

# Overview

Pairwise, initially developed with [a generous grant from Token House to be a voting app for RetroPGF 3](#) has demonstrated its potential as a tool for facilitating community engagement in funding decisions. This document serves to check Pairwise's performance in RetroPGF 3 and outline our vision for its enhanced role in RetroPGF 4 and beyond.

## Purpose

The goal is to assess Pairwise's impact, gather community feedback, and prepare for a more significant and refined role in future iterations.

## Development Journey

### Initial Phase

Funded by Token House, with General Magic fronting costs, the initial phase of Pairwise's development was a journey of discovery and innovation, originally poised to be a voting application. We completed the features planned for the voting application to complete our grant, but we couldn't use many of these features in RetroPGF 3 as we were not selected as a voting application. We pivoted from this initial phase to be a list generation tool.

### Team Dynamics

The development team exhibited strength and independence, successfully navigating the challenges of a tight timeline and evolving needs of RetroPGF 3, setting a strong foundation for future growth.

## Challenges and Improvements

Our biggest challenges were derived from pivoting to become a list creation tool as opposed to an end-end voting tool for RetroPGF. We are proud of our efforts to conform our application to this use case, despite the long nights and weekends needed to revise our initial product to these requirements. We also could have had more frequent engagement with both the foundation and the community to improve the product and development efficiency. Future development cycles will focus on better communication with both the community and the foundation so we can better anticipate the evolving needs of RetroPGF.

### Community Feedback

- The gamification and choosing between two choices was really fun
- Proposing lists with the tool was intuitive
- The first several pairs were really hard to rank because the badgeholders didn't know the projects yet, ranking projects got a lot easier after they reviewed all the projects in the round once.
- The Pairwise categories were inconsistent with other categories made by other groups, and some projects categorization was controversial.
- Some people wanted to use Pairwise to rank other lists instead of our default categories. We were able to [add a feature to do that](#) near the end of the round.
- Some people wanted to use Pairwise to generate their ballot, but didn't want to create a public list. We added an "export to .xls" option near the end of the round, but there was no way to import this into the ballot, so it still resulted in a lot of manual work.
- It was easy to discover and appreciate unknown smaller name projects with Pairwise because of the categories and being compelled to compare them against other projects.
- Some people requested an Undo button.
- Some people requested the ability to stop a project from being shown in pairings if they didn't want it to receive any OP
- Some people wanted to be able to give 0 OP to a project (we considered 0, an "abstain," so if someone wanted to give a project 0 we suggested they give them 1 OP)[

864×308 57.5 KB

](https://global.discourse-

- The most common and most difficult to manage feedback was that most badgeholders felt compelled to assess all of the projects, or at least vote for the projects they knew across a variety of categories, where as Pairwise was designed for badgeholders to focus on the categories they were experts on or had interest in.
- Many badgeholders created their own assessment system in spreadsheets and Pairwise was difficult to integrate into that system.

## Engagement and Technical Achievement

- The community responded positively, showing excitement about the voting mechanism.
- 185 users used Pairwise, only 27 badgeholders used their badgeholder key to connect to Pairwise and assess projects
- 26 lists were submitted by badgeholders
- 19,484 total Pairwise Rankings were made and 4,838 of those were made by badgeholders signed in with their badgeholder key
- We successfully met the tight deadline and implemented requested features, during the round to respond to badgeholders needs.

The community's positive response and excitement about the voting mechanism were encouraging. The successful engagement of badgeholders with Pairwise and their active participation in list submission highlighted the tool's effectiveness. Meeting the tight deadline and implementing requested features showcased our technical proficiency.

# Pairwise vision for RetroPGF 4.

## TL;DR

Pairwise aims to make voting both straightforward and enjoyable. We employ an algorithm similar to the [Elo rating system](#), which is the dominant way of doing ranking in the default world and it works!

We reviewed the [feedback for RetroPGF 3](#) and are excited to incorporate it to create an epic voting interface for RetroPGF 4. In this new phase, we're focusing on impactful and strategic funding allocation. We would like to propose a design where

- Only a little more than 50% of projects that apply receive funding.
- Badgeholders rank projects as opposed to giving each project an OP amount
- Badgeholders choose categories to review and rank the ~20 projects in each category, as opposed to being expected to rank all the projects.
- After the badgeholders rank the projects within a category, they will rank that category against the other categories they ranked
- We strongly suggest that we reward badgeholders for their engagement and will implement anti-gaming measures to ensure rewards can be fairly distributed.
- We will propose a system to increase transparency and badgeholder feedback to the projects while still maintaining the anonymity of the voting process.

Following this proposal, we are excited to receive feedback, incorporate it and build a user-friendly RetroPGF voting system for the collective.

## Why Pairwise?

Pairwise, is a great tool for making voting straightforward, simple and enjoyable. The UX is simple, two options pop up, and you pick one. The badgeholder does this many times and then that data is used to rank the projects. The algorithm used to rank the projects is very similar to the Elo rating system which is the dominant way of doing this work in the default world (chess, video games, professional and college sports, etc). Pairwise converts all these simple subjective inputs into objective, measurable outputs, minimizing the cognitive burden of voting.

## How will it work?

First let's imagine the user experience and then we will dive deep into the why behind the design choices.

## Our Proposed RetroPGF 4 Badgeholder UX

1. Badgeholder connects their wallet to Pairwise.vote.
2. Badgeholder chooses the category they will rank.
3. Badgeholder sees each project one at a time and reviews their details, choosing if they believe that project should receive RetroPGF or not.
4. The projects the badgeholder believes should receive funding are shown to them 2 at a time and the badgeholder picks the project that should receive more RetroPGF funding.
5. There will be 15-30 projects in each category, the more projects that the badgeholder believes should receive funding, the more assessments they will have to make before an accurate ranking can be shown.
6. Voting progress is saved as they go, so they can take a break at any time and pick it up at any time.
7. Once they have completed a category, they make a comment on their thought process. Comments will be anonymous and will be summarized by AI to obscure the writing style.
8. Badgeholder chooses a new category to rank and the steps above repeat.
9. Once badgeholders have completed multiple categories, they will rank the impact of the categories that they have assessed so far.
10. The ballots will be anonymized and made public along with the badgeholder's comments for each category at the end of the voting period.
11. Round results will therefore be independently verifiable.

## Why do it like this?

### The Power of Categories

One of the most important pieces of feedback given by badgeholders in both RetroPGF 2 & 3 was that there were too many projects to review, and that popular projects have too much of an advantage. Having strong categories can address both these issues.

[

809x672 97.4 KB

](<https://global.discourse-cdn.com/business7/uploads/bc41dd/original/2X/2/276875ed8a2b62eb85de895e0118aee8193d7c89.png>)

Categories simplify the voting process, allowing participants to focus on areas they're passionate about. This not only makes voting simpler but also ensures that each vote is backed by knowledge and genuine interest.

If badgeholders were required to review categories of ~15-30 similar projects, as opposed to being able to review any and every project that applied they would naturally discover less popular projects. They would also be directed to go deep into a few categories as opposed to feeling obligated to assign a number to every project they know. If we tracked how many times each category had been reviewed, badgeholders could effortlessly coordinate on dividing the work up amongst each other.

To make GREAT categories we need to ask specific questions about what category projects fall into during the application process, these questions should be publicly accessible, but do not need to clutter the proposal's application page. The answers to these questions would be used before the voting starts to form categories and validate the placement of projects into categories.

## Who will Categorize?

Having a team dedicated to creating categories and subcategories is crucial for making RetroPGF scalable. It's the only way to effectively scale up. We simply cannot expect a badgeholder to review 600+ projects, we have to break it up into digestible chunks.

[RetroPGF Round 3 Feedback Thread](#)

My biggest takeaway as a badgeholder over two rounds:

There is simply no way to truly assess 643 projects. I want to dive deep on projects I have some sort of expertise on

(governance, DAOs, regen, conferences) and trust others to do the same on stuff I don't (infra, gaming, media, programming, etc)

LauNaMu did a great job making categories during the voting round, but for RetroPGF4 we would love to collaborate to create the best categories possible before the voting begins and include anyone knowledgeable to help. Categories should be validated with the badgeholders, RetroPGF core team and the community before being finalized.

## Enabling Expertise to Shine

With categorization, badgeholders will rank projects in the domains where their expertise and interests lie, ensuring that the ranking of projects is informed by knowledgeable interested people. This approach allows us to maximize the value derived from badgeholder's time and contributions, and for a more effective ranking of projects, as each badgeholder brings their specialized knowledge to the table. If someone wants to support a specific project, they will need to rank all the projects in that category, and in doing so, learn about all the other awesome projects doing similar things.

## Promoting Project Discovery

Keeping the focus on categories will naturally lead to badgeholders discovering the countless public goods projects that are flying under the radar. This is not only good for the small projects that deserve to be rewarded, it is also good for the badgeholders that also happen to be builders, as they can discover public goods that they can build with! We expect the second order effects of ensuring project discovery is prioritized in the design of RetroPGF 4 will be huge for the ecosystem.

## Simplifying Assessment

Another important change we propose is to not allocate a number to quantify a project's impact. Technically Pairwise can very easily determine a number for each project and let the badgeholder adjust it, but we feel this step would unnecessarily complicate the vote. We saw a lot of feedback on average vs median and the pro's and con's to each, but we think the best solution is to sidestep that problem completely and instead have badgeholders rank projects without quantifying exactly how much they relatively like the project. This simplification will also greatly reduce the friction and cognitive overhead of the voting experience.

## Ranking the Categories

When badgeholders finish ranking the projects within a category they will then decide how that entire category's impact ranks against the other categories they have ranked. Even though most voters might only rank 3-5 out of 50 categories, the rankings of many badgeholders can be pieced together to determine a collective ranking of all the categories. As long as each category was ranked by several voters, we can aggregate the votes to get an accurate final rank of the categories.

Having this two-layer approach, where you rank every project in a category and then rank the category against other categories, will turn every ballot into multiple pieces of the RetroPGF puzzle. Each category will tallied independently, so that there is a clear collective ranking for each category as determined the badgeholders that reviewed that category, and the categories themselves will also have a ranking. How this turns into an OP amount is described in the next section.

## Project Funding Calculation

Badgeholders will rank the categories they completed against each other. So if a badgeholder completes 3 categories, their category ranking will have only those 3 categories ranked relative to each other. Pairwise will aggregate all the rankings to derive the category. Every category needs to be ranked at least by a certain number of badgeholders in order to have accurate numbers. Voting should be extended if we don't reach that quorum. The exact number will depend on how many categories there are.

We will split the total funds up into the categories such that the average project in the top category receives ten times more than the average project in the bottom ranked category. We use the average so that it doesn't matter if you are in a large category or a small category, and if the competition in that category is strong the projects will all likely receive more funding than in weaker categories.

50ish% of projects in each category will qualify for rewards. With the first place getting 20x the last place and a linear distribution between.

### For Example:

If there are 21 projects and the category was allocated 115,500 OP by badgeholders. The top 11 projects will be rewarded and the reward distribution would be like this:

1: 20,000 OP

2: 18,100 OP

3: 16,200 OP

4: 14,300 OP

5: 12,400 OP

6: 10,500 OP

7: 8600 OP

8: 6700 OP

9: 4800 OP

10: 2900 OP

11: 1000 OP

12: 0

1. 0

...:

21: 0

## Only 50% of Projects will get Rewarded?!?

In our opinion, RetroPGF 4 should be more competitive regarding which projects receive funding. In RetroPGF 2, everyone that was nominated got rewarded, this likely contributed to why there were ~1500 projects that applied to RetroPGF and only ~650 of them qualified. We believe if we say upfront that only ~50% projects will get rewarded, there will be a much stronger self-selection, greatly reducing spam for the reviewers.

But more importantly, it will better reward the projects that are truly the top performers in the collective, while allowing the projects that are less promising to fail fast. We see RetroPGF as a public goods free market. Everyone is welcome to participate and provide value but if your project is not making enough impact, it should fail, just as a startup that can't get investment does. Rewarding a lot of mediocre projects a little bit is sending mixed signals to the market, and crowds the field. This sounds cold, but the goals of the system will be better achieved by making RetroPGF more competitive.

It is important to note that this competitiveness lives within each category, so 50% for the Etheruem Core Infra will get rewarded, and 50% of the Art NFTs will get rewarded. This creates competition within each category to provide more value to the collective and point to more impact than the other projects in their category.

## Reward participation

### [RetroPGF Round 3 Feedback Thread](#)

1. First, financial compensation for badgeholders is essential. Allocating a part of the RetroPGF budget for this acknowledges their extensive efforts.

In RetroPGF 3 some badgeholders spent a lot of time reviewing projects and others did not, and it is likely that all badgeholders will be rewarded the same amount (if they get rewarded at all). We suggest changing this for RetroPGF4 and making it explicit that badgeholders will be valued and the rewards will scale based on the number of categories that are ranked, but max out once they have ranked 33% of the categories. Of course badgeholders can rank every category if they so choose. We suggest that the rewards max out at 33% because we want to encourage them to focus just on the categories that interest them and that they have expertise in.

Rewarding people based on the number of categories they review will be highly impactful, regardless of how we divide it. We can draft a reward system for participants integrated with Pairwise. This approach aims to reward both the influence and the engagement in the voting process.

Beyond that, the exact rewards distribution for badgeholders will depend on how many categories there are and will be determined before the voting starts.

## Anti-Gaming Measures and Transparency

- The Pairwise voting process is not like snapshot, it will be very easy to track how users use the application using traditional analytics tooling and separate authentic users and people who are just clicking through to promote their own projects.

- If the ballot submission looks like they are only farming the reward or just voting for their own projects without putting in an authentic effort to judge other projects, we will not give the reward to that badgeholder, but we will still count the ballot.
- We will maintain complete transparency in the reward calculation and distribution process. Everything will be public.

## **Simplifying Voting via Stronger Submission Guidelines**

Having stronger submission guidelines for RetroPGF is essential to improving the program. As a team we would be able to provide a better experience to the community having an open process to create the submission guidelines, but our initial thoughts on requirements to be able to participate are:

- Individuals can not apply, unless they frame themselves as projects.
- Impact stated in an application can only be cited once. If projects collaborate on impact, they have to decide who gets the credit for the impact stat.
- If it is a development project, it must be Open Source
- Cannot have a token with a circulating supply market cap of over 10 million
- Projects had impact they can point to that happened 1 month before the application period starts
- Projects provide a free good or service to the Optimism community (Freemium models are ok)
- Projects cannot have received more than \$2M in VC investment
- Projects cannot have received more than \$1M in Revenue and or Grants in the past 2 years (excluding OP grants)

We are also confident that we should make badgeholders' experience as simple as possible by removing how much funding and revenue they received from their consideration. Instead we could subtract amounts from the amount they receive in RetroPGF 4, e.g. if a project received a grant from Optimism, or were part of the last RetroPGF round we can simply deduct the amount they were rewarded from their RetroPGF 4 rewards, and these subtracted funds can go back to the grant council and/or the next RetroPGF round. This approach makes it much easier for badgeholders to focus on impact and forget about extraneous details, like who got a grant, vs who didn't, which we believe was under appreciated by many badgeholders in RetroPGF 3.

## **Expanding the Badgeholders Group**

Let's open up the badgeholder group more, to the point where anyone who is a delegate could easily be a badgeholder. One of the cool things about public goods is that they're accessible to everyone. The more people we have judging, the better as more people can discover these great projects. Having just 150 people to distribute this amount of money isn't sufficient.

One of the best parts of participating as a project in a QF round on Gitcoin or Giveth is the marketing exposure. Donors in QF often discover so many awesome projects that they can integrate with, and we believe this would be true for RetroPGF too.

Having a small badgeholder set limits the second order network effects that can be gained by expanding the group. We believe that we can use Pairwise to prevent gaming that could occur if we reward people for doing the work behind RetroPGF. If we both expanded the badgeholder group dramatically and rewarded the badgeholders for voting, a lot of people can discover these projects and integrate them into the systems they are building. Public goods are there to be useful, let's maximize their reach!

## **Thank you for reading this!**

Thank you for reading our rant/retrospective/suggested spec for next round. We appreciate many of these suggestions are strong deviations from the current direction of the RetroPGF program, and we don't expect all of the suggestions to be widely accepted and applied in the next round. We would love the chance to be a larger part of the voting experience in the next round in anyway possible, whether our suggestions are adopted or not. We believed illustrating them in the form of a spec would be easier to understand than just saying our feedback about how to improve the round.

This post was co-authored by Griff and I