

Whose Content, Whose Benefit? A Creator Value Retention Audit Framework for AI Wenyiwen Search Summaries on Xiaohongshu

Jialin Gao *

University of Leeds, Leeds, The United Kingdom

* Corresponding author: (Email: gaojialin_@outlook.com)

Abstract: Generative search can present an answer before a user reaches the creators whose posts supplied the underlying knowledge. This paper introduces the Creator Value Retention Audit (CVRA), a reproducible computational framework for examining interface-level retention without claiming unobserved economic harm. Friction redistribution serves as an interpretive mechanism, not a measured outcome: the experiment records source interface layers and visible creator pathways but does not measure user effort. The framework combines rendered attribution salience, 1-indexed source interface-layer depth, creator affordances, relational cues, identity representation, source diversity, and lexical coverage in a transparent Creator Value Retention Index (CVRI). We executed a controlled browser experiment with 40 matched scenarios and 80 synthetic local HTML stimuli at a fixed 1280 by 900 viewport. The original-note-first condition achieved a mean CVRI of 97.92, compared with 57.07 for the AI-summary-first condition. Both conditions retained visible attribution, while summary-first stimuli placed the source at layer two instead of layer one, retained 25% of creator affordances, retained 40% of relational cues, and compressed visible text to 36.38% of source length. Equal-weight and component-ablation analyses preserved the gap direction. The experiment demonstrates internal metric sensitivity to prespecified synthetic differences, not real-world Xiaohongshu traffic, income, attention, creator loss, or perceived friction.

Keywords: Algorithmic auditing, creator labour, frictionlessness, generative search, platformisation, Xiaohongshu.

1. Introduction

Generative search changes more than the form of an answer. It changes the route through which information, attribution, interaction, and value become connected. Platformisation research shows that platforms reorganize cultural production through changing infrastructures, markets, and data practices [1]. Web platformisation also makes content and activity processable through standardized technical interfaces [2]. An AI summary intensifies both dynamics. It does not merely rank creator posts. It can retrieve, combine, compress, and present their informational content as a platform-authored answer. The resulting surface may remain useful and properly attributed while still altering the path from information discovery to creator recognition. This paper therefore asks a computational question: how much creator-facing value remains visible and reachable when an AI-style summary becomes the first access surface?

The question matters because creator content is simultaneously expressive work, searchable information, and a basis for relationship formation. Terranova described how voluntary cultural production becomes productive within the commercial Internet [3]. Grohmann later distinguished visible content labour from data labour that maintains platform infrastructures and algorithmic systems [4]. On lifestyle platforms, these categories overlap. A detailed note may establish expertise, but it may also generate metadata, retrieval material, and behavioral signals. Creator identity adds another layer. Branded-self research explains how visibility and reputation require continuous self-presentation [5], while aspirational-labour research emphasizes uncertain returns from sustained creative investment [6]. A summary

that preserves facts but removes follow, profile, comment, or message pathways can therefore retain informational utility while thinning the mechanisms through which a creator may obtain recognition.

Frictionlessness provides an interpretive mechanism connecting interface design to possible redistribution. Platform ecosystems seek seamless movement, rapid fulfillment, and reduced interruption for users [7]. Usability engineering similarly treats extra steps, unclear navigation, and unnecessary effort as costs imposed on task completion [8]. These principles matter when the task is obtaining an answer. Yet friction may be relocated rather than eliminated. A summary-first design can present information earlier while placing sources, creator context, or creator-directed actions on a deeper interface layer. We call this possible friction redistribution. The experiment measures layer structure and visible pathways, not task effort, perceived convenience, or user motivation. The concept therefore guides interpretation without treating friction as an observed outcome.

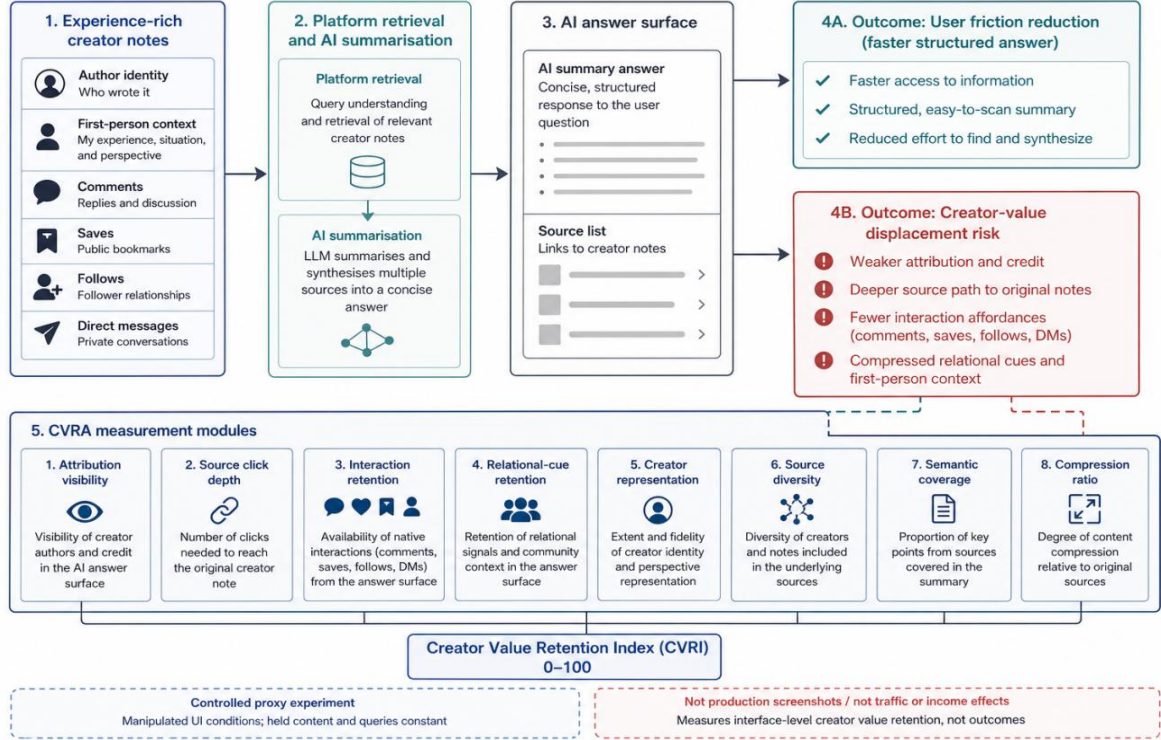
The contribution is an auditable framework rather than a causal estimate. CVRA measures rendered interface proxies within matched synthetic stimuli. It supplies a component vector and a bounded CVRI, but it never equates either with observed labour value. The experiment reports actual browser-rendered geometry, lexical coverage, source layers, and interaction affordances. It does not use creator interviews, production Xiaohongshu screenshots, platform logs, impressions, click-through rates, revenue records, or income observations. Table I establishes these evidence boundaries before any statistical comparison. The study contributes a reproducible algorithmic audit, an interpretive account of possible friction redistribution, and a sensitivity analysis that exposes the normative role of index weights.

Table 1. Evidence Classes and Claim Boundaries

Evidence class	Material used	Permitted claim	Excluded claim
Reported platform facts	Official post, agreement, and journalism	Features or statements were publicly reported	Reported statements prove creator harm
Controlled stimuli	80 synthetic local HTML pages	Two designed access paths differ	Stimuli reproduce production Xiaohongshu
Rendered proxies	DOM geometry, text, links, and affordances	Visible creator-value pathways differ	Proxies equal attention, trust, or income
Statistical summaries	40 matched scenario pairs	Differences characterize this controlled design	Tests establish causal platform effects

Fig. 1 summarizes the CVRA workflow from matched controls through rendering, component extraction, composite scoring, and sensitivity analysis. The framework carries

evidence labels through the pipeline so that reported facts, synthetic design properties, proxy measurements, and untested outcomes remain distinct.

**Figure 1.** Creator value retention audit framework

2. Platform Context and Theoretical Mechanism

2.1. Platformisation and Creator Labour

Platform power operates through infrastructures as well as visible content policies. Plantin and Punathambekar describe digital media infrastructures as arrangements of pipes, platforms, and politics [9]. Gillespie likewise shows that technical intermediaries frame operational decisions as neutral while organizing participation and visibility [10]. In AI-mediated search, the platform becomes an interpretive intermediary. It selects source material, decides which contextual elements survive compression, and determines how users can return to original notes. This intervention differs from ordinary ranking because the platform produces a new answer object. The object can substitute for some reading activity without replacing the source in a legal or technical sense. The relevant issue is therefore not whether the original still exists, but how the interface allocates prominence and reachable action between answer and source.

Xiaohongshu notes may possess relational value beyond textual information, but that proposition requires careful evidentiary wording. Platform society research explains how

datafication, commodification, and selection convert social activity into governable processes [7]. Crawford further treats AI as an assemblage of data, infrastructure, labour, and power rather than a detached information processor [11]. Under this lens, creator notes can support both community interaction and an AI answer service. However, reuse alone does not demonstrate exploitation, loss, or unauthorized training. The narrower testable claim is that CVRA should respond when controlled interface designs retain different creator-related cues and actions.

2.2. Reported Xiaohongshu Facts

Public materials describe AI search functions on Xiaohongshu, but their evidentiary status is limited. The official search account announced an AI search service in June 2025 [12]. A 2026 Jiemian report described voice-based questioning, AI-generated answers drawn from community material, source-note links, and strong search use among platform users [13]. A separate 36Kr report described Wenyiwen as a community-search initiative using community data and reported internal debate about query coverage [14]. These statements establish a reported product context. They do not provide an auditable sample of production summaries, verified creator-level exposure changes, or causal estimates

of engagement. The Generative AI Service Agreement provides a governance document relevant to notice, service operation, and platform responsibilities [15], but contractual language does not reveal actual interface performance.

These distinctions are summarized visually through the experimental retention proxies in Fig. 2. The plot belongs to the controlled experiment, not to production observation. Its two lines compare designed conditions that instantiate an original-note-first route and an AI-summary-first route. The

audit dimensions include attribution, source access, relationship cues, and source visibility. The experiment implemented them in local pages whose contents and structure were known. This separation matters because reported Xiaohongshu facts and synthetic measurements answer different questions. Reports establish why the audit is relevant. Synthetic stimuli establish whether the metric pipeline detects prespecified design contrasts. Neither evidence class demonstrates how real users respond.

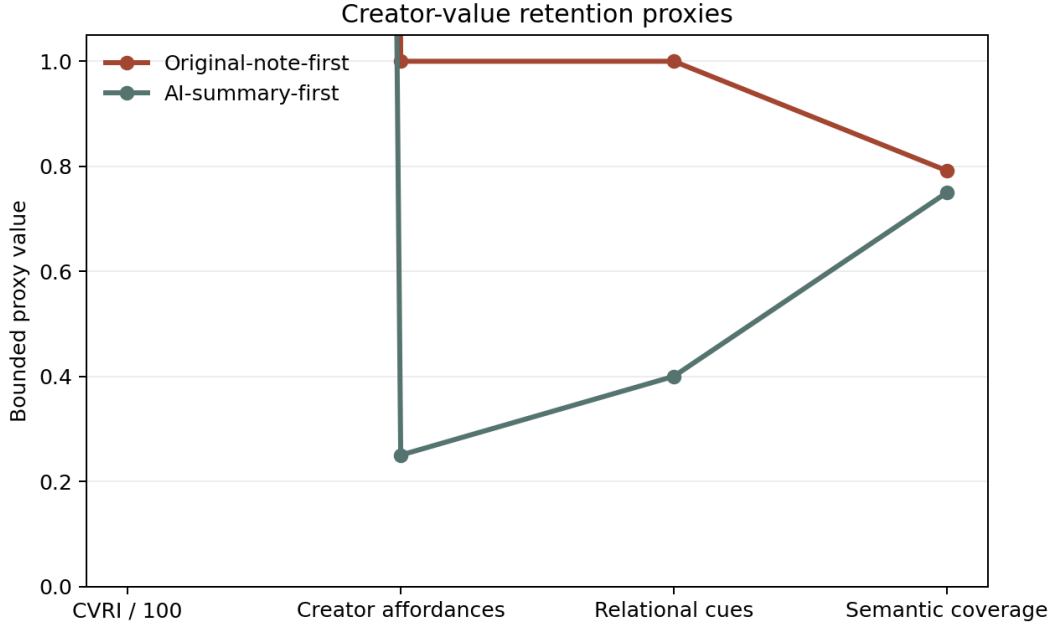


Figure 2. Creator value retention proxies

2.3. Friction Redistribution

Let latent user-side acquisition friction represent effort required to obtain a usable answer. Let latent creator-side return friction represent effort required to reach sources, creator context, and creator-directed actions. Their conceptual difference for condition x is defined in Eq. (1):

$$\Phi(x) = F_c(x) - F_u(x) \quad (1)$$

Equation (1) is interpretive and is not estimated in this study. No task time, perceived effort, completion rate, or answer-seeking behavior was collected. A future larger value would not by itself imply exploitation or economic loss. The model instead states a hypothesis that convenience could be distributed asymmetrically. In the present stimuli, a visible creator name coexists with a deeper source layer, fewer identity fields, reduced source diversity, and missing comment or message affordances. CVRA measures those interface properties separately and uses possible friction redistribution only to interpret their joint arrangement.

3. Creator Value Retention Audit Framework

3.1. Audit Unit and Component Model

CVRA uses a matched scenario as its analytical unit. Each scenario supplies the same domain, creator identity, source-note content, semantic requirements, and expected creator-facing cues to two interface conditions. Only the access path changes. In original-note-first stimuli, creator identity, the full note, source types, relational cues, and four creator-directed actions appear on the first surface. In AI-summary-first stimuli, a concise answer appears first, attribution remains

visible, source access passes through a details intermediary, and fewer creator-related elements remain. The method therefore compares controlled interface states rather than comparing unrelated posts. This matched design removes content topic and creator identity as between-condition explanations within each scenario.

The full CVRA workflow shown in Fig. 1 begins with matched content controls and renders both access paths under an identical viewport. It then extracts seven auditable component scores and reports both components and CVRI. Evidence labels travel with the outputs so that synthetic measurements cannot be mistaken for production observations. Sensitivity analysis follows composite scoring and exposes dependence on the chosen weights.

Rendered attribution salience combines visibility, contrast, and occupied viewport area. Let V denote binary visibility, C the WCAG contrast ratio, and R the attribution-area ratio. The normalization thresholds are 4.5 for text contrast and 0.018 for visible area. Attribution salience A is defined in Eq. (2):

$$A = V \min\left(1, \frac{C}{4.5}\right) \min\left(1, \frac{R}{0.018}\right) \quad (2)$$

Source accessibility converts 1-indexed source interface-layer depth l into a bounded score in Eq. (3). Layer one places source access on the initial content surface. Layer two places it behind one declared details intermediary. The stored values count interface layers rather than observed clicks. The mapping describes architecture and does not quantify behavioral effort.

$$S = \max\left(0, 1 - \frac{l-1}{3}\right) \quad (3)$$

Creator-facing retention components are normalized

against elements deliberately supplied by the source condition. Let f be visible creator affordances out of four, r visible relational cue types out of five, i visible identity fields out of four, and s visible source types out of three. Equation (4) defines their bounded scores:

$$F = \min\left(1, \frac{f}{4}\right), \quad R_c = \min\left(1, \frac{r}{5}\right), \quad I = \frac{i}{4}, \quad D = \min\left(1, \frac{s}{3}\right) \quad (4)$$

The component vector in Eq. (5) combines attribution, source access, affordances, relational cues, identity, source diversity, and lexical coverage. It is the primary CVRA output because each element has a direct interface interpretation.

$$\mathbf{z}_{jx} = [A_{jx}, S_{jx}, F_{jx}, R_{c,jx}, I_{jx}, D_{jx}, Q_{jx}] \quad (5)$$

3.2. Lexical Coverage and Compression Measures

The legacy experiment field `semantic_coverage` stores a recall-style lexical measure. Let T_j be the prespecified audit vocabulary for scenario j , and let $B(j,x)$ be the token set visible on the content surface. A term counts as present when it matches a visible token or its defined token prefix. The scenario vocabulary guides the audit but is not guaranteed to appear verbatim even in a generated full source note after text

assembly and tokenization. Equation (6) therefore measures lexical visibility, not factual completeness, nuance, comprehension, or semantic equivalence.

$$Q_{jx} = \frac{1}{|T_j|} \sum_{t \in T_j} \mathbb{1}[t \in B(j,x)] \quad (6)$$

Compression is measured separately because a short surface can retain required terms while removing narrative context. Let L_{jx} be visible content characters and L_{jO} be original-note characters for the same scenario. Equation (7) defines the compression ratio K :

$$K_{jx} = \frac{L_{jx}}{L_{jO}} \quad (7)$$

Compression is not a CVRI component. Including it would reward longer interfaces regardless of relevance. It instead diagnoses how much textual material remains. Fig. 3 compares rendered geometry, normalized source interface-layer depth, attribution contrast, and compression. Both conditions made attribution readable, while layer depth and content length differed by design. Their zero within-pair variance is therefore reported descriptively rather than converted into inferential claims.

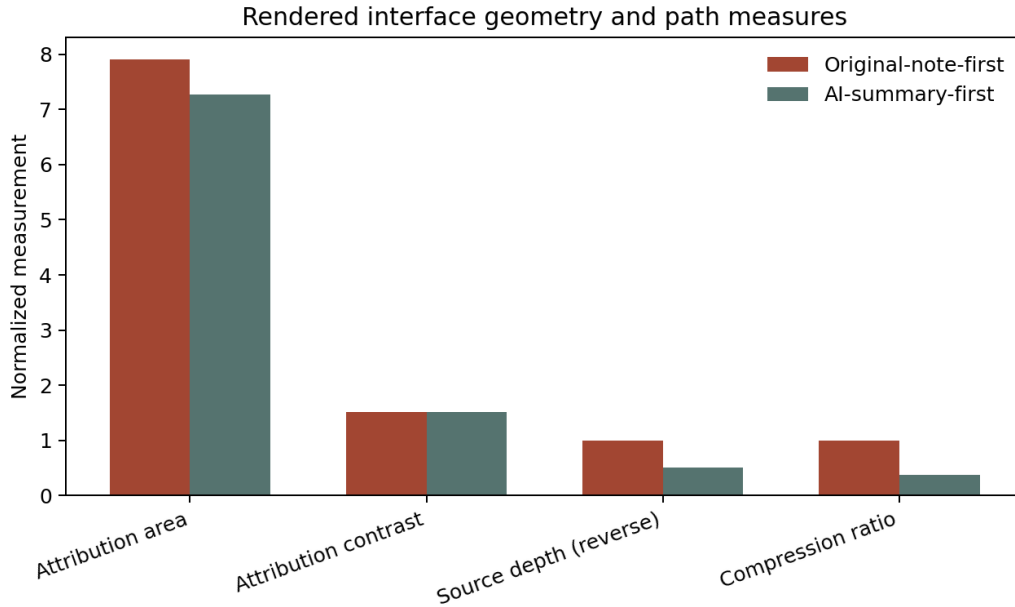


Figure 3. Rendered geometry and path measures

3.3. Creator Value Retention Index

CVRI summarizes the seven component scores when a compact comparison is useful. Table II lists the fixed weights. Attribution receives 0.20, creator affordances 0.18, relational

cues and identity 0.16 each, source access 0.12, lexical coverage 0.10, and source diversity 0.08. These values are normative analytic choices, not parameters learned from creator outcomes. The weights sum to one and remain unchanged across scenarios and conditions.

Table 2. CVRI Components and Fixed Weights

Component	Symbol	Observable basis	Weight
Attribution salience	A	Visibility, contrast, and area	0.20
Source accessibility	S	1-indexed source interface-layer depth	0.12
Creator affordances	F	Follow, profile, comment, message	0.18
Relational cues	Rc	Five source-note cue types	0.16
Creator representation	I	Name, handle, avatar, profile link	0.16
Source diversity	D	Distinct visible source types	0.08
Lexical coverage	Q	Required visible scenario terms	0.10

For nonnegative weights w_k that sum to one, Eq. (8) defines CVRI as the weighted component sum scaled to 100. The bounded score is calculated independently for each scenario-condition pair.

$$CVRI_{jx} = 100 \sum_{k=1}^7 w_k z_{jxk}, \quad w_k \geq 0, \quad \sum_{k=1}^7 w_k = 1 \quad (8)$$

The name “creator value retention” refers only to retention of specified interface pathways. It does not claim that the index measures total creator value. It also does not assign monetary worth to clicks, cues, or identity fields. Researchers should report the component vector with the composite, inspect weight sensitivity, and revise the audit ontology for other platforms. CVRI is most defensible as an auditable summary of design choices, not as a universal welfare score.

4. Controlled Browser Experiment

4.1. Stimuli and Reproducibility

The experiment generated 40 matched scenarios across eight lifestyle domains and rendered two local HTML stimuli per scenario. The resulting dataset contained 80 pages.

Scenario order and content generation used seed 20260901. All pages were rendered through a headless Chromium browser using Playwright at 1280 by 900 pixels with device scale factor one. Python 3.10.18 executed the pipeline. The run recorded 5,000 bootstrap iterations, a SHA-256 hash of the runner, output inventory, timestamp, and the declaration that all stimuli were synthetic. The screenshots stored by the experiment document local stimulus rendering only. They are neither captured Xiaohongshu pages nor replicas asserted to match a production interface.

The browser collected bounding-box geometry, computed styles, DOM-visible source types, relational cue types, creator affordances, declared source interface-layer depth, visible text, audit terms, and original character counts. Contrast followed the standard relative-luminance calculation. Identity representation counted name, handle, avatar, and profile link fields. Each row in the output CSV corresponds to one scenario-condition pair. The summary file aggregates condition means and matched differences. Fig. 4 displays CVRI against lexical coverage for all scenarios, allowing the limited dispersion of the synthetic design to remain visible rather than hiding it behind averages.

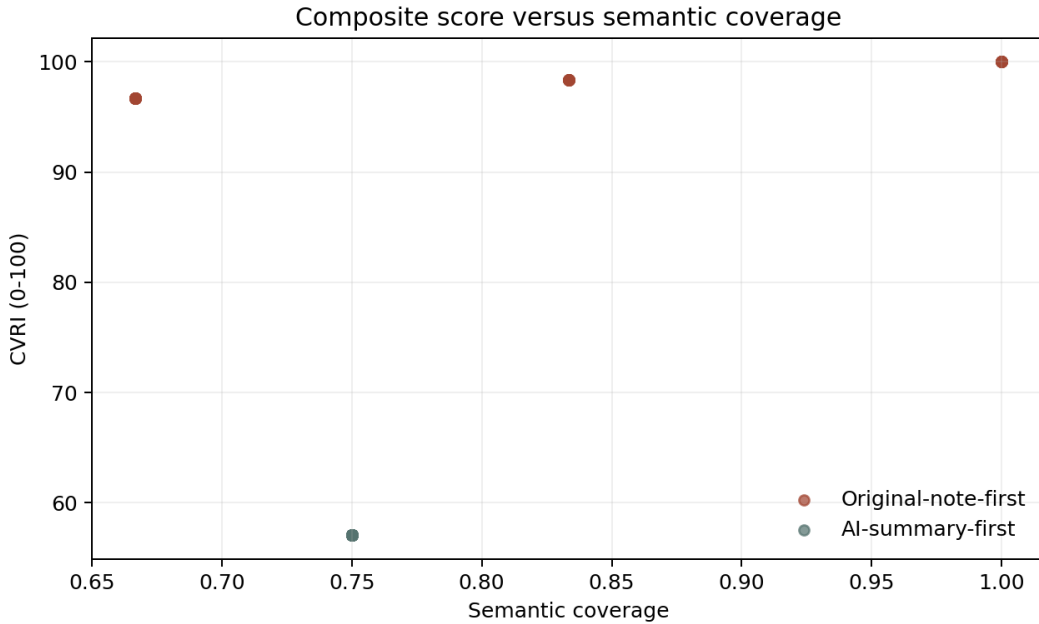


Figure 4. CVRI and lexical coverage

4.2. Paired Analysis

For metric m and matched scenario j , the paired difference subtracts the original-note-first value from the AI-summary-first value. Equation (9) then calculates the mean difference across 40 scenarios.

$$\Delta_{jm} = y_{jm}^{(AI)} - y_{jm}^{(O)}, \quad \bar{\Delta}_m = \frac{1}{40} \sum_{j=1}^{40} \Delta_{jm} \quad (9)$$

The analysis file mechanically includes paired t tests, Wilcoxon signed-rank tests, standardized paired differences, and percentile bootstrap intervals. In this paper, they are generator-resampling diagnostics with no external population target. Attribution visibility, attribution area, contrast, source interface-layer depth, creator affordances, relational cues, creator representation, and source diversity were fixed by the templates. Their within-pair differences consequently had

zero variance. Infinite t statistics or zero p values for nonzero constants do not represent population evidence. Those components are reported descriptively. Diagnostics for varying measures characterize only repeated scenarios from this generator, not production Xiaohongshu.

Fig. 5 presents paired differences across the five displayed outcomes. The vertical distance from zero is an interface-design difference under the controlled specification. For lexical coverage, the mean difference was -0.04167 . Its generator diagnostics included a paired t -test p value of 0.0672 , a Wilcoxon p value of 0.1553 , and a bootstrap interval from -0.08333 to approximately zero. Compression differed by -0.63615 , with a generator-resampling interval of $[-0.64191, -0.63002]$. These values summarize stability under synthetic scenario resampling. They are not confidence statements about real summaries or users.

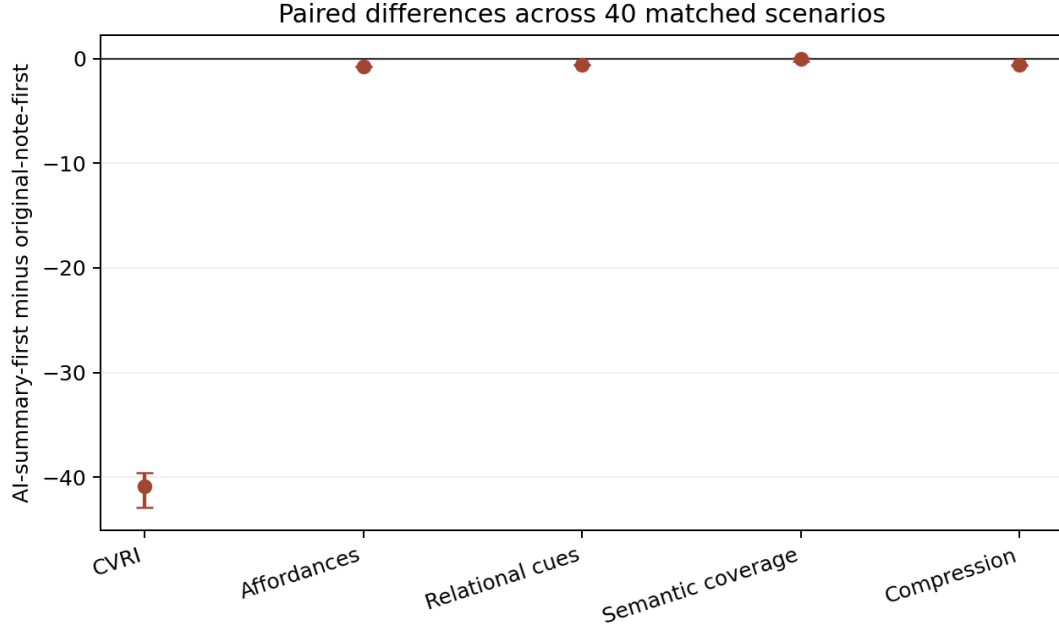


Figure 5. Matched scenario paired differences

4.3. Sensitivity Protocol

Weight sensitivity is evaluated in two ways. First, all seven components receive equal weight. Second, each component is removed in turn, with remaining weights renormalized. For removed component h , Eq. (10) defines the leave-one-out CVRI:

$$CVRI_{jx}^{(-h)} = 100 \sum_{k \neq h} \frac{w_k}{1 - w_h} z_{jxk} \quad (10)$$

This procedure tests directional dependence on the ontology, not statistical uncertainty. It answers whether the observed controlled gap exists only because one component received a favorable weight. It cannot establish the socially correct weights. Such weights would require stakeholder deliberation, creator research, and outcome validation.

5. Results and Discussion

5.1. Component Results

Table III reports the actual condition means. Attribution remained visible in every stimulus, with identical mean contrast of 15.04376. The attribution area occupied 7.91458% of the viewport in original-note-first pages and 7.27778% in summary-first pages. Both ratios exceeded the salience normalization threshold, so attribution scored one in both conditions. This result prevents a simplistic conclusion that summaries necessarily erase attribution. In these stimuli, attribution survived. The prespecified contrasts concerned interface-layer depth, creator actions, relational cues, identity fields, source diversity, and compression.

Table 3. Controlled Condition Mean Measurements

Metric	Original-note-first	AI-summary-first	Summary minus original
Attribution visibility	1.00000	1.00000	0.00000
Attribution area ratio	0.07915	0.07278	-0.00637
Attribution contrast ratio	15.04376	15.04376	0.00000
Source interface-layer depth	1.00000	2.00000	1.00000
Creator affordance retention	1.00000	0.25000	-0.75000
Relational cue retention	1.00000	0.40000	-0.60000
Creator representation	1.00000	0.50000	-0.50000
Source diversity	1.00000	0.33333	-0.66667
Lexical coverage	0.79167	0.75000	-0.04167
Compression ratio	1.00000	0.36385	-0.63615
CVRI	97.91667	57.06667	-40.85000

Creator affordance retention fell from one to 0.25 because the original condition exposed four expected actions while the summary condition exposed one. Relational cue retention fell from one to 0.40, creator representation from one to 0.50, and source diversity from one to one third. The source occupied interface layer one in the original condition and layer two in the summary condition. These are exact stimulus-template

properties, not observed click counts. Fig. 6 shows the measured path-related component differences interpreted through possible friction redistribution. The answer-first surface preserved a readable creator header while placing a declared intermediary between answer and source. Possible friction redistribution remains an interpretation of that arrangement, not a measured outcome.

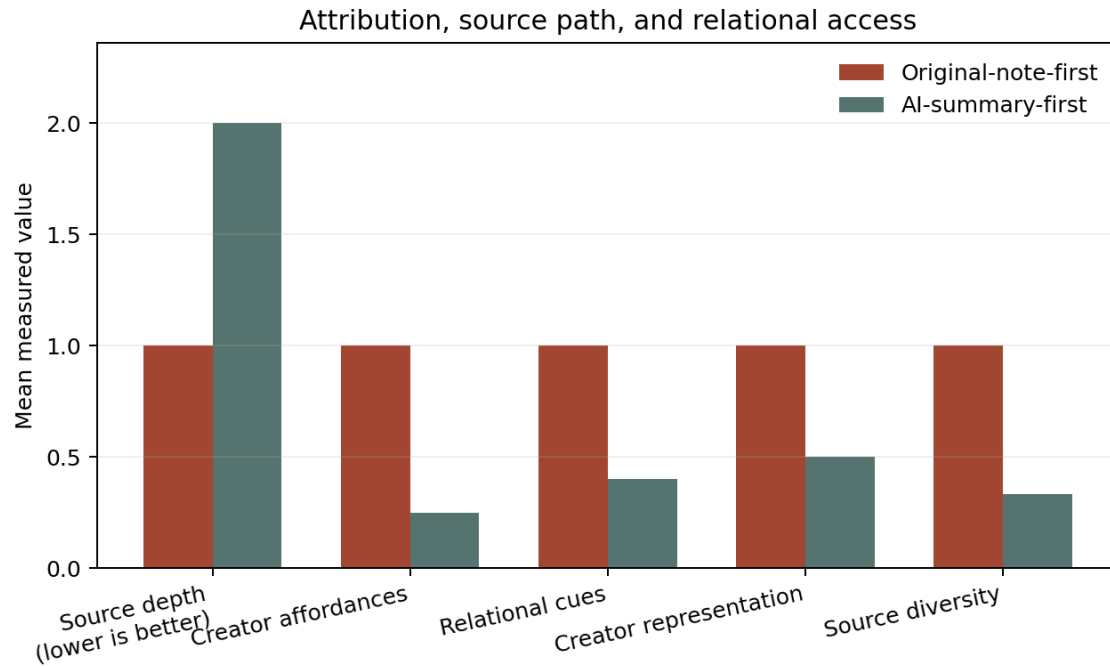


Figure 6. Source path and relational access

Lexical coverage changed much less than relational measures. Original-note-first pages averaged 0.79167, while summary-first pages averaged 0.75000. Original values were 0.66667, 0.83333, or 1.00000, whereas all generated summaries scored 0.75000. The original mean was below one because scenario terms formed a prespecified audit vocabulary and were not guaranteed to appear verbatim in every assembled full note. The metric therefore cannot establish factual completeness. Summary pages retained many audit terms while containing 36.385% of original visible characters on average. Compression ratios ranged from 0.33985 to 0.40056. This result demonstrates lexical retention under designed compression, not preserved meaning or user understanding.

5.2. Composite Results and Sensitivity

The original-note-first condition obtained a mean CVRI of 97.92, with scenario values from 96.67 to 100.00. The AI-summary-first condition obtained 57.07 in every scenario because all its CVRI components were fixed by design. The mean matched gap was -40.85 points. Generator-resampling produced an interval from -41.31 to -40.43, a Wilcoxon p value of 1.94 times 10 to the power of -8, and a standardized difference of -29.19. These diagnostics have no external population target. Their extreme values reflect a deliberately consistent generator, not a huge real-world effect. Fig. 4 discloses the constrained dispersion, while Fig. 5 keeps the matched design contrasts visible.

The equal-weight CVRI produced means of 97.02 and 55.71, a gap of -41.31 points. Leave-one-component-out analysis preserved a negative gap in every configuration. The gap ranged from -33.35 points when creator affordances were removed to -51.06 points when attribution was removed. Removing attribution enlarges the gap because attribution had identical full scores and therefore moderated the base difference. Removing creator affordances reduces the gap because that component carried a large designed contrast. Table IV makes these dependencies visible. The direction is stable within this ontology, while magnitude remains weight-sensitive.

Table 4. CVRI Sensitivity and Component Ablations

Weight specification	Mean CVRI gap
Prespecified weights	-40.85
Equal weights	-41.31
Remove attribution salience	-51.06
Remove source accessibility	-41.88
Remove creator affordances	-33.35
Remove relational cues	-37.20
Remove creator representation	-39.11
Remove source diversity	-38.61
Remove lexical coverage	-44.93

5.3. Theoretical Interpretation

The results support a narrow internal-sensitivity claim. CVRA responded when controlled summary-first pages retained fewer pathways to creator identity, relationships, and actions. Friction redistribution provides one interpretation of that pattern. The summary stimulus presented an answer before the full source, while the creator-oriented route occupied a deeper layer and contained fewer cues. However, the study did not establish that users experienced greater convenience or source friction. Frictionlessness is therefore treated as a hypothesis about how effort may be allocated among tasks and stakeholders, not as a property demonstrated by these measurements.

This mechanism also clarifies why content labour and data labour should not be collapsed. The experiment does not reveal whether the synthetic “source notes” resemble any particular production pipeline. It instead shows how a platform could preserve content utility as machine-organized answer material while reducing the surrounding relations that helped make the source valuable to its creator. The source remains named, yet fewer creator actions and identity cues remain available. The framework therefore treats datafication as a change in functional arrangement, not as proof of unauthorized training or uncompensated appropriation.

5.4. Governance and Design Implications

CVRA can inform design review because each component

maps to an inspectable intervention. A platform can provide direct source links, preserve creator handles and avatars, expose profile and comment actions, show multiple source types, and disclose the number or diversity of underlying notes. It can also distinguish machine-generated interpretation from creator-authored statements. These changes do not require a ban on summaries. They seek reciprocity between convenient answer acquisition and reachable creator context. In high-risk domains, stronger source diversity and uncertainty cues may also improve user protection. The present experiment does not test those interventions, but its metrics provide an evaluation protocol for future controlled comparisons.

Governance should also distinguish attribution from meaningful control. A visible name does not answer whether creators received notice, could opt out, could correct a summary, or could understand reuse conditions. Contractual analysis can identify formal provisions, but interface auditing cannot verify informed consent. Conversely, a missing interface cue does not prove a contractual violation. CVRA is designed to complement policy analysis, creator interviews, and platform accountability research. It should not replace them.

5.5. Limitations and Future Validation

The strongest limitation is ecological validity. Every stimulus was local, synthetic, and generated by a controlled script. No page was captured from production Xiaohongshu. The audit did not observe rankings, query coverage, personalized recommendations, user scrolling, clicks, dwell time, task effort, comprehension, trust, creator reach, brand cooperation, revenue, or income. It also did not test whether a real Xiaohongshu summary contains the same cues or source interface-layer depth. Accordingly, the experiment cannot estimate perceived friction, traffic displacement, engagement decline, creator compensation, or causal benefit transfer.

The component ontology and weights introduce construct-validity limits. Four actions, five relational cues, four identity fields, and three source types are reasonable audit anchors, but they are not universal. Their importance may vary by creator type, content category, device, disability, or commercial model. Lexical coverage relies on required-term matching rather than contextual similarity or factual correctness. Character compression ignores images, video, tone, and narrative sequencing. Finally, the matched tests operate on scenarios generated from one script, so conventional population inference is weak. Future work should preregister platform samples, archive production interfaces lawfully, recruit users and creators, validate weights through stakeholder methods, and connect proxy changes to observed behavior without conflating correlation with income loss.

6. Conclusion

This paper introduced CVRA, a computational framework for auditing how AI-summary-first interfaces retain creator-facing pathways. Possible friction redistribution supplies an interpretive mechanism rather than a computed outcome. The study did not measure effort for answer acquisition or source exploration, and it did not estimate Φ in Eq. (1). Its component model separates attribution, source interface-layer accessibility, creator affordances, relational cues, identity representation, source diversity, and lexical coverage. CVRI

provides a transparent bounded summary, while sensitivity analysis prevents one weight specification from appearing natural or inevitable.

The executed browser experiment compared 40 matched scenarios through 80 synthetic local stimuli. Summary-first pages retained visible, high-contrast attribution and many prespecified audit terms. They also placed sources at layer two instead of layer one, retained one quarter of expected creator affordances, retained 40% of relational cues, and compressed text to 36.38% of original length. Mean CVRI declined from 97.92 to 57.07. Equal-weight and component-ablation analyses preserved the controlled gap direction. These values demonstrate internal sensitivity to synthetic differences, not Xiaohongshu's production impact.

The main implication is methodological. Debate about AI search and creator labour needs measures that can identify interface redistribution without inventing economic evidence. CVRA offers such a bridge, provided its boundaries remain explicit. Reported facts motivate the Xiaohongshu case, controlled stimuli test the framework, and theory interprets possible consequences. Claims about actual traffic, attention, revenue, consent, or creator welfare require separate evidence. The reproducible component outputs also permit platforms, regulators, and researchers to disagree about weights without disputing the underlying observations. That separation is essential for accountable audit practice because normative judgments remain visible rather than being hidden inside one score. Future research should combine lawful longitudinal interface sampling, behavioral studies, creator participation, and platform data access to determine when friction redistribution becomes material creator harm.

References

- [1] Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292. <https://doi.org/10.1177/1461444818796915>
- [2] Helmond, A. (2015). The platformization of the web: Making web data platform ready. *Social Media + Society*, 1(2), 1–11. <https://doi.org/10.1177/2056305115593001>
- [3] Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social Text*, 18(2), 33–58. https://doi.org/10.1215/01642472-18-2_59-33
- [4] Grohmann, R. (2024). Free labor and data labor in the digital economy. In T. Venturini, A. Acker, J.-C. Plantin, & T. Walford (Eds.), *The SAGE handbook of data and society* (Ch. 7). SAGE.
- [5] Hearn, A. (2008). Meat, mask, burden: Probing the contours of the branded self. *Journal of Consumer Culture*, 8(2), 197–217. <https://doi.org/10.1177/1469540508090409>
- [6] Duffy, B. E. (2017). (Not) getting paid to do what you love: Gender, social media, and aspirational work. Yale University Press.
- [7] van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.
- [8] Nielsen, J. (1993). *Usability engineering*. Morgan Kaufmann.
- [9] Plantin, J.-C., & Punathambekar, A. (2019). Digital media infrastructures: Pipes, platforms, and politics. *Media, Culture & Society*, 41(2), 163–174. <https://doi.org/10.1177/0163443718818673>

- [10] Gillespie, T. (2010). The politics of platforms. *New Media & Society*, 12(3), 347–364. <https://doi.org/10.1177/1461444809342738>
- [11] Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- [12] Xiaohongshu Official Search Account. (2025, June 12). Xiaohongshu AI search is now live! Come and ask any questions! Xiaohongshu. <http://xhslink.com/o/82CghVXkJ2i>
- [13] Zhao, S. (2026, February 24). Behind the launch of voice Ask AI: Has Xiaohongshu's answer logic changed? Jiemian News. <https://m.jiemian.com/article/14030994.html>
- [14] Xiao, S. (2026, May 7). Why has Xiaohongshu shifted from AI restraint to sudden acceleration? 36Kr. <https://36kr.com/p/3799028783439111>
- [15] Xiaohongshu. (2024). Xiaohongshu generative AI service agreement. <https://agree.xiaohongshu.com/h5/terms/ZXXY20240103002/-1>