

Since one of our Collective Values is Iterative Innovation, we are continuing to refine retroactive public goods funding (retro funding) with round 4. The focus here is on collaboratively curating impact metrics, guided by the principle of [Metrics-based Evaluation](#). Unlike previous Retro Funding rounds, Retro Funding 4 shifts from voting on individual projects to engaging badgeholders in selecting and weighting metrics that measure various types of impact.

This following runthrough will showcase the stages of this process—from the initial surveys and workshop insights to post-workshop adjustments—and how we're integrating your feedback to shape a metrics framework that accurately reflects the impact we aim to reward.

## Retro Funding 4's New Direction: Metrics-based evaluation

The core hypothesis of Retro Funding 4 is that by using quantitative metrics, our community can more accurately express preferences for the types of impact they wish to see and reward. This approach marks a significant shift from earlier Retro Funding iterations where badgeholders voted directly on projects. In RF4, the emphasis is on leveraging data to power impact evaluation, ensuring that metrics are robust, verifiable, and truly representative of the community's values.

One important thing to keep in mind is that this is one of many experiments that we'll be running this year and you can find all the details of what's to come [right here](#).

### The Initial Survey:

Establishing the baseline

The initial survey we sent out was a means of gauging community perspectives on qualitative and quantitative measurement approaches and gathering input on a potential first draft list of metrics.

Key findings included:

- Balance in metrics: Badgeholders showed a relatively balanced interest in metrics assessing both quantitative growth and qualitative impact, indicating a need for an evaluation framework that incorporates both dimensions.
- Innovation and Open Source: Metrics related to innovation and open source scored significantly higher, reflecting a community preference for collaborative efforts.
- Concerns about gaming metrics: A notable concern was the potential for gaming the system, highlighting the need for metrics that are difficult to manipulate and that genuinely reflect impactful contributions.
- Engagement and a focus on quality: There was a strong preference for metrics that measure sustained engagement and ongoing quality vs one-time actions, highlighting the value placed on long-term impact.

[

1600×763 107 KB

](<https://europe1.discourse-cdn.com/bc41dd/original/2X/3/319c8b72e08e73c76d0fb4d84ac3daf79c95a48b.png>)

### The Workshop: Data Deep Dive and Refining Metrics

Building on the survey insights, the workshop we ran on May 7th enabled badgeholders to provide more in-depth feedback on the data-driven aspects of the impact metrics system as well as take a pass at selecting the most meaningful metrics grouped around:

Network growth

These are metrics designed to quantify the expansion and scalability of the network. They focus on observable, direct measurements and aim to capture high level activity.

Network quality

This aims to evaluate the trustworthiness and integrity of interactions on the network. They focus on depth and reliability.

User growth

The focus here is on measuring new, active and consistent participation in the network.

User quality

The quality measurements are aimed at assessing the value and consistency of user contributions, not just their numbers.

For a detailed look at the survey results, key moments and discussions, you can review the [workshop space here](#).

[

1600×870 114 KB

](https://europe1.discourse-cdn.com/bc41dd/original/2X/5/576b42fd7b4ec6a01e766fe52277fcb7c7584de4.jpeg)

You can also explore [this blog post](#) by Open Source Observer that recaps the work they've done so far to map up the metrics. OSO is collaborating with the Optimism Collective and its Badgeholder community to develop the impact metrics for assessing projects in RF4.

## Post-Workshop Survey: Refining metrics

Following the workshop discussions and feedback, and since not all Badgeholders were able to attend, a detailed survey was sent out to get further feedback on refined metrics, divided into the four identified segments:

- Network Growth: Feedback suggests a need for simple, intuitive, and directly measurable metrics like gas fees and transaction counts
- Network Quality: There is general support for metrics that consider the quality and trustworthiness of transactions, with suggestions to refine how 'trusted transactions' are defined and measured
- User Growth: A preference for metrics that accurately capture genuine user engagement emerged, with calls for clearer definitions of what constitutes an 'active' or 'engaged' user
- User Quality: Responses highlighted the challenge of measuring user quality without arbitrary metrics, suggesting a need for ongoing discussion and refinement to develop quality metrics that truly reflect contributions to the ecosystem

## Refining and implementing feedback

The insights from the surveys together with the discussions in the workshop are leading us to the next stage of metrics refinement. A key focus across the workshop and the second survey has been the granularity of the trusted users element so to respond to and integrate the insights, Open Source Observer (OSO) is working on the following:

Trusted user model

Based on some good suggestions and ideas for how to improve upon the "trusted user" model coming from the second survey shared, OSO are exploring several additional sources of data for the model.

Important Note:

Open Source Observer is only considering sources that leverage existing public datasets (no new solutions), are privacy-preserving, and can be applied to a broad set of addresses (>20K).

Action items:

- Extending the trusted user model: The current model relies on three trust signals (Farcaster IDs, Passport Scores, and EigenTrust by Karma3 Labs). OSO will now move forward with a larger set of potential trust signals such as:
  - Optimist NFT holders
  - Sybil classifiers used by the OP data team
  - Social graph analysis
  - Pre-permissionless Farcaster users
  - ENS ownership
- OSO will be running some light scenario analysis and choose the best model(s). They will also model a variety of options and share a report on the results.

Small fixes:

In the second survey, badgeholders also made a number of straightforward suggestions for how metrics could be standardized and implemented more consistently. These include:

- Limiting all relevant metrics to the 6 month Retro Funding 4 evaluation window
- Creating versions of some metrics on both linear and logarithmic scales
- Creating versions of "user growth" metrics that look at a larger set of addresses vs just the "trusted user" set

- Copy-editing to make descriptions clearer

Action items:

- OSO will proceed directly with making these changes

New metrics:

There were several metrics that had previously been considered and were proposed again in the feedback survey. These include looking at gas efficiency; power users; a larger set of addresses than just trusted users; and protocols that are favored by multisig users.

Action items:

- OSO will implement these new metrics where feasible

## **Next Steps: Testing metrics out**

The next step in this process will be to take the metrics set for a test drive. In the week of June 3rd, Badgeholders will have the chance to test out the metrics through an interface that simulates IRL voting. Badgeholders will also have the opportunity to submit extra feedback per metric as they test each one out.

Once this final round of feedback is collected and analyzed, we will be able to come back together in a session to discuss and decide on the final set of metrics to be used in Retro Funding 4.