

On The Record: Toward a Formal Epistemology of Journalistic Attestation

Vince Gonzalez

ORCID 0009-0005-3640-014X

2026-04-18

v1.0.0 · 2026

Abstract

The phrase "on the record" is journalism's foundational epistemic claim. It means, at minimum, that a statement or observation is attributable, verifiable, and accountable. Yet no journalism standard specifies what verifiable means formally — by what procedure a third party, at a later time, could confirm that an attributed observation was causally connected to the event it purports to document. This paper identifies this gap as structural rather than procedural and argues that it is now critical: in an environment where AI-generated content is indistinguishable from observational content by inspection, the informal epistemology underlying "on the record" is insufficient. The paper proposes that the three conditions of "on the record" — attribution, verifiability, and accountability — each have formal counterparts in the Modulign Standard's observer annotation (%OBS), append-only observational ledger (GSI-ID + ^EVID), and Classification Decision Protocol (CDP). A journalistic observation classified under the Modulign Standard is not merely claimed to be on the record; it is formally on the record in a sense that is auditable, reproducible, and resistant to post-hoc manipulation. The paper does not argue that existing journalistic practice is epistemically deficient — it argues that existing journalistic norms, properly formalised, already demand exactly the architecture the Modulign Standard provides.

Keywords: journalistic epistemology, on the record, observational attestation, provenance, Modulign Standard, DAG-OR, journalistic verification, AI-generated content, chain of custody

1. The Phrase That Does More Work Than It Can Bear

"On the record" is one of the most consequential phrases in journalism. A source who speaks on the record accepts attribution. A reporter who publishes on the record accepts accountability. An editor who publishes on the record makes a verification claim — that what is reported was observed, that what was observed was accurately described, and that someone whose name can be attached to that claim stands behind it.

These three conditions — attribution, verifiability, accountability — constitute the informal epistemology of journalistic attestation. They are not merely conventions. They are the conditions under which a journalistic publication can be distinguished from speculation, fabrication, or error. Without them, journalism is opinion; with them, it is, in some meaningful sense, a record of what happened.

The problem is that none of these conditions is formally specified. Attribution means the observer's name is attached — but not what the observer's qualifications are, or how their account was produced. Verifiability means a third party could, in principle, check the claim — but not by what procedure, against what standard, or with what access to the original observational evidence. Accountability means someone can be held responsible — but not that the underlying observation is cryptographically anchored to a specific time, place, and classification act that

cannot be retroactively altered.

These gaps were tolerable when journalism operated in an environment where fabrication required effort, where forgery was detectable by experts, and where the causal chain from event to publication was reconstructable from physical evidence. That environment no longer reliably exists. AI-generated images, synthetic video, and large language model outputs have severed the previously implicit connection between publication and observational grounding. A photograph published today may or may not have a causal connection to the event it depicts. The formal apparatus for distinguishing these cases — for establishing, not merely claiming, that an observation is on the record — does not exist in current journalistic practice.

This paper asks what it would mean to formalise "on the record" — to give its three conditions precise content that is auditable, reproducible, and resistant to post-hoc manipulation. It argues that the answer is already implicit in what journalism's best epistemic norms demand, and that the Modulign Standard's architecture provides the formal machinery to make those demands explicit.

2. Attribution: From Name to Observer Annotation

Attribution in journalism means that a specific person, identifiable by name, accepts responsibility for a specific claim. Its function is to create an epistemic anchor: the reader knows who is asserting what, and the asserter knows they can be held accountable for the assertion's accuracy.

What attribution does not specify is the observational basis of the assertion. A reporter who attributes a claim to a named source has documented the provenance of a statement, not the provenance of an observation. The source may have witnessed the event directly, heard about it secondhand, or inferred it from incomplete evidence. The attribution conveys none of this. It tells the reader who said it, not what epistemic relationship the sayer had to the thing said.

This is not a critique of attribution — it is an accurate description of what attribution does and does not establish. Attribution is a social and legal instrument. It creates accountability for assertions. It does not, by itself, establish the causal connection between assertion and observable reality that distinguishes observational knowledge from hearsay, inference, or fabrication.

The Modulign Standard's observer annotation (%OBS) formalises what attribution leaves implicit. The %OBS segment encodes: observer type (human, automated, hybrid, advisory), domain competence level (:T trained, :E expert, :A authority), consensus count (the number of independent classifiers who reached the same classification), and — through the observer registry — a permanent, verifiable record of the classifying agent's identity and qualifications. A Modulign-classified observation does not merely attribute the classification to a name; it encodes the epistemic credentials of the observer, the conditions under which the classification was made, and whether independent consensus was reached.

The formal counterpart of attribution is not a name on a byline. It is an observer annotation that specifies who classified, what their qualifications were, and whether their classification was independently corroborated. These are the conditions journalism's best norms already implicitly require — the Modulign Standard makes them explicit and machine-verifiable.

3. Verifiability: From "Could Check" to Append-Only Ledger

Verifiability in journalism means that a third party could, in principle, confirm what was reported. In practice this means: the reporter's notes exist, the source is contactable, the footage has not been edited, the document has not been altered. Verifiability is a hypothetical — it describes what could be done, not what has been done or what infrastructure exists to make it possible.

The hypothetical nature of journalistic verifiability was not a serious problem when the physical properties of evidence made retroactive alteration detectable. Film negatives, audio recordings, physical documents — each left traces that forensic analysis could examine. Digital evidence does not have these properties. A digital image file can be altered without leaving any trace detectable by standard inspection. A video file can be edited, re-encoded, and re-distributed with metadata intact. The hypothetical verifiability of digital journalistic evidence is weaker than it appears, because the physical constraints that made retroactive alteration difficult no longer apply.

The Modulign Standard's observational ledger provides a formal architecture for verifiability that is not hypothetical. When an observation is classified under the Standard, a Global Stream Identifier (GSI-ID) — a UUID-v4 permanent unique identifier — is assigned at the moment of classification. The GSI-ID anchors an append-only ledger: every subsequent classification of the same observation stream is appended to the ledger, never deleted, with a cryptographic chain linking each entry to its predecessor via SHA-256 content hashing. Modification of any entry in the ledger changes its hash, making tampering detectable with a collision probability of approximately 2^{-256} (Gonzalez, V. 2026b, Theorem 3.1).

This is not a claim that all journalistic evidence should be classified under the Modulign Standard. It is a demonstration that what journalism means by verifiable — that a third party could confirm the observational record — has a formal counterpart that the Standard already provides. The GSI-ID and its associated ledger constitute a verifiable observational record in a sense that is not hypothetical: it is structurally tamper-evident, permanently anchored, and independently auditable by any party with registry access.

The ^EVID meta-flag formalises the highest verifiability standard. A classification bearing ^EVID has been produced by: a named observer with documented domain competence; confirmed by independent consensus (`consensus_count` ≥ 2 from :E-level or :A-level observers); with confidence $\kappa \geq 0.80$; timestamped in ISO 8601 format; and cryptographically chained in an append-only ledger. A journalistic observation classified at ^EVID grade is not merely claimed to be verifiable — it is verifiable in a formally specified, auditable, and reproducible sense.

4. Accountability: From Editorial Process to Classification Decision Protocol

Accountability in journalism operates at two levels. Individual accountability holds the reporter responsible for the accuracy of their observations and the faithfulness of their reporting. Institutional accountability holds the publication responsible for the editorial processes through which claims are verified before publication.

Both levels of accountability are procedural. They specify who is responsible and what processes they must follow. What they do not specify is the epistemic procedure by which an observation is connected to observable reality — the steps that must be executed, in what order, with what inputs, to produce a justified classification of what was observed. Editorial accountability governs the publication of claims; it does not govern the production of observations.

This distinction matters because the production of an observation — the act of classifying what is

happening at a specific place, time, and scale — is the epistemic foundation on which all subsequent claims rest. A reporter who observes a protest and describes it as a riot has made a classification decision. A photographer who frames a crowd to include or exclude a specific element has made a classification decision. A camera operator who selects which footage to transmit has made a classification decision. Editorial accountability does not govern any of these decisions; it governs what happens after they are made.

The Modulign Standard's Classification Decision Protocol (CDP) is an eight-step procedure for producing a valid classification from observable features. It specifies, in order: domain resolution, subdomain resolution, geographic resolution, node assignment, scale annotation, medium annotation, observer annotation, and confidence scoring. Every step uses only input from the observable features of the phenomenon being classified. No step may introduce assumptions not traceable to those features (Gonzalez, V. 2026a, §§3–8). The protocol is public, documented, and reproducible: any competent observer applying it to the same observable features must reach the same classification.

This is the formal counterpart of accountability at the observation level. The CDP does not replace editorial accountability — it operates beneath it, specifying the epistemic conditions under which an observation qualifies as a justified classification of observable reality. A journalistic observation produced by executing the CDP is not merely editorially approved; it is epistemically grounded in a procedure whose steps are documented, auditable, and reproducible by any qualified party.

The CDP's intersubjectivity condition is the key. The Modulign Standard specifies that intersubjectivity in DAG-OR is procedural, not metaphysical: two trained classifiers applying the CDP to the same observable features must reach the same classification because the procedure is defined in terms of observable properties, not subjective interpretation (Gonzalez, V. 2026a, §4.2). Empirical confirmation of this claim has now been published: inter-rater reliability testing across $N = 234$ live observation streams demonstrated 100% domain-level agreement for semantically unambiguous domains, with systematic disagreement concentrated at domain boundaries identified as targets for specification amendment (Gonzalez, V. 2026g). The CDP's reproducibility is not merely claimed — it is empirically demonstrated.

5. The AI Content Problem as a Clarifying Crisis

The argument of this paper does not depend on the AI content problem. The formal gap in journalistic attestation — the absence of a specified procedure connecting "on the record" claims to observable reality — predates AI-generated content and would exist even if no synthetic content existed. But the AI content problem makes the gap visible in a way it was not before, and it is worth examining why.

The distinguishing property of AI-generated content, for the purposes of this argument, is not that it is false — AI-generated content may accurately depict things that happened — but that it lacks observational grounding. A synthetic image of a building on fire may look identical to a photograph of the same building on fire. The photograph has a causal connection to the observable event: light reflected from the fire, passed through a lens, was recorded on a sensor at a specific time and place. The synthetic image has no such connection. It was produced by a generative model from a statistical distribution over training data. Its resemblance to the event is not evidence that it depicts the event.

The Modulign Standard's VR/·:SYN address primitive encodes this distinction formally. Content

classified as VR/:SYN is synthetic — it carries a realm prefix indicating virtual/generated origin and a :SYN medium code. Content classified as SR/ carries an Earth surface realm prefix, indicating that the observation was made from a specific location in observable physical space. The distinction is not about content; it is about causal history. SR/ content has a documented causal connection to observable reality. VR/:SYN content does not (Gonzalez, V. 2026c).

For journalism, the significance of this distinction is that "on the record" has always implicitly required SR/ grounding — a causal connection to the observable event being reported. The AI content problem makes explicit that this requirement was never formally specified. The appropriate response is not to develop new detection tools for synthetic content — detection is a losing arms race — but to formalise the existing requirement. Content that is on the record must have a documented observational grounding. Content that lacks such documentation is not on the record, regardless of how it appears.

This is not a new standard. It is the existing standard, formally stated.

6. What "Formally On the Record" Looks Like

A journalistic observation that is formally on the record, in the sense developed in this paper, satisfies the following conditions:

Attribution: The observer is registered in the Modulign observer registry with a documented identity, observer type, and domain competence level. The observer annotation (%OBS) in the classification record permanently associates the observation with the classifying agent.

Verifiability: The observation has a GSI-ID assigned at the moment of classification. The GSI-ID anchors an append-only observational ledger that cannot be retroactively altered without cryptographic detection. Any third party with registry access can verify the complete classification history of the observation at any subsequent time.

Accountability: The classification was produced by executing the eight-step Classification Decision Protocol, using only observable features as input. The reasoning record — encoded in the mandatory JSON output — documents the inputs and outputs of each step, making the classification procedure auditable by any qualified observer.

These three conditions correspond precisely to the three conditions of "on the record" as journalism currently uses the phrase. They do not add new requirements; they formalise existing ones. A journalistic organisation that adopted Modulign classification for its observational evidence would not be doing something different from what journalism already requires — it would be doing what journalism already requires with formal, auditable, reproducible documentation.

The practical threshold for this is not high. A reporter observing an event applies the CDP — domain, subdomain, location, scale, medium, observer, activity, confidence — and registers the classification. The entire procedure adds minutes to the reporting process. What it adds to the evidentiary record is substantial: a permanently anchored, cryptographically secured, observer-annotated classification that any party can verify against the public registry at any subsequent time.

7. Objections

7.1 Journalism is not law — formal standards are inappropriate

The most common objection to formal epistemic standards in journalism is that journalism operates under different norms than law or science. Journalism's value lies partly in its speed, its accessibility, and its independence from institutional gatekeeping. Requiring formal classification procedures would slow publication, create new gatekeepers, and privilege technically sophisticated outlets over those with fewer resources.

This objection mistakes the scope of the proposal. The argument of this paper is not that all journalism must be formally classified. It is that journalism's existing concept of "on the record" already implicitly demands the formal conditions described here, and that the Modulign Standard provides a way to satisfy those demands explicitly for observational evidence that is intended to bear significant epistemic weight — evidence introduced in legal proceedings, evidence underlying major investigative claims, evidence whose authenticity is likely to be contested. The formal standard is not a replacement for journalistic judgment; it is a specification of what journalistic judgment, applied rigorously, already requires.

7.2 The Standard is privately controlled

The Modulign Standard is authored and currently controlled by an independent scholar. A privately controlled standard for journalistic verification raises legitimate concerns about access, governance, and the concentration of epistemic authority.

Two responses are relevant. First, the Modulign Standard is published as an open specification under an open-access license. Any party may implement it, audit it, critique it, or build upon it without license fees or restrictions (Gonzalez, V. 2026d). The specification is not proprietary; only the trademark is. Second, the Standard's governance roadmap commits to multi-stakeholder governance following first institutional adoption, with a Standards Advisory Council and eventual transfer to a formally constituted standards body analogous to ANSI or ISO accreditation (Gonzalez, V. 2026e). The current single-author control is appropriate for a specification in its founding phase; it is not the intended permanent structure.

7.3 Classifiers can be wrong

The CDP produces classifications with a confidence score, and classifications can be incorrect. A formally classified observation that is incorrectly classified is not epistemically superior to an informally documented observation that is correctly described.

This objection misunderstands the architecture. The Modulign Standard does not claim that CDP classifications are always correct. It claims that the procedure by which they are produced is documented, auditable, and reproducible — and that incorrect classifications are detectable as such through the $\wedge$ PROV provisional flag (emitted when confidence $\kappa < 0.60$), the $\wedge$ CONT flag (emitted when confidence is in the review range $0.60 \leq \kappa < 0.80$), and the consensus mechanism (:Ex2 or :A required for $\wedge$ EVID). The system is explicitly designed to surface uncertainty rather than suppress it. A classification that acknowledges its own uncertainty — formally, in a machine-readable, permanently recorded format — is epistemically superior to a publication that presents uncertain observations as established facts.

8. Conclusions

"On the record" is journalism's most consequential epistemic claim. It asserts attribution, verifiability, and accountability for observed facts. These three conditions are not merely conventions — they are the epistemological foundation that distinguishes journalism from opinion, inference, and fabrication. They are also, in current journalistic practice, informally specified: no procedure defines what verifiable means, no protocol specifies what accountable observation requires, and no architecture makes attribution more than a name on a byline.

The AI content problem has not created this gap. It has revealed it. When AI-generated content is visually indistinguishable from observational content, the informal epistemology of "on the record" is insufficient to bear the weight placed on it. The appropriate response is formalisation — specifying what "on the record" already implicitly requires and building the infrastructure to satisfy those requirements explicitly.

The Modulign Standard provides that infrastructure. Its observer annotation formalises attribution. Its append-only observational ledger formalises verifiability. Its Classification Decision Protocol formalises accountability. A journalistic observation classified under the Standard is not merely claimed to be on the record — it is on the record in a sense that is auditable, reproducible, tamper-evident, and verifiable by any third party with registry access.

This paper does not argue that journalism needs new epistemic norms. It argues that journalism's existing norms, properly formalised, already demand exactly the architecture the Modulign Standard provides. The phrase "on the record" has always meant more than it could formally deliver. It can now deliver what it has always meant.

References

Gonzalez, V. 2026a. Modulign Standard: A Dimensional Address Grammar for Observable Reality. Zenodo. <https://doi.org/10.5281/zenodo.19348703>

Gonzalez, V. 2026b. The Formal Logic of Modulign. Zenodo. <https://doi.org/10.5281/zenodo.19350847>

Gonzalez, V. 2026c. Synthetic Content as Epistemic Category. Zenodo. <https://doi.org/10.5281/zenodo.19351055>

Gonzalez, V. 2026d. Modulign Implementation License v1.0. modulign.org. <https://modulign.org/license>

Gonzalez, V. 2026e. Modulign Privacy and Governance. Zenodo. <https://doi.org/10.5281/zenodo.19559689>

Gonzalez, V. 2026f. Modulign as Evidence: A Legal Epistemology Treatise. Zenodo. <https://doi.org/10.5281/zenodo.19351250>

Gonzalez, V. 2026g. Inter-Rater Reliability of Automated Modulign Standard Classifiers. Zenodo. <https://doi.org/10.5281/zenodo.19643322>

© 2026 Vincent Gonzalez · modulign.org · The Modulign Standard is published as an open specification.