

Archively Research Note 01

When Does an AI Proposal Become Archival Description?

A human reviewer is necessary, but not sufficient. Professional control depends on who may accept a machine proposal, what context and criteria they use, what choices they can actually make, and what evidence and correction path the institution keeps afterwards.

Saad Muhammad

6 September 2026

Version 1.0

About this note: abstract, scope and review

Machine systems now produce text that looks like archival description: transcripts, candidate dates, titles, scope and content notes, subject terms, biographical statements. The standard assurance is that a person reviews the output. This note argues that the assurance is too cheap to be worth much, and proposes an institutional acceptance boundary in its place. A machine produces a proposal. An institution produces a description. Between the two sit one or more recorded acceptance events, governed by institutional roles and rules, taken against the evidence available at the time and inspectable afterwards. The note sets out six proposed conditions for evaluating whether review functions as an institutional control, argues that review depth should be set primarily by intended use and consequence of error and that a confidence measure cannot by itself establish fitness for acceptance, shows what retained evidence for one accepted statement looks like, and converts the argument into questions that name no product.

This note is the author's synthesis of the published standards and guidance. The framework it proposes is the author's own.

Author

Saad Muhammad, Founder, Archively

Saad Muhammad is the founder of Archively, which sells software to archives.

Review

Published for open review. Sources, version history and substantive corrections remain visible. Send corrections to researchnotes@archively.ai.

Editorial note and method

Editorial note

Saad Muhammad is the founder of Archively, which sells software to archives.

This note is a critical synthesis of published archival standards, professional codes and guidance, and governance frameworks from outside the sector. It reads these sources against the way machine-assisted description is framed in product and procurement discussions. The institutional acceptance boundary and six conditions are the author's proposed framework.

Sources are ranked by their standing in this argument. The ledger at the foot of the page makes that ranking visible and distinguishes archival standards, professional guidance and frameworks from other sectors. Every citation links to a ledger entry.

The note is published for open review. Sources, version history and substantive corrections remain visible. Send corrections to researchnotes@archively.ai.

The disagreement is real

Whether a machine should draft descriptive text is a real disagreement, and it is not a disagreement between people who have understood these systems and people who have not. Three positions are worth separating. Each can be held firmly by a serious person, and the divisions below are the author's way of organising the argument rather than a report of what the profession currently believes.

The first is refusal. Description is an interpretive act, performed by a qualified person, on behalf of an institution that can be held to it afterwards. A system with no professional formation and no answerability cannot perform that act, and putting its output in the institution's voice is a category error rather than an efficiency. The profession's own ethics do not settle that either way. The Code of Ethics treats description as an exercise of professional judgement and asks the archivist to be transparent about that role by documenting descriptive work and collections policy decisions, and it says nothing at all about machines [Society of American Archivists](#). Refusal is compatible with it, and so is the third position below. It is not a stage institutions pass through on the way to agreeing with suppliers, and this note treats it as a legitimate outcome rather than as resistance to be managed.

The second is constrained experiment. Use these systems where the output stays close to the source and an error is cheap, visible and reversible. Transcription, handwriting recognition, candidate dates, format and language identification. Keep interpretation manual. Revisit in two years, when there is more evidence and fewer sales calls.

The third is conditional provisional use. Let a system propose descriptive text, keep the proposal marked as provisional, and require a person with the right authority to accept it before a reader ever sees it.

This note argues from the third position, and it is worth being blunt about what that costs. Conditional use puts a new load on the institution. Somebody has to define the conditions. Somebody has to hold the authority. Somebody has to have the hours. An institution that cannot fund all three is better served by the first position than by a version of the third that exists only in a policy document.

The three positions are further apart on method than on principle. Each of them can be held while accepting that a statement published under the institution's name is the institution's statement, whatever produced the first draft. The argument is about what has to happen in between.

One distinction has to be made before any of that can be argued, because collapsing it is what makes the rest of this note necessary. Review can fail while the proposed text is perfectly acceptable. The quality of what a system produces and the quality of the control exercised over it are separate variables, and a good rate on the first says nothing about the second. What the control is worth depends on whether the role holding it has the context, the authority and the hours to inspect an ordinary item, reject it, and send it somewhere else. This note has no evidence about how those roles are staffed in current implementations, so it treats capacity as something an institution has to specify and test rather than as something the author claims to have seen.

“Human in the loop” is too cheap a claim

The phrase does an enormous amount of work for very little money. It asserts that a person was present. It says nothing about whether that person could have changed the outcome, and it is offered in procurement conversations as though those were the same fact.

Presence is not control. A control has parts, and every part can be missing while a person remains, technically, in the loop. Does the reviewer hold the authority to reject, or only to edit? Do they see the context needed to judge the proposal, or only the proposal itself? Do

they have the competence for this particular decision, which may turn out to be a rights decision wearing a descriptive costume? Do they have the time to disagree, given the queue behind them? Are the criteria written down, or improvised per item and forgotten by Thursday? Is there any record afterwards of what they decided and why they decided it?¹

Governance work outside the sector has been explicit about this for years, and archivists should feel free to borrow the vocabulary without adopting the worldview. The European regulation treats oversight as a set of capacities held by a person: to interpret the output, to stay alert to the tendency to over-rely on it, to disregard, override or reverse it, and to interrupt or stop the system. The separate requirement that whoever exercises oversight hold the necessary competence, training, authority and support sits not in that article but in the one that binds the organisation deploying the system European Union. A cross-sector risk framework does the same work for the organisation around the person, treating documented roles and responsibilities, training, human oversight, monitoring, feedback, appeal and override as arrangements an organisation makes rather than as properties the software has NIST.

None of that is archival doctrine, none of it was written with a fonds in mind, and nothing here claims that either instrument reaches archival description. Both are imported for one narrow purpose, which is that they give names to the difference between a person being present and a person being in control.

The profession has since said something of its own. The ARA Code asks that any use of these systems in the creation, processing or presentation of records that materially affects their informational or evidential value be made transparent to users, requires such processing to be undertaken with appropriate human oversight and quality control, and leaves the member responsible for the accuracy of whatever was generated on their instruction Archives and Records Association. That is binding on the association's individual members rather than advisory, which is worth saying plainly. What it does not do, and does not set out to do, is say what oversight has to consist of before an institution is entitled to call it oversight. Appropriate is left to the institution, which is the right place for it and also the place this note's argument starts.

There is a second thing the phrase never says, which is who the person is. It names no role, no level of authority and no quantity of time, and none of those follow from the fact that somebody was present. Who receives the queue, and with what authority and capacity, varies by institution. This note treats the role, its authority and its capacity as things an institution has to specify and test, which is the argument for writing them

down: a control that has not been assigned to anybody in particular is not a control, it is an expectation.

Key point

A person who can only accept or edit is not an oversight mechanism. They are a throughput mechanism with a signature attached. The test is whether refusing is a real, supported and survivable option under the rate and the staffing the institution has actually committed to, rather than one that exists in the policy and nowhere else.

Extraction is not description

Treating machine output as one category is the root mistake in this area. A transcript and a biographical note are not the same kind of claim. Reviewing them the same way wastes effort in one direction and invites harm in the other, and it leaves the institution with a risk register in which every line says “AI output”.

The useful axis is distance from the source. At one end the system is reading marks that are physically present on an object. At the other it is asserting things about people, events and significance that no amount of looking at the object would settle. Appraisal sits beyond the far end, and nothing in this note proposes moving it.

Output	What the system is doing	Distance from the source	What review has to catch
Transcript text	Reading marks on a page	Very close	Misreadings, dropped lines, silent normalisation
Dates and named entities	Matching patterns against reference sources	Close	Confident matches to the nearest famous name
Image caption	Naming what is visible	Middle	Things described that are not in the frame
Subject or authority links	Attaching a record to a controlled vocabulary	Middle	Plausible links to the wrong entity, at scale
Title and scope note	Summarising material it has not appraised	Far	Invented events, wrong emphasis, borrowed voice
Biographical or administrative history	Asserting facts about people and bodies	Far	Unsupported claims about real individuals

Output	What the system is doing	Distance from the source	What review has to catch
Sensitivity or rights signal	Predicting harm or legal status	Far	Errors in both directions, silently

The right-hand column is the point. Those are different failure modes and they need different readers.² A misread character is caught by anyone who can read the original. A wrong authority link is caught by someone who knows the vocabulary and the collection. An invented biographical claim may not be caught inside the institution at all, because it is fluent, plausible and about a person no member of staff has any reason to know.

Titles deserve a specific note, because they are where the temptation is strongest. A title in a finding aid is not a headline. It is a required element with rules behind it: it separates a formal title, which appears on the material itself, from a devised one, and it asks for professional judgement in deciding when a title has to be devised rather than transcribed DACS, Title. The standard does not say who or what may draft the words, so a machine-drafted title is not ruled out by the rules. It is owed the same assessment against them, and against the collection the material sits in, that a devised title gets from the archivist who wrote it. The output can be grammatical, evocative and wrong at exactly the level a reader cannot check.

Published guidance for the sector goes a certain distance and then stops. Candidate titles, dates and short summaries, suggested subject terms and proposed groupings are to be treated as non-authoritative suggestions, and all outputs are to be reviewed by archivists Colavizza and Jaillant. What that leaves open is which of those outputs an institution may release provisionally, at what level of description, and against what verification burden. Those are institutional policy choices, and the rest of this note is an argument about how to make them inspectable rather than an attempt to make them on anybody's behalf.

The institutional acceptance boundary

Most accounts of machine-assisted description have one gate. The system writes something, a person approves it, it appears in the catalogue. One gate hides four separate decisions, and the confusion between them is where institutions get hurt.

A proposal is not a description. It is a candidate produced by a system that holds no authority. This note proposes that it be retained as generated and kept apart from whatever the accepted text becomes, so that anyone can later see what was actually proposed rather than what survived editing. No standard cited here requires that. The

reason for asking anyway is that the moment a proposal is silently overwritten by its accepted form, the institution has destroyed the only evidence that would let it audit its own review.

A disposition (*the recorded outcome of a review decision: accept, modify then accept, reject, defer, escalate or stop*) is what a person does with that candidate. Only the first two produce an institutional statement. The rest are outcomes too, and a system that cannot record a rejection is not recording review. It is recording throughput and calling it review, which is the same substitution the whole note is about.

Acceptance is not publication. An accepted working description can still fail quality review, and it can still be held back by a rights or access decision made by a different role for different reasons. Collapsing those into one act is how a sensitivity problem becomes a public problem.

Publication is not finality either. Correction is a state in the model, not an embarrassment outside it. This part is not new and does not need arguing. The conceptual model the profession has been building treats agents, the activities they carry out, the rules they apply and the dates on which they act as entities in their own right; it treats an archival description as a record resource in its own right, produced by a describing activity, attributed to an agent, dated, revised over time, and supported by sources cited either for the whole description or for a single statement within it; and it already counts a mechanism as a kind of agent Records in Contexts. The ontology built on that model declares the mechanism a subclass of agent, held apart from a person and from a position, and describes it as something that may generate or modify records Records in Contexts Ontology. All of that supports representing what a system did. None of it hands the system any authority. The encodings go further than most people notice. A maintenance event carries its type, its date, a description of what changed and the agent responsible, with the agent type recorded as human, machine or unknown EAD3. Preservation metadata has recorded events, agents and their roles for years, software included. It comes from preservation rather than from description and says as much itself, and it is borrowed here as a working pattern rather than as a requirement that anything be evidenced PREMIS.

What the boundary names, then, is not a moment of transformation and not one universal click. A machine output becomes part of an institution's description through one or more recorded acceptance events, governed by institutional roles and rules. Such an event may cover a single statement, a whole record, or a defined batch handled under a stated rule. It may be a disposition recorded item by item, or acceptance into a working description with quality control following separately, or a decision taken by a role rather

than by a person whose name a reader ever sees, or one event in a revision history that carries on for decades. Description is cumulative, distributed, inherited and repeatedly revised, and a boundary that only worked for one archivist accepting one sentence would not describe the work.

The purpose of the event is narrow. It marks when the institution assumed responsibility for the use of an output, and it preserves enough evidence for that decision to be inspected and revised afterwards. It does not make the description final and it does not make it correct. To say that an institution has accepted a statement is to say that it will answer for it, and nothing more. It means that when a researcher writes to ask where the date came from, there is an institution on the other end of the letter rather than a system that has since been replaced. Every archivist reading this knows a catalogued statement can be wrong for fifty years. The boundary is about who owns the mistake, not about preventing it.

Figure 1

The institutional acceptance boundary

Evidence and context			
Source object		Collection context	
Relevant local policy		Authority or reference source	
Uncertainty or missing evidence			
Machine proposal			
No one has asserted this			
“Letters relating to the harbour dedication, 1962”			
Type	Generated	Status	Retained
Scope and content note	4 March 2026	Provisional	Exactly as generated
No status in the catalogue			
Authorised disposition			
• Accept		• Modify, then accept	
Reject		Defer pending evidence	
Escalate to another role		Stop or exclude the task	
Institutional assertion begins here			
Controlled record states			
1	Accepted working description	Internal. Not yet published.	
2	Quality review	A separate decision.	
3	Rights and access clearance	A separate decision.	
4	Publication	Public archival description.	
5	Correction or new version	Publication is not finality.	
↑	A correction returns to state 1 and is decided again.		
Retained decision history: what was proposed, what was decided, by whom, and what changed			

Source: Author’s proposed framework. The boundary and the dispositions are the author’s synthesis, not a requirement of the sources named: DACS 8.1 Description Control, Records in Contexts on documenting description, the AI Preparedness Guidelines for Archivists by Colavizza and Jaillant, published by the Archives and Records Association, and NIST oversight and override concepts. Example material is fictional. Verified 30 July 2026 archively.ai/research-notes/01

A machine output should become an institutional statement only through an authorised, evidenced decision. Acceptance is not the same as publication.

A single review gate folds two different decisions into one. Separating them is the author’s proposal, and it makes four distinctions visible: a proposal is not description, acceptance is not publication, quality review and rights clearance are distinct decisions, and publication is not finality.

Caveat

It is worth being exact about what is being claimed. The parts are drawn from the standards and codes cited; the arrangement is the author's. Nor are the standards silent about machines. That would be the easier argument to make, and it is not true. The descriptive standard provides an element, classified as added value rather than as a minimum requirement, for recording the sources used, the rules or conventions applied, the agent responsible, and the dates of creation and revision, and its rule on that last point names a generative artificial intelligence system alongside a person as something that may create or revise a description DACS, Description Control. The profession can already record that a machine took part. What no standard sets out is the institutional condition under which the output of one may be accepted, and that narrower gap is what this note is about. The boundary described above, and the conditions in the next section, are the author's proposed framework and not a requirement established by any source cited for their parts. If this reads as ordinary quality assurance with a new label, the request is the same: apply the ordinary quality assurance, in writing, and say who signed.

Six proposed conditions for meaningful review

If presence is not control, something has to replace it. What follows is the author's proposed test for whether review is doing institutional work. Its parts draw on archival descriptive control and on governance practice from outside the sector, and several of them have support in the sources cited elsewhere in this note, but no standard, professional body or framework has adopted these six as a set or called them a minimum. They are a proposal, and they are offered to be argued with.

The proposed conditions

1. **Defined scope and non-use boundary.** The institution states what the system may propose, for what material, for what intended use, and which decisions stay outside its reach entirely. A non-use boundary is a positive statement and someone has to sign it, because a scope defined only by what the software can do is not a scope.

2. **Relevant source and collection context.** The reviewer can see what they need in order to judge the proposal: the object, the collection context, the local rules, the authority sources, and what the system did not have access to. A proposal presented without its context is a request for agreement, not a request for judgement.
3. **Competent and authorised reviewer.** The role holds appropriate knowledge for this particular decision and delegated authority to accept, modify, reject, defer or escalate it. Competence for a transcript is not competence for a sensitivity signal, and one generic reviewer role covering every class of output is a sign that the classes have not been thought about.
4. **Real decision power and sufficient capacity.** The reviewer is not confirming a default. There are usable alternatives, and there are hours: a stated number of items per reviewer per day that somebody has actually costed, held against a stated staffing level. A rate set by the size of the queue rather than by the work is not capacity, it is a deadline wearing capacity's clothes.
5. **Criteria proportionate to use and consequence.** Review depth follows the type of output, its intended use, the likelihood and cost of an error, and whether correction is possible later. A confidence measure may help decide what to look at first, and professional guidance already contemplates thresholds for routing low-confidence output to a person Program for Cooperative Cataloging, but no such measure establishes by itself that an output is fit for an institution to accept. The criteria are written down before the queue starts, in a form a second reviewer could apply to the same item and reach the same answer.
6. **Retained evidence, monitoring and correction.** The institution can reconstruct what was proposed, what informed the decision, who made it, what changed between proposal and accepted text, and how a later error was reported and fixed. Monitoring means someone looks at the aggregate: rejection rates, edit distances, and which classes of proposal keep coming back wrong.

Figure 2

A person present, or review as an institutional control

AREA	WEAK HUMAN-PRESENCE CLAIM	MEANINGFUL INSTITUTIONAL CONTROL	EVIDENCE A COMMITTEE CAN REQUEST
Scope	A person reviews AI output.	The permitted tasks, materials, uses and exclusions are written down.	Approved use-case and non-use statement.
Context	The source file is visible.	Collection context, local rules and supporting sources are at hand.	Context and source references.
Reviewer	A staff member approves it.	The role holds the competence and the delegated authority for this decision.	Role definition, training, delegation.
Decision power	Accept, edit and reject buttons exist.	The reviewer can defer, escalate, stop the task and challenge the default.	Decision options and escalation path.
Capacity	Every output is reviewed.	Workload and interface permit attention rather than routine confirmation.	Staffing assumption, review time, override and edit data.
Criteria	Low-confidence results are checked.	Review depth follows intended use and consequence, not the confidence score.	Field-specific rules and tolerances.
Evidence and correction	An audit log is kept.	Proposal, evidence, decision, revision and correction can be reconstructed.	Decision history and correction record.

Review becomes a control only when the person has context, competence, authority, capacity, real options, and a recorded path to correction.

Seven areas where the phrase “human in the loop” is usually doing more work than the arrangement behind it. The middle column is the claim worth making. The right column is what a director or a procurement committee can ask to see, which is the difference between a claim and a control.

Author’s proposed framework, a synthesis rather than a standard. Areas informed by the Description Control element in DACS and by NIST governance concepts of roles, training and human oversight.

Last verified 30 July 2026

Condition four is the one most easily satisfied on paper, and its failure leaves nothing behind that an interface could display. The European regulation identifies over-reliance on the output of a system as a risk that oversight arrangements have to address, which is why it asks whoever exercises oversight to remain alert to automation bias (*the tendency to accept a system's suggestion because it is the suggestion in front of you*) European Union.

That is a governance framework naming a risk, which is not the same as research measuring one. This note does not have the measurement, and it will not pretend that a regulation citing a phenomenon is evidence of its size in an archival reading room.²

Two things these conditions do not settle. The first is who decides what risk is acceptable. Not the supplier. The institution decides, through its practitioners, with the specialists a particular collection requires, with the communities a record describes where they can be reached, and inside whatever legal and ethical obligations already bind it. A supplier that offers a default risk setting is offering to make that decision on the institution's behalf, and it should be declined in those words.

The second is ethics. These are process conditions. An institution could satisfy all six and still publish a description that is accurate, authorised, fully evidenced and indefensible, because whether a record should be described in public at all, and in whose terms, is not a workflow question. The conditions make it possible to find out who decided. They do not make the decision good.

Key point

Most of what is in this list is ordinary descriptive quality practice under different names. That is the argument, not a hole in it. What is different is that the describing agent is now a machine proposal plus an institutional acceptance, and a standard that records who created or revised a description does not say on what conditions an institution may accept what a machine drafted.

Review depth follows use and consequence

A wrong character in a transcript inside an internal discovery aid costs a search result. The same wrong character inside a quotation on a public collection page is a misquotation attributed to the institution, and it will be cited by someone who has no way of knowing it was machine-produced. The output has not changed. The use has.

Work the case through properly, because it is the case in which the distinction does the most work. An institution runs handwriting recognition across a series of nineteenth-century letters. Accuracy is good, not perfect. Used internally, the transcripts make an unsearchable series searchable for the first time, the errors are visible to anyone who opens the image next to the text, and the cost of an error is a search that misses. That is a defensible use of an imperfect output, provided the limitations are stated and nobody quotes from it.

Now take one of those transcripts and put a sentence from it on the public page as the text of the letter. Nothing about the transcript changed. Everything about the institution's exposure did. The reader has no image to check against, no reason to doubt it, and every reason to quote it in something else. The verification that use requires is a person reading the passage against the original, and that is a different job with a different cost, done by a different person, at a different point in the workflow.

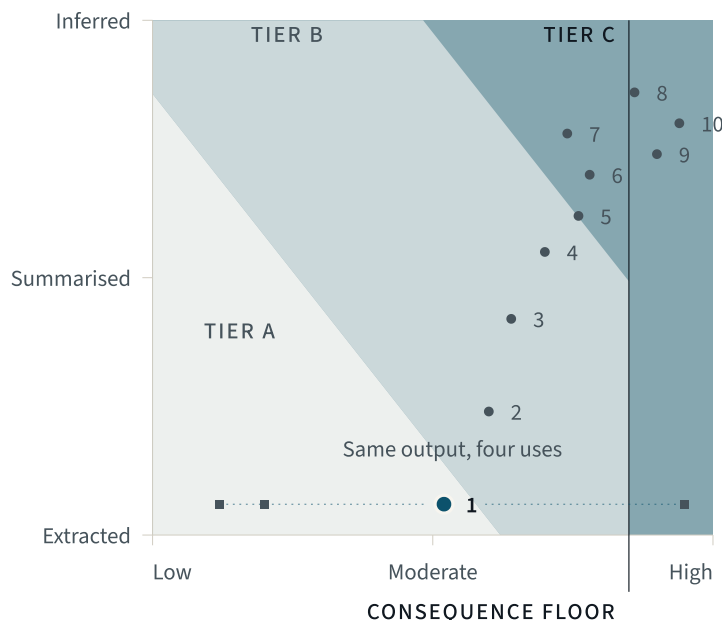
A third case is worse than either, because it does not look like a decision at all. A system detects a personal name in a letter and links it to an authority record. The link is plausible. It is also a claim that this person, in this collection, is that person, with that life and those relationships, and it will propagate: into the authority file, into aggregated discovery, into somebody's research, into a family's understanding of their own history. Nobody reviewed a sentence, because there was no sentence. There was a link, offered beside the name with a tick next to it, and accepting it took a quarter of a second.

This is why a confidence measure cannot carry the control on its own. What such a measure reports varies from system to system, and may be a token likelihood, a calibrated estimate, the weight of whatever evidence was retrieved for the answer, or something else again that the number itself does not announce. Whatever it reports, it is a property of the output. It knows nothing about whether the sentence is going into a staff spreadsheet or onto a published record, and the institution's exposure is set almost entirely by that second thing. A measure that has been validated against the institution's own material can help decide what to look at first. What sets how hard the looking has to be is the intended use and the cost of getting it wrong.

Figure 3
Review depth changes with use and consequence

Static view: OCR or transcript text, used as public archival description.

INTERPRETIVE DISTANCE FROM THE SOURCE



OUTPUT CLASS

TIER

1	Transcript	A
2	Dates, entities	B
3	Captions	B
4	Title, scope note	B
5	Authority links	C
6	Relationships	C
7	Admin history	C
8	Biographical note	C
9	Sensitivity signal	C
10	Rights decision	C

CONSEQUENCE OF AN ERROR

TIER A	Provisional or internal	TIER B	Accepted public description	TIER C	High consequence or contested
--------	-------------------------	--------	-----------------------------	--------	-------------------------------

Positions are assigned by the author, not measured.

Source: Author's proposed framework. Every coordinate here was assigned by the author, not measured: the axes have no units and the figure is not evidence. Tiers follow $0.42 \times \text{interpretive distance} + 0.58 \times \text{consequence}$, and anything past the consequence floor is Tier C. It synthesises DACS description control, Colavizza and Jaillant's AI preparedness guidelines, and NIST oversight and override concepts. Illustrative framework. Control follows intended use and consequence, not model type or confidence score alone. Verified 30 July 2026 archively.ai/research-notes/01

The same machine output can need light review or specialist review, and which one it needs is decided by the use, not by the model or its confidence score.

An example state, not a recommendation.

1 **OCR or transcript text** used as public archival description

Tier A: Provisional or internal

Evidence needed

Enough to spot a clear error: the passage as imaged, beside the transcribed text.

Who reviews it

Any trained staff member working within the approved scope.

Decisions available

Accept, modify, reject, or leave provisional.

Second review or escalation

Not routinely indicated.

What users are told

The record states that part of the description began as a machine proposal, and gives a correction route.

What is retained

The proposal, the disposition, and who made it.

A model may propose

A model may propose. An authorised person disposes.

Ten classes of machine output, positioned by how far each sits from the source and what an error would cost. The intended use decides where a class sits on the consequence axis: an imperfect transcript is a minor nuisance in an internal finding aid and a misquotation in a published record. The tiers describe review depth, not permission.

The tiers describe review depth, not permission. Nothing here says a Tier C output may not be produced. It says that if it is produced and then used publicly, the review that accepted it has to have been proportionate to that use, and that for some decisions the proportionate answer is that the machine does not participate at all.

Rights and access are the clearest case at the far end. A rights decision is a legal and ethical judgement about a specific person's material, taken under whatever legal and ethical obligations already bind the institution, by a role that carries the authority to take it. An automated signal can put material in front of that role, and routing material for specialist attention is a reasonable thing to ask a system to do. What such a signal cannot do is determine the decision or stand in for the authority to make it. A score that sends a file to the rights officer is doing useful work. A score that clears the file is doing the rights officer's job without holding the rights officer's mandate.

Three design questions remain open inside the framework proposed here. They are the author's own, and none of them is a position held by anybody else. The first is that biographical and administrative history are better treated as two classes of output than as one. The subjects differ and so do the consequences: an administrative history of a body wound up decades ago carries a different risk from a biographical note about a living person's father, and the table in section 3 folds the two into a single row in a way that understates the difference. The second is that provisional public access needs an explicit scope, a stated review point and a correction rule, all decided when the material is released rather than when somebody complains about it. Internal is a statement of intention rather than a property of a system, and it should not be assumed to hold indefinitely. The third is that a rights or access decision should be recorded separately from the acceptance of descriptive text and taken by the role authorised for it, because a decision offered in the same interface as a descriptive proposal, at the same moment, is being invited to inherit that proposal's disposition. None of the three is settled here, and any product that ignores them, the author's included, deserves the question.

What proof looks like

The controls above are worth nothing if nobody can inspect them afterwards. So what does the evidence for one accepted descriptive statement actually look like, and how much of it is there?

Less than people expect. The proposal, retained as generated. The evidence available at the time of the decision, including what was missing. The rule, convention or local policy applied. The role that decided, the disposition, the timestamp. The exact difference between the proposal and the accepted text, and the reason recorded then rather than reconstructed for an auditor two years later. Then what happened afterwards: quality review, rights clearance, publication, and any correction event with its date and its trigger.

None of that set is required by anything cited here. No standard asks an institution to keep the proposal as generated, or the field-level difference between the proposal and the accepted text, or the reason the reviewer held at the time. Those three are the author's proposal, and they are asked for because they are what makes a review inspectable afterwards rather than merely recorded. How much of the set is worth keeping for any one statement follows from the consequence of getting that statement wrong and the cost of holding the evidence.

Most of the categories already have somewhere to live. The descriptive standard provides an element for the sources used, the rules applied, the agent responsible and the dates of creation and revision, and it has provided for them since long before anyone needed them for this, although it offers them as added value rather than as a minimum DACS, Description Control. The older general standard set out an archivist's note, the rules or conventions used, and the date of the description ISAD(G), which is evidence of a long-standing model of descriptive control rather than evidence that every institution filled it in. What is different now is only that the agent is two things: a proposal from a mechanism, and an acceptance by a role the institution can name.

The honest complication is cost. Field-level evidence for every accepted statement can make an assisted workflow slower than doing the work by hand, which would be a real argument against the whole enterprise. The answer is not to skip the evidence. It is to stop pretending every field needs the same amount of it. A batch of transcripts accepted under one stated rule needs one decision record for the batch and a note of the exceptions. A biographical statement about a named person needs its own.

There is a limit to what any of this can capture, and it is worth stating before somebody sells a completeness claim on the back of it. A decision record shows what the reviewer was given. It cannot show what they read. It cannot show that they opened the image, or that they knew the series well enough to notice that the date was implausible for that correspondent. The evidence is a record of the conditions of the decision, not of the attention paid to it. Aggregate monitoring is the only partial answer: if one reviewer accepts everything and another rejects a fifth of it, the difference is visible even though neither reviewer's attention is.

Figure 4

Evidence ledger for one accepted descriptive statement

Constructed worked example. No real repository or holding is represented.

Record CEA/PH/1962/0114

Machine proposal, not yet asserted

Photographs of the ~~harbour~~ dedication ceremony, 1962

Modified, then accepted by an authorised reviewer

The programme in the same folder names the reading room. No harbour works appear in the 1962 Council minutes.

Photographs of the East Harbor reading room dedication, 1962

The record kept at each stage

1 Evidence available Before 11 March 2026

Source object	Six gelatin silver prints, folder 3
Collection context	Constructed Example Archive, Public Works photographs
Sources consulted	Dedication programme, same folder; Council minutes, 14 May 1962
Rule applied	DACS 2.3 title; local rule LR-04, event names

2 Decision 11 March 2026, 09:42

Reviewer role	Descriptive archivist, delegated authority for series PH
Accepted version	v1, internal working description
Quality review	Passed 19 March 2026

3 After publication From 2 April 2026

Published	2 April 2026
Rights and access	Cleared, no restriction
Correction	Reported by a reader, 17 June 2026; place name corrected as v2

Evidence that a decision occurred is not evidence that the decision was correct. It makes the decision inspectable, attributable and corrigible.

Source: Author's constructed example. The repository, record, people, dates and places are invented.

Verified 30 July 2026 archively.ai/research-notes/01

What a committee can inspect: the wording as proposed, the phrase a reviewer changed, the evidence available at the time, and what happened after publication.

Constructed worked example. The repository, record, people, dates, place names and correction event are fictional: what it illustrates is the evidence a decision leaves behind, not an institutional case. The categories are not new. Sources used, rules applied, the agent responsible and the revision date are description control under DACS 8.1, which names generative artificial intelligence systems, alongside people, as creators or revisers of a

description. The description, the describing activity and the agent responsible are entities in their own right in Records in Contexts. Review by a person before acceptance, with machine output treated as a non-authoritative suggestion, is common to the AI Preparedness Guidelines for Archivists by Colavizza and Jaillant, published by the Archives and Records Association, and to the Program for Cooperative Cataloging task group report on artificial intelligence and machine learning. Documented roles and responsibilities, post-deployment monitoring and a route to appeal and override come from the NIST risk management framework, which is cross-sector rather than archival. What changes is that the describing agent is a machine proposal and an institutional acceptance rather than a single act of authorship.

The example is deliberately unglamorous. A plausible place name, wrong, corrected three months after publication by a reader who knew the building.³

Caveat

An audit trail does not make a description true. It does not make it ethical either. Everything in Figure 4 could be complete, signed and timestamped around a statement that is simply wrong, or that describes people in terms they would refuse. Retained evidence supports answerability, inspection and correction. It is not evidence of correctness, and any claim that it is should be rejected.

What users should be told

Internal decision evidence and public disclosure are different things, and conflating them produces both a privacy problem and a comprehension problem. A reader has no use for the ledger. What this note proposes they be told is four things, in language that does not require them to care about any of this.

That a machine contributed to a description, and roughly what it contributed. That the description is provisional, if it is. What the known limitations are, including which parts of the collection this applies to and which it does not. And how to tell the institution it is wrong, in a way that produces a response rather than a receipt.

The last one is the easiest of the four to get wrong. A correction route that nobody is funded to read is worse than no correction route, because it converts a reader who knows something into a reader who has been ignored. If the institution cannot staff the route, the honest disclosure says who to write to and how long it usually takes, which is at least true.

Coverage is the other thing to settle. A single site-wide notice saying the institution uses these tools cannot be resolved to the description in front of the reader, and read strictly it covers every description in the catalogue, including the ones written by people over the

last forty years. For some institutions that will be the level they can actually maintain, and it is better than silence. Where machine assistance has materially affected a particular description, this note proposes disclosure at the finest level the institution can keep true, which is usually the record.

Proportion matters in the other direction too. What a reader can use is that a machine drafted the scope note and that a member of staff accepted it under a stated rule. Whether the individual who accepted it is named in public is a separate question, and it has different answers in different descriptive environments: some traditions have signed descriptions for a century, and in other settings naming one member of staff moves correspondence that belongs to the institution into a particular person's inbox. The distinction underneath is the one to hold to. Internal evidence names a person, because answerability inside an organisation requires it. Public disclosure names the institution, because that is who is answerable outside it.

Key point

Record-level disclosure is harder than a site-wide notice, and where machine assistance has materially affected the description this note recommends it anyway. It is a recommendation and not a rule: no descriptive standard or professional code cited here requires it. The argument for it is that a reader who cannot tell which description was machine-assisted cannot weigh the one in front of them, and a notice covering the whole catalogue tells them something about the institution rather than something about the record.

A procurement test

The argument reduces to questions that can be asked in a meeting by an archivist, a director, a rights officer or a technical reviewer. They name no product, and they apply as well to a system an institution builds for itself as to one it buys. They should be asked of every supplier, including the one that published this note.

Questions you can ask any vendor, including this one

- What exactly may the system propose, for what material, and what is explicitly out of scope?
- Is the proposal retained as generated, separately from the accepted text, and for how long?
- What decision options does a reviewer have, and can they defer, escalate or stop the task rather than accept or edit?
- What context is on the screen at the moment of decision, where does it come from, and what is the reviewer not being shown?
- Who holds the authority to accept each class of output, how is that role recorded, and who can change it?
- How many items per reviewer per day does your implementation assume, and what happens to quality when that number is doubled?
- What is retained for one accepted statement, and can it be exported and read without your software in ten years?
- Show us one documented example or controlled demonstration, end to end, including a proposal that was wrong, the evidence available to the reviewer and the recorded disposition.

Ask them in writing and keep the answers. The answers that should worry a committee are the confident ones with no detail. That the system is designed with a human in the loop. That everything is logged. That accuracy is high. None of those answers any question above.

The last question is the test. The first seven can be answered from a slide deck. The eighth requires that somebody kept the evidence, that a wrong proposal was recorded rather

than tidied away, and that whoever built the system is willing to walk an institution through a case where its own output was wrong. Where there is no permission to show a customer's material, a controlled demonstration answers it honestly. What cannot be substituted is the wrong proposal, what the reviewer had in front of them at the time, and the disposition that was recorded.

Aside

The same questions apply to work built in-house, where nobody is selling anything and the pressure comes from a grant deadline instead of a quarter. The difference worth noticing is procedural: a purchase brings a process that puts the answers in writing, and an internal build has to arrange one for itself.

Conclusion

Human judgement does not make description correct. It never has. Archivists have been wrong about dates, names, provenance and significance for as long as there have been finding aids, and what the standards built in answer to that was a way of documenting who described what, under which rules, and when, so that the next person can find the error and fix it. That the means exist is not the same as saying every institution has used them, and the standards make no such claim either.

That is the whole claim in this note. Judgement does not confer accuracy. It locates responsibility in an institution that can be asked to explain itself and argued with by the people a record affects. A machine proposal is not description. It becomes description when an authorised role accepts it, on the record, with the evidence retained, and it stays contestable afterwards.

The framework here is the author's proposal, published in a form that invites checking. Where it is wrong about practice, the practice is right, and the note lets that be established in public rather than settled in private.

- 1 Seven questions here, six conditions in section 5, seven rows in Figure 2. The numbers are not significant and nobody should try to align them. Each list is what survived attempts to collapse it.
- 2 The weakest part of the evidence base, put here rather than buried. The ledger holds standards and specifications, two professional codes, two pieces of current guidance for the sector, and two governance frameworks written outside it. It holds no institutional case study, no measured study of an implementation and no interview evidence, so nothing in this note rests on an observed outcome. Section 3's ordering and the automation point in section 5 are reasoning rather than evidence. Institutions that have run this work at volume hold the evidence, and it is wanted.

- 3 The record, the institution and the correction in Figure 4 are fictional. The failure it shows, a confident and plausible place name that was wrong, is the ordinary case rather than the dramatic one, and the one that survives review most easily, because nothing about it looks like an error.

Open review

This note is a synthesis rather than a professional standard, and it is published to be checked. If something in it is wrong, or imprecise, or missing a piece of context a practitioner would treat as essential, the address is researchnotes@archively.ai. Substantive corrections are recorded in the change log with the date the change was made, and credited by name where the person who sent them is content to be named.

Cite this note

APA Saad, M. (2026). When does an AI proposal become archival description? (Archively Research Note 01, Version 1.0). Archively.

Chicago Saad, Muhammad. "When Does an AI Proposal Become Archival Description?" Archively Research Note 01, version 1.0. Archively, 2026.

Version and corrections

Publication history

Version 1.0 · 6 September 2026 · First published edition.

Canonical web edition: archively.ai/research-notes/01. This PDF is a snapshot of version 1.0, published 6 September 2026.

Corrections: researchnotes@archively.ai. Substantive changes are dated in the publication history and issued under a new version number.

Source ledger

Sources are grouped by type. Each group states how its sources are weighted in this argument, distinguishing archival standards, professional guidance and frameworks from other sectors.

Sources last verified 7 August 2026.

Normative sources, carrying the most weight in this note.

Describing Archives: A Content Standard https://saa-ts-dacs.github.io/dacs/06_part_1/09_chapter_08/01_description_control.html

Standard · Technical Subcommittee on Describing Archives: A Content Standard, Society of American Archivists; revisions approved by the SAA Council. Version 2022.0.4, released 14 July 2026; on a continuous revision cycle since 2013, with no fixed edition date.

Part I, Chapter 8, element 8.1 Description Control (Added Value), General Rules, Archivist and Date, rule 8.1.5

Cited for the element that records the sources used, the rules or conventions applied, the agent responsible, and the dates of creation and revision. Rule 8.1.5 expressly names generative artificial intelligence systems, alongside people, as possible creators or revisers of a description, and the element carries a worked example in which a named person uses such a system to revise a biographical note and compile an inventory. So the standard is not silent about machine participation. The element is Added Value rather than a minimum requirement, and nothing here should be read as a claim that it mandates the evidence set proposed in section 7.

Describing Archives: A Content Standard https://saa-ts-dacs.github.io/dacs/06_part_1/03_chapter_02/03_title.html

Standard · Technical Subcommittee on Describing Archives: A Content Standard, Society of American Archivists; revisions approved by the SAA Council. Version 2022.0.4, released 14 July 2026; on a continuous revision cycle since 2013, with no fixed edition date.

Part I, Chapter 2, element 2.3 Title (Required), Purpose and Scope and rules 2.3.1 to 2.3.3

Cited because the title is a controlled element with rules behind it, not a free text field that any competent writer can fill. It is a required element, it separates a formal title from a devised one, and it asks for professional judgement in deciding when a title must be devised rather than transcribed. What it does not say is who or what may draft the words.

Records in Contexts: Conceptual Model (RiC-CM) <https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf>

Standard · International Council on Archives, Expert Group on Archival Description. Version 1.0, November 2023.

Section 6, Documenting Description (6.1, 6.4 and 6.5, with Figures 6, 7 and 8); section 2.1, entity hierarchy; section 2.2.3 Agent, with entity RiC-E13 Mechanism; section 5.5, relation attribute RiC-RA05 Source of Relation

Cited for a model in which agents, the activities they carry out, the rules they apply and the dates on which they act are entities in their own right, and in which a mechanism is already a kind of agent that may generate or modify records. The model also treats an archival description as a record resource in its own right: produced by a describing activity, attributed to an agent holding a position, dated, revised over time, and supported by sources cited either for the whole description or for a single statement within it. That reading belongs to the model rather than to this note. What the model does not supply is the institutional condition under which an output may be accepted.

Records in Contexts Ontology (RiC-O) https://www.ica.org/standards/RiC/RiC-O_1-1.html

Standard · International Council on Archives, Expert Group on Archival Description. Version 1.1, dated 22 May 2025.

Class Mechanism, declared a subclass of Agent and distinct from Person and from Position, with its scope note and its stated correspondence to entity RiC-E13 of the conceptual model; class Agent

Cited separately from the conceptual model, because a class statement belongs to the ontology and folding the two together would misattribute both. Here a mechanism is declared a kind of agent, held apart from a person and from a position, and described as something that may act in the world and generate or modify records. This supports representing a system's participation in description. It does not hand the system institutional authority, which is the thing this note is about.

ISAD(G): General International Standard Archival Description

https://www.ica.org/app/uploads/2024/01/CBPS_2000_Guidelines_ISADG_Second-edition_EN.pdf

Standard · International Council on Archives, Committee on Descriptive Standards. Second edition, published Ottawa 2000, adopted at Stockholm in September 1999.

Paragraph I.11, the seventh of the areas of descriptive information; 3.7 Description Control Area, with 3.7.1 Archivist's Note, 3.7.2 Rules or Conventions and 3.7.3 Date(s) of descriptions

Used for historical comparison rather than as a principal current source. The expectation that a description says who prepared it, under what rules, and when, is decades older than machine assistance and was not invented for it. It is evidence of a longstanding descriptive control model, not evidence that every institution recorded those elements.

Encoded Archival Description Tag Library, version EAD3 1.1.2 <https://www.loc.gov/ead/EAD3taglib/EAD3-TL-eng.html>

Standard · Technical Subcommittee for Encoded Archival Standards, Society of American Archivists; documentation hosted by the Library of Congress. Version 1.1.2, released June 2023, superseding the 1.1.1 tag library of 2019 and the 1.1.0 tag library of 2018.

control: maintenancehistory, maintenanceevent, eventtype, eventdatetime, agent, agenttype, eventdescription

Cited because the encoding already provides somewhere to put this: each maintenance event carries its type, its date and time, a description of what changed, and the agent responsible, with the agent type required to be recorded as human, machine or unknown. That is a place to record the fact. It is not a requirement to retain the original proposal, the comparison against it, or the reason a reviewer accepted it.

PREMIS Data Dictionary for Preservation Metadata, version 3.0 <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>

Standard · PREMIS Editorial Committee; maintained by the Library of Congress. Version 3.0, June 2015, revised November 2015, still the current version.

Event entity: 2.2 eventType, 2.3 eventDateTime, 2.4 eventDetailInformation with 2.4.1 eventDetail, 2.5 eventOutcomeInformation, 2.6 linkingAgentIdentifier with 2.6.3 linkingAgentRole; Agent entity: 3.2 agentName, 3.3 agentType, 3.4 agentVersion

From preservation rather than description, and the dictionary says so itself: descriptive metadata is named in its own list of things outside its scope. Cited for one working convention this note borrows, that an event records what happened, when it happened and which agent did it, with software carried as a type of agent and the role recorded on the event rather than on the agent. It is a metadata pattern worth learning from, not a requirement to evidence acceptance.

Professional guidance

4 entries

Codes, guidance and reports from professional bodies and programmes. Codes bind members where specified.

SAA Core Values Statement and Code of Ethics <https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics>

Guidance · Society of American Archivists; both documents approved and revised by the SAA Council. Core Values of Archivists approved May 2011, revised August 2020 and August 2025; Code of Ethics for Archivists approved February 2005, revised January 2012, August 2020 and August 2025.

Core Values of Archivists: Accountability, Responsible Stewardship. Code of Ethics for Archivists: Judgment, Authenticity

Cited for the professional position that description is an exercise of judgement for which the archivist and the institution are answerable, and for the duty to be transparent about the archivist's own role by documenting descriptive work and collections policy decisions. Neither document mentions machines, artificial intelligence or automated description anywhere, so it cannot be read as settling whether machine-drafted description is acceptable, and it is not cited here as though it did. The refusal position in section 1 is treated as legitimate on its own terms, not as something the Code prescribes.

ARA Code of Ethics <https://www.archives.org.uk/ara-code-of-ethics>

Guidance · Archives and Records Association (UK and Ireland); interpreted and periodically revised by its Ethics Panel, with enforcement under the association's bye-laws. Updated May 2026.

Section C, Accuracy, Transparency and Accountability, clause 28; clause 23 on decisions to digitise or to use such tools; Section B, clause 14, machine learning among threats to impartiality; Glossary, "Access"

Binding on the association's individual members and used to guide disciplinary procedures, which is a different kind of force from a statement of values and worth saying plainly. Clause 28 asks that any use of AI in the creation, processing or presentation of records which materially affects their informational or evidential value be transparent to users, requires AI-assisted processing to be undertaken with appropriate human oversight and quality control, and leaves the member responsible for the accuracy of anything generated on their instruction. The transparency duty is written as a should and the oversight duty as a must, and this note does not flatten the two into one.

AI Preparedness Guidelines for Archivists <https://www.archives.org.uk/ai-preparedness-guidelines-for-archivists>

Guidance · Giovanni Colavizza and Lise Jaillant; published by the Archives and Records Association (UK and Ireland) as an output of the FLAME project, which the association funded. February 2026.

“Draft descriptive metadata and subject enhancement”, covering candidate titles, dates and short summaries, suggested topical keywords and candidate groupings; the section on sensitivity review; the checklist item on human-in-the-loop workflows for reviewing and approving outputs

The authors are Colavizza and Jaillant. The association funded and published the work, so it must not be cited as though the association wrote it. Cited for two lines: that all outputs must be reviewed by archivists and treated as non-authoritative suggestions, and that such tools help archivists focus attention but do not replace their legal or ethical judgement. The guidelines define no status called a provisional output, and nothing in this note may be cited to them as though they did.

PCC Task Group on AI and Machine Learning in Cataloging and Metadata: Final Report

<https://www.loc.gov/aba/pcc/taskgroup/AI-and-Machine-Learning-TG-final-report.pdf>

Guidance · Program for Cooperative Cataloging, Task Group on AI and Machine Learning for Cataloging and Metadata; approved by the PCC Policy Committee. The file is hosted by the Library of Congress, which supplies the programme’s secretariat and holds one seat on the task group, and it is not that library’s policy. Submitted 16 October 2025, approved and the task group discharged 11 December 2025.

Deliverables, Statement of principles, the Guiding Principles for Use of AI and Machine Learning Technologies in Cataloging and Metadata Work; Appendix 3, Best Practices, sections on Recording Data Provenance, on Assessment and Human Review with its subsection on confidence score implementation and decision-making, and the appended checklist

A professional report from a cooperative cataloguing programme, not policy of the library whose site holds the file, and mislabelling it as the latter would inflate its standing. Cited for its recommendation that AI-generated content be thoroughly reviewed by a human cataloguer for accuracy, appropriateness and potential bias or harm, and for its treatment of data provenance, including a record-level note naming the tool used and the date. Its status is uneven and the unevenness matters: the data provenance sections are headed as still to be reviewed, updated and approved by the programme’s standards committee and carry unfilled placeholders, while the section on assessment and human review carries no such marker and does define confidence thresholds for deciding when output may be accepted automatically and when it must be routed to a person. That is a position this note has to argue with rather than ignore.

Cross-sector governance framework

2 entries

Frameworks from outside the archival sector, cited for operational concepts.

Artificial Intelligence Risk Management Framework (AI RMF 1.0) <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
Governance · National Institute of Standards and Technology, United States Department of Commerce. NIST AI 100-1, January 2023.

GOVERN 2.1 and 2.2, documented roles, responsibilities, lines of communication and training; GOVERN 3.2, roles and responsibilities for human-AI configurations and oversight; MAP 3.5, processes for human oversight; MANAGE 2.4, mechanisms to supersede, disengage or deactivate; MANAGE 4.1, post-deployment monitoring including appeal and override

A cross-sector risk framework, not archival doctrine, and it is labelled that way wherever it appears. Cited only to operationalise scope, delegated authority, reviewer capacity, monitoring, appeal and override. It does not use the words escalation or recourse at all, so neither is presented here as its vocabulary, and the delegated acceptance model proposed in this note is an archival adaptation rather than anything the framework requires. Nothing in it was written with archival description in mind.

Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

Governance · European Parliament and Council of the European Union. Adopted 13 June 2024.

Article 14, human oversight; Article 26(2), obligations of deployers

Cited only for how it treats oversight. Article 14 concerns oversight measures for high-risk systems: interpreting the output, staying alert to the tendency to over-rely on it, disregarding, overriding or reversing it, and interrupting or stopping the system. The separate requirement that oversight be assigned to natural persons with the necessary competence, training and authority, and the necessary support, is in Article 26(2) rather than Article 14, and the two are kept apart here. This note borrows that vocabulary and makes no claim that archival description systems are high-risk systems under the regulation, nor that any archive falls within its scope.

archively.ai/research-notes/01