# Peer-to-Peer Consensus based on Content Authoring Reputation

Francisco Sant'Anna Department of Computer Science, Rio de Janeiro State University

**Abstract**—Content publishing in public Internet forums and social media suffers from excess and abuse, such as low quality posts and fake news. Centralized platforms employ filtering algorithms and anti-abuse policies, but impose full trust from users. We propose a publish-subscribe peer-to-peer protocol to model public content dissemination without centralized control. The central idea of the protocol is a reputation system that moderates content and, at the same time, supports network consensus. We trace a parallel with Bitcoin: consolidated posts create reputation (vs proof-of-work); likes and dislikes transfer reputation (vs transactions); and forks are ordered by reputation (vs longest chain). The reputation system depends solely on human work to create and rate content, preventing content excess and abuse and imposing consensus on a peer-to-peer setting.

| Index Terms—distributed consensus, | peer-to-peer, publish- | subscribe, reputation s | ystem |
|------------------------------------|------------------------|-------------------------|-------|
|                                    |                        | _ 🛦                     |       |

# 1 Introduction

CONTENT publishing in public Internet forums and social media platforms is increasingly more centralized in the hands of a few companies [1]. On the one hand, these companies offer free storage, friendly user interfaces, and robust access. On the other hand, they concentrate more power than required to operate, since they collect and control our data, "algorithmize" our consumption, and yet obstruct portability with proprietary standards. Peer-to-peer alternatives [2] eliminate intermediates and push to end users the responsibility to manage data and connectivity. However, due to decentralization of authority and network infrastructure, some new challenges arise to deal with malicious users and enforce overall state consistency.

In a scenario of interest, suppose users want to discuss a political event in the Internet. In an ideal system, all posts (i) would reach even temporarily disconnected users; (ii) would be delivered in a consistent order; (iii) would be respectful and on topic. In a centralized system, items (i) and (ii) are trivially achieved assuming availability and delivery order in the service, while for item (iii) users have to trust the service to moderate content (e.g., removing SPAM and fake news). In a decentralized setting, however, none of these demands are easily accomplished. A common approach in gossiping protocols is to replicate the whole conversation in all peers and disseminate it proactively until all users receive it [2]. However, this approach does not guarantee consensus since posts can be received in conflicting orders in different peers. As an example, antagonistic messages such as "X is final" vs "Y is final" might be sent concurrently, preventing the network to decide as a group between *X* and *Y*.

Bitcoin [3] proposes the first successful permissionless consensus protocol founded on scarce virtual assets, the *bitcoin tokens*. The only way to create new bitcoins is to work towards consensus in the network by proposing a total order among all transactions in the system. This way, Bitcoin prevents double spending [3], which is analogous to conflicting

posts in public discussions: deciding between X and Y as a group is the same as deciding between transferring to X and Y with insufficient funds for both. However, Bitcoin's basic operation is to blindly transfer bitcoins between users, with no subjective judgment that could affect transactions. In contrast, our challenge is on assessing human content based on social interactions between them.

In this work, we propose a consensus algorithm based on authoring reputation. Inspired by Bitcoin, content authors accumulate tokens named reps, which serve as currency to rate posts in the network. Users can rate posts with likes and dislikes, which transfer reps between them. In the algorithm, work is manifested as new posts, which rewards authors with reps, but which are still judged by other users. This way, like Bitcoin, token generation is expensive, while verification is cheap and made by multiple users. However, unlike Bitcoin, both creation and verification are subjective, based on human creativity and judgement, which match our target domain of content publishing. Posts and likes are linked as blocks in a Merkle DAG that persists the whole conversation and is disseminated in the network with gossiping. To reach consensus, branches with more reputed authors (i.e., with more work) take priority. Conflicting operations in concurrent branches, such as likes with insufficient reps, are rejected along with further posts. The proposed consensus algorithm is integrated in Freechains<sup>1</sup>, a peer-to-peer publish-subscribe content dissemination protocol [4].

In Section 2, we introduce the basic functionalities of Freechains to create, rate, and disseminate posts. In Section 3, we describe the consensus algorithm with its goals, incentives, and possible attacks and mitigations. In Section 4, xxx. In Section 5, yyy.

#### 2 FREECHAINS

Freechains is an unstructured peer-to-peer topic-based publish-subscribe system [4]. Each topic, or *chain*, is organized as a *Merkle DAG*, i.e., a directed acyclic graph immune

1. http://www.freechains.org

| Туре            | Prefix | Arrangement  | Behavior  | Examples  |
|-----------------|--------|--|---|---|
| Private Group   | \$     | Private $1 \leftrightarrow 1$ , $N \leftrightarrow N$ , $1 \hookleftarrow$ | Trusted groups (friends or relatives) exchange encrypted messages among them. The communication can be in pairs $(1 \leftrightarrow 1)$ , groups $(N \leftrightarrow N)$ or even individual $(1 \leftrightarrow)$ . | - E-mails<br>- WhatsApp groups<br>- Backup of documents.                  |
| Public Identity | @      | Public 1 → N, 1 ← N  | A public identity (person or organization) broadcasts authenticated content for a target audience $(1 \rightarrow N)$ with optional feedback $(1 \leftarrow N)$ .   | - News sites<br>- Streaming services<br>- Public profiles in social media |
| Public Forum    | #      | Public N ↔ N   | Participants with no mutual trust communicate publicly.   | - Q&A forums<br>- Chats<br>- Consumer-to-consumer sales                   |

Fig. 1. The three types of chains and arrangements in Freechains.

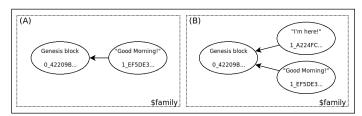


Fig. 2. Two DAG configurations. (A) has a single head pointing to the genesis block. (B) has a fork with two heads pointing to the genesis block.

to modifications. The chain DAG is disseminated peer by peer in the network with gossiping. This way, as an author posts to a chain, other users subscribed to the same chain eventually receive the message. Freechains supports multiple types of chains with different arrangements of public and private communication, which are detailed in Figure 1. In this section, we operate a private group to describe the basic behavior of chains. At the end of the section, we also exemplify a public identity chain. In Section 3, we detail the behavior of public forums, which involve untrusted communication between users and require the proposed reputation and consensus mechanisms.

All Freechains operations go through a *daemon* (analogous to Bitcoin full nodes) which validates posts, links them in the Merkle DAGs, persists the chains in the disk, and communicates with other peers in the network to disseminate the graphs. The command that follows starts a daemon to serve further operations:

```
> freechains-daemon start '/var/freechains/'
```

The actual chain operations use a separate client to communicate with the daemon. The next sequence of commands (i) creates a shared key, (ii) joins a private group chain (prefix \$), and (iii) posts a message into the chain:

```
> freechains crypto shared 'strong-password' # (i)
A6135D.. <- returned shared key
> freechains '$family' join 'A6135D..' # (ii)
42209B.. <- hash of chain
> freechains '$family' post 'Good morning!' # (iii)
1_EF5DE3.. <- hash of post</pre>
```

A private chain requires that all participants use the same shared key to join the group. A *join* only initializes the DAG locally in the file system, and a *post* also only modifies the local structure. No communication occurs at this point. Figure 2.A depicts the state of the chain after

the first post. The genesis block with height 0 and hash 42209B.. depends only on the arguments given to *join*. The next block with height 1 and hash EF5DE3.. contains the posted message. As expected from a Merkle DAG, the hash of a block depends on its payload and hash of previous block.

Freechains adheres to the *local-first* software principle [5], allowing networked applications to host their own data and work locally while offline. Except for synchronization, all other operations in the system only affect the local replica. In particular, joining a chain with the same arguments in another peer results in the same genesis state, even if the peers have never met before. Hence, before synchronizing, others peers have to initialize the example chain with the same steps:

```
> freechains-daemon start '/var/freechains/'
> freechains crypto shared 'strong-password'
A6135D..
> freechains '$family' join 'A6135D..'
42209B..
```

Synchronization is explicit, in pairs, and unidirectional. The command *recv* asks daemon in *localhost* to connect to daemon in *remote-ip* and receive all missing blocks from there:

```
> freechains '$family' recv '<remote-ip>'
1/1 <- one block received from <remote-ip>
```

The complementary command *send* synchronizes the DAGs in the other direction. Note that Freechains does not construct a network topology or synchronize peers automatically. There are no preconfigured peers, no *root servers*, no peer discovery. All connections happen through the *send* and *recv* commands which have to specify the peers explicitly. In this sense, the protocol only gives basic support for communication in pairs of peers and any further automation requires external tools.

The next sequence of commands checks the hash(es) of the block(s) at the head of the local DAG (the latest blocks), and then reads the payload of the single head found:

```
> freechains '$family' heads
1_EF5DE3..
> freechains '$family' payload '1_EF5DE3..'
Good morning!
```

Now, the new peer is in the same state as the original peer in Figure 2.A. However, since the network is inherently concurrent and users are encouraged to work locally, typical graphs are not lists, but DAGs with multiple heads. As an example, suppose the new peer posts a new message "I'm here!" before the *recv* above, when the local DAG is still in its genesis state. In this case, as illustrated in Figure 2.B, the resulting graph after the synchronization now contains two blocks with height 1. Note that forks create ambiguity in the order of messages, which is a fundamental obstacle to reach consensus. In private chains, which we use in this example, we can apply simple methods to reach consensus, such as relying on the timestamps of blocks. However, in public forums, malicious users could modify their local time to affect new block timestamps and manipulate the order of messages.

To conclude the basic chain operations, users can rate posts with *likes* and *dislikes*, which can be consulted later:

```
> freechains '$family' like '1_EF5DE3..'
2_BF3319..
> freechains '$family' reps '1_EF5DE3..'
1 # post received 1 like
```

In private groups, likes are unlimited and behave much like typical centralized systems. In public forums, however, likes are restricted, have to be signed by users, and are at the core of the consensus algorithm.

For the sake of completeness, Freechains also supports public identity chains (prefix @), which relies on public-key cryptography to attach an identity to a chain and to verify the authenticity of posts:

In the example, a public figure creates a pair of public/private keys and joins an identity chain attached to his public key. Every post in this chain needs to be signed with the associated private key to be accepted in the network.

# 3 THE CONSENSUS ALGORITHM

In the absence of moderation, peer-to-peer public forums are impractical. At the root of the problem lies Sybil attacks, which use large numbers of fake identities to abuse the system. For instance, it takes a few seconds to generate thousands of public/private key identities and SPAM million of messages into the system. For this reason, we propose a reputation system that works together with a consensus algorithm to mitigate Sybil attacks and make peer-to-peer public forums practical.

## 3.1 Overall Design

In the proposed reputation system, users can spend tokens named *reps* to post and rate content in the chains: a *post* operation initially penalizes authors until it consolidates and counts positively; a *like* operation is a positive feedback that helps subscribers distinguish good content amid excess; a *dislike* operation is a negative feedback that revokes content when crossing a threshold. However, without restrictions, post and like operations alone are not satisfactory, specially considering hypothetical Sybils abusing the system.

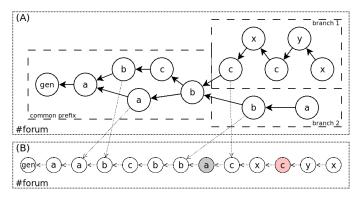


Fig. 3. (A) A public forum DAG with a common prefix and two branches. (B) Hypothetical corresponding total order between blocks.

Therefore, *reps* must be subject to some sort of scarcity that demands non-trivial work immune to automation.

Bitcoin employs CPU proof-of-work as a solution for Sybil attacks. But in the context of content publishing, we understand that authoring is already an intrinsic human work that we can take advantage. Creating new content is hard and takes time, but comparatively easy to verify and rate. In our system, in order to impose scarcity, content authoring is the only operation that can generate *reps*, while likes and dislikes only transfer *reps* between users. Still, scarce post and like operations are not enough as they call for consensus in the network: now we need to track the reputation of users to check if they are allowed to post new messages or spend likes. Since posts and like operations are concurrent in the network, we need to figure out how to order them consistently across all peers so that they converge to the same state.

Our solution is to preserve the order of messages that is seen by the users constituting the majority of reputation in the network. This way, in order to attack and manipulate the order of messages, malicious users first need to cooperate with good content to gain reputation, which is a non-trivial endeavor. Figure 3.A illustrates the reputation criterion. A public chain DAG has a common prefix with signed posts from users a, b, and c. Let's assume that within the prefix, users a and b have contributed with more content and have more reputation combined than c has alone. After the prefix, the chain forks in two branches: in branch 1, only user c remains active and we see that new users x and y, with no previous reputation, generate a lot of new content; in branch 2, only users a and b exist but with less activity. Nonetheless, branch 2 takes priority because, at the forking point, a and b have more reputation than c, x, and y combined. In this example, user c represents a malicious user trying to cultivate fake identities x and y in separate of the network to accumulate reps. However, the whole branch 1 becomes vulnerable when rejoining since branch 2 takes priority and can invalidate c completely as discussed further.

Figure 3.B indicates the consensus order between blocks in the chain. All operations in branch 2 are accounted before any operation in branch 1. Note that the common prefix also had a fork that favored user a. At any point in the consensus timeline, if an operation fails, all remaining blocks are removed from the chain DAG. As an example, suppose the last post by a (in gray) is a dislike to user c, which

| Type Rule number name |      | ıle   | Description  | Observations  |  |
|-----------------------|------|---|--|---|--|
|                       |      | name  | Description  |   |  |
| Emission              | 1.a  | pioneer   | Chain join counts +30 reps to pioneer.   |   |  |
|                       | 1.b  | old post  | Consolidated post counts +1 rep to author.   | A post takes 24 hours to consolidate. An author accounts at most 1 post per day.  |  |
| Penalty               | 2    | new post  | New post counts -1 rep to author.  | The discount period depends on the sum of authors' reps in subsequent posts and varies from 0 to 12 hours (50% to 0% reps). |  |
| Tranfer 3.a           | like | Like counts -1 rep to origin and +1 rep to targets. | Origin is the user signing the operation.<br>Targets are the referred post and its corresponding author. |   |  |
|                       | 3.b  | dislike   | Dislike counts -1 rep to origin and -1 rep to targets.   | If a post has at least 3 dislikes and more dislikes than likes, then its contents are hidden.                               |  |
| Constraints           | 4.a  | min   | Author requires at least +1 rep to post.   |   |  |
|                       | 4.b  | max   | Author is limited to at most +30 reps.   |   |  |
|                       | 4.c  | size  | Post size is limited to at most 128Kb.   |   |  |

Fig. 4. Reputation rules for public forum chains.

decreases its reputation. Then, it's possible that the last post by c (in red) is rejected together with all posts by y and x in sequence. Unlike Bitcoin, forks are not only allowed but encouraged due to the local-first principle. The resulting linked list is only a view of the primary DAG structure. However, the longer a peer remains disconnected, the more conflicting operations it may perform, and the higher is the probability of rejection when rejoining.

Some other considerations about merging forks: Peers that saw branches with less reputation first (*branch* 1) need to reorder all blocks starting at the forking point. This might even involve removing content in the end user software. This behavior is similar to chain reorganization in Bitcoin when a peer detects a new longest chain and disconsiders old blocks. Likewise, peers that saw branches with more reputation first (*branch* 2) just need to put the other branch in sequence and do not need to recompute anything. This normal behavior is expected to happen in all peers connected with the majority of the network.

### 3.2 Public Forum Chains

The public forums of Freechains adopt the proposed reputation system. We now present the concrete rules that track the reputation of users in order to enforce consensus and support content moderation. Authors have to sign their posts in order to be accounted by the reputation system and operate in the chains. The reputation rules are detailed in Figure 4 and discussed as follows. The next example creates an identity whose public key is assigned as the pioneer in the public chain (prefix #):

```
> freechains crypto pubpvt 'pioneer-password'
4B56AD.. DA3B5F.. <- public and private keys
> freechains '#forum' join '4B56AD..'
10AE3E..
> freechains '#forum' post --sign=DA3B5F.. \
    'The purpose of this chain is...'
1_CC2184..
```

The *join* command in rule 1.a bootstraps a public chain, assigning 30 reps to the pioneer referred in the public key. The pioneer shapes the initial culture of the chain with its first posts and likes, while gradually transferring reps to

other authors, which may also transfer to other authors, expanding the community. In this regard, the *post* command above, signed by the pioneer, indicates the purpose of the chain to other users joining. Note that chains with the same name but different pioneers are incompatible because the hash of genesis blocks also depend on the pioneers' public keys.

Our most basic concern in public forums is to resist Sybils spamming the chains with new posts. Fully peer-to-peer systems cannot rely on identity logins or CAPTCHAs due to the lack of a central authority. Other alternatives include (i) building social trust graphs, in which users already in the community vouch for new users, or (ii) imposing economic costs for new posts, such as proof of work.

We propose a mix between trust graphs and economic costs. Rule 4.a imposes that authors require at least 1 rep to post, while rule 2 imposes a cost of 1 rep to post, even for existing users. These rules combined prevents Sybils to abuse the system. Rule 3.a allows an existing user to like a new user's post, which unblocks the post and vouches for the new user, but also at the cost of 1 rep. Note that the pioneer rule 1.a solves the chicken-and-egg problem imposed by rule 4.a.

The next commands illustrate the reputation rules with numbers. A new user joins the same public chain as before and posts a new message, which is welcomed with a like signed by the pioneer:

Figure 5 illustrates the chain DAG up to the like operation. The pioneer starts with 30 reps (rule 1.a) and posts an initial message, reducing his balance to 29 reps (rule 2). Then, a new user with 0 reps tries to post a message (hash C3A40F..), which is initially blocked (rule 4.a), as

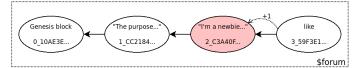


Fig. 5. The like approves the newbie message into the #forum DAG.

the red background highlights. Then, the pioneer likes the blocked message, which reduces himself to 28 reps and increments new user to 1 rep (rule 3.a), which is immediately consumed by the blocked post to become accepted. At the end, the pioneer has 28 reps and the new user remains with 0 reps but with his post accepted. Note that the overall balance after the commands above is -2 reps due to the penalties of the two post operations (rule 2). After at most 12 hours, these penalties are dismissed and the 2 reps are recovered (rule 2): the pioneer goes to 29 reps and the new user to 1 rep, which sum up to the pioneer's original 30 reps.

With no additional rules to generate *reps*, the initial 30 *reps* would constitute the whole "chain economy" forever. For this reason, rule 1.b awards authors 24 hours after a new post with 1 *rep*. This rule stimulates content creation and grows the economy of chains. The 24-hour period allows other users to judge the post before awarding the author, and also regulates the growth speed of the chain. Therefore, after the commands above complete 1 day, the pioneer accumulates 30 *reps* and the new user 2 *reps*, growing the economy in 2 *reps* as result of the two consolidated posts. Note that rule 1.b awards at most one post per day per author, i.e., with 10 posts in 7 days an author collects at most 7 *reps*. Note also that rule 4.b limits authors to at most 30 *reps*, which provides incentives to spend likes and thus decentralize the network.

Likes and dislikes (rules 3.a and 3.b) serve three purposes in the chains: (a) welcoming new users, (b) measuring the quality of posts, and (c) censoring abuse (SPAM, fake news, illegal content). The reputation of a given post is the difference between its likes and dislikes, which can be used for end-user software for filtering and highlighting purposes. The quality of posts is subjective and is up to users to judge then with likes, dislikes, or simply abstaining. On the one hand, since reps are finite, users need to ponder and avoid indiscriminate expenditure. On the other hand, since *reps* are limited to at most 30 reps per author (rule 4.b), users also have incentives to spend them. Hence, the scarcity and limits work together towards the quality of the chains. Note that a dislike shrinks the chain economy since it removes reps from both the origin and target. As detailed next, the contents of a post may become hidden if it has at least 3 dislikes and the number of dislikes is higher than the number of likes (rule 3). However, considering that reps are scarce, dislikes should be more directed to combat abuse, but not much to eliminate divergence of opinion.

A post has three possible states: *BLOCKED*, *ACCEPTED*, or *HIDDEN*. Figure 6 specifies the transitions between the states. If the author has reputation, its new post is immediately *ACCEPTED* in the chain. Otherwise, it is *BLOCKED* and requires a like from another user. Blocked posts are not considered part of the chain DAG in the sense that new posts do not link back to it. Peers are not required to hold

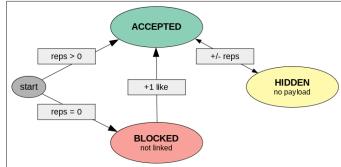


Fig. 6. State machine of posts: *BLOCKED* posts are not linked in the DAG. The payload of *HIDDEN* posts are not retransmitted. *ACCEPTED* posts are linked and retransmitted.

blocked posts and neither retransmit them to other peers. However, if blocked posts do not reach other users, they will never have the chance to be welcomed with a like. A reasonable policy for blocked posts is to hold them in a temporary bag and retransmit them for some visibility in the network. Rule 4.c limits the size of posts to at most 128Kb to prevent DDoS attacks using gigantic blocked posts. Once accepted, a post becomes part of the chain and can never be removed again since Merkle DAGs are immutable by design. If the number of dislikes of a post exceeds the threshold, its payload becomes HIDDEN and is not retransmitted to other peers. Since the Merkle DAG depends only on its hash, removing the actual payload is harmless. Also, the DAG itself contains the dislikes that prove the hidden state to other peers. Later, if the post receives new likes and changes its state, it means that the payload is still known somewhere and peers can request it when synchronizing again.

#### **3.3 TODO**

- users that work twice, bots

Terminar com "@!xxx" dizendo que vale reps mas que owner pode "intervene"

- summary - parallel w/ bitcoin - ponte entre reputacao -; sybil -; consenso - no problem w/ forks

We reach a similar solution to Bitcoin but adapted to our domain - we need consensus to make the reputation system work - we need reputation system to reach consensus - same virutous cycle as bitcoin

The more CPU work is done, the stronger becomes the proposal, the more peers follow it, the more tokens are mined. There is a strong association between work, profit and consensus that enables Bitcoin as a peer-to-peer cash system.

- reputation system to rate messages - positively (to distinguish from excess), or - negatively (to block SPAM, fake news, illegal content) - some sort of scarcity (work) - you like you loose - you work you get - otherwise sybil, "likes" abuse - incentives for continuous, good quality posts - consensus to ???

#### 4 RELATED WORK

Bitcoin employs CPU proof-of-work as a solution, but did not solve the centralization issue completely because the cost to perform work is not equally distributed (depends on costs of equipment and energy) - bitcoin demands connectivity to avoid forks - forks ok in freechains since operations are not critical and can be reverted - 50+1 attack easy in first nodes if not all nodes keep participating - freechains requires checkpoints

- related work public forums in scuttlebutt are permissioned (you decide who participates). each user has a completely different view of #chat. not really public although some intersections - you can see an answer but not a question, or a question w/ no answers even if the best answer exists. no content discovery

## 5 CONCLUSION

Like *bitcoins*, *reps* are scarce, hard to generate, and easy to verify. Unlike *bitcoins*, *reps*.

With this design, we retarget, the balance between work, profit and consensus also applies

attack 50+1 but not as cite local-first software

(crypto proof & authoring), while verification is cheap requires work but evaluation

, but are evaluated by other user The system also imposes incentives to rate posts and

Unlike bitcoin

Since likes and dislikes spend reputation, *creps* are scarce resources Incentives to post and rate content.

There is a strong association between work, profit and consensus that enables Bitcoin as a peer-to-peer cash system.

distinguish SPAM from legitimate Another problem CPU

content.

(e.g., posts on social media, public conversations based on the reputation of users in the network.

biggest difference: work is subjective as is the evaluation by other users

with consensus since whoever proposes the ordering will choose one of the operations arbitrarily

which is similar to the example above (buy X vs buy Y) spent

with work

this gives consensus with total order, solves double spend, which is equivalent to solving X/Y is decided above we borrow token, scarcity

that requires work to Only one purpose

- incentive security
- Last-Write-Wins
- just all tokens are the same, no subjective judgment, for example on why token is being transferred

and xx timestamps.

of the service and , and trust from users to deal

In a decentralized setting, - spam - abuse - on topic - disconnections - order

Bitcoin

responses would

n user posts

- discovery - also consensus, ensure that participants receive all data in a consistent order - bitcoin, discovery even disconnected, consensus, but not quality

Regardless of the Internet growth over the years,

# 5.1 Subsection Heading Here

Subsection text here.

5.1.1 Subsubsection Heading Here

Subsubsection text here.

#### 6 CONCLUSION

The conclusion goes here.

# REFERENCES

- [1] J. Zittrain, "Fixing the internet," 2018.
- [2] S. A. Theotokis and D. Spinellis, "A survey of peer-to-peer content distribution technologies," *ACM Comput. Surv.*, Dec. 2004.
- [3] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Tech. Rep., 2019.
- [4] F. Sant'Anna, F. Bosisio, and L. Pires, "Freechains: Disseminação de conteúdo peer-to-peer," in *Workshop on Tools, SBSeg*'20.
- [5] M. Kleppmann, A. Wiggins, P. van Hardenberg, and M. Mc-Granaghan, "Local-first software: you own your data, in spite of the cloud," in *Proceedings of Onward'19*, 2019, pp. 154–178.

Michael Shell Biography text here.

PLACE PHOTO HERE

John Doe Biography text here.

Jane Doe Biography text here.