ChatGPT

Designing CREW for Research Integrity: Mitigating Citation Metrics and Bias

Introduction. The CREW platform must foster a culture of integrity in research evaluation and publication. Rather than relying on simplistic citation counts or journal prestige, CREW will use evidence-based, community-driven mechanisms. In contrast to traditional closed peer review, CREW envisions an open, community-oriented review workflow. *Figure: Traditional vs. community peer review workflows. Traditional review relies on a few anonymous reviewers; community review invites broader, transparent feedback.* By broadening review and feedback loops, CREW helps expose and correct bias, as noted by Alves *et al.* [1]. Our goal is to **prevent** systemic distortions before they occur, using built-in technical, structural, and cultural safeguards.

Challenges: Citation-Based Metrics and Human Bias

- **Overreliance on citation-based metrics.** When resource allocation or prestige is tied to raw publication or citation counts, researchers are incentivized to "publish or perish" and cite in self-promoting ways. Critics note that administrators often use simple counts of papers and citations to distribute grants or jobs, even though ideal evaluation should "be based on merit and scientific integrity" [2]. This can encourage quantity over quality, citation cartels, and neglect of negative or replication studies. In CREW, no single numeric score (such as citation tally) will determine reputation or reward. We explicitly **decouple** user profiles and node reputation from citation totals, preventing gaming by citation accumulation.

- **Social and demographic biases.** Human reviewers can be swayed by an author's gender, nationality, institution or seniority. Studies show these factors often skew review outcomes [1]. For example, women and researchers from underrepresented regions face disproportionate disadvantage, and anonymity can sometimes shield unfair criticism [1]. CREW's design must therefore minimize such biases: by anonymizing sensitive attributes where possible, diversifying reviewer selection, and auditing decision patterns for demographic skew.

- **Hierarchical and groupthink bias.** Traditional peer review is hierarchical (editor ► reviewers ► author) and tends to reinforce established authority and common opinions. This can suppress novel ideas or interdisciplinary work. CREW implements **non-hierarchical review mechanisms**: any qualified participant can contribute feedback on a manuscript or dataset. Reviews and comments are collected in a flat, distributed manner rather than controlled by a single editor. As Alves observes, decentralized models "diversify participation" and shift evaluation into a more public space, breaking closed "in-group" dynamics [1].

- **Lack of incentives for review and reproducibility.** Reviewers are often unpaid and uncredited, leading to reviewer fatigue and superficial reviews [3]. Important tasks like replication or data sharing have little reward. CREW addresses this by building in **incentive tokens and credentials**: reviewers earn verifiable credit (e.g. a signed "review certificate") for thorough reviews [3] [4].

Validators of replication studies or data curation likewise receive recognition. This ensures the "behind-the-scenes" labor of peer review and reproducibility is visible and acknowledged [4].

Technical Mechanisms

- **Rich metadata and provenance.** Every CREW submission carries machine-readable metadata: author ORCID, affiliations, funding sources, data/code links, and keywords. Authors explicitly record their contributions (concept, experiment, analysis, writing) in structured form. This provenance metadata is cryptographically signed (e.g. by the authors' private keys) and timestamped. An immutable record links each manuscript version to its data, code, and review history. As Morales-Alarcón *et al.* note, blockchain-style ledgers can "securely record all stages of the process," ensuring data cannot be surreptitiously altered [5]. CREW nodes will implement tamper-evident logs (using, e.g., append-only ledgers or blockchain backends) so that paper submissions, review events, and publication decisions are traceable in full.

- **Pseudonymous review with accountability.** To reduce social bias yet maintain trust, reviews in CREW are submitted under persistent pseudonyms linked to verified identities. Each reviewer has a cryptographic ID (tied to ORCID or institutional account) but publishes reviews under a handle. The system tracks reviewer performance anonymously: excellent reviews earn public "reputation badges" without revealing the reviewer's real name. Finke & Hensel's proposal highlights issuing digital certificates and tracking reviewer reputation in a distributed system [3]. CREW adopts this idea: every completed review yields a signed certificate (verifiable by any node) that updates the reviewer's profile. This punishes low-effort reviews (by withholding certificates) and encourages high-quality feedback. Technical implementation can use smart contracts or similar mechanisms to govern certificate issuance [6].

- **Audit trails and transparency.** All actions (submission, review, decision, editorial comments) are appended to a secure audit trail. CREW nodes intercommunicate to replicate these logs, preventing any single node from tampering with history. We leverage **cryptographic audit trails**: for example, each review's hash can be chained so any change is detectable. As one review noted, blockchain ensures that "reviews, revisions, and data are permanently recorded, enabling complete transparency and accountability" [7]. In practice, CREW's blockchain or ledger stores hashes of manuscripts and review reports. Authorized users can retrieve a full timeline of each paper's processing. This auditability deters misconduct: if a conflict of interest affected a decision, the records will expose it.

- **Smart-contract governance.** Automated rules (encoded as smart contracts or workflow engines) enforce fairness. For example, conflicts-of-interest prompts are mandatory: if a reviewer is from the same institution as an author, the contract flags a COI and may reassign. Review assignments can be randomized or lottery-based to prevent editor bias. Reward distribution (tokens or badges) is automatically triggered by contract once a review is completed and signed. Smart contracts thereby "automate and ensure the execution of agreements" (from rewards to IP management) "guaranteeing transparency and integrity in all academic transactions" [6]. All policies (e.g. anonymity guidelines, revision deadlines, appeal processes) are transparent and encoded, leaving less room for arbitrary human decision.

Structural Design Choices

- **Decoupling reputation from citations.** In CREW, user reputation is multi-dimensional and **not determined by citation counts**. Instead, reputation scores integrate peer feedback quality, review contributions, data-sharing practices, and community endorsements. For instance, a researcher gains merit by reviewing data reproducibility, mentoring early-career scholars, or publishing thorough negative results – none of which are captured by citations. Evaluation algorithms explicitly exclude raw citation metrics. As experts warn, overreliance on one-dimensional metrics reinforces inequities [1]; CREW's reputation model is designed to counter that by construction.

- **Non-hierarchical, collaborative review workflow.** CREW replaces single-editor decisions with a **community consensus** model. When a paper is ready, it is posted for open peer review: any qualified member (often via a token stake or invitation) can submit a review or comment. Decisions (accept/reject/major revision) emerge from an open vote or deliberation by the review community, not just a small editorial panel. This flattens power structures and reduces gatekeeping. All participants see each other's comments (possibly after anonymization), allowing counterpoints and consensus-building. Such decentralized review expands inclusion (early-career researchers can participate) and dilutes individual biases, as suggested by the move toward "decentralized, community-driven peer review" [8].

- **Continuous post-publication review.** CREW allows ongoing commentary even after formal acceptance. Registered users can annotate, replicate, or dispute published results, in context. Post-publication review is transparent and added to the audit trail. This not only crowdsources quality control but also discourages confirmation bias: since criticism is visible, authors and readers are aware that any flaws will be openly challenged.

- **Incentives and recognition redesign.** Rather than grant tenure or funding based on citation benchmarks, CREW ties rewards to community-defined contributions. Nodes or institutions gain reputation by facilitating collaboration (e.g. hosting data, organizing workshops) and upholding standards, not by citation quantity. Token mechanisms (digital or reputational) are used: for example, small tokens can be granted to reviewers, data curators, or reproducibility validators. These tokens are convertible to access or privileges within CREW (e.g. increased voting weight on decisions), aligning incentives with service to science. As one analysis notes, token-based incentives "ensure a rigorous review process" and acknowledge contributors behind the scenes [4].

Cultural Affordances

- **Transparency rituals.** CREW embeds cultural practices that normalize openness. For instance, every decision (accept, revise, reject) must be accompanied by a public rationale: a structured statement explaining the reasoning. Decision statements are linked to evidence (e.g. "Experiment 3 lacked control; see detail in report"). These become part of the record. CREW nodes can schedule regular "integrity audits" where review logs and metadata are randomly sampled for consistency checks. Transparency also means any conflicts of interest and funding disclosures are explicitly recorded before review begins. These rituals – making process visible at each step – help inculcate a norm of accountability.

- **Review and citation disclosure prompts.** The interface actively reminds authors and reviewers to reflect on bias. For example, when recommending citations, a popup asks: *"Have you cited recent work that challenges your findings?"* or *"Are any potential conflicts of interest present in these references?"* Reviewers see prompts to consider diversity (e.g. "Check if at least one reference is from an early-career scholar"). Such *nudges* promote ethical behavior. Users must complete disclosure forms (e.g. funding sources, affiliations) as part of submission. By building these prompts into CREW's workflow, we make ethical reflection routine.

- **Community governance and training.** CREW nodes foster a culture of integrity through shared norms. Each node maintains clear guidelines on research ethics, open data, and proper credit. New users undergo orientation on bias mitigation and proper citation practices. CREW's open discussion forums and "code of conduct" tutorials reinforce peer accountability. Regular community forums (virtual or in-person) allow members to discuss integrity cases and update best practices. By interweaving cultural cues (norms, rituals, education) with the technical platform, CREW ensures that preventing misconduct is seen as everyone's responsibility, not just a bureaucratic requirement.

Conclusion

By **designing out** incentives for misconduct and bias, CREW sets a new standard for scholarly platforms. Technical measures (immutable metadata, audit trails, cryptographic credentials) work in concert with structural and cultural policies (open review workflows, decoupled reputations, transparency rituals) to make integrity the path of least resistance. As experts conclude, embracing open, transparent, and diverse review systems leads to a more equitable and effective research ecosystem [8]. CREW's preventive architecture thus ensures that trust and rigor are built in from the start, rather than being an afterthought.

Sources: The strategies outlined draw on recent proposals for decentralized review and open science [3] [5] [1] [4] [6], adapting them into a cohesive design tailored for CREW's implementers.

---

[1] [8] Diversification and Decentralization of Peer Review: Part 1—Initiatives at the Forefront - Science Editor
https://www.csescienceeditor.org/article/diversification-and-decentralization-of-peer-review-part-1/

[2] [3] Decentralized Peer Review in Open Science: A Mechanism Proposal
https://arxiv.org/html/2404.18148v1

[4] [7] How Decentralized Science (DeSci) Improves Research
https://www.ulam.io/blog/how-decentralized-science-is-revolutionizing-research

[5] [6] Blockchain and Its Application in the Peer Review of Scientific Works: A Systematic Review
https://www.mdpi.com/2304-6775/12/4/40