Decentralized Contributions System

Jonathan Kung   Craig Calcaterra   Leonhard Horstmeyer

DFG ♦   DFG ♦, MSU♣   Lakat♠, CSH⋆

♠ Lakat - Permissionless ScientificPublishing, ♦ DAO Governance Framework, ⋆ Complexity Science Hub
♣ Metro State University

Abstract

We are exploring how to improve power allocation and resource distribution for a healthy knowledge ecosystem assuming AI will be ubiquitous using a combination of web2.0 and web3.0 tools. We are layering a weighted-directed acyclic graph (WDAG) onto existing wiki data as a meta-layer citation graph and its parameters will be managed by wiki contributors using a novel voting mechanism. We will simulate the system in cadCad and we will conduct a comparative study between Mediwiki's current contribution metrics with our own.

Introduction

The current internet is also an attention-based economy. Though useful for profit and spreading misinformation and disinformation, DGF + Lakat aim to create a knowledge/truth economy which we are currently calling "intellectual mycelial networks". We will start with doing research to update the "wiki" with an improved tech stack that also incorporates the lessons that WikiMedia has learned over the last 2+ decades. The Wikimedia ecosystem has been able to democratize knowledge, but the process that governs the creation of knowledge, i.e. scientific publishing, is still in a state of disarray and would profit immensely from the Wikimedia principles of an open, pluralistic and community driven approach.

We propose to investigate to what extent a decentralized autonomous organization (DAO) can be developed for wikis, leveraging the existing infrastructure of the meta-wiki and the mediawiki. The outcome would be to link notions of contributorship with public goods funding in an autonomous way without centralized control. We will be using web 3.0 tools to manage a governance layer to both moderating conversations, provide novel voting infrastructure, and fairly attribute community contributions. Succinctly, we want to use web 3.0 tools and wikimedia data to determine novel power allocation and resource distribution techniques that better addresses Arrow's Impossibility Theorem and Goodhart's Law even if AI is ubiquitous.

Date: We propose to start on June 1, 2024 and conclude by February 27, 2025.

Related work

Calcaterra et. al. have developed several dynamically readjusted metrics for measuring reputation on weighted DAGs, accounting for the importance of foundational contributions [1]. These DAGs could be the citations, the hyperlinks or the edits. These metrics have been implemented in a typescript library [2].

Methods

Using several metrics inspired by the Page Rank, we associate scores to contributions, activity, engagement of users and also to

articles. The guiding idea is that the impact of contributions percolates through the tree of references. A fair metric ought to emphasize the value of foundational contributions, more than a simple counting of references would. So we recalculate the score of any new contribution using a metric developed in [1][3] that is akin to a Page Rank for DAGs. Contributions that turn out to have flaws or are revoked, can have a negative impact on the score and in this sense the score is dynamically adjusted. We will develop a REST-API that calculates the scores. To obtain some empirical data we envisage to use contribution data from a less represented and smaller wiki. We will download either the wiki or use one of the public api endpoints. We will then simulate the results in cadCad.

Expected output

We expect three key outputs:
- Insights into how the combination of web 2.0 and web3.0 tools can help manage a healthy decentralized knowledge forum.
- Simulations results of system performance using cadCad.
- Quantitative analysis of how new contributions score compares to new contributions score.

The primary audience for the analysis is the wiki community itself and researchers who are interested in quantitative studies of composability and reusability of the largest open knowledge database in human history.

Risks

As with any kind of quantitative research there is a possibility that the data is not as conclusive as was anticipated. We are willing to take this into account. We do not see any infringement of people's rights.

Community impact plan

This project would greatly profit from contributors with insights into data science and an understanding of the wikimedia data model. We would like to reach out to the Wikimedia volunteer developers. Some of us are members in the Wikimedia discord server and the Wikimedia Hackathon Telegram group. We would like to present our project there as well and furthermore would like to involve the community in both the process of research as well as the dissemination of findings.

Evaluation

Our research can be evaluated on two grounds. First on a quantitative basis, regarding the statistically significant results, the publication thereof and the soundness of the methods deployed therein. Secondly the project may be evaluated by the engagement of the Wikimedia community and the potential ramifications for projects emerging out of the research, such as a hackathon, a discussion about testing implementations of fine-grained version control systems for wikis.

Budget

We envisage that the salary is split between compensation for long hours of coding, data analysis and data cleaning, i.e. salary, and possibly open access publishing costs. We are a team of 3 working for 9 months, so we suggest $15000 for salary and $2000 for open access publishing fees. We would prefer however to publish without predatory fees.

Prior contributions

We are all quite passionate about data, protocols, quantitative and network analysis. Horstmeyer has conducted research in the field of network systems, analyzing metrics for the knowledge graphs of arxiv.org [5] and network

dynamics [6] [7] and higher-order connectivities [8]. Calcaterra has been investigating Organizational and Societal Structures [9] and Network-based metrics for reputation systems [10][11]. Kung and Calcaterra are both working on protocol research and DAOs, investigating metrics for reputation systems on DAGs such as git or blockchains [12]. That research is coordinated through a mediawiki [13]. Horstmeyer has also worked on a decentralized version-control protocol[14]. Bose is a researcher in cosmology and Kuzam a php and mediawiki-extension developer.

References

1. Calcaterra, Craig. "On-chain governance of decentralized autonomous organizations: Blockchain organization using Semada." *Available at SSRN 3188374* (2018).

3. gitlab.com/dao-governance-framework/

3. Zhu, Zhibo, Qinke Peng, Zhi Li, Xinyu Guan, and Owais Muhammad. "Fast PageRank computation based on network decomposition and DAG structure." IEEE Access 6 (2018): 41760-41770.

4. Calcaterra, Craig, Wulf A. Kaal, and Vlad Andrei. "Blockchain infrastructure for measuring domain specific reputation in autonomous decentralized and anonymous systems." *U of St. Thomas (Minnesota) Legal Studies Research Paper* 18-11 (2018).

5. Calcaterra, Craig, and Wulf Kaal. *Decentralization: Technology's impact on organizational and societal structure*. Walter de Gruyter GmbH & Co KG, 2021.

6. Horstmeyer, Leonhard. "Lakat: An open and permissionless architecture for continuous integration academic publishing." *arXiv preprint arXiv:2306.09298* (2023).

7. https://daogovernanceframework.com

8. https://github.com/Lakat-OS/lakat-py