization is represented by a portable symbolic qualification that can support downstream decisions without reproducing the full upstream history.

Credential compression is therefore a special case of the social compression developed in Section 2. Its distinctive feature is institutional authorization. A degree is not merely a sentence asserting that a person possesses competence. It is issued through an organization with admission rules, curricula, assessment procedures, completion thresholds, record systems, and reputational consequences. A professional licence similarly condenses examinations, supervised practice, continuing requirements, or regulatory authority into a compact sign that can travel across contexts. Institutionalization can increase the reliability of a symbol because the issuing body bears some responsibility for maintaining the procedures that give the sign its social meaning.

Table 9: Credential surfaces and the upstream histories they compress. The final column identifies the type of ambiguity that generative mediation can introduce without implying that the credential thereby loses all value.

Credential surface	Compressed upstream history	Typical downstream inference	Generative-era ambiguity
Degree	admission, coursework, assessment, completion, institutional authorization	broad preparation, persistence, field exposure, threshold attainment	assessed artifacts may involve heterogeneous forms of AI mediation across courses and years
Professional licence	formal training, examination, supervised practice, regulatory approval	minimum professional competence and permission to practise	the relation between assessed performance and independent capability may vary with changing assessment conditions
Certificate	course participation, competency test, provider verification	bounded skill or knowledge claim	low-friction production can expand certificates while widening variation in assessment rigor and provider credibility
Grade or honours classification	repeated evaluation within an institutional regime	relative achievement under specified standards	comparability depends on mediation rules, task design, and institutional adaptation to generative tools
Institutional affiliation	selection, appointment, membership, continuing association	access to organizational screening, resources, or recognized standing	affiliation may carry increasing evidential weight even when the individual’s current contribution remains difficult to infer
Award or distinction	nomination, comparison, adjudication, institutional recognition	exceptional achievement or standing within a reference group	the rewarded artifact may be easier to produce while the adjudication process acquires greater importance

Table 9 separates the credential surface from its compressed history. The same qualification can support several downstream interpretations at once. A degree may be read as evidence of skill acquisition, as a screening signal, as evidence of persistence, as proof of institutional selection, or as a legally relevant prerequisite. These interpretations need not move together. A technological change that weakens one inference can leave another relatively intact. Generative assistance may complicate what a degree indicates about unaided writing while leaving its evidence of programme completion, institutional affiliation, or long-term participation largely unchanged. The symbolic-crisis framework therefore treats credential meaning as multidimensional rather than assuming that a credential either remains valid or becomes empty.

11.1 Credential Surfaces and Institutional Histories

Let K_i denote a credential held by person i . Let H_i^{inst} denote the institutional history compressed into that credential, including admission, participation, assessment, progression, and completion. Let L_i represent the underlying capability profile that downstream actors may wish to infer. Let M_i represent the mediation conditions under which assessed performances were produced. A downstream decision

maker observes the credential-related information collected in Equation 42.

$$K_i, I(K_i), C_t, \quad (42)$$

In Equation 42, $I(K_i)$ denotes the identity and perceived reliability of the issuing institution and C_t denotes the current decision context. The resulting inference is represented in Equation 43.

$$K_i, I(K_i), C_t \rightsquigarrow \hat{L}_i. \quad (43)$$

The institutional history H_i^{inst} remains only partially visible. Credential compression is useful precisely because downstream actors generally lack the time or access required to inspect that history directly.

This compression creates substantial coordination advantages. Neugebauer, Heisig, and Bol’s recent factorial survey experiment with German employers found that otherwise similar university graduates received markedly higher interview probabilities than non-completers, while employers associated graduation particularly with occupation-specific and non-cognitive skills (Neugebauer et al. 2026). The study does not establish one universal interpretation of degrees, yet it illustrates that credential completion can exert an independent effect even when employers receive other applicant information. The credential is therefore not simply redundant with the visible performance record.

Proposition 11.2 (Institutional compression value). *A credential can retain substantial coordination value even when the symbolic achievements from which it was partly constructed become less individually diagnostic, because the credential also compresses institutional selection, repeated assessment, completion, and authorization.*

This proposition is deliberately limited. Institutional compression does not guarantee accurate inference. It describes why credentials can remain socially powerful under conditions that destabilize other symbolic surfaces. Institutions aggregate across time and often impose multiple thresholds. A degree typically reflects more than one essay, one examination, or one presentation. Repetition can reduce dependence on any single artifact. At the same time, if the underlying assessment regime changes systematically, repeated compression can reproduce the same ambiguity at larger scale rather than eliminate it.

11.2 Institutional Diagnosticity

The diagnosticity framework from Section 4 can be extended from individual artifacts to institutions. Institutional diagnosticity is denoted in Equation 44.

$$D_{\text{inst}}(K; L, C) \quad (44)$$

Equation 44 denotes the diagnostic value of credential K for capability property L in context C . The quantity is conceptual rather than presumed directly measurable. It draws attention to three distinct relations: the quality of the issuing process, the stability of the inference from that process to the capability of interest, and the relevance of that capability to the downstream decision.

Definition 11.3 (Institutional diagnosticity). *Institutional diagnosticity is the extent to which an institutionally issued credential provides reliable evidence about a specified property because the issuing process constrains, aggregates, or verifies the histories capable of producing that credential.*

Institutional diagnosticity can be high for one property and modest for another. A medical licence may be strongly diagnostic of having satisfied formal regulatory requirements while providing less information about bedside communication. A university degree may be strongly diagnostic of sustained participation in a recognized programme while providing variable evidence about a particular software

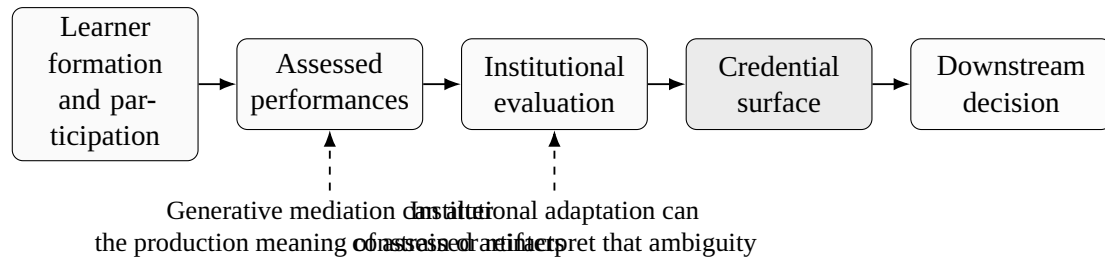


Figure 11: Institutional compression chain. Credentials can stabilize inference by aggregating and authorizing upstream processes, yet ambiguity introduced at the performance layer can propagate into the credential unless assessment and interpretation adapt.

skill. A certificate may be informative about completion of a bounded course yet weak evidence of long-term transfer. The appropriate interpretation therefore remains property-specific.

Generative systems affect institutional diagnosticity primarily through the assessment chain. If an institution continues to treat essays, reports, code, presentations, or take-home exercises as evidence of individual capacities while the production pathways for those artifacts become more heterogeneous, then the credential can inherit uncertainty from the upstream assessments. The problem is illustrated in Figure 11.

Figure 11 shows why the credential cannot be analysed independently of its production infrastructure. A degree remains a compact sign, but the social meaning of that sign depends on how institutions respond to changes in the diagnosticity of the performances they assess. Institutional adaptation can include redesigned assessment, oral verification, supervised practice, process documentation, longitudinal evaluation, or explicit recognition of mediated work. The present paper does not prescribe a preferred response. Its diagnostic claim is that the credential’s evidential meaning is partly endogenous to institutional adaptation.

11.3 Credential–Performance Divergence

Credentials can also diverge from current performance in the opposite direction. A person may possess a strong credential while performing poorly on a particular task, or lack a conventional credential while demonstrating high capability. Such divergence existed long before generative AI. Collins’s credentialist analysis emphasizes that credentials can function in social closure and status allocation in ways that are only partially reducible to productive skill (R. Collins 2019). Arrow’s filtering account similarly permits credentials to operate through information about underlying ability rather than direct skill formation (Arrow 1973). Generative abundance adds another layer by changing the cost and interpretability of the performances against which credentials might otherwise be checked.

Definition 11.4 (Credential–performance divergence). *Credential–performance divergence* is a condition in which the institutional status represented by a credential and the observable performance of its holder provide materially different evidence about the property relevant to a downstream decision.

The divergence can run in several directions. A highly credentialed actor can generate mediocre work. An uncredentialed actor can produce excellent work. A weakly credentialed actor can use generative systems to produce a surface equivalent to that of a highly trained actor. A highly trained actor can use the same systems to extend capability while preserving substantial human judgment. These cases weaken any simple mapping among credential, artifact quality, and competence.

The resulting environment may increase the importance of institutional metadata even when direct performances remain visible. If two applicants present similarly polished portfolios, decision makers may assign more weight to degrees, licences, institutional affiliations, verified work histories, or other

signals whose issuance is harder to replicate through one-shot generation. This is one reason symbolic abundance need not lead to dec credentialization. It can produce the opposite tendency in some domains.

Working Claim 11.5 (Credential-dependence paradox). When isolated symbolic performances become less diagnostic of underlying capability, downstream actors may increase reliance on institutionally issued credentials as lower-cost verification devices. If the credentialing institutions themselves continue to rely on increasingly ambiguous performance surfaces, greater dependence on credentials can coexist with growing uncertainty about what those credentials substantively certify.

This paradox is central to the symbolic crisis. The social response to weakening individual symbols may be stronger institutional compression. Yet institutional compression is only as informative as the processes it compresses and the properties it was designed to represent. The result can be a temporary period in which credentials become more valuable as coordination devices while becoming less transparent as capability indicators.

11.4 Verification Premiums and Institutional Stratification

The credential-dependence paradox has distributive implications. Institutions differ in resources, reputational capital, selectivity, assessment intensity, regulatory standing, alumni networks, and capacity to make their procedures credible to outsiders. These differences already matter in labour and academic markets. Headworth and Freese, for example, document a strong relationship between doctoral institution and placement in the U.S. academic sociology market, while also finding that institutional differences operate together with publications and awards rather than through prestige alone (Headworth and Freese 2016). The finding illustrates a broader point: institutional identity can become part of the evidential package through which individual capability is inferred.

Under generative abundance, this institutional layer may acquire what can be called a verification premium.

Definition 11.6 (Institutional verification premium). An *institutional verification premium* is the additional evidential or allocative value attached to a credential because the issuing institution is believed to possess credible procedures for selection, assessment, authorization, and accountability under conditions in which direct symbolic performance is comparatively easy to produce.

The concept does not imply that high-status institutions are necessarily more accurate. A verification premium is a social valuation attached to perceived process credibility. It can therefore reflect genuine procedural quality, inherited prestige, network advantage, exclusion, or some combination of these. This distinction matters because a symbolic crisis can redistribute trust toward institutions without automatically improving the justice or accuracy of allocation.

One possible consequence is intensified institutional stratification. If employers, funders, publishers, or professional bodies become less confident in isolated demonstrations of competence, they may lean more heavily on institutional provenance. Actors with access to highly trusted credentialing institutions then receive a lower verification burden, while others may face repeated demands to demonstrate competence directly. The symbolic crisis can therefore redistribute the cost of proving credibility.

This possibility should be held together with evidence on credential inflation. Collins's account emphasizes cycles in which expanding educational participation can produce rising credential requirements and declining distinctiveness of common qualifications (R. Collins 2019). More recent research similarly finds that the value of credentials varies with credential level, skill, and the expansion of tertiary education rather than moving uniformly (Araki and Kariya 2022). Generative abundance adds another axis to this problem: the institutional value of a credential may depend increasingly on whether outsiders trust the process by which the institution evaluates work produced under new mediation conditions.

Proposition 11.7 (Verification-burden redistribution). *If direct performance becomes cheaper to produce while institutional verification remains scarce, the burden of establishing credibility can shift toward access to trusted credentialing processes. This shift can increase the allocative importance of institutional position even when individual productive capability becomes more widely distributed.*

The proposition is a hypothesis about institutional response rather than a forecast that applies uniformly across sectors. Some domains may move toward skills-based assessment, open portfolios, repeated work trials, or other forms of direct verification. Others may strengthen licensing, credential requirements, or institutional provenance. The symbolic-crisis framework predicts heterogeneity because the relative costs of direct evaluation and institutional reliance differ across domains.

11.5 Inherited Ambiguity in Credential Chains

The deepest issue is that institutional compression can conceal uncertainty as effectively as it compresses information. Downstream users usually see the final credential rather than the sequence of assessments, mediation rules, exceptions, accommodations, revisions, and institutional judgments that produced it. This is precisely why credentials are useful: their portability depends on leaving much of the upstream history unobserved. Yet when production conditions change rapidly, that same compression can slow social adaptation because inherited meanings remain attached to a stable credential surface.

Definition 11.8 (Inherited credential ambiguity). *Inherited credential ambiguity* is the uncertainty transmitted into a credential when the institution continues to compress upstream performances whose relation to the capability of interest has changed faster than the credential's public interpretation.

Inherited ambiguity can persist even when every participant acts legitimately. A student may use permitted generative assistance. An instructor may assess according to stated rules. A university may award the degree according to established standards. An employer may then interpret the degree using an older model of what the assessed work required. No single act needs to be fraudulent for the downstream inference to become miscalibrated. The symbolic crisis can therefore emerge through temporal mismatch among production practice, institutional assessment, and social interpretation.

This point also clarifies why credential verification alone cannot resolve the broader crisis. Cryptographic verification can establish that a credential was issued by a particular institution. It cannot by itself determine what capability the credential should evidence, whether the assessment construct remains stable, or how much of the upstream work was delegated. Verification secures the authenticity of the institutional sign; it does not automatically secure the validity of every inference drawn from it.

Proposition 11.9 (Credential-authenticity distinction). *The authenticity of a credential and the validity of the capability inference drawn from that credential are analytically distinct. Stronger verification can secure issuance while leaving substantive interpretation dependent on evolving institutional practices and assessment validity.*

The distinction parallels the earlier separation between disclosure, agency legibility, authenticity, and legitimacy. A credential can be genuine while being interpreted too broadly. It can be institutionally legitimate while providing only weak evidence about a specific capability. It can remain highly useful for coordination while compressing histories that have become more heterogeneous than downstream users realize.

The symbolic crisis therefore reaches beyond individual authorship or educational misconduct. It enters the architecture through which modern societies distribute trust across institutions. Credentials remain powerful precisely because they reduce verification costs. Generative abundance can increase the demand for that reduction while simultaneously making the compressed content harder to interpret. The next section turns to a different institutional domain—culture and luxury—where symbolic scarcity itself participates directly in value formation and where generative abundance can displace scarcity toward provenance, access, participation, and relational exclusivity.

12 Culture, Luxury, and Scarcity Displacement

The preceding sections examined the evidential and institutional consequences of generative abundance. The present section turns to a domain in which scarcity participates directly in value formation: culture, luxury, and symbolic distinction. Its role is diagnostic. It asks how expanding capacity to reproduce styles, narratives, images, motifs, and branded visual surfaces changes the location of scarcity within cultural markets. The section proceeds in six steps. First, it treats scarcity as a multidimensional profile rather than as a single property of an object. Second, it develops the concept of *scarcity displacement*, in which value-bearing constraint can migrate from symbolic form toward provenance, materiality, process, access, participation, temporality, and recognition. Third, it examines the relation between cultural distinction and reproducible style. Fourth, it considers biography and provenance as resources for singularization. Fifth, it analyzes relational and experiential scarcity. Finally, it uses recent findings on AI-mediated fashion and luxury advertising to show that generative abundance can reweight authenticity and effort cues without producing one uniform consumer response.

Luxury provides a useful analytical site because its value has long involved more than functional utility. Veblen’s account of conspicuous consumption emphasized the social visibility of expenditure and status differentiation (Veblen 1899). Bourdieu later placed taste and cultural distinction within broader relations of social position and cultural capital (Bourdieu 1984). Contemporary luxury research continues to identify rarity and exclusivity as important, while also distinguishing several forms of rarity and several kinds of perceived value. Wang, Sung, and Phau, for example, find that perceived natural rarity predicts functional, social, and emotional value across several luxury categories, while virtual rarity and exclusivity show more uneven effects (Wang et al. 2024). These literatures already suggest that scarcity is heterogeneous. Generative systems intensify the importance of that heterogeneity because they can sharply reduce the scarcity of some symbolic forms while leaving other constraints largely intact.

Definition 12.1 (Scarcity profile). For a cultural good, practice, or experience z , a *scarcity profile* is the vector of constraints that limit access to, production of, participation in, or recognized possession of z across dimensions such as symbolic form, material resources, skilled process, provenance, place, time, social access, and institutional recognition.

A schematic profile is stated in Equation 45.

$$\mathbf{S}(z, t) = (S_{\text{form}}, S_{\text{material}}, S_{\text{process}}, S_{\text{provenance}}, S_{\text{access}}, S_{\text{participation}}, S_{\text{temporal}}, S_{\text{recognition}})_t. \quad (45)$$

The components need not be independent. A rare material may require a specialized process; access to a private event may depend on institutional recognition; provenance may acquire force only when a community recognizes the history as significant. The profile therefore represents a relational configuration rather than a list of intrinsic object properties.

Table 10 clarifies why the phrase *symbolic abundance* should not be generalized into a claim of universal abundance. A generative model can make a plausible image of a hand-finished object abundant without making the relevant material, workshop capacity, craft lineage, or ownership relation abundant. It can produce imagery associated with an invitation-only event without multiplying seats in the room. It can generate a stylistically persuasive representation of a place without creating additional physical access to that place. Cultural value can therefore reorganize around the dimensions that remain constrained.

12.1 Scarcity Displacement

The concept needed here concerns relative movement in the value-bearing role of scarcity. Let $V(z, t)$ denote the socially perceived value of some cultural object or experience. A simple decomposition would

Table 10: Dimensions of cultural scarcity and their possible reconfiguration under generative abundance. The final column states a pressure rather than a predicted outcome.

Scarcity dimension	Typical source of constraint	Pressure under generative abundance
Symbolic form	difficulty of producing a distinctive image, style, text, or narrative	imitation, variation, and stylistic recombination become cheaper and faster
Material	rare ingredients, physical resources, location-specific inputs, limited capacity	comparatively unchanged by symbolic generation, though representation may become easier to imitate
Process	craft skill, time, embodied practice, specialized coordination	process may remain scarce even when a convincing surface representation becomes abundant
Provenance	continuous history of origin, ownership, participation, or creation	historical trace can acquire greater relative importance when surfaces are easier to reproduce
Access	restricted entry to spaces, events, institutions, or networks	access may become a stronger differentiator when public symbolic imitation becomes cheap
Participation	being present in, contributing to, or co-producing an event or relation	participation remains bounded by time, attention, and social capacity
Temporal	one-time occurrence, age, duration, sequence, irrecoverable moment	retrospective symbolic reconstruction cannot fully reproduce prior participation or temporal priority
Recognition	socially accepted status, membership, authorship, connoisseurship, or standing	recognized position can become more valuable while also becoming a site of intensified gatekeeping

be misleading because the dimensions interact, yet the schematic relation in Equation 46 remains useful.

$$V(z, t) = F(\mathbf{S}(z, t), A_t, R_t, C_t), \quad (46)$$

In Equation 45, A_t denotes aesthetic evaluation, R_t denotes relevant social relations and recognition, and C_t denotes the wider cultural and institutional context. Generative abundance can lower one component of $\mathbf{S}$, especially S_{form} , while leaving others relatively constrained. The resulting change in V depends on whether other scarcity dimensions can substitute for, complement, or become newly salient relative to the abundant symbolic form.

Definition 12.2 (Scarcity displacement). *Scarcity displacement* is a change in the relative location of value-bearing constraint when abundance in one dimension reduces its differentiating force and cultural or economic value becomes more strongly associated with other constrained dimensions of the scarcity profile.

Scarcity displacement is neither a conservation law nor a prediction that every lost scarcity will be replaced elsewhere. At least three outcomes are possible. Value may decline because a previously scarce form becomes commonplace. Value may reorganize around another constraint such as provenance or access. Value may also increase if abundant reproduction expands attention, participation, or recognition around a cultural form. The concept therefore describes a shift in the structure of scarcity, while leaving the direction and magnitude of value change open to empirical analysis.

Proposition 12.3 (Non-conservation of scarcity value). *A reduction in symbolic-form scarcity does not imply an equal increase in the value of provenance, access, participation, or any other dimension. Scarcity displacement identifies a possible reweighting of constrained dimensions rather than conservation of value across them.*

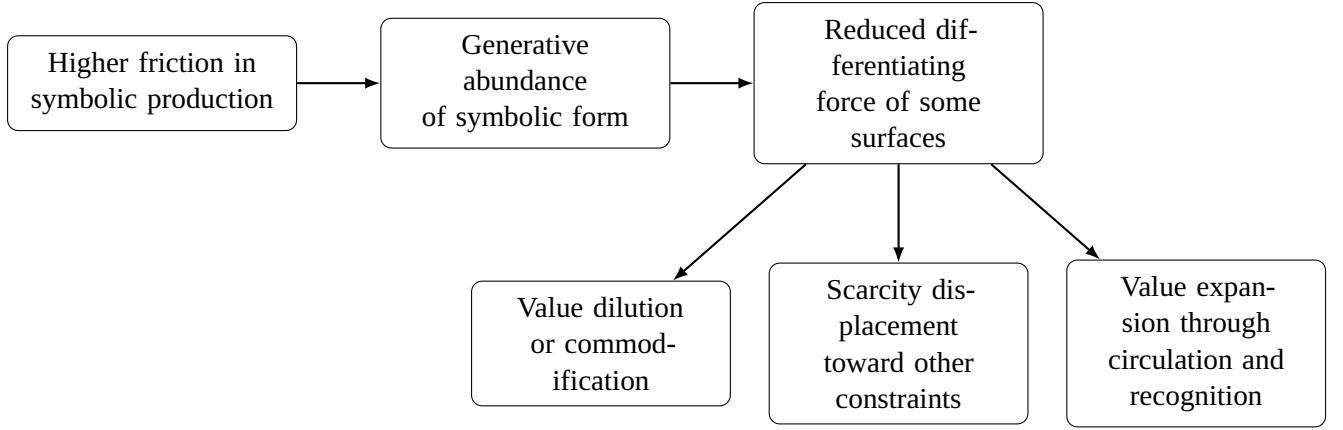


Figure 12: Generative abundance can reduce the differentiating force of some symbolic forms while producing several possible value trajectories. Scarcity displacement is one possibility among value dilution, reorganization, and value expansion.

This qualification matters because narratives of cultural change often move too quickly from technological abundance to a new presumed locus of authenticity. The framework developed here instead asks which dimensions remain difficult to reproduce, which communities recognize those dimensions as significant, and which institutions are capable of certifying or controlling access to them.

Figure 12 places scarcity displacement among several possible outcomes. The figure is intentionally non-teleological. Generative abundance can weaken a distinction, intensify attention to another distinction, or expand the cultural field in ways that create new forms of recognition. The argument concerns changing diagnostic and scarcity relations rather than a general law of cultural decline.

12.2 Reproducible Style and Cultural Distinction

Bourdieu's analysis is useful here because cultural distinction depends on socially organized judgment, classification, and competence rather than on the visual rarity of an isolated object alone (Bourdieu 1984). Generative systems can make stylistic surfaces more accessible. A person can request a visual language, textual register, interior aesthetic, or fashion concept that once required extensive technical competence or access to specialized production. This expansion can reduce the production cost of a recognizable style. It does not automatically reproduce the social relations through which that style acquires distinction.

This gap can be expressed by distinguishing *formal reproducibility* from *recognized distinction*. Formal reproducibility concerns the capacity to generate a surface with relevant perceptual or stylistic features. Recognized distinction concerns the social processes through which an object, practice, or person is accepted as possessing cultural significance, legitimacy, status, or connoisseurial standing. Under generative abundance, the contrast between formal reproduction and recognized distinction is represented in Equation 47.

$$C_{\text{formal reproduction}} \downarrow \quad \text{while} \quad C_{\text{recognized distinction}} \quad (47)$$

may remain high, move elsewhere, or change institutional form.

Definition 12.4 (Distinction displacement). *Distinction displacement* is the movement of socially consequential differentiation away from easily reproducible symbolic surfaces toward less reproducible relations such as provenance, access, recognized participation, institutional standing, or historically accumulated judgment.

Distinction displacement can increase inequality as well as reduce it. Lower-cost stylistic production can democratize aesthetic experimentation and weaken monopolies over some forms of expression.

At the same time, if distinction relocates toward invitation, membership, private access, recognized networks, or expensive embodied experience, the new scarcity regime may become more relationally exclusive. The relevant political-economic question is therefore not whether symbolic abundance democratizes culture in the abstract. It is how it changes the distribution of access to the dimensions that continue to carry recognized value.

12.3 Provenance, Biography, and Singularization

Kopytoff’s analysis of the cultural biography of things provides a useful bridge from commodity exchange to historically situated value. His account emphasizes that commoditization and singularization are culturally organized processes and that the social status of a thing can change across its biography (Kopytoff 1986). This perspective becomes especially important when symbolic surfaces are cheap to reproduce. Two objects can become visually similar while differing sharply in origin, sequence of ownership, conditions of production, relation to an event, or recognized place within a cultural history.

Generative abundance can therefore increase the relative salience of *biographical difference*. A generated image can reproduce the appearance of an archive object; it cannot retrospectively become the object that passed through the archive’s history. A newly manufactured artifact can imitate a visual pattern; it cannot acquire an earlier participation in the events through which another artifact became historically significant. These observations do not make provenance intrinsically valuable. They identify a dimension whose scarcity is generated by temporal sequence and relational history rather than by surface form alone.

Proposition 12.5 (Biographical asymmetry). *When symbolic form becomes highly reproducible, histories that depend on temporal sequence, prior participation, and externally anchored relations can retain differentiating force because later reproduction cannot retroactively reproduce the same biography.*

The proposition also has institutional consequences. Provenance systems, archives, authentication practices, and recognized chains of custody can acquire greater importance because they connect present objects to histories that surface inspection alone cannot establish. Yet provenance is itself symbolically represented through records, labels, certificates, databases, and testimony. It therefore remains vulnerable to the diagnostic problems developed earlier in this paper. Provenance does not stand outside the symbolic crisis; it becomes one field in which institutions attempt to stabilize a relation between present representation and historical trajectory.

12.4 Access, Participation, and Relational Scarcity

The scarcity of direct participation is structurally different from the scarcity of symbolic production. A system can generate millions of personalized representations while human time, attention, presence, and relational capacity remain bounded. Luxury and culture already make extensive use of access constraints: private previews, invitation-only events, limited seating, personal appointments, member spaces, curated networks, and geographically specific experiences. Recent luxury research treats rarity and exclusivity as distinct constructs and finds that naturally constrained rarity can carry strong value even when artificially induced exclusivity has more variable effects (Wang et al. 2024). The broader implication is that scarcity attached to real capacity constraints may remain socially consequential even when symbolic representations of exclusivity proliferate.

This produces what can be called *relational scarcity*. Let B_h denote human relational bandwidth: the finite amount of direct attention, reciprocal interaction, and co-presence that persons and institutions can sustain. Generative systems can increase symbolic output much faster than they increase B_h . The resulting relation is stated in Equation 48.

$$\frac{\text{symbolic personalization}}{B_h} \quad (48)$$

can rise dramatically. A personalized luxury narrative, invitation-style visual, or one-to-one brand voice may therefore become abundant while actual access to a designer, artisan, curator, community, or restricted event remains scarce.

Definition 12.6 (Relational scarcity). *Relational scarcity* is the constraint created by finite human attention, co-presence, reciprocal participation, and socially recognized access, especially when symbolic representations of those relations can be produced at much larger scale.

Relational scarcity helps explain why generative abundance may shift some value from possession toward participation. A representation of having attended an event can be produced cheaply; actual participation still requires access at a particular time and place. A generated account of collaboration can imitate its linguistic surface; sustained participation in the relevant relation remains bounded by attention and history. This does not imply that experience automatically becomes more valuable. It identifies a resource whose supply is governed by different constraints from symbolic production.

12.5 Effort, Craft, and Authenticity Cues

The relation between effort and value becomes particularly unstable under generative abundance. Traditional luxury often foregrounds craft time, rare materials, artisanal skill, and elaborate production narratives. Generative systems can reproduce the visual signs of such effort much more cheaply than the underlying processes themselves. The symbolic crisis appears when recipients continue to infer costly process from a surface whose production route has become heterogeneous.

Recent studies show that this issue is already visible in consumer response. Lee and Kim find that participants generally evaluated human-designed fashion more favorably than AI-designed fashion in their experiments, with perceived authenticity and expected quality mediating the difference; opportunities for consumer input reduced the negative response by increasing perceived authenticity (Lee and Kim 2024). To, Wu, Kianian, and Zhang report that disclosed AI-generated luxury advertising can be evaluated more negatively because it is perceived as requiring less effort, which in turn reduces perceived brand authenticity; highly creative imagery attenuated the effect in their studies (To et al. 2025). These findings are context-specific. They do not establish a permanent human-authenticity premium. They show that effort remains one of the latent histories consumers can infer from luxury surfaces and that disclosure of a lower-friction production path can alter that inference.

Definition 12.7 (Effort-surface decoupling). *Effort-surface decoupling* occurs when a symbolic form that historically functioned as evidence of substantial production effort becomes reachable through production paths with widely varying effort profiles.

Effort-surface decoupling creates several possible adaptations. Consumers may continue to value effort and demand stronger process evidence. Communities may place more weight on aesthetic outcome than on production difficulty. Institutions may differentiate between human craft, human-AI co-creation, and largely automated generation. New conventions may emerge in which some forms of AI use themselves become signs of technical or curatorial competence. The current empirical evidence is therefore best interpreted as evidence of an unsettled recalibration rather than a stable hierarchy of production modes.

12.6 Artificial Scarcity and the Limits of Displacement

Scarcity can also be strategically produced. Limited editions, timed releases, restricted memberships, artificial supply limits, and controlled distribution have long created scarcity without depending entirely on natural constraints. Generative abundance does not remove the organizational capacity to create such limits. It can even increase incentives to do so when symbolic forms are otherwise easy to reproduce.

The distinction between natural and virtual rarity in current luxury research is useful here. Wang, Sung, and Phau find substantially different effects across these forms, with natural rarity carrying broader positive associations in their data than virtual rarity (Wang et al. 2024). The result cautions against assuming that scarcity has a uniform value simply because it is scarce. Recipients interpret the source and legitimacy of the constraint. A deliberately limited digital item, a geographically constrained experience, a rare material, and an invitation-only social relation can all be scarce while carrying very different meanings.

This motivates a final distinction between *scarcity quantity* and *scarcity legitimacy*. Scarcity quantity concerns how difficult access or reproduction actually is. Scarcity legitimacy concerns whether the constraint is interpreted as meaningful, authentic, fair, manipulative, arbitrary, or culturally appropriate. Generative abundance can therefore produce a second-order contest over scarcity itself: institutions may seek to manufacture new limits precisely because old symbolic limits have weakened, while audiences evaluate whether those limits deserve recognition.

Proposition 12.8 (Scarcity-legitimacy distinction). *The economic or cultural effectiveness of a scarcity mechanism depends partly on how the source of the constraint is interpreted. A technically effective restriction can have weak or negative value effects when recipients regard the scarcity as artificial, manipulative, or culturally unconvincing.*

The section's broader conclusion is that culture and luxury should not be described through a simple transition from scarcity to abundance. Generative systems primarily expand abundance in particular representational and production dimensions. Other dimensions remain constrained, and some can become more salient precisely because surface imitation is easier. The relevant transformation is therefore a reorganization of scarcity profiles. Symbolic form, material process, provenance, access, participation, temporality, and recognition may be reweighted differently across domains.

This diagnosis also points toward the political economy developed in the next section. Scarcity displacement changes where rents, gatekeeping power, verification costs, and status advantages can accumulate. If symbolic forms become easier to produce while recognized provenance, private access, trusted institutions, and relational participation remain constrained, economic value may migrate toward actors who control those infrastructures. The symbolic crisis thus concerns not only interpretation. It can reorganize the distribution of value and power around the dimensions that remain difficult to reproduce.

13 Political Economy under Symbol Inflation

The preceding section argued that generative abundance can displace scarcity away from symbolic form toward provenance, access, participation, temporality, and recognition. The present section develops the political-economic implications of that displacement. Its role is diagnostic. It asks where value, rents, verification burdens, and gatekeeping power can accumulate when the marginal cost of producing plausible symbolic artifacts falls faster than the cost of attending to, ranking, verifying, and socially recognizing them. The section proceeds in seven steps. First, it distinguishes symbolic abundance from economic abundance. Second, it treats attention and verification as bottlenecks that can become more consequential when production expands. Third, it examines platforms and ranking infrastructures as institutions that organize access to scarce visibility. Fourth, it develops the concepts of *verification rent* and *credibility infrastructure*. Fifth, it analyzes the redistribution of labour from production toward curation, checking, and post-generation work. Sixth, it considers how symbolic abundance can intensify rather than dissolve hierarchy. Finally, it connects these dynamics to the next section's analysis of provenance and its limits.

The basic economic intuition is straightforward. Information and symbolic artifacts can be highly reproducible even while the resources needed to receive, interpret, compare, verify, and act upon them

Table 11: Political-economic bottlenecks under symbolic abundance. The final column identifies a possible locus of rent or gatekeeping rather than a universal prediction.

Bottleneck	Constraint under generative abundance	Possible locus of value or power
Attention	recipients retain finite time and cognitive capacity while symbolic supply expands	ranking, recommendation, curation, notification, and audience access
Verification	claims can be generated faster than institutions can check provenance, accuracy, competence, or authorization	auditing, authentication, fact-checking, trusted intermediaries, liability-bearing institutions
Visibility	abundant production competes for limited positions in feeds, search results, catalogues, journals, and professional attention	platforms, search and recommendation systems, editorial institutions, gatekeepers
Provenance	similar surfaces can correspond to heterogeneous generative histories	trusted records, identity infrastructure, archives, signatures, credentialing and provenance systems
Recognition	social standing and accepted authority remain relationally and institutionally bounded	brands, professional institutions, communities, membership systems, reputational networks
Relational bandwidth	direct attention, co-presence, mentorship, collaboration, and access remain finite	exclusive access, high-touch service, professional networks, communities, interpersonal trust
Responsibility	scalable generation can separate output production from accountable judgment	organizations willing and able to endorse, insure, review, correct, or legally stand behind outputs

remain constrained. Goldhaber formulated an early version of this problem in the language of the attention economy, arguing that information abundance increases the relative importance of scarce attention (Goldhaber 1997). Platform economics adds another layer: intermediaries can acquire power by structuring interactions between distinct groups and by determining the conditions under which users, advertisers, producers, or complementary services encounter one another (Rochet and Tirole 2003; Kenney and Zysman 2016). Generative systems intensify both dynamics because they can increase the volume, personalization, and apparent competence of symbolic production without proportionally expanding human attention or institutional verification capacity.

Definition 13.1 (Political economy of symbol inflation). The *political economy of symbol inflation* concerns configurations in which symbolic production and strategic symbolic engineering expand while economically consequential constraints increasingly concentrate in attention, ranking, authentication, provenance, access, verification, recognition, and relational participation.

The definition does not imply that production ceases to matter. Compute, data, intellectual labour, model access, energy, infrastructure, and specialized expertise remain costly. The claim concerns relative bottlenecks. As some classes of symbolic output become easier to generate, value can become increasingly sensitive to the institutions that select which outputs are seen, believed, certified, purchased, or treated as socially consequential.

Table 11 shows why a fall in production friction need not flatten economic hierarchy. Some bottlenecks move downstream. The ability to produce a plausible text may become widespread while the ability to place that text before a valuable audience, certify its reliability, connect it to a trusted identity, or obtain consequential institutional uptake remains scarce. The political economy of symbol inflation therefore concerns control over the remaining constraints as much as access to symbolic production itself.

13.1 Attention Scarcity and Symbolic Overproduction

Let Q_t denote the volume of potentially relevant symbolic artifacts available to an actor at time t , and let A_t denote the actor's effective attention capacity. A simple pressure indicator is defined in Equation 49.

$$\Lambda_t = \frac{Q_t}{A_t}. \quad (49)$$

The ratio is only schematic. Symbolic artifacts vary greatly in relevance and length, and attention is not homogeneous. Its purpose is to express a structural asymmetry: generative capacity can increase Q_t rapidly while A_t remains comparatively bounded. As Λ_t increases, the economic problem shifts from producing information toward deciding what deserves attention.

This shift has at least three consequences. First, ranking acquires greater importance. An artifact that is never surfaced can have little social effect regardless of its quality. Second, selection becomes a form of productive labour. Curation, filtering, reviewing, summarization, and recommendation become necessary precisely because generation is abundant. Third, attention itself can become more aggressively contested. Personalized symbolic production can target increasingly narrow audiences, increasing the supply of messages competing for the same finite cognitive bandwidth.

Definition 13.2 (Attention bottleneck). An *attention bottleneck* occurs when the feasible supply of symbolic artifacts expands faster than the capacity of relevant recipients to notice, compare, interpret, or evaluate them.

The attention bottleneck helps explain why lower production costs can coexist with higher expenditure on distribution, audience acquisition, recommendation, branding, or privileged placement. The economic scarcity has moved. The producer may spend less on generating a particular message and more on ensuring that the message enters a consequential attentional channel.

Proposition 13.3 (Downstream bottleneck displacement). *When symbolic production becomes less constraining while attention, verification, and recognized access remain limited, the relative economic importance of downstream selection and validation infrastructures can increase even if the average cost of producing an individual artifact falls.*

This proposition remains contingent. Some domains may instead experience broad value dilution. Others may develop public or decentralized filtering institutions that reduce concentration. The point is that abundance at one stage of a symbolic process can increase the strategic importance of constraints at another stage.

13.2 Ranking Infrastructures and Platform Power

Abundant symbolic production becomes socially consequential through infrastructures that order visibility. Search engines, recommendation systems, social platforms, marketplaces, journals, credential systems, and professional networks perform different functions, yet all can participate in selecting which artifacts receive scarce attention. Platform economics is useful here because it emphasizes that intermediaries do not merely connect users; they can shape the terms of interaction between different sides of a market (Rochet and Tirole 2003). Kenney and Zysman likewise describe digital platforms as structures that reorganize economic activity and the distribution of resulting profits (Kenney and Zysman 2016).

Generative abundance can deepen this dependence on ranking infrastructures. When supply expands, unranked abundance becomes difficult to navigate. The platform or institution that determines visibility can therefore become more important even when production is increasingly decentralized. A creator may possess abundant generative capacity while remaining dependent on a small number of channels for discovery, reputation, monetization, or institutional uptake.

Generative systems can deepen this role because platform infrastructures may increasingly combine several functions that were previously more separable: generation, distribution, ranking, identity management, monetization, and credibility signaling. The economic significance of a platform can therefore move beyond distribution alone toward partial control over the conditions under which synthetic content is produced, ordered, and treated as consequential.

Definition 13.4 (Visibility rent). A *visibility rent* is an economic advantage obtained from controlling or preferentially accessing a scarce channel through which abundant symbolic production becomes visible to consequential audiences.

Visibility rent can arise through advertising markets, ranking privileges, editorial access, subscriber networks, reputation, search placement, or institutional standing. It should not be understood as automatically illegitimate. Ranking is necessary when supply exceeds attention. The political-economic issue concerns who controls ranking, how criteria are set, whether alternative channels remain available, and how the costs of visibility are distributed.

13.3 Verification Markets and Credibility Infrastructure

A second bottleneck concerns verification. Generative systems can reduce the cost of producing a claim faster than the cost of establishing whether the claim is accurate, authorized, original, competent, or responsibly endorsed. The resulting asymmetry resembles a familiar problem in information economics. Akerlof's analysis of quality uncertainty showed how markets can be disrupted when buyers cannot reliably distinguish quality and when institutions capable of reducing uncertainty are absent or weak (Akerlof 1970). The present setting differs substantially, but the same general insight remains relevant: uncertainty about hidden quality or provenance can change market structure and create demand for counteracting institutions.

In symbolic-abundance environments, verification can itself become a market. Actors may pay for authenticated identity, trusted certification, professional review, provenance records, reputation systems, warranties, insurance, auditing, or institutional endorsement. The value of these services derives partly from their capacity to reconnect a symbolic surface to a history, competence, process, or responsibility structure that the surface alone no longer establishes.

Definition 13.5 (Verification rent). A *verification rent* is value captured by an actor or institution because it controls a trusted and scarce capacity to reduce uncertainty about the history, quality, authorization, or responsibility associated with an abundant symbolic artifact.

Verification rent can be socially productive when it funds costly checking, accountability, or maintenance of trustworthy records. It can also become extractive when verification is monopolized, access is expensive, standards are opaque, or the verifying institution uses its position to exclude competitors or subordinate groups. The concept therefore describes a possible source of value and power rather than a normative judgment.

Figure 13 summarizes the central displacement. The figure should not be read as implying that production has become costless or that all value migrates downstream. It identifies a change in relative constraint. In some sectors, model providers can themselves capture substantial rents through control of compute, data, or proprietary systems. In others, open or commoditized generation can make distribution and verification more important. The relevant empirical question is therefore where the binding bottlenecks lie in a particular domain.

13.4 Trust Arbitrage and Verification Lag

Verification does not necessarily adjust immediately when fabrication becomes cheaper. A recent economic model develops the concept of *trust arbitrage*: when generative AI lowers fabrication, targeting,

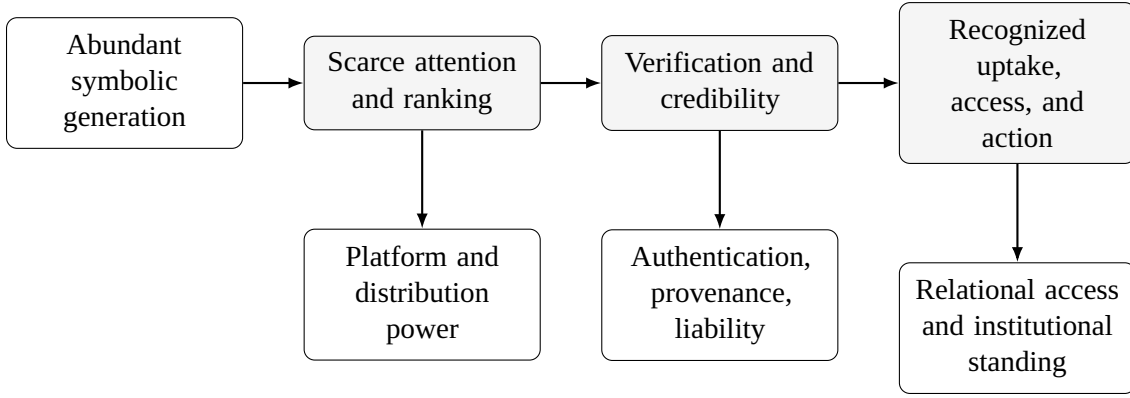


Figure 13: A schematic political economy of symbolic abundance. Falling production friction can shift relative bottlenecks toward attention, ranking, verification, credibility, recognized access, and institutional uptake. The figure represents possible sites of value concentration rather than a deterministic sequence.

and verification costs at different rates, high-trust environments can temporarily become attractive to deception because established trust delays verification (Yin and Wen 2026). The model is narrow, but it captures a broader temporal possibility that fits the symbolic-crisis framework. Institutional expectations formed under earlier production conditions can persist after the cost structure of symbolic fabrication has changed.

This lag creates a political-economic interval in which actors can capture value from inherited credibility before verification practices adapt. The relevant resource is not only the capacity to fabricate. It is access to environments in which recipients continue to grant relatively high prior trust. The symbolic crisis can therefore create temporary rents from exploiting old interpretive conventions.

Definition 13.6 (Interpretive lag rent). An *interpretive lag rent* is value captured when an actor benefits from a persistent social inference or trust convention after the production conditions that originally made that inference reliable have changed.

Interpretive lag rent is structurally temporary because widespread exploitation can itself trigger adaptation. Yet the transition may be prolonged when verification is costly, institutions update slowly, or recipients face coordination problems. The concept links the paper’s earlier temporal-mismatch thesis to economic incentives: outdated interpretive conventions can become exploitable assets.

13.5 Post-Generation Labour and Burden Redistribution

Symbolic abundance can also redistribute labour rather than simply eliminate it. Time saved in first-draft production can reappear elsewhere as checking, prompt construction, source verification, integration, editing, attribution, risk review, formatting, curation, and responsibility-taking. The underlying mechanism is broader than any one sector: lower generation cost can increase the amount of downstream work required to make an output trustworthy, context-appropriate, or institutionally usable.

Let L_g denote labour devoted to initial generation and L_v labour devoted to verification, curation, repair, and responsibility. A technology can reduce L_g while increasing or redistributing L_v . The net effect depends on task structure, error costs, quality expectations, and who bears each component. Productivity gains therefore do not by themselves establish how labour burdens are distributed across workers, recipients, reviewers, institutions, and consumers.

Proposition 13.7 (Verification-burden redistribution). A reduction in symbolic production labour can coexist with an increase in verification, curation, and accountability labour, and these burdens can be transferred to actors who did not receive the original productivity gain.

This proposition is especially important for institutions. A firm that automates customer communication may save drafting time while customers spend more effort determining whether a message reflects an actual human decision. A publisher may lower content-production costs while editors and readers face higher screening burdens. An educational institution may permit extensive AI assistance while assessment staff incur additional verification costs. Symbolic abundance therefore has incidence effects: the location of saved labour and the location of new labour can differ.

13.6 Abundance, Hierarchy, and Infrastructural Concentration

Generative abundance is sometimes associated with democratization because more actors can produce polished text, images, code, or analysis. That possibility is significant. Lower production barriers can expand participation, reduce dependence on specialized intermediaries, and allow individuals or small organizations to create outputs previously requiring larger teams. Yet lower production barriers do not imply equal access to attention, trusted distribution, high-quality models, compute, proprietary data, recognized credentials, legal protection, or influential networks.

The resulting political economy can combine broad productive capacity with concentrated infrastructural power. Many actors may be able to generate, while comparatively few actors control the systems that rank, authenticate, distribute, monetize, or legitimate what is generated. This produces a distinction between *production democratization* and *consequence democratization*. The first concerns the capacity to create a symbolic artifact. The second concerns the capacity to make that artifact socially consequential.

Definition 13.8 (Consequence asymmetry). *Consequence asymmetry* is the unequal capacity of actors with comparable symbolic production ability to obtain attention, verification, recognition, institutional uptake, or material effects from what they produce.

Consequence asymmetry explains why symbolic abundance can coexist with persistent or intensified hierarchy. If the scarce resources move from creation toward distribution and recognition, actors with existing institutional position may retain substantial advantages. Established brands can attach trusted provenance to abundant content. Platforms can control visibility. Professional institutions can certify competence. Wealthy organizations can purchase human review, legal responsibility, and premium relational access. The new abundance can therefore change the basis of hierarchy without necessarily reducing hierarchy itself.

Working Claim 13.9 (Infrastructural concentration under symbolic abundance). Where symbolic generation becomes widely accessible while ranking, verification, recognition, and consequential access remain concentrated, economic power can migrate toward the infrastructures that organize these downstream bottlenecks.

The claim is conditional and should be tested sector by sector. Some decentralized systems may weaken gatekeeping. Open standards can reduce dependence on a single verifier. Public institutions can supply trusted infrastructure without direct commercial rent extraction. Competitive markets can also constrain platform power. The purpose of the claim is to identify a plausible political-economic tendency, not to treat concentration as inevitable.

13.7 Credibility Infrastructure and the Distribution of Trust Costs

The concepts developed above can be combined under the heading of *credibility infrastructure*. Credibility infrastructure includes the institutional and technical arrangements through which claims become connected to identities, histories, standards, liabilities, reputations, or verification procedures. Examples include professional licensing, editorial review, cryptographic provenance, authenticated identity, brand warranties, audit systems, archives, and legal responsibility.

Definition 13.10 (Credibility infrastructure). *Credibility infrastructure* is the set of institutional, technical, and relational arrangements that reduce uncertainty about whether a symbolic artifact can be connected to a sufficiently reliable history, agent, process, or responsibility structure for a particular use.

Credibility infrastructure can lower transaction costs by allowing recipients to rely on intermediaries rather than reconstructing every generative history themselves. Yet it also allocates power. A verifier can define admissible evidence. A platform can determine which identities count as authenticated. A credentialing institution can decide which histories qualify as expertise. A brand can convert reputation accumulated over decades into a premium attached to newly generated symbolic material.

The distributive question therefore concerns not only who can generate symbols, but who pays to make them credible. Let C_p denote production cost and C_v denote the cost of verification, authentication, and credibility maintenance. Generative systems can lower C_p for some actors while leaving C_v unchanged or raising it for recipients. A symbolic-abundance economy can therefore exhibit a widening separation between production cost and social-trust cost.

Proposition 13.11 (Trust-cost incidence). *Changes in symbolic production cost and changes in the social cost of establishing credibility need not accrue to the same actors. Lower producer costs can coexist with higher verification costs for recipients, institutions, or third-party intermediaries.*

This incidence problem is politically important because verification capacity is unequally distributed. Large organizations can purchase legal review, provenance infrastructure, professional auditing, and reputational insurance. Individuals and small organizations may face greater difficulty demonstrating credibility even when their substantive outputs are strong. Conversely, trusted incumbents may be able to generate large volumes of low-cost symbolic material while continuing to benefit from historically accumulated credibility. The symbolic crisis can therefore create a premium on inherited trust capital.

13.8 Political-Economic Diagnosis

The political economy of symbolic abundance can now be stated more precisely. Generative systems expand productive possibility, but economic scarcity does not disappear. It can move toward attention, visibility, verification, recognized provenance, relational access, responsibility, and institutional uptake. These bottlenecks can support new markets and socially useful infrastructures. They can also become sites of rent extraction, gatekeeping, and inequality.

The relevant historical change is therefore not a simple transition from scarcity to abundance. It is a reorganization of the cost structure through which symbols become consequential. Production can become cheap while credibility remains expensive. Representation can become abundant while recognition remains scarce. Personalization can scale while human relational bandwidth does not. Expertise-like performance can spread while institutions struggle to identify where robust competence resides. These asymmetries create new incentives for platforms, verifiers, credentialing bodies, brands, and other intermediaries.

Proposition 13.12 (Political economy of symbolic abundance). *As the marginal cost of some symbolic production falls, the distribution of value and power increasingly depends on the remaining bottlenecks that determine attention, verification, provenance, recognition, responsibility, and consequential access. The direction of redistribution is institution-specific and cannot be inferred from production abundance alone.*

This proposition prepares the next analytical problem. Provenance appears to offer one answer to diagnostic instability because it reconnects a present symbolic surface to a traceable history. Yet provenance is itself represented symbolically, can be partial, can authenticate origin without establishing quality or sincerity, and can become another infrastructure over which economic and institutional power is exercised. The next section therefore examines provenance as a necessary but limited response to the symbolic crisis.

14 Provenance and the Limits of Provenance

The preceding sections have shown why symbolic abundance increases pressure on the infrastructures that connect present artifacts to credible histories. Provenance appears to offer a direct response to this pressure. If a symbolic surface no longer carries enough information about how it came to exist, a provenance record can attach additional information about origin, transformation, participating agents, tools, times, and prior versions. The role of the present section is diagnostic rather than prescriptive. It clarifies what provenance can restore, what kinds of inference it can support, and which parts of the symbolic crisis remain unresolved even when provenance is technically valid and institutionally recognized.

The concept has a substantial history in data and information systems. Buneman, Khanna, and Tan characterize data provenance in terms of where data came from and why particular outputs exist (Buneman et al. 2001). The W3C PROV framework later generalized provenance into interoperable descriptions of entities, activities, and agents involved in producing a data item or artifact (Groth and Moreau 2013). Contemporary content-provenance infrastructures extend this logic into digital media. The C2PA technical specification, for example, provides a standardized architecture for cryptographically bound assertions concerning asset creation, transformations, ingredients, and signers (Coalition for Content Provenance and Authenticity 2026). These systems are important because they add structured historical information where the visible surface alone has become increasingly ambiguous.

Yet provenance should be treated as a differentiated informational layer rather than as a universal substitute for trust. A verified trace can establish some relations between an artifact and its recorded history while leaving other questions open. It may identify which software processed an image without establishing whether the image depicts a true event. It may record that a person signed or authorized a claim without establishing the quality of the judgment behind that claim. It may show that an AI system participated in a workflow without revealing how much interpretation, revision, or responsibility remained with the human principal. Provenance therefore narrows some uncertainties while preserving others.

Definition 14.1 (Provenance profile). For a symbolic artifact z , a *provenance profile* is the structured set of traceable relations connecting z to its recorded origin, transformations, agents, tools, temporal sequence, ingredients, attestations, and integrity conditions.

A schematic profile is stated in Equation 50.

$$\mathbf{P}(z) = (P_{\text{origin}}, P_{\text{process}}, P_{\text{agent}}, P_{\text{tool}}, P_{\text{time}}, P_{\text{transform}}, P_{\text{attestation}}, P_{\text{integrity}}). \quad (50)$$

The profile is deliberately broader than a single authorship label. A record that says “AI-generated” can still be thin if it omits the originating intention, model or tool, source ingredients, human review, subsequent edits, and the authority under which publication occurred. Conversely, some uses require only a small subset of the full profile. Provenance adequacy is therefore purpose-relative.

Table 12 makes the section’s central distinction visible. Provenance can strengthen the evidential relation between an artifact and a recorded history, yet the meaning of that history remains an interpretive problem. A cryptographically protected assertion can be authentic as an assertion while the proposition asserted remains contestable. A trace can faithfully record that a workflow occurred while providing limited evidence about whether the workflow produced a good decision. The value of provenance therefore depends on the inference target.

14.1 Provenance as Diagnostic Infrastructure

The symbolic crisis was defined earlier as a weakening and redistribution of the relation between symbolic surfaces and the generative histories recipients infer from them. Provenance partially responds by making some of those histories explicit. Instead of asking the visible surface to carry all evidential work, a recipient can inspect an auxiliary trace.

Table 12: Dimensions of provenance and residual ambiguity. A trace can improve one class of inference while leaving adjacent questions unresolved.

Provenance dimension	Information the trace can support	Residual ambiguity
Origin	source device, account, organization, or initial asset	truth, competence, intention, or legitimacy of the source
Process	recorded workflow or sequence of activities	unrecorded work, off-system activity, quality of each activity
Agent	identified participant, signer, or delegated actor	degree of attention, understanding, endorsement, or responsibility
Tool	software, model, device, or service used in production	quality of use, prompt structure, human judgment, hidden upstream dependencies
Temporal	ordering and timestamps of recorded events	continuity of participation, undocumented prior development, historical meaning of the sequence
Transformation	edits, derivations, ingredients, or recombinations	semantic significance of an edit and whether the resulting representation is misleading
Attestation	signed assertions or institutional statements	accuracy of the asserted proposition beyond integrity of the attestation mechanism
Integrity	tamper evidence and binding between record and artifact	factual truth, moral legitimacy, social value, or substantive reliability

Let $S(z)$ denote the symbolic surface of artifact z , let $H(z)$ denote its full generative history, and let $P(z)$ denote the provenance information available to a recipient. Under higher-friction conditions, recipients often relied on the practical inference represented in Equation 51.

$$S(z) \rightsquigarrow \hat{H}(z) \quad (51)$$

Generative abundance widens the set of histories compatible with a similar $S(z)$. Provenance adds the auxiliary inference channel represented in Equation 52.

$$(S(z), P(z)) \rightsquigarrow \hat{H}(z). \quad (52)$$

The added channel can reduce posterior ambiguity when $P(z)$ is relevant, credible, and sufficiently complete for the inference at issue.

Proposition 14.2 (Provenance as partial diagnostic restoration). *When symbolic abundance weakens the diagnostic relation between surface form and generative history, sufficiently relevant and credible provenance information can restore part of the lost diagnostic relation by constraining the set of histories compatible with the observed artifact.*

The proposition is intentionally partial. Provenance constrains possible histories rather than reproducing the full history itself. It is a representation of selected relations among entities, activities, agents, and claims. The W3C PROV family was designed precisely to represent and exchange such information across heterogeneous systems (Groth and Moreau 2013). C2PA similarly creates tamper-evident structures for assertions associated with digital assets and their transformations (Coalition for Content Provenance and Authenticity 2026). In both cases, the practical achievement is substantial: histories that would otherwise remain implicit become inspectable, interoperable, and in some settings machine-verifiable.

This can improve several classes of judgment. Provenance can distinguish an original asset from a derivative, connect a media object to a known publisher, identify declared AI involvement, expose an

editing sequence, preserve temporal priority, and make later modification detectable. It can also reduce the recipient’s verification burden by allowing a trusted infrastructure to maintain and validate parts of the chain.

Yet the improvement depends on what is actually represented. Provenance systems formalize selected aspects of history. The selection itself is consequential. A system optimized for media authenticity may record capture devices and edit actions while saying little about the social conditions of production. A scientific workflow may record datasets and scripts while omitting informal reasoning that influenced variable selection. A communication system may record that a principal authorized a message while omitting whether the principal read the recipient’s earlier message closely. Provenance therefore creates a structured window into history rather than transparent access to history as such.

Definition 14.3 (Provenance reach). The *provenance reach* of a record is the set of historical relations and inference targets that the record can support with adequate specificity for a given use.

This concept is useful because provenance systems differ greatly in reach. A timestamp has narrow reach. A version-control history has broader process reach. A cryptographically signed content credential can have strong integrity reach while still having limited reach concerning factual truth or human participation. A full provenance graph can be structurally rich and still omit morally or epistemically relevant events that occurred outside the recorded workflow.

14.2 Integrity, Truth, and Quality

The first major limit concerns the distinction between integrity of provenance and truth of content. Current C2PA specifications state this distinction explicitly. Their design verifies that assertions are associated with an asset, correctly formed, and resistant to tampering; the specification does not assign a positive or negative value judgment to the underlying content (Coalition for Content Provenance and Authenticity 2026). A valid provenance chain can therefore accompany a false claim, a misleading image, a low-quality analysis, or an ethically objectionable artifact.

The three analytical levels are listed in Equation 53.

$$\begin{array}{l} \text{provenance integrity} \\ \text{content accuracy} \\ \text{normative acceptability} \end{array} \quad (53)$$

The levels can interact, yet they should remain analytically distinct. Provenance integrity concerns whether the trace is correctly bound to the artifact and has remained intact. Content accuracy concerns whether the represented proposition or state of affairs is reliable. Normative acceptability concerns whether the creation, use, or effects of the artifact satisfy relevant ethical, legal, or institutional standards.

Proposition 14.4 (Integrity–truth separation). *A valid provenance record can increase confidence about the origin and recorded history of an artifact without establishing the truth, quality, fairness, or legitimacy of the artifact itself.*

This separation matters because the symbolic crisis can create a strong demand for technical signs of authenticity. Once surface cues become less diagnostic, a validated credential can acquire substantial persuasive weight. Recipients may then overextend what the credential establishes. The result is a new form of diagnostic mismatch: provenance integrity is used as evidence for properties outside its reach.

Definition 14.5 (Provenance overreach). *Provenance overreach* occurs when a recipient treats verified provenance information as evidence for a property that the provenance record does not adequately support.

Examples include treating known authorship as evidence of factual correctness, known AI involvement as evidence of low quality, a continuous edit history as evidence of sincerity, or an institutional signature as evidence that the institution’s underlying judgment was sound. Provenance can improve judgment while still becoming another symbolic cue that recipients interpret too broadly.

14.3 Participation and Responsibility

A second limit arises from the difference between artifact history and participatory history. Sections 5 and 6 showed that communication can distribute goal formation, interpretation, generation, authorization, monitoring, and responsibility across several actors. A provenance record can identify some of those actors while still leaving the quality and intensity of their participation unclear.

Suppose a message is recorded as generated by model M , authorized by person A , and sent by organization O . This already improves agency legibility. Yet several questions remain. Did A formulate the substantive judgment? Did A read the incoming communication? Did A inspect the generated reply? Was authorization a deliberate act or a standing rule? Who would recognize an error? Who can revise the commitment if the recipient challenges it? These questions concern participation profiles rather than mere actor presence.

Definition 14.6 (Provenance–participation gap). The *provenance–participation gap* is the difference between recording that an actor, tool, or institution appears in an artifact’s generative history and establishing the substantive form, depth, and consequences of that actor’s participation.

The gap becomes especially important in agentic systems. An autonomous agent can produce a rich trace of operations while the human principal participates only at the level of initial configuration or ex post monitoring. Conversely, a human can engage deeply with an AI-assisted workflow even when the provenance trace records only the tool invocation. A provenance graph and a participation profile therefore answer different questions.

The same distinction applies to responsibility. A provenance system can make attribution easier by recording which actors were associated with which activities. Attribution supports responsibility practices because it narrows uncertainty about involvement. Yet responsibility also depends on authority, foreseeability, control, role obligations, institutional rules, and the possibility of answerability. These properties are only partly encoded by a record of participation. A signed authorization can be relevant to responsibility without exhausting it.

Working Claim 14.7 (Artifact provenance and responsibility structure). Artifact provenance can support the reconstruction of responsibility relations, while the normative allocation of responsibility requires additional information about authority, control, expectations, institutional roles, and the consequences of participation.

This distinction prevents technical provenance from being treated as a complete moral ledger. It can provide evidence within responsibility assessment while leaving the assessment itself open.

14.4 Completeness and Off-Chain Histories

A third limit concerns completeness. Every provenance infrastructure has boundaries. Some events enter the record and others remain outside it. The resulting trace can be internally valid while still being historically incomplete.

Let the full generative history of an artifact be $H(z)$ and the recorded provenance be $P(z) \subseteq H(z)$. The substantive problem is not simply the size of $P(z)$ but the selection mechanism that determines which events enter the trace. If omitted events are irrelevant to a given use, the record may be adequate. If omitted events systematically contain the very information recipients need, the record can create false confidence.

Definition 14.8 (Provenance completeness gap). The *provenance completeness gap* is the practically relevant difference between the full history needed for a particular judgment and the subset of that history represented in the available provenance record.

Off-chain histories can include informal conversations, undocumented prompts, prior drafts, external datasets, tacit expertise, private deliberation, outsourced labor, institutional pressure, or earlier conceptual development. In cultural and scholarly work, an artifact’s formal creation history may begin long after the intellectual or relational trajectory that made the artifact possible. In organizational communication, the provenance of a sent message may omit the internal decision process that determined what could be said.

This limit is structural rather than merely technical. A system that attempted to record every relevant antecedent would face severe costs in storage, privacy, interpretive complexity, and governance. Relevance itself changes with the question being asked. The provenance required to establish temporal priority differs from the provenance required to establish informed consent, expertise, cultural lineage, or legal responsibility.

Proposition 14.9 (Purpose-relative completeness). *No provenance record is complete in the abstract. Adequacy depends on whether the recorded history contains the relations required for the specific inference, accountability practice, or decision in which the record is used.*

This proposition also limits the usefulness of a binary label such as “provenance available.” The more important questions concern which history is recorded, which history remains outside the trace, who selected the representation, and what inference the recipient intends to draw.

14.5 Trust Anchors and Institutional Dependence

Provenance infrastructure can reduce uncertainty only when recipients have reasons to rely on the actors and institutions that generate, sign, store, interpret, or validate the trace. C2PA’s trust model makes this institutional dependence explicit through signing identities, certificate chains, trust lists, and other trust signals (Coalition for Content Provenance and Authenticity 2026). A valid signature identifies a signer and protects the integrity of the associated claim; the recipient still needs a basis for deciding how much weight to place on that signer and the claim.

The recursive trust structure is summarized in Equation 54.

$$\text{artifact uncertainty} \longrightarrow \text{provenance infrastructure} \longrightarrow \text{trust in provenance infrastructure.} \quad (54)$$

Provenance can relocate rather than eliminate trust. The recipient moves from relying primarily on the symbolic surface to relying partly on a standards ecosystem, identity system, archive, platform, institution, or signer. Section 13 identified this possibility as part of credibility infrastructure. Provenance systems therefore participate in the political economy of symbolic abundance: they can lower verification costs while also creating new gatekeeping positions.

Definition 14.10 (Provenance trust stack). A *provenance trust stack* is the layered set of technical and institutional dependencies through which a provenance claim becomes interpretable and credible, including the artifact binding, assertions, signer identity, trust anchors, repositories, validation software, and recipient-side standards of acceptance.

Failures can occur at several layers. The asset can lose attached metadata. A signer can be unknown or untrusted. A repository can become unavailable. A recipient can misunderstand the user interface. A technically valid trace can be overinterpreted. A widely recognized trust list can privilege some institutions over others. These are distinct failure modes and should not be collapsed into one notion of authenticity.

The recursive structure also creates a governance issue. If provenance becomes a prerequisite for credibility in high-stakes domains, actors without access to recognized signing infrastructure, stable identity systems, archival capacity, or compliant production tools may face higher verification burdens. The social value of provenance can therefore depend on whether its infrastructures are broadly accessible and institutionally plural.

14.6 Privacy, Anonymity, and Unequal Verifiability

Provenance expands traceability, yet traceability can conflict with privacy, anonymity, safety, and legitimate opacity. Contemporary provenance standards explicitly recognize these tensions. C2PA's design goals include privacy and harms mitigation, and its security guidance discusses risks created when signing identifiers or certificates make creators more identifiable than intended (Coalition for Content Provenance and Authenticity 2026). The problem is especially important for journalists, dissidents, whistleblowers, vulnerable communities, and participants whose safety depends on selective disclosure.

The tension between the two socially valuable capacities is stated in Equation 55.

$$\text{verifiability} \quad \text{and} \quad \text{protected opacity.} \quad (55)$$

A system optimized entirely for traceability may increase the evidential value of artifacts while imposing unacceptable exposure on some creators. A system optimized entirely for anonymity may preserve safety while leaving recipients with fewer resources for origin verification. The appropriate balance is context-dependent.

Definition 14.11 (Unequal verifiability). *Unequal verifiability* is a condition in which actors differ systematically in their capacity to produce socially recognized provenance without incurring excessive financial, institutional, privacy, or safety costs.

Unequal verifiability can reinforce existing institutional asymmetries. A major publisher can maintain signing infrastructure, persistent archives, organizational identities, and dedicated compliance staff. An independent researcher, small cultural organization, anonymous source, or precarious worker may have fewer such resources. If provenance becomes a dominant credibility convention, the absence of a recognized trace may be interpreted as suspicious even where the absence reflects resource constraints or legitimate privacy needs.

The current C2PA framework explicitly cautions that absence of Content Credentials should not by itself be treated as evidence that an asset is untrustworthy, and that validated provenance does not itself establish truth (Coalition for Content Provenance and Authenticity 2026). These cautions are important because social conventions can become stronger than the technical claims of the standard. Once a provenance mark acquires symbolic authority, users may infer more from its presence or absence than the infrastructure was designed to support.

14.7 The Provenance Paradox

The analysis now reaches a recursive problem. Provenance is introduced because the visible symbol has become a weaker guide to its generative history. Yet provenance information itself is represented through additional symbols, credentials, interfaces, labels, signatures, and institutional assertions. The symbolic crisis therefore cannot be escaped simply by adding another symbolic layer. The new layer can be stronger because it is technically bound, institutionally governed, and historically structured, but its interpretation remains socially mediated.

Definition 14.12 (Provenance paradox). The *provenance paradox* is the condition in which symbolic ambiguity creates demand for provenance, while provenance itself must be communicated through symbolic and institutional forms whose meaning, scope, and trustworthiness require interpretation.

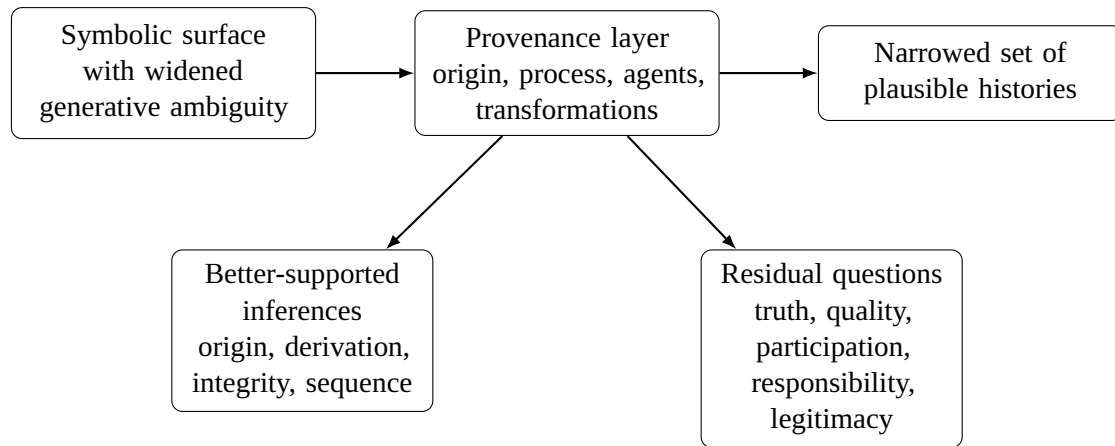


Figure 14: Provenance as partial diagnostic restoration. A provenance layer can narrow the set of plausible generative histories while leaving adjacent epistemic, relational, and normative questions unresolved.

The paradox does not make provenance futile. It identifies why provenance functions as infrastructure rather than as a final epistemic foundation. Provenance improves the architecture of inference by constraining histories, preserving temporal relations, and exposing some transformations. It still operates inside a social field of trust, standards, institutions, interfaces, and interpretive habits.

Figure 14 summarizes the boundary developed throughout this section. Provenance can recover some of the historical structure lost when symbolic surfaces become cheap to generate. Its strongest contribution lies in traceability: origin, sequence, transformation, attribution, and integrity can become more legible. Its limits appear where social inference requires properties that exceed the recorded chain: competence, sincerity, understanding, responsibility, legitimacy, cultural meaning, and the depth of human participation.

Proposition 14.13 (Provenance boundary). *Provenance can strengthen the connection between symbolic artifacts and recorded histories, while the symbolic crisis persists wherever socially important judgments depend on unrecorded participation, contested interpretation, normative evaluation, or historical relations that exceed the provenance record.*

This boundary prepares the next part of the diagnosis. Even strong provenance infrastructures operate within inherited perceptual habits. Social actors may continue reading polished language as attention, institutional affiliation as competence, personalization as care, and visual realism as evidence of occurrence long after production conditions have changed. The next section therefore turns from the architecture of traces to the historical timing of perception itself: the temporal mismatch between changing symbolic production and slower-changing interpretive habits.

15 Temporal Mismatch in Symbolic Perception

The preceding sections have described several domains in which symbolic surfaces become less diagnostic of the histories, capacities, relations, and forms of participation that recipients previously inferred from them. Yet the symbolic crisis does not follow automatically from technological change alone. A production regime can change rapidly while the social habits through which people interpret its outputs change much more slowly. The role of this section is to analyze that temporal asymmetry. The central problem is a mismatch between the speed at which symbols can be generated and the speed at which social actors, institutions, and evaluative practices recalibrate the meanings assigned to those symbols.

This problem has recognizable precedents in social theory. Ogburn's account of cultural lag described maladjustment that can arise when different components of culture change at different rates (Og-

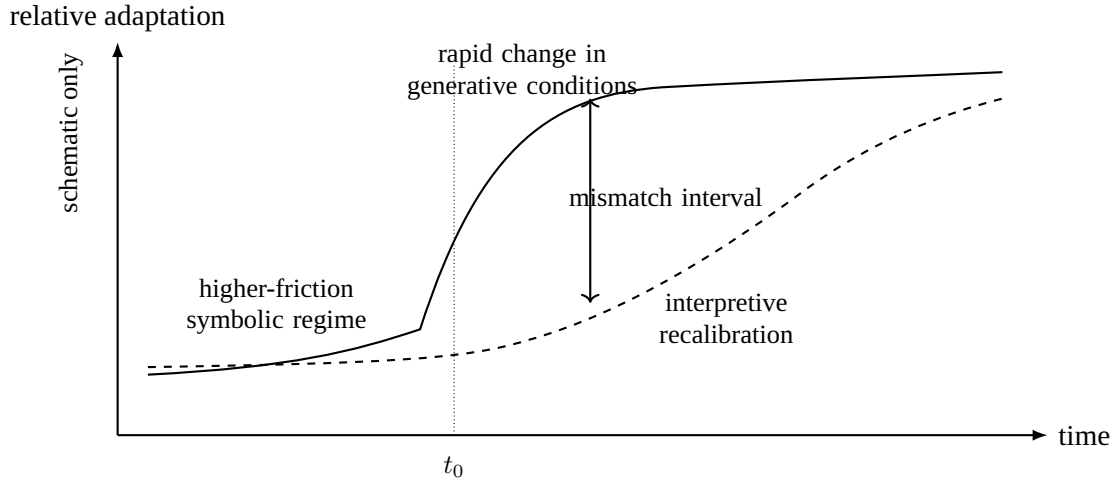


Figure 15: Temporal mismatch between symbolic production and symbolic perception. The schematic depicts a rapid change in generative conditions at t_0 followed by slower recalibration of interpretive expectations. The gap is conceptual rather than a claim about a universal measurable curve.

burn 1922). Bourdieu’s theory of habitus likewise emphasizes the durability of embodied dispositions formed under prior conditions, allowing practices and expectations to persist as objective circumstances change (Bourdieu 1977). The present argument draws on these traditions while narrowing the object of analysis. The relevant lag is not simply a general delay between technology and culture. It concerns the persistence of *symbolic inference rules*: learned expectations connecting visible form to authorship, effort, competence, attention, scarcity, sincerity, and social commitment.

Definition 15.1 (Temporal mismatch in symbolic perception). A *temporal mismatch in symbolic perception* occurs when the generative conditions of symbolic production change faster than the interpretive expectations through which recipients infer social meaning from symbolic surfaces.

The mismatch can be represented schematically by distinguishing a generative regime G_t from an interpretive regime I_t . The generative regime includes the technical, institutional, and organizational conditions under which symbolic artifacts can be produced. The interpretive regime contains the expectations, heuristics, conventions, and institutional routines through which those artifacts are read. A symbolic crisis becomes more likely under the temporal asymmetry stated in Equation 56.

$$\Delta G_t \gg \Delta I_t, \quad (56)$$

This condition describes substantial production change alongside comparatively stable inference rules attached to visible form.

This notation is illustrative rather than metric. The claim does not require a common scale on which technological and interpretive change can literally be measured. It specifies a temporal relation: symbolic production may enter a new regime before the practices of symbolic interpretation have undergone comparable recalibration.

Figure 15 depicts the central temporal structure. The period of greatest instability can occur after the generative regime has changed but before interpretive practice has caught up. During this interval, the recipient continues to make inferences that were more reliable under the previous regime. The surface has changed less visibly than the pathway by which the surface can be produced.

15.1 Inherited Inference Rules

Symbolic interpretation depends on accumulated social learning. A polished report may have been read as evidence of preparation because producing such a report usually required substantial human time. A personalized message may have indicated attention because personalization was costly to perform at scale. A fluent explanation could support an inference of domain familiarity because producing it generally required either knowledge or substantial search. A realistic image could support an inference of prior visual capture because photorealistic fabrication was relatively costly and specialized.

These relations were never perfect. Ghostwriting, forgery, scripted customer service, advertising, institutional delegation, and strategic impression management long predate generative AI. The diagnostic value of symbols has always depended on context and convention. The present change concerns the breadth and ordinary accessibility of alternative generative pathways. When high-quality symbolic surfaces become reachable through dramatically lower-friction processes, the old inference rule can persist after its reliability has weakened.

Definition 15.2 (Inherited inference rule). An *inherited inference rule* is a socially learned association by which a recipient treats a symbolic feature as evidence for an underlying property because that association was sufficiently reliable under prior production conditions.

Examples include treating extensive prose as evidence of extensive writing time, stylistic fluency as evidence of language mastery, personalized wording as evidence of direct attention, or elaborate presentation as evidence of substantial preparatory labour. These rules need not be consciously articulated. They can function as perceptual shortcuts embedded in everyday practice.

The temporal problem appears under the conditional mismatch stated in Equation 57.

$$P(H \mid S, G_{t-1}) \neq P(H \mid S, G_t), \quad (57)$$

Under this condition, recipients continue to behave as though the earlier conditional relation still holds. Here S denotes a symbolic surface, H a hypothesized generative history, and G the production regime. The change in G modifies the evidential meaning of S even if the visible surface itself remains familiar.

Proposition 15.3 (Interpretive carryover). When symbolic production changes faster than learned inference practices, recipients can continue assigning evidential weight to cues whose diagnostic relations were formed under earlier production conditions.

Interpretive carryover helps explain why the symbolic crisis can feel psychologically disorienting. The problem is not simply that people lack factual knowledge about generative AI. A person may fully know that AI can generate polished language and still experience polished language as effortful, attentive, or authoritative. Declarative knowledge about production conditions and embodied habits of interpretation can move at different speeds.

15.2 Interpretive Hysteresis

The persistence of old inference rules after a change in generative conditions can be described as a form of *interpretive hysteresis*. The term is used here analogically and should not be confused with a claim that social perception obeys a literal physical hysteresis curve. The concept draws on the broader idea that dispositions formed under earlier structural conditions can remain active after those conditions have changed, a problem closely related to Bourdieu's account of durable habitus (Bourdieu 1977).

Definition 15.4 (Interpretive hysteresis). *Interpretive hysteresis* is the persistence of symbolic expectations, evaluative habits, or institutional inference rules after the production conditions that originally supported those expectations have materially changed.

Layer	Inherited inference	Source of lag	Possible mismatch
Individual perception	polished form indicates effort, attention, or competence	durable habits and bounded verification time	over-attribution of human participation or expertise
Interactional norms	personalized response indicates reciprocal engagement	etiquette changes more slowly than production tools	relational meaning exceeds actual participation
Organizations	documents and outputs evidence work performed by identifiable roles	legacy workflows, accountability structures, assessment routines	formal outputs conceal redistributed agency
Markets	scarcity of form supports differentiation and price	brand histories, status conventions, platform incentives	symbolic abundance coexists with old scarcity premiums
Education and credentials	completed artifact indicates learner formation	assessment systems and credential chains adjust slowly	artifact quality outpaces evidential value about learning
Public institutions	official symbolic forms support authority and procedural trust	legal continuity, procedural inertia, legitimacy requirements	new generative pathways exceed inherited verification rules

Table 13: Illustrative layers of temporal mismatch. The table does not imply that every domain experiences the same direction or magnitude of change.

Interpretive hysteresis matters because adaptation is costly. People cannot continuously reconstruct the generative history of every message, image, credential, report, or cultural artifact they encounter. Heuristics exist precisely because direct verification is expensive. Institutions likewise encode earlier assumptions into assessment formats, documentation requirements, professional norms, and trust procedures. When production conditions shift, these routines cannot be rewritten instantly without creating new coordination costs.

The consequence is not merely error. Old inference rules can remain socially functional even when their diagnostic strength declines. A polished report can still communicate clearly. A degree can still coordinate labour-market sorting. A courteous email can still facilitate interaction. A brand can still organize expectations. The mismatch therefore persists partly because symbolic conventions continue to perform coordinative functions even while some of their evidential foundations weaken.

This duality makes adaptation difficult. If a symbolic form had simply become useless, abandonment would be comparatively straightforward. The harder condition is selective degradation: the same form remains useful for coordination while becoming less reliable for certain inferences. Section 4 described this as target-relative diagnosticity. Temporal mismatch adds a historical dimension: the social reweighting of diagnostic targets can lag behind the technological change that made reweighting necessary.

15.3 Uneven Adaptation Across Social Domains

Interpretive recalibration is unlikely to occur at one speed across society. Different domains face different error costs, institutional capacities, production technologies, and incentives to update. A casual social message, a legal filing, an academic examination, a luxury advertisement, a public statement, and a medical report can therefore enter the symbolic crisis at different rates.

Table 13 emphasizes that adaptation has several loci. Individual perception can change while organizational procedures remain fixed. Institutions can require new provenance practices while everyday recipients continue to rely on old surface cues. Younger cohorts can normalize mediated communication while professional cultures preserve prior authenticity norms. Some sectors can develop strong verification infrastructures while others remain dependent on interpersonal heuristics.

This produces a patchwork transition rather than a single historical break. The term *symbolic crisis*

should therefore describe a family of mismatches whose intensity differs across domains. The crisis is strongest where three conditions coincide: production pathways change rapidly, the relevant symbolic cue previously carried substantial evidential weight, and social adaptation remains slow or contested.

Definition 15.5 (Mismatch intensity). For a symbolic cue S and inference target H , *mismatch intensity* refers to the degree of divergence between the evidential weight currently assigned to S and the evidential weight warranted by the current production regime.

This concept is intentionally relational. It does not classify a symbolic form as globally obsolete. The same cue can exhibit high mismatch for one target and low mismatch for another. A fluent message may remain highly informative about what proposition has been transmitted while being weak evidence of how much direct attention the nominal sender devoted to composing it.

15.4 Windows of Symbolic Arbitrage

Temporal mismatch creates opportunities for strategic exploitation. If recipients continue to interpret a signal using older production assumptions while producers gain access to cheaper generative pathways, an actor can obtain social benefits associated with the old cue at a reduced production cost. Section 13 described one economic form of this problem as interpretive lag rent. The present section generalizes the temporal mechanism.

Definition 15.6 (Symbolic lag arbitrage). *Symbolic lag arbitrage* is the strategic capture of benefits from an inherited symbolic inference after the cost structure or generative conditions that previously sustained that inference have changed.

The concept includes, but is broader than, deception. An actor may exploit inherited assumptions without making an explicitly false statement. Automated personalization can generate the felt appearance of attention. Highly polished assistance can borrow the prestige associated with expertise. Synthetic abundance can reproduce stylistic markers previously associated with rarity. The recipient's inference supplies part of the value.

The existence of lag arbitrage also explains why symbolic recalibration is politically and economically contested. Actors who benefit from old inference rules may have incentives to preserve ambiguity. Others may gain from establishing new disclosure requirements, provenance systems, or verification conventions. The pace of adaptation is therefore shaped by power and institutional interest as well as by learning.

Proposition 15.7 (Temporal arbitrage window). *A rapid reduction in production friction can create a temporary interval in which inherited symbolic inferences remain socially valuable after the generative costs that supported them have fallen, allowing rents or advantages to accrue to actors positioned to exploit the mismatch.*

The duration of such a window is uncertain. Some cues may recalibrate quickly once users repeatedly encounter counterexamples. Others can persist because institutions stabilize them, because verification remains costly, or because their social value depends partly on convention rather than empirical diagnosticity. The history of symbolic adaptation therefore need not converge toward one perfectly informed equilibrium.

15.5 Generalized Suspicion and Overcorrection

Adaptation can also overshoot. Once recipients learn that a polished surface can be generated cheaply, they may cease to treat the surface as evidence of anything beyond immediate content. This response can protect against over-attribution, yet it can also erase legitimate distinctions among generative histories.

A human author may have invested years in developing an argument even if an AI system assisted with phrasing. A personalized response may genuinely reflect careful human judgment even if the final wording was machine-generated. A luxury artifact may possess documented process and provenance even if its surrounding imagery uses generative tools. The symbolic crisis therefore creates a risk of *generalized suspicion*: a broad discounting of symbolic evidence that treats heterogeneity of production as equivalent to absence of human participation, competence, or commitment.

Definition 15.8 (Generalized symbolic suspicion). *Generalized symbolic suspicion* is an interpretive response in which awareness of heterogeneous generative pathways produces indiscriminate discounting of symbolic cues across histories that remain substantively different.

This response reproduces the same structural error in the opposite direction. The inherited regime over-infers from the symbolic surface; generalized suspicion under-infers by refusing to distinguish meaningful histories behind superficially similar outputs. Both responses compress heterogeneity too aggressively.

The diagnostic task is therefore neither to restore old inference rules intact nor to discard symbolic inference altogether. It is to understand how inferential weight is redistributed across surfaces, provenance, institutional guarantees, longitudinal records, relational experience, and other forms of evidence. The companion paper on response directions will examine such possibilities in greater detail. The present paper remains concerned with the transitional condition itself.

15.6 Institutional Time and Perceptual Time

The temporal mismatch also involves several clocks. Technical capabilities can change on the scale of months. Organizational procedures may change over years. Legal standards, professional norms, educational systems, and credential regimes may change more slowly. Individual perception can adapt quickly in one context while remaining deeply habitual in another. Cultural meaning can persist for generations even after its original material basis has changed.

Ogburn's concept of cultural lag is useful precisely because it rejects the assumption that the components of social change move synchronously (Ogburn 1922). Yet symbolic perception adds a further complication: interpretation is not only an institutional rule but also a lived, affective, and embodied practice. Bourdieu's account of habitus helps explain why dispositions can reproduce historically acquired expectations without requiring explicit deliberation each time they are enacted (Bourdieu 1977).

For the symbolic crisis, this means that information campaigns alone may have limited reach. Telling recipients that "AI can generate this" does not immediately reconstruct the emotional and social meaning of polish, personalization, fluency, or rarity. Recalibration can require repeated encounters, new norms, new institutions, and changes in what forms of interaction people come to experience as trustworthy or meaningful.

Proposition 15.9 (Multiple adaptation clocks). *Symbolic recalibration proceeds through multiple temporalities: technical, organizational, legal, institutional, interpersonal, and perceptual. A change in production capability therefore does not imply a synchronized change in social meaning.*

This proposition helps explain why the symbolic crisis should be understood as an era-scale problem rather than an individual decision problem. Individual actors can modify their own practices, but they cannot unilaterally rewrite the shared meanings on which communication, credentialing, trust, and valuation depend. New conventions become effective only as they are stabilized through repeated social use, institutional adoption, and reciprocal expectation.

15.7 Historical Transition in the Location of Warrant

The temporal-mismatch argument can be expressed in a broader historical form. Symbolic systems have long allowed societies to place epistemic and social warrant in portable representations: documents,

signatures, credentials, titles, brands, measurements, photographs, reports, and other formalized signs. These systems reduced the need for direct access to generative histories. Their usefulness depended partly on institutions that stabilized relations between signs and the processes they represented.

Generative abundance does not remove symbolic mediation. It changes the amount of inferential work that a visible surface can carry unaided. During the mismatch interval, inherited symbols continue to command recognition even as their generative foundations diversify. Over time, warrant may become redistributed across additional layers: provenance, institutional assurance, historical record, observable participation, or other infrastructures not yet fully stabilized. The present paper does not assume which configuration will prevail.

Working Claim 15.10 (Temporal redistribution of warrant). The symbolic crisis can be interpreted as a historical interval in which the production conditions of symbolic evidence change faster than the social conventions that allocate epistemic and relational warrant to symbolic form.

This formulation avoids predicting a post-symbolic society. Symbols remain indispensable for large-scale coordination because direct access to every relevant history is impossible. The historical question concerns how much warrant remains concentrated in surface form and how much is redistributed to other traces, institutions, and relations.

15.8 Diagnostic Implications

Temporal mismatch clarifies why the symbolic crisis can produce contradictory social reactions at the same time. Some recipients continue to trust familiar surfaces. Others become suspicious of them. Institutions add provenance requirements. Organizations automate communication while maintaining human-branded voices. Educational systems retain artifact-centered assessment while questioning what the artifacts demonstrate. Markets make symbolic production cheaper while intensifying competition over attention, verification, access, and recognition.

These are not necessarily inconsistent responses. They can be understood as different positions within a historical recalibration whose endpoint remains uncertain.

Proposition 15.11 (Temporal-mismatch diagnosis). *The symbolic crisis is intensified when generative conditions change faster than the interpretive, institutional, and perceptual practices through which symbolic surfaces acquire evidential and relational meaning. The resulting mismatch can produce over-attribution, strategic arbitrage, generalized suspicion, and uneven recalibration across social domains.*

The diagnostic value of this proposition lies in shifting attention from a static comparison between “human” and “AI” artifacts toward a historical relation between production and perception. The same technology can create different social consequences depending on whether interpretive practices have adapted, whether institutional infrastructures have stabilized new meanings, and whether actors share expectations about legitimate mediation.

The next section extends this temporal argument. If symbolic practices cannot be recalibrated instantaneously, the relevant question becomes historical rather than merely technical: how trust practices themselves change, stabilize, and become ordinary over time. The next section therefore turns to the historical adaptation of trust practices and the slow formation of new conventions under changing symbolic conditions.

16 Historical Adaptation of Trust Practices

Section 15 described the symbolic crisis as a temporal mismatch between changing generative conditions and slower-moving interpretive expectations. The present section develops the next diagnostic step. Its role is to examine how trust practices themselves change historically when inherited cues lose diagnostic

force. The section treats trust adaptation as a distributed social process involving routines, institutions, verification arrangements, professional norms, relational histories, and second-order assurances. The objective is neither to prescribe a completed post-crisis trust architecture nor to identify one universally superior mode of trust. It is to clarify the mechanisms through which trust practices can migrate, layer, stabilize, and become ordinary over time.

The historical literature already provides an important starting point. Zucker's analysis of trust production distinguished process-based, characteristic-based, and institution-based sources of trust and situated their relative importance historically (Zucker 1986). Shapiro's analysis of impersonal trust showed that large-scale societies rely on procedural constraints, professional guardians, policing mechanisms, and institutional assurances when trust cannot remain embedded entirely in direct personal relations (Shapiro 1987). Möllering later emphasized reason, routine, reflexivity, and the suspension of uncertainty as distinct dimensions of trust (Möllering 2006). More recently, Lumineau, Schilke, and Wang argue that technological intermediation can shift organizational trust toward more impersonal and system-based forms and can alter both the production and target of trust (Lumineau et al. 2023). These works collectively support a historical interpretation of trust: trust is reproduced through social arrangements whose forms vary with institutional complexity, technological mediation, and the availability of credible evidence.

Definition 16.1 (Trust practice). A *trust practice* is a repeatable social arrangement through which actors become willing to rely on persons, institutions, systems, or claims under conditions of incomplete verification. A trust practice can combine symbolic cues, prior experience, institutional guarantees, monitoring, sanction, recourse, reputation, routine, and shared expectations.

The definition shifts attention away from trust as a purely internal attitude. A person may experience trust psychologically, yet the conditions that make reliance possible are often produced socially. Signatures, credentials, professional licensing, escrow, warranties, references, reputational records, due-process procedures, institutional affiliations, and repeated interaction all distribute the work of reducing uncertainty. Trust therefore has a practical infrastructure.

Definition 16.2 (Historical trust adaptation). *Historical trust adaptation* is the distributed process through which a society, institution, profession, or community changes the cues, routines, assurances, and relational practices by which reliance becomes acceptable after the evidential environment has changed.

Historical adaptation is slower than a change in technical capability because the relevant object is shared expectation. An individual can decide to disregard one symbolic cue tomorrow. A profession cannot instantaneously create a new credentialing convention with common meaning. A legal system cannot immediately establish a new liability structure. A community cannot rapidly make an unfamiliar verification ritual feel ordinary. Trust practices become scalable when actors develop reciprocal expectations about how those practices should be interpreted.

16.1 Trust Production across Historical Conditions

Zucker's typology is useful because it separates several ways in which trust can be produced. *Process-based trust* draws on the history of interaction and evidence generated through prior exchange. *Characteristic-based trust* draws on socially recognized attributes or category membership. *Institution-based trust* draws on formal structures, certifications, rules, intermediaries, or background assurances (Zucker 1986). These modes can coexist, and their relative importance varies by social setting.

The symbolic crisis pressures this mixture. When polished symbolic performance becomes easier to produce, a cue that previously contributed to characteristic- or performance-based inference can lose some diagnostic value. Actors can then place more weight on prior interaction, third-party verification,

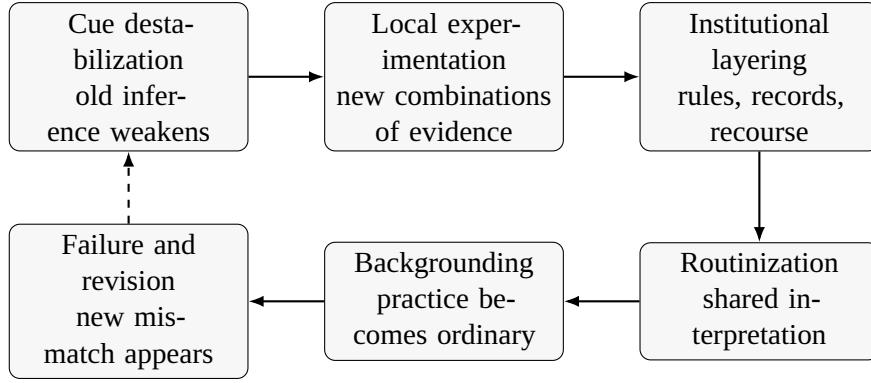


Figure 16: A schematic cycle of historical trust adaptation. The figure represents a possible sequence from cue destabilization through experimentation, institutional layering, routinization, and later revision. The sequence is neither universal nor irreversible; stages can overlap, stall, or proceed differently across domains.

institutional standing, or other evidential channels. Alternatively, they may continue using the inherited cue because it remains cheap, familiar, and coordinatively useful. Historical adaptation therefore involves changes in *weight*, not necessarily wholesale replacement of one trust mode by another.

A trust practice at time t is represented schematically in Equation 58.

$$\Pi_t = \langle w_P, w_C, w_I, w_S, w_R \rangle, \quad (58)$$

In Equation 58, w_P represents weight placed on process history, w_C on characteristic or categorical cues, w_I on institutional assurance, w_S on immediate symbolic performance, and w_R on recourse or repair mechanisms. The profile is descriptive rather than metric. Its purpose is to represent the possibility that the social production of trust changes by reweighting multiple evidential sources.

Definition 16.3 (Trust-production profile). A *trust-production profile* is the context-dependent configuration of process history, categorical cues, institutional assurance, symbolic performance, and recourse through which reliance is socially supported.

Proposition 16.4 (Historical reweighting). *Changes in symbolic production conditions can alter the relative importance of different trust-production modes without eliminating the continued use of symbolic mediation.*

This proposition is consistent with the argument developed by Lumineau, Schilke, and Wang that technological mediation can shift trust toward more impersonal and system-based arrangements while changing the modes through which trust is produced (Lumineau et al. 2023). The present paper narrows that claim to generative abundance: as symbolic surfaces become reachable through more heterogeneous histories, actors can gradually relocate trust toward other sources of evidence or assurance.

Figure 16 emphasizes that adaptation can be iterative. New trust arrangements can later become the inherited arrangements exposed by another technological or institutional change. A trust regime is therefore historically provisional even when it feels settled to participants.

16.2 Experimentation before Institutionalization

New trust practices often begin in an unsettled form. Participants experiment with additional disclosure, stronger verification, repeated interaction, more explicit authorization, reputation systems, technical provenance, human review, or different expectations about acceptable delegation. Many such experiments disappear. Others become locally successful routines and are later codified through professional standards, organizational procedures, technical protocols, contractual language, or law.

The distinction between local experimentation and institutionalization matters because the former can function without broad agreement. Two collaborators can decide that all AI-assisted drafts will be reviewed synchronously before sending. A larger profession requires shared expectations about what the review signifies, which actions it authorizes, and which responsibilities remain with the human participants. Scale changes the trust problem.

Definition 16.5 (Trust-practice institutionalization). *Trust-practice institutionalization* is the process through which a locally contingent method of establishing, transferring, or repairing trust acquires wider recognition, repeated use, and relatively stable interpretive meaning.

Institutionalization can lower verification costs. Once a practice has recognized meaning, actors need less negotiation each time it is used. A professional licence, standardized audit, or recognized warranty does not eliminate uncertainty; it compresses a larger background of procedures into a portable assurance. The same mechanism can emerge around generative mediation, although the final forms remain historically open.

This compression recreates a familiar tension. The new assurance becomes another symbol. Its usefulness depends on the processes and institutions that keep it coupled to what it claims to represent. Historical adaptation therefore does not escape symbolic mediation. It builds new symbolic and institutional layers around a changed evidential environment.

Working Claim 16.6 (Recompression after adaptation). Historical trust adaptation tends to generate new forms of social compression. Practices that initially require explicit explanation can become standardized symbols, routines, or institutional assurances once their interpretation stabilizes.

This claim clarifies why the symbolic crisis is unlikely to produce a permanently de-symbolized society. Direct verification and deep co-experience are too costly for large-scale coordination. Instead, societies can develop new compressed interfaces whose credibility depends on newly stabilized histories, institutions, and recourse structures. Whether those interfaces solve earlier problems or reproduce them in altered form remains an empirical question.

16.3 Routinization and the Perceptual Layer

Institutionalization alone does not complete historical adaptation. A rule can exist before people experience it as ordinary. Möllering's distinction among reason, routine, and reflexivity is useful here because trust can depend on explicit deliberation in one setting and background routine in another (Möllering 2006). During a transition, actors may need to reason explicitly about mediation practices that later generations encounter as unremarkable infrastructure.

This transition from reflexive attention to routine matters for the symbolic crisis. At present, an interaction may prompt questions such as whether a message was AI-generated, whether the apparent sender reviewed it, or whether the institution stands behind an automated commitment. A stabilized trust practice can redistribute these questions into routine procedures. The recipient may cease asking them consciously because authorization, recourse, or provenance has become an accepted part of the environment.

Definition 16.7 (Trust routinization). *Trust routinization* is the gradual movement of a trust-supporting practice from explicit, reflexive evaluation toward a background expectation that participants can ordinarily rely on without reconstructing its full justification in each interaction.

Routinization creates efficiency, yet it also creates vulnerability to later mismatch. Once a practice becomes background infrastructure, people can stop examining the conditions that sustain it. A new technological change can then weaken those conditions while the routine persists. The process described in Section 15 can begin again.

Proposition 16.8 (Routinization–lag cycle). *Successful trust adaptation can eventually create the pre-conditions for a later temporal mismatch because routinized practices persist after the conditions that made them reliable begin to change.*

This proposition gives the symbolic crisis a recursive historical structure. Social learning produces useful routines; useful routines reduce cognitive cost; reduced reflexivity makes those routines durable; later environmental change can then expose a gap between inherited practice and current conditions.

16.4 Guardians, Assurance, and Second-Order Trust

Large-scale trust increasingly depends on intermediaries. Shapiro’s analysis of impersonal trust emphasizes the importance of guardians, procedural constraints, policing mechanisms, entry restrictions, and insurance-like arrangements for relationships that lack dense personal embedding (Shapiro 1987). Yet these mechanisms create a recursive problem: the guardian of trust is also an object of trust.

This recursion becomes especially salient under symbolic abundance. A recipient may distrust the diagnosticity of an individual symbolic surface and therefore rely on a provenance provider, accreditation body, platform reputation system, institutional reviewer, identity service, audit process, or legal guarantee. The resulting arrangement shifts the trust question upward rather than eliminating it.

Definition 16.9 (Second-order trust). *Second-order trust* is reliance placed in the actors, institutions, protocols, or infrastructures that certify, monitor, constrain, or repair a first-order trust relation.

Second-order trust can scale coordination. It can also concentrate power, create new opacity, and generate dependence on infrastructures whose operation ordinary users cannot directly inspect. This is consistent with the provenance trust stack developed in Section 14: verification systems require their own signers, validators, standards, repositories, and institutions.

Proposition 16.10 (Guardian recursion). *Every institutional mechanism that reduces the verification burden of first-order trust creates a corresponding second-order question concerning the reliability, authority, incentives, and accountability of the trust-supporting infrastructure.*

The proposition does not imply an infinite regress that makes trust impossible. Social systems routinely stop verification at practical points. The diagnostic implication is narrower: historical adaptation redistributes vulnerability. A new assurance architecture can make one uncertainty manageable by concentrating reliance elsewhere.

16.5 Layered Adaptation and Mixed Trust Regimes

Historical change rarely replaces one trust regime completely with another. Older practices can persist alongside newer ones. A professional community may rely simultaneously on credentials, personal reputation, long-term collaboration, institutional affiliation, technical provenance, contractual recourse, and direct inspection. Different sectors can stabilize different combinations.

For this reason, the symbolic crisis is likely to produce *mixed trust regimes*. In one domain, personal process history may become more valuable. In another, institution-based assurance may dominate. In a third, platform-level system trust may substitute for direct knowledge of counterparties. Lumineau, Schilke, and Wang’s analysis of technological transformation similarly emphasizes shifts in the form, production, and targets of trust rather than disappearance of trust itself (Lumineau et al. 2023).

Definition 16.11 (Mixed trust regime). A *mixed trust regime* is a historically situated configuration in which several trust-production modes coexist and are assigned different weights across tasks, risks, institutions, and relationships.

Table 14: Mechanisms of historical trust adaptation and their characteristic tensions. The table describes possible functions rather than ranking the mechanisms.

Adaptation mechanism	Primary function	Characteristic tension
Process accumulation	Builds evidence through repeated interaction and observed conduct	High relational cost and limited scalability
Institutional assurance	Compresses verification through rules, credentials, audit, recourse, or certification	Dependence on guardians and second-order trust
Category recalibration	Revises the meaning assigned to familiar symbolic or social cues	Risk of stereotyping, overgeneralization, and slow perceptual change
Technical verification	Adds provenance, identity, integrity, or process information	Infrastructure opacity, privacy cost, and unequal access to verification
Routinization	Reduces cognitive and coordination costs after a practice stabilizes	Later vulnerability to interpretive hysteresis
Mixed-regime layering	Combines several trust sources across contexts and exposure levels	Complexity, inconsistency, and uneven institutional capacity

Mixed regimes help explain why the same society can display apparently contradictory responses. People can be skeptical of AI-mediated intimacy while accepting highly automated financial settlement. They can distrust a generated academic essay as evidence of student formation while trusting an institutionally verified automated calculation. They can demand human review in medicine while accepting autonomous filtering in low-stakes administrative communication. These patterns can be coherent once trust target, function, exposure, and institutional context are separated.

Table 14 highlights a recurring theme: every adaptation mechanism has costs. Process history is information-rich but expensive. Institutional assurance scales but creates second-order reliance. Technical verification can increase traceability while creating privacy and infrastructure concerns. Routinization lowers transaction costs while increasing the possibility of future lag. Historical adaptation therefore redistributes uncertainty rather than abolishing it.

16.6 Coordination Thresholds and Collective Meaning

A trust practice becomes socially powerful when participants can anticipate that others will interpret it similarly. This creates a coordination threshold. A personal convention can work bilaterally; a social convention requires enough mutual expectation that the practice carries stable meaning across interactions.

Consider a hypothetical marker indicating that a message has been reviewed and endorsed by the named sender. Its technical implementation could be trivial. Its social effect depends on shared understanding: what level of review occurred, what commitment endorsement creates, what responsibility follows if the message is false, and which institution can enforce the claim. Without those expectations, the marker is merely another symbol.

Definition 16.12 (Trust coordination threshold). A *trust coordination threshold* is the level of shared recognition at which a trust-supporting practice becomes sufficiently interpretable across actors that it can reduce verification and negotiation costs beyond isolated relationships.

The threshold need not involve majority adoption. A specialized profession can stabilize a convention within a bounded field. The important property is reciprocal expectation. Actors trust the practice partly because they expect other relevant actors to understand and enforce it in a similar way.

This helps explain why individual action, although meaningful, is insufficient for an era-scale symbolic transition. An individual can maintain a longitudinal record, disclose delegation, create richer

provenance, or privilege direct interaction. These practices become broadly trust-supporting only when recipients know how to interpret them and when social institutions create consequences for misuse. Historical adaptation therefore moves through collective stabilization.

Proposition 16.13 (Collective stabilization). *The scalability of a new trust practice depends on shared interpretation and enforceable expectation, not solely on the informational richness of the practice itself.*

This proposition also limits purely technical solutions. A technically sophisticated signal can remain socially weak if recipients lack shared conventions concerning its meaning. Conversely, a comparatively simple practice can become powerful once embedded in institutions, law, professional routine, and repeated social use.

16.7 Power and Unequal Adaptation Capacity

Trust regimes are also political-economic arrangements. Actors differ in their capacity to create, certify, and benefit from new trust practices. Large institutions can maintain audit systems, identity infrastructure, reputation databases, legal teams, or proprietary verification layers. Smaller organizations and individuals may rely on cheaper, more exposed symbolic channels. Historical adaptation can therefore alter the distribution of credibility.

The transition can also privilege incumbent institutions. When uncertainty about isolated symbolic performance rises, recipients may place greater weight on recognized brands, established universities, licensed professions, major platforms, or other institutional anchors. Section 11 described a related credential-dependence paradox. Historical trust adaptation can reinforce that tendency when institutional assurance substitutes for direct evidence.

At the same time, new technologies can lower entry barriers for verification or enable alternative communities to build trust infrastructures that were previously unavailable. The direction of change is therefore not predetermined. The analytical point is that adaptation capacity itself is unequally distributed.

Definition 16.14 (Adaptation-capacity inequality). *Adaptation-capacity inequality* is the unequal ability of actors to create, access, interpret, or enforce emerging trust practices during a transition in the evidential environment.

Working Claim 16.15 (Credibility redistribution). Historical adaptation to symbolic abundance can redistribute credibility toward actors who control recognized verification, institutional assurance, relational access, or the infrastructures through which new trust conventions are stabilized.

The claim links this section to the political economy developed in Section 13. Symbolic abundance can reduce some production barriers while increasing the strategic importance of scarce trust infrastructure. The historical formation of new trust practices is therefore also a contest over who can credibly certify, recognize, and enforce social meaning.

16.8 Historical Openness of the Emerging Trust Regime

The preceding analysis does not identify a single destination for trust after generative abundance. Several futures remain plausible. Institutions can strengthen verification and system trust. Communities can assign more weight to long-term relational history. Professional fields can redefine acceptable mediation. Some interactions can become increasingly automated while others become more deliberately human-intensive. New combinations can emerge that are difficult to describe using present categories.

This uncertainty is theoretically important. A historical transition should not be described as though its endpoint were already known. The symbolic crisis is occurring while the conventions that may eventually stabilize it are still being produced.

Proposition 16.16 (Historical openness of trust adaptation). *The adaptation of trust practices under generative abundance is path-dependent, institutionally mediated, and historically open. Present disruptions identify pressures on inherited trust arrangements more clearly than they determine the mature trust regime that will follow.*

This proposition also defines the boundary between the present diagnostic paper and the companion inquiry on response directions. The present paper can identify mechanisms through which trust practices historically adapt—reweighting, layering, institutionalization, routinization, and second-order assurance. It cannot infer from those mechanisms which future practice should be normatively preferred or which configuration will become socially dominant.

16.9 Diagnostic Implications

Historical trust adaptation reframes the symbolic crisis as a problem of institutional and perceptual transition. Symbols lose or change diagnostic force; actors experiment with new combinations of evidence; successful arrangements acquire shared meaning; institutions compress those arrangements into scalable assurances; and routinization can eventually make them background expectations. The process is slow because trust practices coordinate vulnerability, responsibility, and recourse across multiple actors.

Proposition 16.17 (Historical-adaptation diagnosis). *The symbolic crisis should be analyzed as a period of trust-regime adaptation in which inherited trust-production profiles are being reweighted and layered through new routines, institutions, technical assurances, and relational practices. The transition remains uneven because these practices acquire social force only through collective stabilization and because each new assurance creates additional sites of reliance, power, and possible failure.*

The section therefore sharpens the era-scale character of the symbolic crisis. Individual actors can respond locally, but the social meaning of a trust practice emerges historically through repeated interaction, institutional recognition, and shared expectation. The next section turns to the conceptual boundary of the diagnosis itself. If symbolic change, trust adaptation, and technological mediation are widespread, the category of *symbolic crisis* requires clear limits. Section 17 therefore examines the conditions under which the concept is informative and the conditions under which simpler descriptions should be preferred.

17 Boundaries of the Symbolic-Crisis Concept

The preceding sections have used the term *symbolic crisis* across communication, trust, expertise, education, credentials, culture, provenance, and political economy. The breadth of those applications creates a conceptual risk of its own. A category intended to diagnose a historically specific instability can lose explanatory value if every form of technological mediation, uncertainty, distrust, automation, or cultural change is redescribed as part of the same crisis. The structural role of this section is therefore restrictive. It specifies the conditions under which the symbolic-crisis concept adds explanatory value, distinguishes nearby phenomena that require different concepts, and states conditions under which the diagnosis should be narrowed or withdrawn.

The paper has consistently treated the crisis as a change in the relation between symbolic surfaces and the histories, capacities, commitments, relations, or forms of participation inferred from those surfaces. This relation is selective. A symbolic artifact can continue to communicate content, coordinate action, perform an institutional function, or support a valid transaction even while one historically familiar inference from the artifact becomes less reliable. The diagnostic object is consequently neither symbolic production in general nor artificial intelligence in general. It is a mismatch among generative pathways, inherited inference rules, and socially consequential reliance.

Table 15: Boundary conditions for a symbolic-crisis diagnosis. Each condition concerns a different part of the relation between production, interpretation, and consequence.

Boundary	Condition supporting the diagnosis	Alternative interpretation when the condition is weak
Generative change	Similar surfaces become reachable through materially broader or altered production histories	Ordinary symbolic ambiguity or stable mediation
Inherited inference	Recipients continue using the surface to infer effort, authorship, competence, attention, scarcity, sincerity, or related properties	Symbolic use without historical inference
Diagnostic mismatch	The inherited inference becomes less discriminating under the new production regime	Stable cue–property relation
Consequential reliance	The mismatch affects trust, evaluation, coordination, legitimacy, allocation, or value	Low-stakes representational abundance
Interpretive instability	Shared expectations and institutional conventions remain unsettled or contested	Restabilized convention under a new production regime
Domain specificity	The mechanism can be identified in a bounded practice, institution, or field	Overgeneralized civilizational diagnosis

Definition 17.1 (Symbolic-crisis diagnosis). A *symbolic-crisis diagnosis* applies to a domain when a material change in symbolic production broadens or alters the generative histories compatible with an observable symbolic surface, inherited practices continue to use that surface as evidence for underlying properties, and the resulting diagnostic mismatch has consequential effects on trust, coordination, valuation, assessment, legitimacy, or social recognition while interpretive conventions remain unsettled.

This definition contains several conditions that should remain analytically distinct. Let G_d denote the generative regime in domain d , I_d the prevailing interpretive rule, M_d the mismatch between plausible generative histories and the histories recipients infer, C_d the social consequences of that mismatch, and A_d the degree to which interpretive practices have restabilized. A symbolic-crisis claim concerns the configuration represented in Equation 59, rather than any one variable in isolation.

$$SC_d = \langle G_d, I_d, M_d, C_d, A_d \rangle. \quad (59)$$

The notation is classificatory rather than metric. It prevents the term *crisis* from being inferred directly from the presence of generative AI, from a decline in production cost, or from a single observed trust effect.

Table 15 makes the restrictive logic explicit. Symbolic abundance is relevant when it changes the evidential or relational work performed by a symbolic surface. A domain in which the generative history of an artifact carries little social significance may experience rapid automation without a corresponding symbolic crisis. Conversely, a domain can experience a strong crisis even when the absolute quantity of generated content remains modest if the affected symbolic cue carries unusually high institutional or relational weight.

17.1 Semantic Function and Historical Inference

The first boundary separates the semantic or performative function of a symbol from the historical inference attached to it. A message can accurately state a meeting time regardless of whether its wording was composed by a person, generated by a model, or produced through a template. A digitally signed document can continue to authorize a transaction even if other stylistic features of the document become less diagnostic of human effort. A credential can remain an authentic institutional record while the relation between that credential and some underlying capability becomes more contested. Austin's analysis of performative language already demonstrates that symbolic acts can derive force from conventions and institutional conditions rather than from the psychological history of the utterance alone (Austin 1962). The symbolic-crisis concept therefore concerns selected inferential functions rather than the continued existence or utility of symbols as such.

Proposition 17.2 (Selective diagnosticity). *A symbolic crisis in one inferential dimension does not entail a collapse of the semantic, coordinative, performative, or institutional functions of the same symbolic form.*

This proposition is important for avoiding a post-symbolic interpretation of the paper. Large-scale societies continue to require compression devices. Documents, credentials, brands, reports, signatures, scores, and standardized messages permit coordination among actors who lack direct access to one another's histories. The changing issue is the amount and type of additional inference that can reasonably be attached to those surfaces.

A highly fluent explanation illustrates the distinction. Its propositional content can be correct. Its pedagogical usefulness can be substantial. Yet its evidential relation to the speaker's unaided expertise can weaken when comparable fluency becomes reachable through heterogeneous human–AI configurations. The symbolic crisis lies in the altered inference from performance to capability, not in the mere existence of fluent explanation.

17.2 Boundary with Deception and Misinformation

The symbolic crisis is also distinct from deception and misinformation. Deception concerns an actor's attempt to induce a false belief or misrepresentation under some account of intent and audience. Misinformation concerns false or misleading content regardless of whether a generative transition has occurred. The symbolic crisis can arise when every proposition communicated is true.

Consider an automatically generated customer-service message that accurately states a refund policy and politely addresses the recipient by name. The message can be factually correct while recipients infer a degree of direct human attention that the production process did not involve. The relevant instability concerns the generative history inferred from the personalized surface. Conversely, a false handwritten letter can mislead without exhibiting the production-perception mismatch studied here.

Source-monitoring research is useful at this boundary because it distinguishes memory for information from memory for its source and generative context (Johnson et al. 1993). The present framework extends that distinction socially. The content of a symbol and the process from which recipients believe the symbol emerged are separable epistemic objects.

Working Claim 17.3 (Content-history separation). Symbolic-crisis effects can occur under truthful content when recipients draw unstable or outdated inferences about the history, agency, effort, participation, or institutional process behind that content.

This boundary keeps the framework from becoming a general theory of information disorder. Deep-fakes, propaganda, fraud, fabricated evidence, and synthetic misinformation may intersect the symbolic crisis, but their analysis also requires theories of deception, media systems, strategic communication, security, and political power that cannot be replaced by the present concept.

17.3 Boundary with General Trust Decline

The symbolic crisis should likewise remain distinct from a universal decline in trust. Earlier sections showed that AI-mediated communication can produce different trust effects across contexts, targets, and functions. Experimental findings of disclosure penalties or lower perceived authenticity in some settings (Schilke and Reimann 2025; Sahebi, Formosa, and Bankins 2026) provide evidence of unsettled norms; they do not establish a general social trajectory toward distrust.

Trust can be redistributed. A recipient can reduce relational trust in a personalized message while increasing operational trust in the organization that reliably resolves the underlying problem. An institution can place less evidential weight on a submitted essay while placing more weight on supervised performance, longitudinal work, or verified process records. A market can discount one form of stylistic rarity while intensifying trust in authenticated provenance. The symbolic crisis therefore concerns a reallocation of evidential weight as much as a decline in it.

Definition 17.4 (Trust redistribution). *Trust redistribution* is a change in the targets, cues, institutions, or relations through which reliance is organized after inherited symbolic indicators lose or change diagnostic force.

Trust redistribution can produce higher trust in some infrastructures while lowering trust in particular surfaces. This possibility is essential to the paper’s historical argument. A crisis of established symbolic inference can coexist with the construction of new assurance mechanisms and new objects of reliance.

17.4 Boundary with Authenticity Discourse

Authenticity is another neighboring concept with only partial overlap. Section 8 distinguished disclosure, authenticity, and legitimacy because each responds to a different problem. A recipient may judge an AI-assisted message authentic when the named sender genuinely endorses its content and regards the mediated expression as their own. Another recipient may reserve authenticity for direct human production. These disagreements matter, but the symbolic crisis extends beyond them.

A technically authentic credential can still become less informative about underlying capability. A culturally authentic artifact can exist in a market where stylistic scarcity has changed. A truthful and sincerely endorsed report can still conceal a delegation architecture that matters for judgments of expertise or responsibility. Authenticity is therefore one target of symbolic inference among several.

Proposition 17.5 (Authenticity boundary). *Authenticity judgments can form part of the symbolic crisis, while the crisis also includes changes in diagnosticity concerning competence, effort, attention, authorship, institutional process, scarcity, responsibility, and relational participation.*

This distinction also prevents a nostalgic theory of the human-made artifact. Human production can be formulaic, delegated, deceptive, or institutionally scripted. AI-assisted production can be deeply considered, endorsed, and embedded in a long history of human judgment. The diagnostic problem concerns the mapping between visible form and hidden generative structure rather than a categorical hierarchy between human and machine production.

17.5 Technological Causation and AI-Specific Scope

Generative AI is the historically salient catalyst for this paper because it expands the range and accessibility of symbolic transformations at unusual scale and speed. Yet the symbolic-crisis concept should not be defined by one technology. Templates, industrial reproduction, photography, recorded media, mass advertising, outsourced authorship, algorithmic personalization, automated translation, and institutional bureaucracy have all altered relations between symbolic surface and generative history in earlier periods.

The distinctive contemporary condition lies in the combination of breadth, quality, accessibility, speed, personalization, and delegation. Generative systems make it possible for many actors to reach symbolic surfaces previously associated with much narrower production histories. AI-mediated communication also allows intelligent agents to modify, augment, or generate messages on behalf of communicators (Hancock et al. 2020), while increasingly agentic systems can participate in interpretation and decision processes as well as surface generation.

Proposition 17.6 (Technological modesty). *Generative AI should be treated as a major contemporary accelerator of symbolic diagnostic instability rather than as the conceptual essence of the symbolic crisis.*

This formulation creates two safeguards. First, historical comparison remains possible. Second, causal claims must remain domain-specific. A change in trust, credentialing, cultural value, or communication should not be attributed to generative AI when platform incentives, organizational restructuring, economic inequality, regulatory change, or older forms of mediation provide a better account.

17.6 Institutional Anchoring and Stable Symbolic Relations

Some symbolic relations remain comparatively stable because institutions anchor them through procedures that generative abundance does not directly replace. A cryptographically verifiable signature can establish that a credential was issued by a particular authority. A legal registry can establish ownership or authorization. A supervised examination can narrow the range of possible production histories for a performance. Content-provenance systems can establish selected facts about the history and transformation of a digital asset (Coalition for Content Provenance and Authenticity 2026; Groth and Moreau 2013).

These infrastructures do not restore every inference attached to the symbol, as Section 14 argued. They can nevertheless stabilize particular relations between representation and history. A symbolic-crisis diagnosis should therefore identify which layer remains unstable after institutional anchoring rather than treating every mediated surface as equally ambiguous.

Definition 17.7 (Anchored symbolic relation). An *anchored symbolic relation* is a relation between a symbolic representation and a specified underlying fact that is stabilized through institutional procedure, technical attestation, controlled observation, or another assurance mechanism sufficiently independent of surface appearance.

Anchoring can substantially reduce one kind of ambiguity while leaving another intact. A digitally signed report can authenticate the issuing organization while revealing little about which employees or agents generated the analysis. A degree can authenticate institutional conferral while leaving the learner's current capability partly uncertain. A provenance credential can record transformation history while leaving truth, sincerity, or legitimacy unresolved. The relevant boundary is therefore inferentially granular.

17.7 Domain Scale and Uneven Crisis Intensity

The symbolic crisis can be local, sectoral, institutional, or cross-domain. Evidence of instability in academic writing does not establish an equivalent crisis in legal signatures. Changes in luxury advertising do not directly establish changes in scientific credibility. AI-mediated interpersonal messages may alter relational expectations more quickly than formal bureaucratic documents whose force depends on codified procedures.

The concept should therefore be indexed to a domain and inference target whenever possible. Let $SC_{d,p}$ denote a symbolic-crisis claim in domain d concerning property p . This notation reminds the analysis that the same surface can have different diagnostic profiles for different properties and that the same technology can produce different effects across institutions.

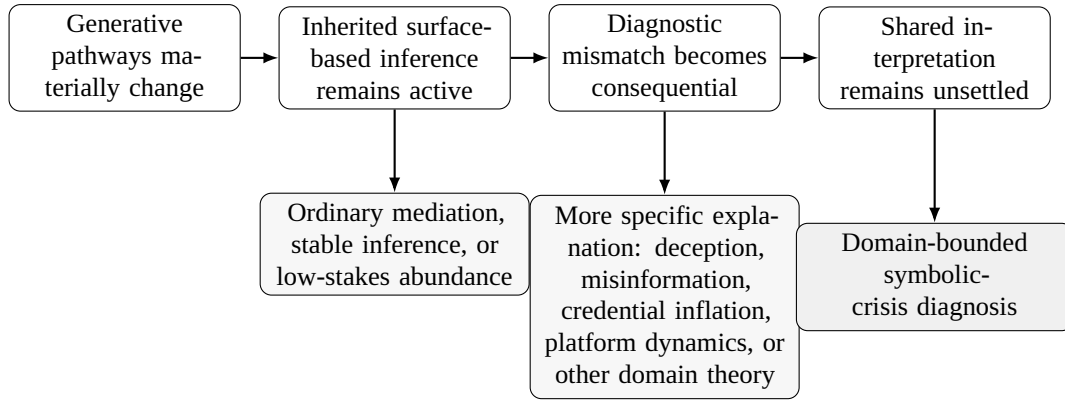


Figure 17: Conceptual filter for a symbolic-crisis diagnosis. The term is most informative when changed generative pathways, inherited inference, consequential diagnostic mismatch, and unsettled interpretation occur together. When these conditions are weak, more specific or ordinary descriptions should be preferred.

Proposition 17.8 (Domain-indexed diagnosis). *Evidence for symbolic diagnostic instability in one domain or for one inferred property does not by itself warrant a society-wide symbolic-crisis claim.*

A broader historical diagnosis becomes more plausible when similar surface–history mismatches appear across several domains and can be related to a shared change in generative conditions. The sections of this paper document such possible commonalities while retaining the need for domain-specific empirical work.

Figure 17 summarizes the concept as a restrictive filter. The diagram does not imply a deterministic sequence. Its purpose is methodological: the symbolic-crisis label should be earned by evidence about changed production, inference, consequence, and adaptation rather than inferred from technological novelty alone.

17.8 Transitional Status and Interpretive Restabilization

The term *crisis* also carries a temporal boundary. A society can make extensive use of generative systems after interpretive conventions have restabilized. If recipients, institutions, and producers share reliable expectations about what a symbolic form does and does not evidence, high levels of mediation need not constitute an ongoing crisis in the sense developed here.

Section 16 described trust adaptation as a process of experimentation, institutional layering, routinization, and second-order assurance. That analysis implies a possible exit from crisis. Once new inference rules, responsibilities, and verification practices become sufficiently stabilized, generative abundance can become an ordinary background condition. The symbolic environment may remain technologically dynamic while a particular mismatch recedes.

Definition 17.9 (Interpretive restabilization). *Interpretive restabilization* occurs when shared expectations, institutional practices, and assurance mechanisms sufficiently recalibrate to changed generative conditions that a symbolic surface again supports reasonably stable domain-specific inferences or is routinely interpreted with recognized limits.

Restabilization does not imply a return to prior conventions. A future regime may distribute warrant across surface form, provenance, institutional guarantees, longitudinal records, and relational history in ways that differ substantially from earlier practices. The diagnostic claim concerns the reduction of mismatch, not restoration of an older symbolic order.

The possibility of restabilization also prevents the paper from treating symbolic crisis as a permanent metaphysical condition. Generative technology can continue to evolve, and recurrent innovation

can generate repeated periods of mismatch. Whether contemporary societies face a temporary recalibration, a sequence of recurrent crises, or a more durable condition in which generative change persistently outpaces interpretive adaptation remains an open empirical question.

17.9 Normative Boundary of the Diagnosis

A final boundary concerns normativity. The symbolic-crisis concept describes instability in social inference. It does not independently establish that earlier high-friction production regimes were ethically preferable, that low-cost generation should be restricted, or that human-only production deserves privileged status.

Lower production friction can expand access to communication, translation, education, design, and professional presentation. It can reduce barriers associated with language, disability, time, or specialized technical skill. It can also increase ambiguity, shift verification burdens, and redistribute rents toward infrastructures controlling visibility and trust. The diagnosis is compatible with both sets of effects.

Proposition 17.10 (Diagnostic-normative separation). *Evidence of symbolic diagnostic instability establishes a need to examine changed inference conditions; it does not by itself establish a normative preference for restoring prior production frictions or limiting generative mediation.*

This separation is especially important because nostalgia can itself misdescribe the historical baseline. Earlier symbolic systems were never transparent windows into human effort, intention, or competence. They contained ghostwriting, institutional delegation, unequal access, scripted interaction, credential privilege, forgery, branding, and strategic performance. Generative AI changes the scale and distribution of these possibilities rather than introducing mediation into a previously unmediated social world.

17.10 Alternative Explanations and Conceptual Parsimony

A broad concept earns its place only when it explains a relation that narrower concepts leave fragmented. The symbolic-crisis framework adds value when several phenomena share the same structural pattern: a familiar surface remains socially legible while the generative histories compatible with that surface widen faster than the interpretive practices used to evaluate it. When that structure is absent, more specific theories should take priority.

Credential inflation can be studied as credential inflation. Platform concentration can be studied through platform political economy. Fraud requires theories of deception, incentives, and enforcement. Educational validity problems require assessment theory. AI-mediated communication requires communication theory. The symbolic-crisis concept should connect these domains only where the surface–history mismatch contributes explanatory leverage across them.

Research Obligation 17.11 (Conceptual parsimony). Applications of the symbolic-crisis concept should compare the diagnosis against narrower domain-specific explanations and retain the broader concept only when the common mechanism of generative diversification, inherited inference, and consequential mismatch explains otherwise disconnected observations.

This obligation also protects the empirical program developed in the next section. A theory that classifies every outcome as evidence of symbolic crisis becomes difficult to test. Domain-specific alternatives make comparative predictions possible.

17.11 Conceptual Exit Conditions

The boundaries developed above support explicit conditions under which the concept should be narrowed or withdrawn. A symbolic-crisis interpretation loses force when generative pathways remain substantially unchanged, when recipients no longer rely on the affected surface to infer the property at issue,

when the surface–history relation remains stable despite mediation, when the mismatch has little social consequence, or when interpretive conventions have restabilized around the new production regime.

The concept should also contract when observed effects are better explained by mechanisms that do not depend on symbolic diagnostic instability. If trust declines because an institution repeatedly violates commitments, the relevant explanation can remain institutional performance. If a luxury market changes because of a recession, scarcity displacement may play little role. If educational assessment fails because of poor test design independent of generative assistance, assessment validity remains the more direct account.

Proposition 17.12 (Conceptual exit). *The symbolic-crisis concept should be narrowed, localized, or withdrawn when the hypothesized mismatch between symbolic surface and generative-history inference does not discriminate among cases, does not predict consequential differences, or is consistently outperformed by simpler domain-specific explanations.*

Conceptual exit is part of theory construction rather than a concession after failure. A historically situated diagnosis should specify the conditions under which its own historical relevance ends.

17.12 Diagnostic Implications

The boundaries developed in this section clarify the paper’s scope. The symbolic crisis is a domain-sensitive and historically contingent instability in the social inference performed through symbols. It can coexist with accurate content, effective coordination, authentic endorsement, legitimate institutions, and useful generative assistance. Its intensity varies across properties, practices, and institutional settings. The diagnosis is strongest where changed generative pathways widen the range of histories behind a familiar surface, inherited inference rules remain active, consequences are socially significant, and collective interpretation has not yet restabilized.

Proposition 17.13 (Bounded symbolic-crisis diagnosis). *The symbolic crisis should be treated as a bounded diagnostic category for consequential mismatches among generative conditions, inherited symbolic inference, and unsettled social interpretation. Its explanatory value depends on identifying the specific domain, inferred property, institutional context, and temporal interval in which that mismatch operates.*

This bounded formulation prepares the empirical task. The next section asks what evidence would support, differentiate, or weaken the framework across domains. An empirical research program must measure changes in generative pathways, diagnosticity, trust targets, institutional response, temporal adaptation, and competing explanations while preserving the possibility that the symbolic-crisis hypothesis is locally absent or historically temporary.

18 Empirical Research Program

The preceding sections have developed the symbolic crisis as a bounded diagnosis of changing relations among generative pathways, symbolic surfaces, social inference, and institutional adaptation. The structural role of this section is empirical. It converts that diagnosis into a research program that can distinguish strong from weak cases, compare domains, identify mechanisms, and withdraw the framework when the proposed mismatch fails to appear. The objective is therefore not to produce a single index of “symbolic crisis.” It is to specify observable relations among production, interpretation, consequence, and adaptation that can be tested separately and in combination.

The empirical problem is unusually demanding because the framework concerns relations that are often hidden behind the final artifact. A polished email does not display how much attention, delegation, editing, or model assistance produced it. A degree certificate does not display the assessment ecology

Table 16: Core constructs for an empirical program on symbolic crisis. The proposed observables are examples rather than universal metrics.

Construct	Empirical object	Possible observables	Evidence weakening the claim
Generative-pathway diversity	Histories capable of producing a similar surface	Human-only, assisted, delegated, autonomous, templated, institutional, and mixed workflows	Production pathways remain substantively stable or produce visibly distinguishable surfaces
Inherited inference	Property recipients infer from a cue	Effort, authorship, attention, competence, sincerity, scarcity, participation, endorsement	Recipients already discount the cue or interpret it with stable limits
Diagnosticity	Relation between cue and inferred property	Posterior accuracy, calibration, discrimination, source attribution, confidence	Cue–property discrimination remains stable across changed production regimes
Consequential reliance	Decisions made through the inference	Trust, hiring, grading, purchase, delegation, legitimacy, allocation, cooperation	Miscalibration has little effect on decisions or welfare-relevant outcomes
Interpretive adaptation	Change in inference rules and assurance practices	Disclosure norms, verification routines, provenance use, supervised assessment, institutional redesign	Rapid and durable recalibration eliminates the mismatch
Burden redistribution	Location of new verification and interpretation costs	Time spent checking, audit costs, credential dependence, platform fees, review labor	Production savings do not shift meaningful costs to recipients or intermediaries

through which it was earned. A luxury image does not display the production history, social access, or institutional recognition that recipients may infer from it. Measurement must consequently reconstruct both sides of the symbolic interface: the generative history behind the surface and the inference made from the surface.

Definition 18.1 (Empirical symbolic-crisis unit). An *empirical symbolic-crisis unit* is the tuple defined in Equation 60.

$$\mathcal{U} = \langle d, s, p, i, t \rangle, \quad (60)$$

In Equation 60, d denotes a bounded domain, s a symbolic surface or cue, p the property inferred from that cue, i the institutional setting, and t the historical interval in which the inference is made.

The tuple prevents domain drift. A fluent paragraph can be highly diagnostic of semantic clarity while becoming weak evidence of unaided writing skill. A signed credential can remain strong evidence of institutional issuance while becoming a more contested indicator of a particular capability. A personalized message can remain operationally useful while becoming less diagnostic of direct human attention. Empirical claims should therefore identify the property being inferred rather than assigning a global credibility score to the symbol.

Table 16 makes a central methodological point: the framework can fail in several different places.

Generative pathways may diversify without changing recipient inference. Inference may change without producing consequential decisions. Institutions may recalibrate quickly enough that the mismatch becomes short-lived. Verification burdens may remain modest. These possibilities should be treated as empirical outcomes rather than anomalies to be explained away.

18.1 Generative-Pathway Reconstruction

The first empirical task is to observe the hidden production side of symbolic artifacts. Studies that classify artifacts only as “human” or “AI” risk reproducing the very compression problem analyzed in this paper. Section 6 showed that mediation can distribute goal formation, interpretation, judgment, generation, authorization, monitoring, and responsibility across different actors. Section 9 likewise showed that strong performance can arise from heterogeneous capability configurations. A useful empirical design should therefore record a production profile rather than a binary label.

A production profile can include the initiating goal, number and type of model interactions, delegation depth, human revision, external verification, authorization, elapsed time, institutional constraints, and the identity or role of accountable actors. In instrumented studies, these data can be collected directly from controlled workflows. In organizational field research, they can be reconstructed from version histories, prompt logs where ethically available, workflow software, interviews, or process observation. Provenance standards can support selected parts of this reconstruction, while Section 14 established that provenance itself has limited reach and should not be treated as a complete record of participation or responsibility (Buneman et al. 2001; Groth and Moreau 2013; Coalition for Content Provenance and Authenticity 2026).

Research Obligation 18.2 (Generative-history adequacy). Empirical studies of symbolic diagnosticity should measure the production history at the level required by the inference under study. A coarse “AI used” indicator is insufficient when the research question concerns attention, judgment, participation, authorization, or responsibility.

This obligation implies that different research questions require different histories. A study of authorship attribution may need a record of text generation and revision. A study of relational trust may need evidence of who attended to the recipient, who interpreted the incoming message, and who remained available for repair. A study of expertise may need repeated performance under changed cases rather than a complete writing log. Methodological adequacy depends on the target property.

18.2 Matched-Surface and Matched-History Designs

The strongest experimental designs should separate symbolic surface from generative history. A *matched-surface design* holds the artifact constant while varying the history participants are told or can verify produced it. Such a design can estimate how beliefs about effort, authorship, authenticity, trust, or legitimacy change when surface information remains fixed. Existing disclosure experiments already demonstrate the importance of this distinction: identical or comparable communication can be evaluated differently when AI involvement is disclosed (Schilke and Reimann 2025; Sahebi, Formosa, and Bankins 2026; Kirk and Givi 2025).

A complementary *matched-history design* holds the generative arrangement approximately constant while varying the surface quality, style, personalization, or institutional cue. This design asks whether familiar surface features continue to carry evidential weight after recipients know that the same production architecture can generate many different appearances. Together, the two designs permit decomposition of surface effects and history effects.

Let S denote the observable surface, H the generative history, and $\hat{P}$ a recipient’s inferred property. The history-conditioned empirical object is given in Equation 61.

$$P(\hat{P} \mid S, H), \quad (61)$$

The corresponding surface-only object is given in Equation 62.

$$P(\hat{P} \mid S). \quad (62)$$

The difference need not be interpreted as a universal measure of crisis. It estimates how much generative-history information changes inference for a specific property in a specific domain.

Hypothesis 18.3 (Production-diversification hypothesis). *Where similar symbolic surfaces become reachable through increasingly heterogeneous generative pathways, the diagnosticity of the surface for properties tied to production, including direct effort, unaided authorship, or direct attention, will tend to decline unless compensating institutional or contextual cues restore discrimination.*

The hypothesis is intentionally property-specific. It predicts less about semantic accuracy, legal validity, or task usefulness when those properties can be evaluated through other means. A finding that content quality remains stable or improves under generative assistance would therefore be compatible with declining diagnosticity for effort or direct participation.

18.3 Inference Calibration and Source Monitoring

A second empirical task concerns recipient calibration. Source-monitoring research distinguishes memory for content from memory for the conditions under which information was acquired (Johnson et al. 1993). The symbolic-crisis framework extends this problem from memory into social inference. Recipients may correctly understand a message while miscalibrating the amount of effort, expertise, attention, endorsement, or institutional process represented by its surface.

Calibration studies can ask participants to estimate these latent properties and compare their estimates with instrumented production histories. Relevant outcomes include absolute error, discrimination among histories, confidence, and overconfidence. Researchers should also examine whether participants learn from repeated feedback. Rapid recalibration would support a temporary-mismatch interpretation; persistent miscalibration would support stronger claims about interpretive hysteresis.

Definition 18.4 (Diagnostic calibration). For a property p , *diagnostic calibration* is the correspondence between recipients' confidence-weighted inferences about p and the empirically reconstructed production or institutional conditions relevant to p .

Calibration should be studied across demographic, professional, and institutional groups without presuming that one group possesses a privileged interpretation. Experts may better identify selected production traces while relying more heavily on domain conventions that themselves are changing. Novices may discount polished surfaces more broadly and thereby lose useful information. The empirical issue concerns the fit between inference and target property, not adherence to a predetermined interpretive rule.

18.4 Trust Profiles and Reallocation

Section 7 argued that trust is multidimensional and can be directed toward the principal, agent, delegation relation, institution, or message. An empirical program should preserve these targets. Aggregate trust scales can obscure reallocation: relational trust in a sender may fall while operational trust in an institution rises; confidence in content can remain high while authenticity judgments fall; trust in an AI system can increase while trust in the human principal's direct participation decreases.

Experimental and field studies should therefore estimate a trust vector rather than a single score. Existing findings of AI and disclosure penalties illustrate why this matters: lower authenticity or trust judgments under disclosed AI mediation do not establish a general collapse of epistemic or operational reliance (Schilke and Reimann 2025; Sahebi, Formosa, and Bankins 2026). Likewise, effort-bearing

interaction can influence trust through channels that differ from propositional accuracy (Chaudhry and Wald 2022; McEllin and Michael 2022).

Hypothesis 18.5 (Trust-reallocation hypothesis). *Generative mediation will more often reallocate trust across targets and functions than produce a uniform change in trust across the interaction.*

This hypothesis can be tested by measuring epistemic, operational, relational, and legitimacy-oriented trust simultaneously. Evidence of highly correlated movement across all targets and functions would weaken the reallocation account and support a more general trust-effect model.

18.5 Longitudinal Designs and Interpretive Adaptation

The temporal mismatch developed in Section 15 requires longitudinal evidence. Cross-sectional experiments can detect present inference differences; they cannot establish whether a society is entering crisis, adapting to it, or already restabilizing. Longitudinal designs should therefore track both production conditions and interpretive rules.

Several strategies are available. Repeated cohort studies can measure how people infer effort, authorship, attention, and authenticity from the same classes of cues over time. Institutional panels can follow changes in assessment rules, hiring practices, disclosure requirements, provenance adoption, or communication norms. Archival studies can compare the rhetoric surrounding earlier automation or mediation technologies with later routinized practice. Field studies around platform or organizational policy changes can examine whether new labels and verification infrastructures alter inference and whether those effects persist.

Let $I_t(s, p)$ denote the prevailing inference from surface s to property p at time t , and let G_t denote the generative regime. A temporal mismatch exists when change in G_t outpaces adaptation in I_t . The empirical program should therefore estimate both rates of change rather than infer lag from technological chronology alone.

Hypothesis 18.6 (Interpretive-lag hypothesis). *Diagnostic mismatch will be greatest during intervals in which generative pathways change faster than recipient inference rules and institutional assurance practices, and will decline where interpretive conventions restabilize around the new production regime.*

A persistent mismatch despite stable generative conditions would indicate that additional mechanisms are operating. Conversely, rapid interpretive recalibration would narrow the temporal scope of a symbolic-crisis diagnosis.

18.6 Institutional Anchoring and Comparative Domains

The framework predicts substantial domain variation because institutions can anchor the meaning of symbols through rules, audits, legal responsibility, supervised procedures, or material constraints. A medical license, an exam score, a private message, a luxury advertisement, and an online review should not be expected to respond identically to generative abundance.

Comparative studies can exploit this variation. Domains with strong institutional anchors can be compared with domains relying heavily on informal surface cues. Credential systems can be compared across assessment regimes with different levels of supervised performance. Communication contexts can be compared by consequence, relational expectation, and disclosure norm. Luxury and cultural markets can be compared by the relative importance of stylistic form, provenance, material scarcity, access, or recognized participation. Section 12's scarcity profile and Section 11's credential analysis provide construct families for such comparisons.

Hypothesis 18.7 (Institutional-anchoring hypothesis). *Where the socially relevant property is secured by durable institutional procedures that remain valid under changed production conditions, symbolic*

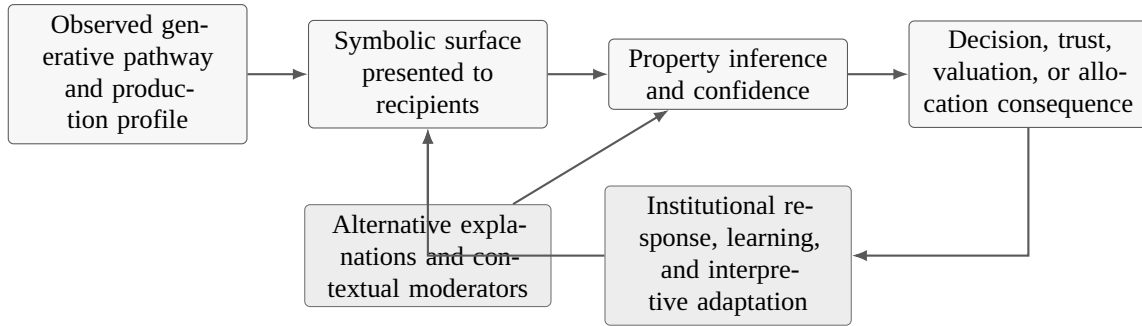


Figure 18: Empirical architecture for studying symbolic crisis. Research should observe production history, symbolic surface, recipient inference, downstream consequence, and subsequent adaptation while testing domain-specific alternative explanations.

diagnosticity will be more stable than where recipients rely primarily on surface cues whose production friction has fallen sharply.

The hypothesis does not assume that institutional anchors are accurate or just. An institution can stabilize an inference while reproducing exclusion, credential privilege, or outdated criteria. The empirical claim concerns stability of inference; normative evaluation remains separate.

18.7 Verification Burdens and Political Economy

Section 13 proposed that lower production costs can move scarcity and rents toward verification, visibility, recognition, and responsibility. This claim requires direct measurement. Studies should estimate who performs additional checking, how much time and money verification consumes, which infrastructures capture fees or attention, and whether these burdens fall unevenly across organizations and social groups.

Organizational studies can measure changes in review time, supervision, audit requirements, interview structure, or provenance infrastructure before and after adoption of generative tools. Platform studies can examine whether ranking, verification, or identity services become more valuable as content supply expands. Labor studies can investigate whether post-generation checking becomes a larger share of work even as drafting time falls. Market studies can test whether institutional affiliation, verified provenance, or exclusive access command larger premiums where stylistic production becomes abundant.

Hypothesis 18.8 (Verification-burden displacement hypothesis). *In domains where generative abundance materially reduces production friction while diagnostic uncertainty rises, part of the saved production cost will tend to reappear as verification, selection, audit, or credibility-management costs elsewhere in the system.*

The hypothesis can fail. Some domains may enjoy both cheaper production and cheaper verification because automated checking improves at similar speed. Others may simply tolerate greater ambiguity. Such cases would limit the political-economy generalization.

18.8 A Multi-Method Research Architecture

The mechanisms studied here are unlikely to be captured well by a single method. A multi-method architecture can combine controlled experiments, instrumented workflow studies, longitudinal panels, organizational field studies, archival reconstruction, platform data, interviews, and comparative institutional analysis. Each method observes a different layer of the surface–history relation.

Figure 18 summarizes the program as a recursive architecture. Production conditions shape symbolic surfaces; surfaces support property inferences; those inferences affect decisions; institutions and recipients learn from the resulting consequences; and that learning changes later interpretation. Alternative mechanisms can enter at several points. The empirical challenge is to identify where the proposed mismatch actually contributes explanatory leverage.

A particularly useful strategy is triangulation across layers. If a controlled experiment finds that AI disclosure reduces perceived authenticity, an organizational field study can test whether the same effect changes behavior under real consequences. Longitudinal evidence can then ask whether the effect persists as norms evolve. Provenance or workflow data can determine whether participants' inferred histories match actual production. The symbolic-crisis interpretation becomes stronger when these layers converge and weaker when effects disappear outside one experimental framing.

18.9 Falsification and Theory Contraction

The boundaries established in Section 17 require explicit empirical exit conditions. The framework should contract when the hypothesized mechanism fails to discriminate among cases. Several findings would weaken a symbolic-crisis interpretation: stable cue–property diagnosticity despite substantial generative diversification; rapid and durable adaptation of recipient inference; negligible consequences of miscalibration; verification costs that remain low or fall; absence of trust reallocation where delegation changes; or superior explanatory performance from narrower domain-specific mechanisms.

Proposition 18.9 (Empirical contraction principle). *The empirical scope of the symbolic-crisis framework should contract when changes in generative pathways fail to produce systematic differences in property inference, consequential reliance, adaptation, or burden redistribution after relevant alternative explanations are considered.*

Theory contraction should occur by domain and property rather than through an all-or-nothing verdict. Evidence may support the framework for direct-attention inference in personalized communication while rejecting it for semantic accuracy. It may support scarcity displacement in selected luxury contexts while rejecting it in markets where provenance and access remain unchanged. It may identify strong educational assessment effects while finding little change in institutionally supervised licensing.

18.10 Ethics of Empirical Investigation

The research program also creates ethical obligations. Studies of symbolic inference can manipulate disclosure, authorship beliefs, effort cues, and trust. In high-stakes domains, such manipulation can affect reputation, grading, employment, financial decisions, or interpersonal relations. Experimental convenience should therefore not reproduce the harms being studied.

Low-stakes simulations, transparent debriefing, synthetic or consented materials, and reversible experimental settings should be preferred where deception would impose meaningful cost. Instrumented workflow research should minimize collection of private prompts, personal correspondence, and proprietary content. Provenance studies should distinguish the value of traceability from an assumption that every participant should expose a complete history of their communicative or intellectual process. Privacy, anonymity, and safety can legitimately constrain verifiability.

Research Obligation 18.10 (Ethical observability). The empirical value of reconstructing generative histories should be balanced against privacy, anonymity, intellectual freedom, and the risk that comprehensive trace requirements create new surveillance or exclusion burdens.

This obligation is particularly important because one possible social response to symbolic ambiguity is to demand ever more documentation. An empirical program should study that tendency without presupposing that maximal observability is desirable.

18.11 Sequenced Research Agenda

A practical research program can proceed in stages. The first stage maps generative pathways and identifies historically important inferences attached to symbolic surfaces. The second stage measures diagnostic calibration under controlled variation in surface and history. The third stage tests decision consequences in field or quasi-field settings. The fourth stage compares institutional anchors and verification burdens across domains. The fifth stage follows adaptation longitudinally to determine whether mismatch persists, recurs, or restabilizes.

These stages need not occur in strict order, but they correspond to the architecture of the theory. Evidence of changed production alone establishes generative abundance. Evidence of changed inference establishes diagnostic instability. Evidence of consequential decisions establishes social importance. Evidence of institutional response and later recalibration establishes historical adaptation. The full symbolic-crisis diagnosis becomes strongest when these components can be observed together.

Proposition 18.11 (Defeasible empirical program). *The symbolic-crisis framework should be treated as a defeasible empirical program whose claims are evaluated at the level of specific symbolic surfaces, inferred properties, institutions, and historical intervals. Its value depends on demonstrating when generative diversification changes social inference and consequence, and equally on identifying domains in which symbolic relations remain stable or restabilize rapidly.*

The empirical program therefore preserves the paper's diagnostic restraint. Generative abundance creates a historically important opportunity to study how societies infer invisible histories from visible forms. Whether that opportunity constitutes a crisis, a temporary recalibration, or an ordinary extension of existing mediation must be established through comparative evidence rather than assumed from technological novelty. The next section turns from research design to the unresolved conceptual and historical problems that remain even after such evidence is gathered.

19 Open Problems in Symbolic-Crisis Research

The preceding sections have treated the symbolic crisis as a bounded and testable diagnosis of changing relations among generative pathways, symbolic surfaces, social inference, and institutional adaptation. The structural role of this section is to identify problems that remain unresolved even if the diagnosis proves empirically useful. These problems concern causal attribution, historical reconstruction, observer effects, norm change, institutional power, privacy, recursive symbolization, and the possibility that the categories used to describe the crisis are themselves altered by the transition. The objective is therefore neither to extend the framework indefinitely nor to convert every uncertainty into a new concept. It is to mark places where further theory, empirical work, or normative analysis is required before strong conclusions are justified.

The open problems also preserve an important distinction between a diagnosis and a solution. Sections 1–18 identify changing symbolic conditions and provide an empirical program for studying them. They do not establish a mature post-crisis trust regime. The possible roles of co-experience, trajectory records, longitudinal credibility, new forms of provenance, and revised educational practice remain topics for a companion inquiry. The present section therefore emphasizes unresolved constraints on interpretation rather than completing that later design work in advance.

Table 17 summarizes the central difficulties. Several are recursive. An institution may respond to ambiguous symbolic surfaces by introducing provenance or longitudinal records, only for those devices to become new symbolic surfaces whose interpretation requires additional trust. A profession may redefine acceptable AI participation, thereby changing what authorship or expertise means in that profession. Empirical observation can therefore alter the object of study by influencing disclosure norms, verification practices, and strategic behavior.

Table 17: Major open problems in symbolic-crisis research. The categories identify unresolved tasks rather than a complete research agenda.

Problem domain	Unresolved difficulty	Risk of premature closure
Historical causality	Separating generative-AI effects from longer changes in platforms, organizations, labor, credentialing, and media	Treating coincident changes as one AI-driven transition
Generative-history reconstruction	Observing hidden mixtures of prompting, delegation, revision, supervision, and institutional workflow	Reducing heterogeneous mediation to a binary AI-use label
Interpretive adaptation	Measuring how norms and cue meanings change across cohorts, professions, and institutions	Freezing current expectations into a permanent theory of trust
Normative co-evolution	Accounting for changes in authenticity, authorship, expertise, and legitimacy concepts during the transition	Evaluating future practices solely with present categories
Institutional power	Identifying who defines valid provenance, credentials, disclosure, and verification infrastructures	Treating assurance standards as politically neutral
Privacy and verifiability	Balancing credible histories with anonymity, confidentiality, and protection from surveillance	Making credibility conditional on pervasive traceability
Recursive symbolization	Understanding how new credibility devices themselves become compressible, gameable, or symbolic	Assuming that a new signal permanently escapes the crisis
Periodization	Determining whether current mismatch is transitional, recurrent, or structurally persistent	Declaring either permanent crisis or imminent restabilization too early

19.1 Historical Causality and Competing Transformations

The first open problem concerns causal attribution. Generative AI enters a social environment already transformed by platformization, remote communication, globalized professional labor, credential expansion, algorithmic ranking, influencer economies, templated customer service, outsourced writing, and longstanding institutional delegation. Many of the phenomena analyzed in this paper therefore have multiple historical causes.

A decline in the diagnostic value of polished prose may reflect generative models, widespread editing assistance, standardized professional templates, or changing educational practices. Rising dependence on credentials may reflect uncertainty created by AI-mediated performance, labor-market congestion, employer risk management, or institutional status competition. Increased demand for provenance may respond to synthetic media while also extending older authentication practices. A strong symbolic-crisis account must therefore discriminate between the contribution of generative abundance and adjacent transformations.

Research Obligation 19.1 (Historical-causal discrimination). Research on symbolic crisis should com-

pare generative-AI explanations with plausible longer-run and domain-specific causes, including organizational delegation, platform mediation, credential expansion, and pre-existing changes in symbolic production.

This obligation matters because the framework concerns a historical acceleration rather than a creation *ex nihilo*. Ghostwriting, advertising, strategic self-presentation, bureaucratic authorship, scripted service, and institutional signaling long predate contemporary generative systems. The empirical claim should concern changes in scale, accessibility, pathway diversity, or interpretive mismatch rather than the first appearance of mediation itself.

19.2 Hidden Histories and Measurement Intrusion

The second problem concerns generative-history reconstruction. The framework repeatedly distinguishes the symbolic surface from the history that produced it. Yet those histories are often difficult to observe without intrusive logging. Prompt archives, revision histories, model telemetry, workplace monitoring, and interaction traces can improve empirical reconstruction while creating privacy, labor, and governance problems of their own.

The measurement problem becomes sharper when the target property is relational. A log can record that a person approved a message without revealing whether the person seriously considered the recipient's situation. A revision history can establish human intervention without establishing judgment quality. An institutional workflow can identify authorization steps without revealing whether responsibility is meaningfully accepted. The most socially important features of a generative history may therefore remain only partially observable.

Proposition 19.2 (History-observability limit). *The generative history relevant to symbolic inference is not equivalent to the technically recordable workflow. Increased observability can reduce uncertainty about selected production events while leaving participation, judgment, attention, and responsibility partly underdetermined.*

This limit places a boundary on provenance-heavy solutions. More trace data can improve some inferences, but complete reconstruction of human participation may be neither possible nor normatively desirable.

19.3 Reflexive Observation and Strategic Adaptation

A third open problem arises because symbolic systems are reflexive. Once people learn that a cue is being used to infer competence, sincerity, effort, or human participation, they can strategically alter that cue. The same problem applies to new assurance mechanisms. If longitudinal records become valuable, actors can optimize for visible continuity. If disclosure labels acquire trust value, organizations can standardize minimally compliant disclosure. If process traces become important, workflow design can be rearranged to produce attractive traces.

This does not make all credibility devices useless. It means their evidential value depends partly on how costly, consequential, and difficult to strategically simulate the relevant history remains. The symbolic crisis can therefore recur at a higher level: a solution introduced to reconnect surface and history can itself become a new symbolic surface.

Definition 19.3 (Recursive symbolization). *Recursive symbolization* is the process through which a device introduced to reveal generative history, participation, or credibility becomes a socially interpreted symbolic surface whose own relation to the underlying history can later become compressed, standardized, or strategically managed.

Recursive symbolization is a central reason the companion paper should avoid presenting any single trust technology as a final solution. The research problem concerns evolving relations among evidence, practice, and institutional interpretation rather than discovery of a permanently ungameable signal.

19.4 Normative Co-evolution of Symbolic Categories

A fourth open problem concerns the categories through which the crisis is presently described. Terms such as *authorship*, *authenticity*, *expertise*, *attention*, and *human communication* carry historically specific expectations. Generative mediation may alter those expectations rather than merely violating them.

Consider authorship. A future scholarly community may regard extensive model use as compatible with authorship when problem selection, argumentative judgment, responsibility, and revision remain humanly situated. A professional organization may treat agentic communication as legitimate if delegation and accountability are institutionally stabilized. Educational systems may revise the relation between independent production and demonstrated learning. In such cases, the evaluative category itself changes.

Let N_t denote the normative-interpretive frame through which symbolic practices are evaluated at time t . Then historical adaptation is represented in Equation 63.

$$(H_t, N_t) \mapsto (H_{t+1}, N_{t+1}), \quad (63)$$

In Equation 63, H_t denotes prevailing generative and institutional practices. The transformation changes both the practice and the criteria through which the practice is interpreted.

Proposition 19.4 (Normative co-evolution). *A symbolic transition can change the evaluative categories used to judge the transition. Analysis should therefore distinguish criticism internal to present norms from claims intended to remain valid across changing conventions of authorship, authenticity, expertise, and legitimate mediation.*

This proposition does not imply relativism. Some concerns, including deception, unaccountable power, exploitation, or preventable harm, may persist across changing conventions. The open problem is to determine which norms concern contingent production expectations and which concern more durable ethical or political relations.

19.5 Institutional Power and the Politics of Restabilization

The fifth problem concerns power. Interpretive restabilization is unlikely to emerge from neutral social learning alone. Institutions can define acceptable authorship, issue credentials, control identity infrastructure, rank visibility, certify provenance, establish disclosure rules, and determine which forms of evidence count as legitimate. The post-crisis allocation of warrant will therefore be politically and economically structured.

Section 13 showed that symbolic abundance can shift value toward infrastructures governing visibility, verification, and consequence. Sections 14 and 16 showed that provenance and institutional assurance generate second-order dependence on validators and guardians. The unresolved problem is how new trust infrastructures distribute authority among states, platforms, professional bodies, educational institutions, employers, standards organizations, and communities.

Research Obligation 19.5 (Power analysis of restabilization). *Studies of interpretive adaptation should identify who gains authority to define valid disclosure, provenance, credentialing, authenticity, and verification, together with the costs imposed on actors who cannot or will not participate in those infrastructures.*

This obligation also prevents the concept of “restabilization” from becoming implicitly celebratory. A stable symbolic regime can still be exclusionary, surveillant, monopolistic, or unjust. Stability of inference and legitimacy of the institutional arrangement remain separate questions.

19.6 Privacy, Anonymity, and Unequal Verifiability

A sixth open problem follows directly. If credibility increasingly depends on richer histories, some actors will be better positioned than others to produce visible and institutionally recognized histories. Public professionals, large organizations, and well-resourced institutions can maintain repositories, version records, certifications, and verified accounts. Others may need anonymity, confidentiality, pseudonymity, or limited traceability for safety.

The tension can be expressed as a trade-off between verification reach and protective opacity. More historical evidence can improve selected trust judgments while increasing exposure to surveillance, reputational permanence, data extraction, or coercive profiling. The problem is especially serious when trust infrastructures become prerequisites for employment, education, finance, or public participation.

Proposition 19.6 (Verifiability inequality). *A trust regime that assigns increasing credibility to persistent and externally anchored histories can create unequal credibility conditions when actors differ in privacy needs, institutional access, historical visibility, or the capacity to maintain recognized records.*

The proposition makes clear why trajectory-based credibility, although promising as a response direction, cannot be treated as a universally benign replacement for symbolic trust. The design problem requires balancing longitudinal evidence with the right to begin again, remain private, or participate without exhaustive historical exposure.

19.7 Cross-Cultural and Cross-Institutional Variation

A seventh open problem concerns variation across symbolic cultures. The same surface can carry different meanings across professions, languages, legal systems, organizations, and communities. The meaning of a signature, polished prose, institutional affiliation, public disclosure, or personal response is culturally and institutionally situated.

The symbolic-crisis framework therefore requires comparative research. It should not assume that production friction, authenticity, personal attention, expertise, or disclosure are weighted similarly across contexts. In some settings, institutional mediation is expected and does little to weaken trust. In others, direct personal authorship is central to the meaning of the exchange. Certain professional domains already normalize collaborative drafting, while others assign strong authorship significance to apparently individual production.

Research Obligation 19.7 (Comparative translation). Cross-domain applications should reconstruct the local symbolic expectations and institutional conventions through which a surface acquires meaning before treating changes in generative history as evidence of a common crisis mechanism.

This obligation also applies to historical comparison. Contemporary symbolic practices should not be evaluated through an idealized baseline in which earlier communication was assumed to be transparent, individual, and unaided.

19.8 Durable Human Scarcity and Agentic Sociality

An eighth problem concerns the future of relational scarcity. This paper has assumed that human attention, co-presence, exposure, and participation remain bounded even when symbolic production becomes abundant. That assumption appears plausible in the present environment, but increasingly autonomous agents may alter the social organization built around those limits.

Long-lived agents can maintain interaction histories, remember preferences, coordinate continuously, and represent persons or organizations across time. Such systems may become objects of operational trust in their own right. The unresolved question is whether this development merely adds another layer of

delegated agency or produces more durable forms of agentic sociality with their own expectations of continuity, responsibility, and relationship.

The paper does not need to resolve the metaphysical status of artificial agents in order to identify the empirical problem. If people begin to rely on persistent agents as stable interaction partners, distinctions among trust in a principal, trust in an agent, and trust in a delegation relation may themselves evolve. The trust topology developed in Section 7 should therefore be treated as historically revisable.

19.9 Periodization and the Duration of Crisis

A ninth open problem concerns the temporal status of the diagnosis. Section 15 treated the present condition as a mismatch between fast-changing production regimes and slower-changing interpretive habits. Section 16 described possible institutional adaptation. Yet the rate of generative innovation may remain sufficiently high that interpretive conventions repeatedly lag behind new capabilities.

Three broad historical patterns remain possible. First, contemporary mismatch may be transitional, followed by substantial interpretive restabilization. Second, societies may experience recurrent local crises as new domains and capabilities destabilize previously settled cues. Third, a more persistent condition may emerge in which the generative space changes faster than common symbolic interpretation can reliably stabilize.

Working Claim 19.8 (Periodization uncertainty). The current framework does not establish whether symbolic crisis is a temporary transition, a recurrent historical pattern, or a durable feature of high-velocity generative systems. This question requires longitudinal evidence about both technological change and institutional adaptation.

The choice among these patterns matters for theory. A transitional diagnosis emphasizes adaptation. A recurrent diagnosis emphasizes resilience of interpretive institutions. A persistent diagnosis would require a more fundamental account of social coordination under continuously shifting generative conditions.

19.10 Relation to the Companion Response Inquiry

A final open problem concerns the boundary between diagnosis and response. Several possible directions have appeared repeatedly: co-experience, longitudinal records, trajectory-based credibility, richer provenance, selective human participation, new educational assessment, and institutions that make delegated agency more legible. These directions remain incomplete because each introduces new costs, inequalities, or opportunities for recursive symbolization.

The analytical sequence rejected here is stated in Equation 64.

$$\text{symbolic crisis} \longrightarrow \text{single replacement signal.} \quad (64)$$

The more plausible open-ended sequence is stated in Equation 65.

$$\text{diagnosis} \longrightarrow \text{plural experiments} \longrightarrow \text{institutional learning} \longrightarrow \text{historically contingent restabilization.} \quad (65)$$

The companion inquiry can examine those experiments without assuming that one practice must replace symbols generally. Symbols remain necessary compression devices for large-scale social coordination. The design problem concerns how their evidential role can be redistributed and supplemented under changed production conditions.

Proposition 19.9 (Open response field). *The diagnosis of symbolic crisis does not determine a unique post-crisis trust regime. Candidate responses should be evaluated as historically situated attempts to redistribute evidential weight across symbols, trajectories, relations, institutions, and forms of participation.*

19.11 Theoretical Implications

The open problems identified here preserve the framework's historical and empirical modesty. The symbolic crisis can be a useful diagnosis only if it remains sensitive to alternative causes, hidden histories, strategic adaptation, changing norms, institutional power, unequal verifiability, cultural variation, agentic evolution, and the possibility of its own historical expiration. Several of these problems are recursive: the instruments created to solve symbolic ambiguity can themselves become symbolic, the institutions that stabilize trust become new trust targets, and the categories used to judge mediation can change through mediation itself.

The paper therefore ends its diagnostic development without claiming a final social settlement. Its strongest contribution is narrower: generative abundance can widen the set of histories compatible with familiar symbolic surfaces faster than shared interpretive practices adapt, producing domain-specific instability in trust, communication, expertise, value, and institutional inference. The next and final section summarizes this argument and clarifies what the concept can contribute without exceeding its evidential reach.

20 Conclusion

This paper has examined a historical problem produced by a widening gap between symbolic surfaces and the generative histories from which those surfaces emerge. Generative systems increase the number, speed, polish, personalization, and recombining of available symbolic artifacts. The resulting change is important because many social practices have long relied on symbolic form as a compressed interface to otherwise inaccessible histories of effort, attention, expertise, authorship, commitment, scarcity, institutional process, and participation. When similar surfaces become reachable through increasingly heterogeneous production pathways, those inherited inferences can become less stable even while the symbols themselves remain useful for communication, coordination, certification, and exchange.

The term *symbolic crisis* has been used here for this diagnostic condition. The concept describes neither a disappearance of symbols nor a universal collapse of trust. It describes a historically uneven reorganization of the relations through which symbolic appearances support social inference. The crisis becomes most visible where production regimes change rapidly, inherited interpretive habits remain active, and the resulting mismatch has meaningful consequences for credibility, communication, expertise, education, value, or institutional legitimacy. Sections 17 and 18 therefore placed explicit limits on the concept. A symbolic-crisis interpretation should contract when simpler explanations are more adequate, when diagnostic relations remain stable, or when new conventions have already achieved substantial restabilization.

20.1 From symbolic abundance to diagnostic instability

The first major contribution of the paper is the distinction between symbolic abundance and symbolic diagnosticity. Section 2 characterized symbols as forms of social compression. A credential, message, report, style, signature, or institutional mark can reduce the cost of coordination by allowing recipients to act without reconstructing the full history behind the surface. This compression has always been lossy, yet its social usefulness depends partly on sufficiently stable relations between the surface and the properties inferred from it.

Section 3 showed how generative systems alter this condition by lowering selected forms of production friction. The resulting abundance is multidimensional. It can reduce the time, skill, iteration, or coordination needed to produce certain symbolic outputs while leaving other constraints intact. Production therefore becomes more heterogeneous rather than uniformly effortless. Similar surfaces can emerge from direct human production, AI-assisted drafting, delegated generation, institutional workflows, or more autonomous agentic systems.

Section 4 translated this heterogeneity into a problem of inference. The relevant issue is property-specific. A polished report may remain highly informative about its stated content while becoming weaker evidence of how much direct attention, unaided expertise, or individual authorship produced it. A personalized message may still coordinate action effectively while supporting a less stable inference about relational investment. The paper therefore treated diagnosticity as a profile rather than a single global quantity.

Working Claim 20.1 (Diagnostic reallocation). Symbol inflation can redistribute the evidential weight carried by symbolic surfaces across different inferred properties. A resulting crisis is strongest where inherited interpretation continues to assign substantial weight to cues whose generative pathways have become substantially more heterogeneous.

This formulation avoids a generalized suspicion toward symbols. Symbolic forms remain indispensable to large-scale social coordination. The analytical task is to identify which inferences remain well supported, which have weakened, and which have migrated toward other forms of evidence.

20.2 Communication, agency, and trust

The second contribution concerns communication. Section 5 argued that successful communication cannot be reduced to correct symbol transmission. Communication can include interpretation, attribution, mutual updating, exposure to response, commitment, and responsibility. A message can therefore remain propositionally successful while becoming relationally thin. The distinction matters because generative systems allow semantic performance to scale independently from some forms of human participation that recipients may previously have inferred from the message itself.

Section 6 developed this problem through delegated agency. AI mediation can distribute goal formation, interpretation, judgment, linguistic generation, authorization, monitoring, and responsibility across several human and artificial participants. One coherent voice on the symbolic surface can therefore correspond to several different agency topologies. The phrase “AI-assisted” is consequently too coarse for many communicative and ethical judgments. What matters is which functions have been delegated and how that delegation is represented to the recipient.

Trust inherits this complexity. Section 7 distinguished trust in the principal, the delegated agent, the delegation relation, the institution, and the propositional content. These targets can move independently. A system may be operationally reliable while remaining relationally thin. A principal may retain moral responsibility while delegating most of the expressive work. A recipient may trust the factual content while remaining uncertain about attention, sincerity, or personal endorsement.

Section 8 then separated disclosure, agency legibility, authenticity, and legitimacy. A simple disclosure statement can improve knowledge about mediation without specifying who interpreted, decided, authorized, or stands behind the resulting communication. Transparency can therefore improve epistemic conditions while leaving the deeper social meaning of the exchange unsettled.

Proposition 20.2 (Communicative decomposition). *Under generative mediation, the symbolic unity of a message increasingly fails to guarantee unity of communicative agency. Trust judgments should therefore be sensitive to the distribution of participation, authorization, responsibility, and relational exposure behind the surface.*

The proposition remains diagnostic. It does not imply that every delegated message requires exhaustive process disclosure. It identifies the source of a new ambiguity whose practical treatment will vary with context, stakes, and institutional norms.

20.3 Expertise, achievement, and institutional compression

A third contribution extends the diagnosis from interpersonal communication to expertise, education, and credentials. Section 9 distinguished epistemic performance from the expertise profile of the person

presenting that performance. Generative assistance can improve the quality of an answer, analysis, or report without producing a proportional change in the user's underlying domain knowledge, calibration, transfer capacity, or error-recognition ability. At the same time, genuine expertise can remain present within AI-assisted work. The relevant analytical problem concerns where competence is located and what the final performance can legitimately evidence about it.

Section 10 developed the parallel distinction between symbolic achievement and learner formation. Essays, presentations, solutions, and scores can remain meaningful products while their relation to independent understanding becomes more context dependent. The educational problem is therefore partly one of construct validity: institutions must decide what an assessment is intended to evidence when the production process includes powerful generative mediation.

Credentials add another compression layer. Section 11 showed that degrees, licenses, titles, and institutional affiliations reduce the cost of verifying long histories of assessment and participation. Symbol inflation may increase reliance on such institutional compression when isolated performances become less diagnostic. Yet the meaning of credentials can itself be affected if the upstream educational and evaluative processes that produced them are changing.

The resulting recursive pattern is summarized in Equation 66.

$$\begin{aligned} \text{performance ambiguity} &\longrightarrow \text{greater reliance on institutional compression,} \\ &\longrightarrow \text{greater attention to the histories that institutions construct and verify.} \end{aligned} \quad (66)$$

The diagnostic reallocation therefore reaches beyond individual artifacts. It can shift attention toward the institutional conditions that make symbolic compression credible in the first place.

20.4 Scarcity, value, and political economy

A fourth contribution concerns cultural and economic value. Section 12 argued that expanding symbolic production can reduce scarcity at the level of symbolic form while leaving other scarcity dimensions intact. Materiality, process, provenance, access, participation, temporality, recognition, and relational history can remain constrained. The resulting process was described as *scarcity displacement*: a reorganization of which constraints carry social and economic significance.

Scarcity displacement is neither automatic nor value preserving. Abundance can dilute previously valuable forms, increase the value of selected constraints, create new categories of distinction, or broaden participation. In luxury and culture, easily reproducible style may increase the relative importance of provenance, access, recognized participation, or biographical continuity. In other domains, abundance can simply reduce rents attached to symbolic form.

Section 13 extended this analysis to political economy through the concept of bottleneck displacement. Lower production costs can coexist with persistent scarcity in attention, verification, distribution, ranking, recognition, and responsibility. Economic power can therefore migrate toward infrastructures that control these remaining constraints. Platforms, credentialing bodies, provenance systems, reputation mechanisms, distribution channels, and validation institutions can become more important even as symbolic production itself becomes cheaper.

Working Claim 20.3 (Consequence bottlenecks). The political economy of symbol inflation is shaped by the location of remaining bottlenecks between symbolic production and social consequence. Abundant symbol generation can coexist with concentrated power over visibility, verification, recognition, and institutional legitimacy.

This claim also clarifies why expanding symbolic productive capacity does not automatically produce democratic equality. Production capacity and consequence capacity can be distributed very differently.

20.5 Provenance, temporal mismatch, and historical adaptation

The fifth contribution concerns the historical character of changing diagnosticity and the narrower conditions under which it becomes a crisis. Section 14 examined provenance as one response to surface ambiguity. Provenance can narrow the set of plausible generative histories and strengthen traceability. Its limits are equally important. Verified provenance does not establish truth, quality, legitimacy, participation, or responsibility by itself. Provenance also creates second-order trust in the institutions, validators, standards, or cryptographic systems that maintain it.

Section 15 placed these problems in time. Production conditions can change faster than shared interpretation. People and institutions continue using inference rules formed under earlier regimes, even after the statistical relations supporting those rules have weakened. The result is interpretive carryover and, in some cases, interpretive hysteresis. This temporal mismatch is central to the historical meaning of symbolic crisis. The condition is produced by the coexistence of new generative possibilities with older habits of symbolic reading.

Section 16 then treated adaptation as a collective institutional process. Trust practices evolve through experimentation, layering, routinization, and the formation of new background expectations. The process is recursive because new assurance mechanisms become new trust targets. A provenance standard requires trust in the validator. A credential system requires trust in its assessment process. A disclosure convention requires shared expectations about what disclosure means.

The paper therefore resists a simple replacement story in which one new signal solves the crisis. Historical adaptation is more likely to produce mixed trust regimes combining symbolic surfaces, institutional assurance, provenance, relational evidence, trajectory information, and domain-specific practices. The composition of these regimes remains open.

20.6 The remaining meaning of symbolic crisis

The framework retains the term “crisis” in a qualified sense. The diagnosis is strongest where a familiar symbolic surface continues to carry inherited social meaning while the generative histories compatible with that surface have expanded substantially. Crisis therefore refers to instability of interpretation and warrant, together with the institutional and economic consequences of that instability. It does not require complete breakdown.

Section 17 specified four conditions that should jointly support the diagnosis: changed generative pathways, inherited inference, consequential diagnostic mismatch, and unsettled interpretation. Section 18 translated these conditions into a research program organized around particular domains, surfaces, inferred properties, institutions, and historical intervals. Section 19 then identified the unresolved questions that remain: hidden histories, strategic adaptation, recursive symbolization, changing norms, unequal verifiability, cross-cultural variation, persistent agents, and uncertainty about the duration of the crisis itself.

Working Claim 20.4 (Symbolic-crisis diagnosis). A symbolic crisis exists where generative change widens the set of histories compatible with socially important symbolic surfaces faster than shared interpretive and institutional practices can recalibrate the inferences drawn from those surfaces. The resulting instability is domain specific, historically situated, and open to later restabilization.

This formulation preserves both the explanatory ambition and the limits of the concept. It recognizes a structural transformation without treating every use of generative AI as evidence of crisis.

20.7 Historical openness and the companion inquiry

The paper has remained primarily diagnostic because the mature social response is still historically open. Several directions already appear plausible: richer provenance, agency legibility, redesigned assessment,

selective co-experience, institutional verification, trajectory records, and new norms of human participation. Each direction also creates new costs. Provenance can increase surveillance. Trajectory records can privilege actors with visible and institutionally recognized histories. Co-experience is resource intensive. Verification can concentrate power in new guardians. Disclosure can improve accuracy while changing perceptions of authenticity. None of these mechanisms can be treated as a universal successor to symbolic trust.

The companion inquiry should therefore study response practices as experiments in redistributing evidential weight. One especially promising concept is trajectory-based credibility: when isolated surfaces become less diagnostic, temporally distributed histories may provide stronger evidence of sustained participation, competence, or commitment. Yet trajectory itself can become symbolized, optimized, surveilled, and unequally accessible. It belongs in an open design space rather than as a final solution.

The broader historical lesson is that societies have repeatedly built large systems of coordination by compressing complex relations into symbols. The present transition reveals the contingency of some of those compression rules. Symbol inflation makes especially visible the difference between a symbolic artifact and the history, participation, judgment, and responsibility that social actors may infer from it. Generative AI is one powerful accelerator of this condition, alongside older and continuing forms of strategic symbolic production. The likely response is therefore a long process of institutional and perceptual adaptation rather than a single technical correction.

The concluding orientation of this paper is one of diagnostic humility. Symbols remain central to social life. Their meanings and evidential functions are changing unevenly. The task is to understand where symbolic surfaces have become poor guides to the histories we continue to infer from them, where institutions can still stabilize those relations, and how new practices may redistribute credibility without reproducing the same ambiguity in another form. The symbolic crisis names the interval in which these relations are becoming visible as historical constructions and are once again available for collective revision.

References

- Akerlof, George A. (1970). “The Market for “Lemons”: Quality Uncertainty and the Market Mechanism”. In: *The Quarterly Journal of Economics* 84.3, pp. 488–500. DOI: 10.2307/1879431.
- Araki, Satoshi and Takehiko Kariya (2022). “Credential Inflation and Decredentialization: Re-examining the Mechanism of the Devaluation of Degrees”. In: *European Sociological Review* 38.6, pp. 904–919. DOI: 10.1093/esr/jcac004.
- Arrow, Kenneth J. (1973). “Higher Education as a Filter”. In: *Journal of Public Economics* 2.3, pp. 193–216. DOI: 10.1016/0047-2727(73)90013-3.
- Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Clarendon Press.
- Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman (2025). “Generative AI without Guardrails Can Harm Learning: Evidence from High School Mathematics”. In: *Proceedings of the National Academy of Sciences* 122.26, e2422633122. DOI: 10.1073/pnas.2422633122.
- Bourdieu, Pierre (1977). *Outline of a Theory of Practice*. Trans. by Richard Nice. Cambridge: Cambridge University Press. DOI: 10.1017/CB09780511812507.
- (1984). *Distinction: A Social Critique of the Judgement of Taste*. Trans. by Richard Nice. Cambridge, MA: Harvard University Press.
- Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond (2025). “Generative AI at Work”. In: *The Quarterly Journal of Economics* 140.2, pp. 889–942. DOI: 10.1093/qje/qjae044.

- Buneman, Peter, Sanjeev Khanna, and Wang-Chiew Tan (2001). "Why and Where: A Characterization of Data Provenance". In: *Database Theory — ICDT 2001*. Vol. 1973. Lecture Notes in Computer Science. Springer, pp. 316–330. DOI: 10.1007/3-540-44503-X_20.
- Chaudhry, Shereen J. and Kristina A. Wald (2022). "Overcoming Listener Skepticism: Costly Signaling in Communication Increases Perceived Honesty". In: *Current Opinion in Psychology* 48, p. 101442. DOI: 10.1016/j.copsy.2022.101442.
- Coalition for Content Provenance and Authenticity (Apr. 2026). *Content Credentials: C2PA Technical Specification, Version 2.4*. Technical specification. URL: <https://spec.c2pa.org/>.
- Collins, Harry and Robert Evans (2007). *Rethinking Expertise*. Chicago: University of Chicago Press. DOI: 10.7208/chicago/9780226113623.001.0001.
- Collins, Randall (2019). *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Columbia University Press. ISBN: 9780231192354.
- Dell’Acqua, Fabrizio, III McFowland Edward, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine C. Kellogg, Saran Rajendran, Lisa Kraye, François Candelon, and Karim R. Lakhani (2026). "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality". In: *Organization Science*. DOI: 10.1287/orsc.2025.21838.
- Fleckenstein, Johanna, Jennifer Meyer, Thorben Jansen, Stefan D. Keller, Olaf Köller, and Jens Möller (2024). "Do Teachers Spot AI? Evaluating the Detectability of AI-Generated Texts among Student Essays". In: *Computers and Education: Artificial Intelligence* 6, 100209. DOI: 10.1016/j.caeai.2024.100209.
- Goffman, Erving (1959). *The Presentation of Self in Everyday Life*. Garden City, NY: Doubleday.
- Goldhaber, Michael H. (1997). "The Attention Economy and the Net". In: *First Monday* 2.4. DOI: 10.5210/fm.v2i4.519.
- Goldman, Alvin I. (2001). "Experts: Which Ones Should You Trust?" In: *Philosophy and Phenomenological Research* 63.1, pp. 85–110. DOI: 10.1111/j.1933-1592.2001.tb00093.x.
- Groth, Paul and Luc Moreau (2013). *PROV-Overview: An Overview of the PROV Family of Documents*. W3C Working Group Note. World Wide Web Consortium, 30 April 2013. URL: <https://www.w3.org/TR/prov-overview/>.
- Hancock, Jeffrey T., Mor Naaman, and Karen Levy (2020). "AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations". In: *Journal of Computer-Mediated Communication* 25.1, pp. 89–100. DOI: 10.1093/jcmc/zmz022.
- Headworth, Spencer and Jeremy Freese (2016). "Credential Privilege or Cumulative Advantage? Prestige, Productivity, and Placement in the Academic Sociology Job Market". In: *Social Forces* 94.3, pp. 1257–1282. DOI: 10.1093/sf/sov102.
- Johnson, Marcia K., Shahin Hashtroudi, and D. Stephen Lindsay (1993). "Source Monitoring". In: *Psychological Bulletin* 114.1, pp. 3–28. DOI: 10.1037/0033-2909.114.1.3.
- Kane, Michael T. (2013). "Validating the Interpretations and Uses of Test Scores". In: *Journal of Educational Measurement* 50.1, pp. 1–73. DOI: 10.1111/jedm.12000.
- Kenney, Martin and John Zysman (2016). "The Rise of the Platform Economy". In: *Issues in Science and Technology* 32.3, pp. 61–69. URL: <https://issues.org/rise-platform-economy-big-data-work/>.
- Kestin, Greg, Kelly Miller, Anna Klates, Timothy Milbourne, and Gregorio Ponti (2025). "AI Tutoring Outperforms In-Class Active Learning: An RCT Introducing a Novel Research-Based Design in an Authentic Educational Setting". In: *Scientific Reports* 15, 17458. DOI: 10.1038/s41598-025-97652-6.
- Kirk, Colleen P. and Julian Givi (2025). "The AI-Authorship Effect: Understanding Authenticity, Moral Disgust, and Consumer Responses to AI-Generated Marketing Communications". In: *Journal of Business Research* 186, p. 114984. DOI: 10.1016/j.jbusres.2024.114984.

- Kong, Chun Yin (2026). "The Delegation Illusion: Why Deploying Autonomous AI Agents Does Not Diminish Principal Responsibility". In: *AI and Ethics* 6, 518. DOI: 10.1007/s43681-026-01383-x.
- Kopytoff, Igor (1986). "The Cultural Biography of Things: Commoditization as Process". In: *The Social Life of Things: Commodities in Cultural Perspective*. Ed. by Arjun Appadurai. Cambridge: Cambridge University Press, pp. 64–92. DOI: 10.1017/CB09780511819582.004.
- Lee, Garim and Hye-Young Kim (2024). "Human vs. AI: The Battle for Authenticity in Fashion Design and Consumer Response". In: *Journal of Retailing and Consumer Services* 77, 103690. DOI: 10.1016/j.jretconser.2023.103690.
- Luhmann, Niklas (1979). *Trust and Power*. Chichester: John Wiley & Sons. ISBN: 0471997587.
- Lumineau, Fabrice, Oliver Schilke, and Wenqian Wang (2023). "Organizational Trust in the Age of the Fourth Industrial Revolution: Shifts in the Form, Production, and Targets of Trust". In: *Journal of Management Inquiry* 32.1, pp. 21–34. DOI: 10.1177/10564926221127852.
- Matthias, Andreas (2004). "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata". In: *Ethics and Information Technology* 6.3, pp. 175–183. DOI: 10.1007/s10676-004-3422-1.
- McEllin, Luke and John Michael (2022). "Sensorimotor Communication Fosters Trust and Generosity: The Role of Effort and Signal Utility". In: *Cognition* 224, p. 105066. DOI: 10.1016/j.cognition.2022.105066.
- Messick, Samuel (1995). "Standards of Validity and the Validity of Standards in Performance Assessment". In: *Educational Measurement: Issues and Practice* 14.4, pp. 5–8. DOI: 10.1111/j.1745-3992.1995.tb00881.x.
- Möllering, Guido (2006). *Trust: Reason, Routine, Reflexivity*. Amsterdam: Elsevier. ISBN: 9780080448558.
- Mühlhoff, Rainer (2026). "From Delegation to Moral Abdication: Classifying Large Language Model Uses by Judgment and Epistemic Control". In: *AI & Society*. DOI: 10.1007/s00146-026-03281-6.
- Neugebauer, Martin, Jan Paul Heisig, and Thijs Bol (2026). "What Does Successful University Graduation Signal to Employers? A Factorial Survey Experiment on Sheepskin Effects". In: *European Sociological Review* 42.2, pp. 165–182. DOI: 10.1093/esr/jcaf028.
- Noy, Shakked and Whitney Zhang (2023). "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence". In: *Science* 381.6654, pp. 187–192. DOI: 10.1126/science.adh2586.
- Ogburn, William F. (1922). *Social Change with Respect to Culture and Original Nature*. New York: B. W. Huebsch.
- Rochet, Jean-Charles and Jean Tirole (2003). "Platform Competition in Two-Sided Markets". In: *Journal of the European Economic Association* 1.4, pp. 990–1029. DOI: 10.1162/154247603322493212.
- Sahebi, Siavosh and Paul Formosa (2025). "The AI-Mediated Communication Dilemma: Epistemic Trust, Social Media, and the Challenge of Generative Artificial Intelligence". In: *Synthese* 205, p. 128. DOI: 10.1007/s11229-025-04963-2.
- Sahebi, Siavosh, Paul Formosa, and Sarah Bankins (2026). "The AI Penalty and Disclosure Paradox: Trust, Authenticity and Knowledge Uptake in AI-Mediated Communication". In: *Computers in Human Behavior: Artificial Humans* 8, 100304, pp. 1–11. DOI: 10.1016/j.chbah.2026.100304.
- Santoni de Sio, Filippo and Jeroen van den Hoven (2018). "Meaningful Human Control over Autonomous Systems: A Philosophical Account". In: *Frontiers in Robotics and AI* 5, p. 15. DOI: 10.3389/frobt.2018.00015.
- Santoni de Sio, Filippo and Giulio Mecacci (2021). "Four Responsibility Gaps with Artificial Intelligence: Why They Matter and How to Address Them". In: *Philosophy & Technology* 34, pp. 1057–1084. DOI: 10.1007/s13347-021-00450-x.

- Schilke, Oliver and Martin Reimann (2025). "The Transparency Dilemma: How AI Disclosure Erodes Trust". In: *Organizational Behavior and Human Decision Processes* 188, p. 104405. DOI: 10.1016/j.obhdp.2025.104405.
- Schnackenberg, Andrew K. and Edward C. Tomlinson (2016). "Organizational Transparency: A New Perspective on Managing Trust in Organization–Stakeholder Relationships". In: *Journal of Management* 42.7, pp. 1784–1810. DOI: 10.1177/0149206314525202.
- Shapiro, Susan P. (1987). "The Social Control of Impersonal Trust". In: *American Journal of Sociology* 93.3, pp. 623–658. DOI: 10.1086/228791.
- Skowronski, John J. and Donal E. Carlston (1987). "Social Judgment and Social Memory: The Role of Cue Diagnosticity in Negativity, Positivity, and Extremity Biases". In: *Journal of Personality and Social Psychology* 52.4, pp. 689–699. DOI: 10.1037/0022-3514.52.4.689.
- Spence, Michael (1973). "Job Market Signaling". In: *The Quarterly Journal of Economics* 87.3, pp. 355–374. DOI: 10.2307/1882010.
- Suchman, Mark C. (1995). "Managing Legitimacy: Strategic and Institutional Approaches". In: *Academy of Management Review* 20.3, pp. 571–610. DOI: 10.5465/amr.1995.9508080331.
- To, Rita Ngoc, Yi-Chia Wu, Parichehr Kianian, and Zhe Zhang (2025). "When AI Doesn't Sell Prada: Why Using AI-Generated Advertisements Backfires for Luxury Brands". In: *Journal of Advertising Research* 65.2. DOI: 10.1080/00218499.2025.2454120.
- Veblen, Thorstein (1899). *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. New York: The Macmillan Company.
- Wang, Xujia, Billy Sung, and Ian Phau (2024). "How Rarity and Exclusivity Influence Types of Perceived Value for Luxury". In: *Journal of Brand Management* 31.6, pp. 576–592. DOI: 10.1057/s41262-024-00359-8.
- Yin, Xieyu and Fenghua Wen (2026). "When Trust Attracts Fraud: AI and Trust Arbitrage". In: *Economics Letters*. Available online 12 September 2026, p. 113234. DOI: 10.1016/j.econlet.2026.113234.
- Zucker, Lynne G. (1986). "Production of Trust: Institutional Sources of Economic Structure, 1840–1920". In: *Research in Organizational Behavior*. Vol. 8. JAI Press, pp. 53–111.