Hey everyone this is my first post on OP! I discussed an interesting idea with@Jonas after his EthCC talk, which he suggested I write up on this forum.

## The Problem

An issue with RPGF is badgeholders getting overwhelmed with the large number of projects they need to review. One of many adverse effects is that projects with a known presence or 'brand' end up getting more votes - precisely what RetroPGF is supposed to guard against (reducing the role of marketing and letting impact speak for itself). These issues are probably going to get worse with time as more projects apply for funding through this mechanism.

## A Solution

The larger conceptual goal should involve moving the RPGF impact evaluation system from where we currently are on the right (a peer review system with experts' qualitative assessment) to the left (computational protocol with humans in the loop reviewing quantitative output).

[

Screenshot 2023-07-30 at 5.33.16 PM

1084×650 63.5 KB

The idea is disbursing money based on composite scores from different LLMs reviewing & scoring project applications (a project with a score of 8/10 gets twice the money as one with a 4/10).

## Pilot

An easy way to get started would be through simulation. Feed project applications from RPGF round 2 to an LLM and ask it to give scores for each of them. Map out how fund distribution via LLM scoring differs from the actual distribution by badgeholders

Over time, the role of badgeholders could evolve into evaluating the scores & justification provided by LLMs ('Yes' if the scoring is on point or 'No' if there is a flaw in reasoning behind a score). If a score has a lot of 'No' votes from badgeholders, the flaws in reasoning are pointed out to the LLM so that it learns from the feedback and re-tabulates all scores.

Would be keen to hear from others in the community and whether this could be a worthwhile working group at the OP Collective!