Technology Feasibility Research

Proof of Humanity (PoH) has been an ongoing field of research for over 20 years. This concept began to materialize with the introduction of one of its earliest forms: the CAPTCHA.

A CAPTCHA is a program that generates and grades tests that most humans can pass, but current computer programs cannot (von Ahn et al., 2003). It leverages the power of computational complexity classes to ensure that a sufficiently advanced AI cannot thwart the test. These types of proof-of-humanity systems have evolved in various forms but still rely on some form of a test that most humans can pass but is very challenging for a computer program to do so.

There are various forms of CAPTCHAs: text-based, image-based, audio-based, video-based, puzzle-based, and motion-based (Hasan, W.K., 2016). Each comes with its own advantages and disadvantages. Some of the most commonly found on the web in the past decade are the text and image versions, although newer forms of puzzle-based versions are arriving. These methods do work in practice, but they are passable with sufficiently advanced AI.

Some key takeaways from CAPTCHAs include the following: humans prefer to spend about 3 to 40 seconds solving a CAPTCHA; younger individuals typically solve CAPTCHAs faster than older individuals; users often favor click-to-pass, slide-to-pass, or rotate-to-pass CAPTCHAs; and CAPTCHAs are only about 80% accurate (Searles et al., 2023).

The next tier of Proof of Humanity (PoH) moves beyond automation via CAPTCHA to include human interaction, regulatory compliance, and on-chain solutions. Companies like Fractal ID and Proof of Humanity now offer advanced methods for PoH, which range from human certification and uniqueness to KYC (Know Your Customer)/AML (Anti-Money Laundering) procedures, webs of trust, biometric verification, behavioral analysis, and verifiable digital credentials. These newer PoH systems typically rely on a combination of these methods as any sufficient complex combination ought to be very difficult to falsify and bypass, but it does require some form of centralization and could potentially compromise user privacy.

In exploring advanced PoH systems, it's essential to understand the diverse methods employed. Human certification and uniqueness involve verifying an individual's identity and ensuring they are distinct within a system, often through manual or semi-automated processes. KYC/AML procedures are regulatory frameworks that identify and verify customers to prevent financial crimes; these are common in financial and legal sectors. Webs of trust rely on networks of trust relationships, where users vouch for each other's authenticity. Biometric verification uses unique physical characteristics like fingerprints or facial recognition for identity verification. Behavioral analysis focuses on patterns in user behavior, such as typing rhythms or mouse movements, to distinguish humans from automated systems. Lastly, verifiable digital credentials involve using digital tokens or certificates that provide cryptographic proof of a user's identity or qualifications. While these methods offer robust ways to establish human presence, integrating them within decentralized platforms like Cardano may require innovative approaches to balance security, privacy, and decentralization.

Most of the newer systems are designed either to upgrade existing systems like OAuth or to extend validation on blockchains such as Ethereum. As of now, there is no Proof of Humanity (PoH) system specifically designed for the Cardano ecosystem. The solution we seek must integrate seamlessly within Cardano and utilize advanced forms of CAPTCHAs. These two criteria will ensure that the project's primary goal is upheld, most notably preserving user privacy. Forcing users to undergo KYC/AML compromises privacy and introduces an inherent bias regarding who can obtain PoH proofs, as not everyone is able to pass KYC/AML procedures. Since Cardano is designed to be decentralized, inclusive, and open, our PoH system should reflect those principles.

Cardano Integration

A Cardano integration can mean many things, depending on the use case. When building on Cardano, the question isn't typically whether the problem can be solved; rather, it becomes a matter of how decentralized—and thus trustless—the desired solution should be.

The position on the spectrum of decentralization will determine the nature of the integration required for a proof of humanity to exist on Cardano. These integrations will range from controlled structures to pure hyperstructures. Structure in this context implies how much has to exist in the off-chain world for the application to exist on Cardano. A pure hyperstructure only requires Cardano to exist, while a controlled structure requires centralized infrastructure outside the blockchain to exist for the application to exist. Consequently, this could lead to the coexistence of various PoH systems, each with its own level of decentralization and mechanisms for proving humanity.

The proof-of-humanity system could operate entirely off-chain, utilizing centralized services to issue certificates of humanity. In this model, the only on-chain element would be the minting, ownership, or reference of the NFT on the Cardano blockchain. This solution can serve as a proof of humanity but requires centralization for the NFT minting process, as this would be the on-chain representation of the certificate of proof of humanity. Typically, controlled structure solutions are easier to implement and work well with already existing third-party PoH applications and protocols.

In the off-chain, centralized model of a proof-of-humanity system, the user experience is significantly influenced by the involvement of centralized services. This approach typically offers a more streamlined and user-friendly process, as centralized systems can provide more intuitive interfaces and quicker verification processes. Users might find it easier to navigate and complete the necessary steps for obtaining their proof of humanity, as these systems often have well-established procedures and customer support.

However, this ease of use comes with certain trade-offs. Users must place a degree of trust in the centralized entity managing the system, which may raise concerns about data privacy and security. The reliance on external services also means that the user experience is subject to the efficiency and reliability of these third-party providers. Additionally, in scenarios where internet connectivity is inconsistent, a users' ability to access these centralized services could be compromised, potentially limiting the inclusivity and accessibility of the system.

Alternatively, the solution could be more integrated with the blockchain, requiring smart contract verification through cryptographic proof generation and validation methods. This Plutus-centric approach may result in higher user costs and longer development time due to the need for direct blockchain interaction, but it could offer enhanced security and unparalleled customization. The closer the solution is to a purely decentralized proof of humanity, the more it aligns with Cardano's ethos. Though it may be difficult to create a hyperstructured PoH system as any PoH system needs to be updated to maintain the requirement that any sufficiently advanced AI cannot thwart the PoH test. In turn, the solution will require some mediation, either in the form of a DAO or a federated party, to upgrade the tests used in the smart contract for the next generation of AI.

In an on-chain, blockchain-integrated PoH solution, the user experience is deeply intertwined with the blockchain. Users who are familiar with blockchain operations may appreciate the transparency, security, and control this system provides, as it allows them to directly verify and engage with the PoH process. However, this approach can be challenging for users who are new to blockchain technology. The necessity of understanding smart contracts, managing digital wallets, and interacting with the blockchain can present a significant learning curve. Additionally, the potential for higher transaction costs and the need for digital currency (to pay for transaction fees on the blockchain) could be barriers to entry for some users. This aspect needs to be carefully balanced to ensure that the PoH system remains accessible and cost-effective for all potential users.

In practice, most solutions end up being some combination of off-chain and on-chain that meets the use case, resulting in a quasi-hyperstructure solution. The application could exist on its own as a program on the blockchain that anyone can use, but there does exist some group that handles any required maintenance to maintain the PoH requirements.

Smart Contract Limitations

Space and time constraints limit Cardano's on-chain validation. The existing protocol parameters, coupled with the absence of bitwise primitives, render certain proof validation methods prohibitively costly. This restricts the range of usable cryptographic proofs on Cardano, as any security parameters of real-world primes would immediately exhaust the capacity of any transaction using the currently available methods in Plutus. There are potentially more cryptographic primitives coming to Cardano in the future with Plutus V3, but the release date is too far into the future to be useful in this project.

The current methods available in Plutus include `ECDSA-SECP256k1`, `Ed25519`, `SCHNORR-SECP256k1`, arithmetic circuits, and customized transaction logic. This limited set of methods makes proving the validity of a signature on data trivial but makes communicating and computing large integers for zero-knowledge proofs unfeasible. Novel primitives like 3-colorings and bulletproofs can be implemented at the cost of development time and the risk of incorrect implementation. To mitigate these implementation risks, rigorous testing, external audits, and leveraging expertise from the broader blockchain development community are essential. By adopting a meticulous and collaborative approach to development, the risks associated with implementing advanced cryptographic techniques can be significantly reduced. When considering just the default primitives available to Cardano, they may limit the development of a PoH system, but we do retain the ability to implement custom validation methods to meet the use case's demands.

Some significant constraints arise from on-chain randomness, on-chain data availability, and the tradability of native assets. A proof-of-humanity test must be generated from true randomness, forcing a trusted oracle to consistently update its random data. It must be non-tradable, necessitating its residence within a permanent lock smart contract. And it must be reference-able, requiring real-time on-chain data. While the perpetual locking of an asset is trivial from a development standpoint, it does demand that the permanent lock contract be a controlled or quasi-hyperstructure, as the ability to remove or void certificates of humanity may be required. Should this token exist within a contract, then any service requiring verification of said token would need access to the latest on-chain information. Although new technologies are continually emerging to optimize the management of this specific data, making the process increasingly feasible, it still represents considerable overhead and centralization for those wishing to utilize this system on Cardano.

In terms of user impact, Cardano's smart contract limitations could affect the efficiency and reliability of a PoH system from the user's perspective. For example, the absence of certain cryptographic primitives might lead to slower verification processes or higher costs for users. This could result in a less intuitive and more time-consuming user experience, particularly for users accustomed to faster, more streamlined processes on other platforms. Addressing these user experience concerns is crucial, as it directly affects user adoption and the practical viability of the PoH system on Cardano. To address these limitations, potential workarounds may include hybrid solutions that combine on-chain and off-chain computations. For instance, complex calculations or data-intensive processes could be handled off-chain, with only essential information and final proofs being recorded on the blockchain. This approach could alleviate the burden on the blockchain while still leveraging its security and immutability.

Implementing Proof Generation and Validation

Getting something to prove humanity on Cardano will take some novel solutions in its current state. Due to the nature of PoH systems, a Cardano PoH will need to be some form of an interacting solution. A potential human will need to prompt the PoH system to start the verification process; they will need to receive the test, solve the test, and then submit the test for a result. The outcome of the test is the validity of the submitted transaction. The validity of the transaction will be based on existing primitives accessible in Plutus and new primitives that are time-based, one-time use, and integrate directly with the test itself. This allows the PoH proof system to be comprehensive without sacrificing security. The general flow leads us towards an advanced CAPTCHA system that can be mapped directly into a cryptographic proof that can be verified on-chain.

For instance, a user seeking to acquire a 'token of humanity'—an NFT serving as a verified proof of humanity—would need to initiate a certification UTxO for a random CAPTCHA state. The user will solve the CAPTCHA, generate the proof, and submit the transaction for verification within a specified time frame. The time frame is long enough that valid transactions will land on the chain, but it is short enough to prevent brute-forcing the solution to the CAPTCHA. If the transaction contains a valid proof within the allotted time for the given CAPTCHA, then the user is acknowledged as human; otherwise, the funds in the certification UTxO are forfeited.

This method integrates various elements of other systems, including time-based, one-time-use CAPTCHAs, monetary stakes, and verification culminating in the minting of the token of humanity. Simultaneously, the system maintains privacy owing to the decentralized nature of the validation.

In designing a PoH system on Cardano, particular attention must be given to the user experience and accessibility. The interface for interacting with the CAPTCHA and submitting proofs must be intuitive and user-friendly, ensuring that individuals with varying degrees of technical expertise can participate without too much difficulty. To support a broad range of users, the system could include step-by-step guides or interactive tutorials. This approach helps in reducing barriers to entry, especially for those who are new to blockchain technology.

Moreover, the system's robustness is crucial, and it must include mechanisms for error handling and dispute resolution. In cases where users encounter issues or contest the invalidation of their transactions, a clear and transparent process should be in place for reviewing and resolving such disputes. This might involve manual review processes or the implementation of secondary verification methods to ensure fairness and maintain user trust. This aspect of the solution is entirely use case dependent.

Economic considerations are also vital, especially regarding the forfeiture of funds in the certification UTxO for failed attempts. This aspect of the system needs careful calibration to avoid unduly penalizing users, particularly those who may lack the financial resources to withstand multiple attempts. A possible solution could be the introduction of a tiered system, where initial verification attempts will require smaller financial stakes, or providing users with a certain number of ‘free’ attempts before any financial penalty is applied. Such measures would contribute to making the PoH system more inclusive and economical, encouraging wider participation without the fear of punitive economic repercussions.

Summary

The landscape of Proof of Humanity (PoH) has undergone significant evolution, transitioning from basic CAPTCHA systems to intricate solutions integrating human interaction, on-chain technology, and advanced cryptographic methods. In the context of Cardano, a blockchain renowned for its commitment to decentralization, the optimal PoH system encompasses more than just advanced CAPTCHAs; it demands a delicate balance of privacy, security, and inclusivity, without depending heavily on centralized processes like KYC/AML or biometrics.

Cardano’s unique architecture, while supportive of a variety of cryptographic methods for on-chain validation, presents specific challenges, particularly in terms of transaction size and computational resource constraints. Despite these challenges, our research indicates the feasibility of implementing a robust PoH system within the Cardano ecosystem. This system would utilize smart contracts to store non-tradable NFTs, symbolizing a user’s proof of humanity.

Envisioning an ideal PoH system for Cardano, we propose a Plutus-centric approach, blending both novel and existing cryptographic primitives. The core of the on-chain system involves user interaction with a smart contract to solve a CAPTCHA within a designated timeframe. Successful completion leads to the minting of a ‘token of humanity’ NFT, while failure results in the forfeiture of funds. This approach integrates time-based, one-time-use CAPTCHAs and monetary stakes, maintaining user privacy and ensuring a decentralized validation.

A crucial aspect of am on-chain PoH system is the user experience, accessibility, error handling, and economic inclusivity. Our design aims to make a PoH system intuitive and accessible to a diverse user base, with mechanisms in place for dispute resolution and error correction. Economic factors, such as the potential penalization of users through fund forfeiture in the certification process, are addressed through a tiered system and ‘free’ attempts, making the system more equitable and user-friendly.

As technology advances, especially within the Cardano ecosystem, the PoH system’s efficiency and security are expected to improve, with better data management and enhanced user interfaces. However, the challenge remains to develop a system that is secure, handles on-chain data effectively, and aligns with Cardano’s ethos of decentralization and user empowerment.

References

1. Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford (2003). *CAPTCHA: Using Hard AI Problems for Security.* E. Biham (Ed.): EUROCRYPT 2003, LNCS 2656, pp. 294–311, 2003

2. Hasan, W.K. (2016). *A Survey of Current Research on CAPTCHA*. International Journal of Computer Science & Engineering Survey, 7, 1-21.

3. Searles, Andrew & Nakatsuka, Yoshimichi & Ozturk, Ercan & Paverd, Andrew & Tsudik, Gene & Enkoji, Ai. (2023). *An Empirical Study & Evaluation of Modern CAPTCHAs.*