

Zapateado Appraisal Protocol: Digital Relevance Realization via Provable Bitcoin Deposits

Siegfried Ludwig

November 2023

Abstract

The introduction of Bitcoin as a value transfer protocol native to the internet opens up entirely new approaches for mapping content relevance to address digital spam and the fraying of the social fabric. Building on the tipping of *zaps* on the decentralized *nostr* social network, the proposed protocol uses provable bitcoin deposits to form trustless maps of relevance for digital content, independent of centralized platforms and privileged information curators. By verifiably binding up liquidity behind digital content, users can indicate both the positive and negative value they derived from their interaction with it. Such user *zaps* added to the positive deposit directly contribute to content discoverability, and can be withdrawn by the content author to claim their monetary value. Attempts to artificially extend the reach of digital content are limited by the opportunity cost incurred when locking up liquidity in the deposit as a form of advertising. The proposed protocol reduces the need for centralized institutions, such as search engines and social media companies, to create maps of relevance for the content on the internet.

1 Introduction

An individual engaging with the indefinite amount of content on the internet faces the continual challenge of separating relevant from irrelevant information. This process, an instance of *relevance realization*, presents a rationally intractable problem [1]. In the context of social media, a first attempt at a solution is to follow the content of other individuals with which one shares some form of trusted relationship. This approach, however, remains limited to the number of relationships one can personally establish and maintain. A very similar situation is encountered in the trade economy of primitive markets, which are equally limited to trusted relationships. The solution in the latter case lies in the emergence of a shared representation of value, concretized as money, which provides a social fabric that allows trade to scale beyond the confines of local human organization [2]. Equally, in the absence of shared representations of relevance for social media content, such as *likes*, follower counts, or identity verification checks, no scaling of the social fabric could take place. All content outside of trusted relationships would be of unknown relevance, with the presence of large numbers of social media bots strongly discouraging interaction beyond the local tribe.

The current solution to the problem of establishing shared representations of relevance for digital content lies in the formation of centralized platforms managed by trusted institutions, keeping track of markers such as *likes*. The analogy in the context of economic trade would be a centrally managed ledger used to keep track of trade balances within a bounded geographical area. This solution suffers from the pitfalls of centralized control, such as susceptibility to undue influence, censorship and the erosion of privacy. Further, the use of centralized platforms leads to walled garden characteristics, where the constructed maps of relevance are not unified between different social media platforms and across other types of content.

The aggregation of *likes* on social media platforms as markers of relevance is based on the assumption that each account is representative of a unique human actor, and

to the degree that this assumption is violated, the social media system of relevance realization is breaking down. In a world where artificial intelligence creates perfect user accounts with curated histories of interactions, posts and thoughts, without limitation, social media companies are forced to react with increasing requirements for identity verification. To escape this trajectory, the exclusive reliance on centralized platforms needs to be replaced with the establishment of native protocols.

The beginnings of a solution native to Bitcoin, the trustless value transfer protocol of the internet [3], can be found in micropayment tips, termed *zaps*, used in the *nostr* decentralized social media protocol¹. This approach draws on the insight that while any number of social media *likes* can be generated via bot farms, this becomes cost prohibitive if each *like* has a monetary value attached.

Nostr zaps can be cheated however, by users tipping their own posts with a second wallet under their control, possibly looping their own funds to indefinitely increase *zap* scores. This self-tipping problem means that the accumulation of such payments is not a trustless sign of value. The more fundamental reason for this can be found in the fact that transactions as such, in this case facilitated by the lightning network [4], have no real constraints, and therefore impose no opportunity cost. Any action without an opportunity cost is ultimately unsuitable as a sign of value. In the futile attempt to limit content authors from tipping their own posts, developers would be forced to analyze tipping traffic to identify malicious actors, which would result in a situation very close to the current legacy systems. Rather than trying to put logical constraints on actors attempting to exploit the relevance realization mechanism through logical vulnerabilities, systemic security is better achieved by imposing a physical cost

¹<https://github.com/nostr-protocol/nostr>

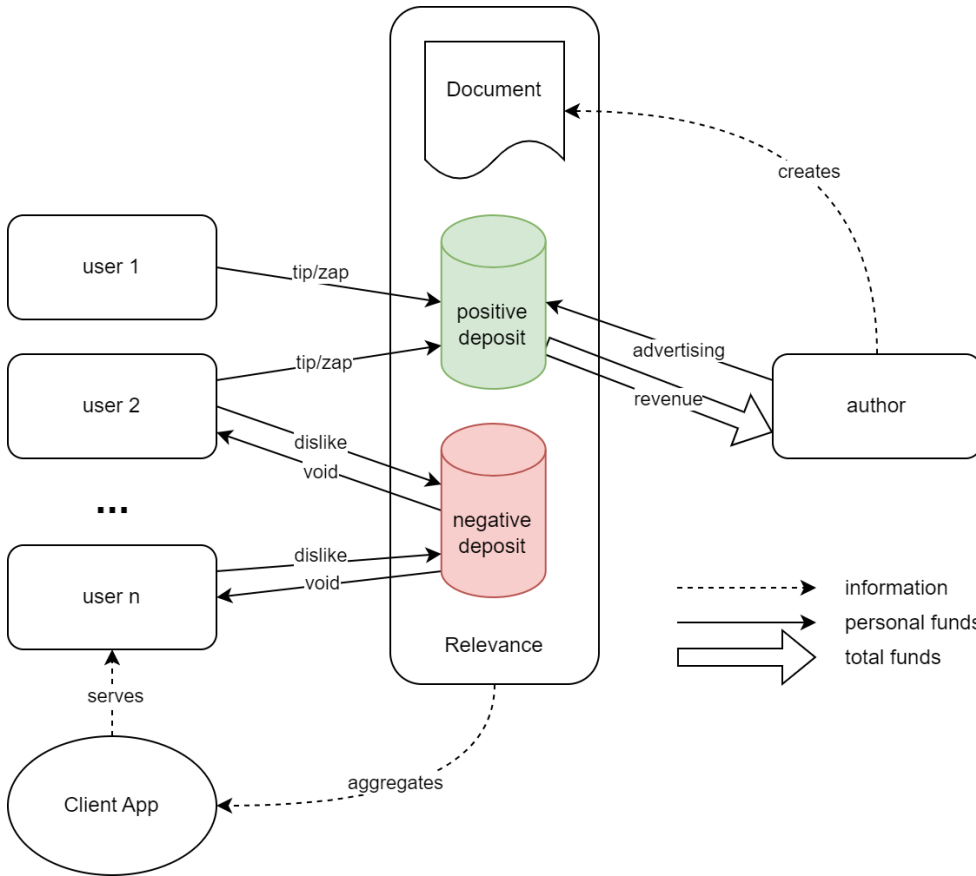


Figure 1: The Zapateado Appraisal Protocol uses provable Bitcoin deposits to attach value to any digital content and thereby quantify its relevance.

on attackers [5].

The required trustless and verifiable constraint physically enforced and native to cyberspace can be found in the bitcoin supply cap of 21 Million coins, further divisible into *satoshis*. The existence of this supply limit automatically makes bitcoin liquidity a limited quantity with attached opportunity cost. Therefore, locking up bitcoin in a verifiable deposit, rather than simply making a transaction, is a direct and trustless signal of value. Such bitcoin deposits are at the core of the Zapateado Appraisal Protocol (ZAP).

Interestingly, the map of relevance constructed with the proposed protocol might serve in the training of artificial intelligence models aligned to a decentralized representation of human values. Given the unavoidable and growing need for separating relevant from irrelevant data, current large language models are trained on datasets imperfectly curated by the respective development teams. This inevitably leads to the alignment of the resulting models with the specific values held by those individuals, rather than alignment with potential maps of relevance generated by all of humanity.

2 ZAP Details

The proposed protocol attaches two verifiable deposits to any piece of digital content, to serve both as positive and negative signs of value. Any user-provided tips, or *zaps*, are directly locked in the positive deposit, and can only be withdrawn by the content author. Conversely, if the user had a bad experience with the given content, they can lock value in the negative deposit, which can only be withdrawn by the same user (see Fig. 1). This setup provides a monetary reward for the content author, while avoiding a situation where the author receives benefits from negative deposits.

Once a deposit is withdrawn, the opportunity cost associated with locking up bitcoin liquidity is lost, and therefore, in the absence of further consideration, the signal of value would be voided. To address this, assuming a trustless timekeeper can be found, such as the bitcoin block height, the integral of deposits over time can be used as a permanent value signal that is not lost when deposits are withdrawn (see Fig. 2). This means that the longer a given amount of liquidity is locked in the deposit, the more valuable the signal becomes, which is directly commensurate with the opportunity cost incurred. The protocol does not necessarily require this to function, but without a trustless implementation of the time integral, any attempt at making value signals permanent after the withdrawal of deposits appears susceptible to corruption.

The use of separate positive and negative deposits allows users to choose varying ratios of each, indicating the specific score they would like to give to a piece of content. While the protocol-level representation of user ratings is simply the size of positive and negative deposits, client applications can display these as arbitrary scales, whichever is preferred by their users. A simple *like* would correspond to a 100:0 positive to negative deposit ratio, a dislike 0:100, 2 of 5 stars would be 40:60 and a score of 79 out of 100 would be 79:21.

The negative deposit can later be withdrawn by the same user, freeing up liquidity

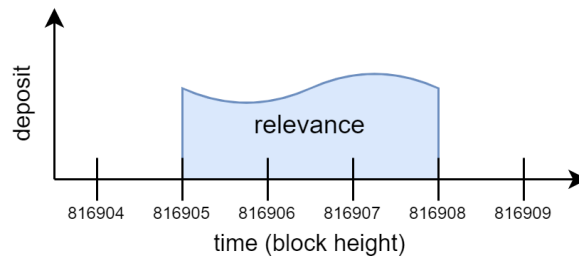


Figure 2: The integral of deposits over time, measured by a trustless timekeeper such as the Bitcoin block height, serves as a permanent marker of relevance.

for other purposes, such as leaving ratings on new content. Client applications could provide frictionless mechanisms for this if the user sets a total liquidity budget for negative deposits, such that older negative deposits automatically get reallocated when funds for new negative deposits are requested (via a simple downvote).

By indicating the relevance of a piece of content, the positive deposits serve as a discoverability mechanism, directly incentivizing the author to keep the coins locked in the deposit. This incentive holds up to the degree that the benefit of the deposit towards discoverability exceeds the opportunity costs the author incurs by keeping the bitcoin locked up. To the degree that a piece of content has run its course over time, the author can draw from the deposit to balance opportunity costs with the value of attracting continued residual interest.

There appears to be no trustless and privacy-preserving way to prevent an author from depositing their own bitcoin, by tipping themselves, to boost the discoverability of their content. While this could be interpreted as resulting in a skewed value signal, it should nevertheless be viewed as an improvement over current advertising mechanisms. The author will be limited in this action by their own available supply and the associated opportunity costs. Boosting the discoverability of their own content is only profitable to the degree that the increased attraction of users results in increased positive tips, ensuring that advertising is only beneficial if the content is valued by users. The profit-seeking actions of the author are then in principle directly aligned with the optimization of the relevance realization process. Instead of spending an advertising budget with uncertain outcome and requiring private user information for optimal targeting, authors are able to reuse the same advertising budget for new content later on, making it superior to conventional advertising.

The proposed protocol arguably allows users with larger pools of capital to exert greater influence on the resulting relevance realization process. It appears however that this is directly analogous to such influence in a classical free market, which we generally judge as the most natural and effective system for discovering a map of value for different goods and services, expressed in prices.

The opportunity ZAP provides does not only apply to original content authors, but also to any comments made by users in response to that content. Just like the original post, any comment can receive positive and negative deposits to facilitate comment discoverability. Similarly, the proposed relevance realization process can in principle be scaled to any digital content beyond social media, thereby constructing an explicit map of relevance with a shared language for the entire internet.

References

- [1] John Vervaeke, Timothy P Lillicrap, and Blake A Richards. Relevance realization and the emerging framework in cognitive science. *Journal of Logic and Computation*, 22(1):79–99, 2012.
- [2] Lyn Alden. *Broken Money: Why Our Financial System is Failing Us and How We Can Make It Better*, chapter 1. Timestamp Press, 2023.
- [3] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.
- [4] Joseph Poon and Thaddeus Dryja. The bitcoin lightning network: Scalable off-chain instant payments. 2016.
- [5] Jason P Lowery. *Softwar: A Novel Theory on Power Projection and the National Strategic Significance of Bitcoin*. self-published, 2023.