

SAN : A Distributed Contribution Graph

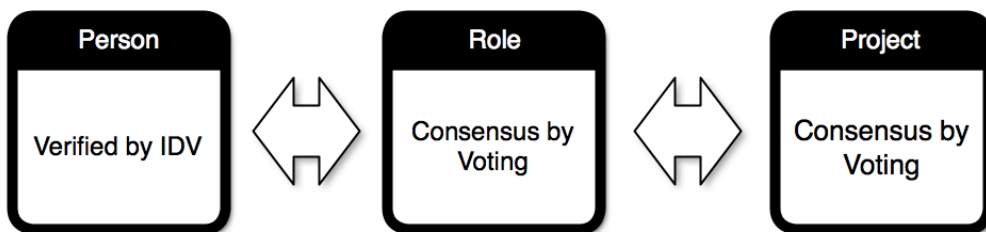
Abstract

SAN aims to build a universal “Contribution Graph”: a verified, distributed & immutable register of projects & who contributed to them. Such a system can answer the following question: Who contributed what, where? e.g. “Linus Torvalds created the Linux Kernel” or “Brian Chesky co-founded AirBnB in 2008”. Such a register can be of tremendous value to the creative professions in the Open Source and Startup communities.

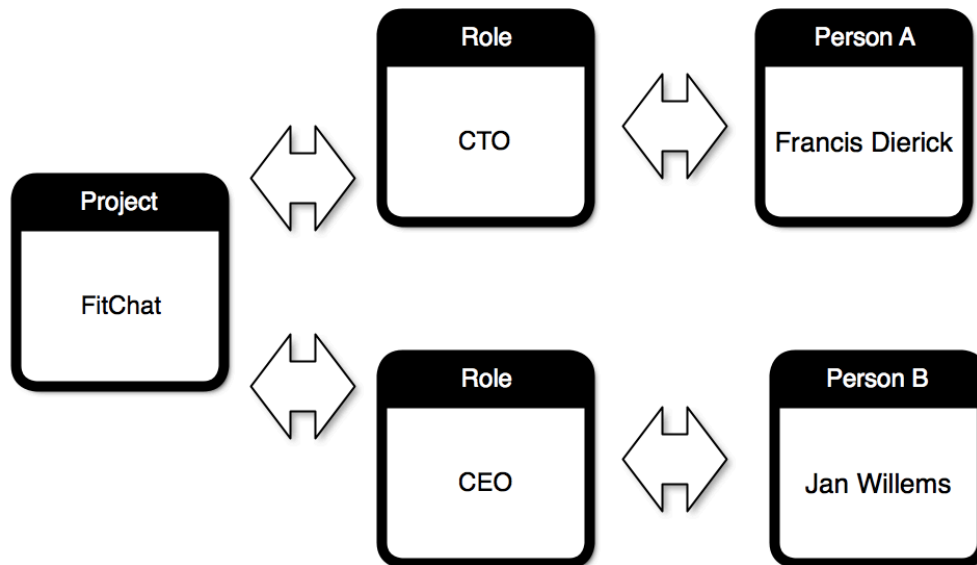
Executive Summary

SAN aims to build a “Contribution Graph”: a system able to answer the following question: “Who contributed what, where?” Being able to answer this question is valuable in hiring & investment decisions in the creative industries where the relative worth of an individual is often more closely linked to the projects they worked on than the companies they worked for.

Such a graph can be used to assess a project contributor’s reputation, assign a fair equity split in early projects or even automatically compensate team members for the work they have done. There are 3 main components required to build such a graph: People, Roles & Projects. Roles (e.g. “CTO”, “Designer”, “Developer”) form the edges linking together People & Project nodes in a graph. Project & Role descriptions are reached by voted consensus among the People participating in a project.



Pre-requisite for such a system is that each Person has a strong, verified, public identity. This is achieved initially through a third-party Identity Verification (IDV) process (e.g. using Trulioo). Long-term this needs to be decentralised & offered by blockchain projects such as Civic. People need to be mapped to real-life identities & past contributions need to be added to a Person's work history. The former can be achieved through a formal IDV process as seen by e.g. AirBnB or Uber. The latter can be achieved by e.g. adding past Github projects to work history.



We propose the introduction of a token called SanCoin (SAN) through an ICO. SanCoin plays the following roles in SAN:

- ICO to fund initial development of the SAN Contribution Graph. (1M USD)
- ICO overselling (if > 10M USD) to subsidise IDV for key members of the existing open source community.
- token to be used as voting currency in the SAN Contribution Graph.

We aim to build the Contribution Graph according to the UNIX philosophy: it is meant to be a small building block on which larger systems can be built. This whitepaper outlines the core system to be built, the roadmap to get there, the ICO process & a few examples of apps that could be built on top of the Contribution Graph.

Why Blockchain?

Blockchain technology offers a couple of interesting properties appropriate for a project like SAN. The following are the core characteristics SAN requires to work.

IMMUTABILITY. Since the Contribution Graph represents historical facts about how people worked together it is desirable to make the graph *immutable* once a consensus among project participants has been reached. Immutability of people's roles in a project can be enforced by storing roles as edges between person & project nodes in a graph. A role's immutable metadata consists of start & stop dates, commitment % & a role identifier. To limit the amount of data stored on the blockchain metadata that is not required for smart contracts will be stored as a reference to a url outside the blockchain.

TRANSPARENCY. Currently people's true roles in e.g. a startup project are not always clearly defined. Common examples are advisors & members who are represented on a startup's team page but in reality no longer contribute. Since all changes to roles in a project are consensus-based and are logged into the blockchain in a distributed fashion there can be no ambiguity about how a consensus was reached. This is merely useful as a reflection of how much each person really contributed for historical projects. For ongoing projects this becomes crucial because e.g. external contracts can use the Contribution Graph for a current project to pay out wages to a contributor based on that person's commitment % to the project.

CREDIBILITY. SAN users will be able to build up a public reputation based on the projects they collaborated on. This reputation is based on the sum total of all the creative projects they have worked on and have registered in the SAN Contribution Graph. Examples of reputation-builders could include Open Source contributions, registered side-projects, startups, job-related projects, hobby projects. For this reputation to be credible it needs to be built on a consensus-based description of the projects & the roles of each project member. This consensus is enforced through voting by team members which uses SanCoin as a voting currency.

DECENTRALISATION. Long-term SAN wants to build a decentralised reputation system. A core component of such a system is a decentralised identity verification system (D-IDV). Decentralised IDV systems are currently being built by third-parties like e.g. Civic & long-term we aim to integrate with third-party IDV providers. To bootstrap the SAN network however we will work with traditional IDV providers & gradually decentralise as more modern IDV systems based on blockchain tech come online.

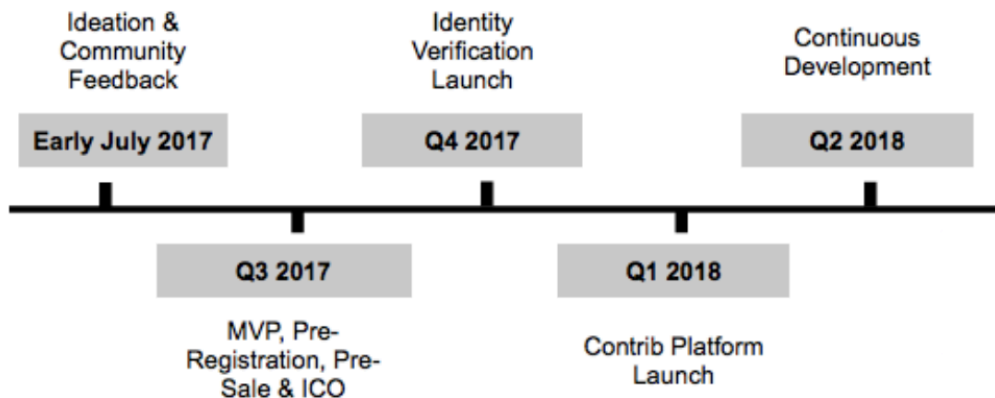
Use Cases

The following section outlines common problems people encounter when working on projects together and how SAN aims to solve them.

- *Problem: Is this startup Dead or Alive?* Answering this question is not always easy because startups tend to fade away into obscurity rather than die outright. As early as the nineties websites like fuckedcompany popped up to track the status of companies. A current example of this is YClist. One solution based on the Contribution Graph would be to setup a contract which looks at a startup's roles: one can expect a living, breathing company to have some fluctuation in the roles associated to it. People come & go, take on different roles, etc ... A purely static Contribution Graph probably signals the startup is dead.
- *Problem: Was advisor X really part of the startup?* Celebrity advisors often join a startup for PR reasons: it simply looks good to have expert advisors onboard. This is especially true right now in the blockchain world. Often these advisors do nothing more than have a single meeting. This has two nefarious side-effects: this presents a false picture of the startup to outside observers like e.g. investors and over time resentment will build among the project member who do provide significant input to the project. Because the Contribution Graph is an immutable representation of all the work account holders have done across all projects there is a strong incentive for a true representation in order not to pollute their reputation.
- *Problem: people start working together on a project & need to agree on a fair equity split.* This discussion usually only happens at the time of incorporation leaving project members in a void about their contributions in the months or years leading up to said incorporation. For some types of projects formal incorporation is even not desirable. (e.g. Open Source) One easy way to decide on a split early on is to record a % distribution by assigning roles for each member of a project & storing this in the SAN Contribution Graph.

Generally speaking the use cases here represent situations that occur in early stage startups or informal projects where legal incorporation has not occurred yet.

Roadmap



Ideation & Feedback

We will present the draft whitepaper & interact with the crypto community before finalising the final version that will be used for ICO.

MVP, Pre-Registration, Pre-Sale & ICO

An MVP of the web interface will be made available where users can pre-register their projects at no cost. Per-Sale for larger buyers will be organised online. After Public ICO token holders will be able to formally claim their accounts & use tokens to request formal IDV. In case of overselling > 10M USD a fund will be set up to subsidise IDV for key members of the open source community.

Identity Verification Launch

Token holders will be able to request Identity Verification where their real-life & project-related identity will be verified. IDV services will be paid in SanCoin by token holders. As a gesture to the Open Source community & to increase the value of the initial SAN network long-time Open Source contributors will be offered free IDV in case of ICO overselling.

Platform Launch

The SAN platform will consist of two main public components. First, a set of open-source command-line tools to allow sophisticated users to interact with SAN in a truly distributed manner. Second, a user-friendly web interface built on top of said tools to allow everyday users to interact with SAN.

Continuous Development

The scope of the initial SAN platform has been kept small on purpose & continuous development will mainly consist of smart contracts that interact with the Contribution Graph. Examples of potential applications are outlined in the “Further Research” section.

Example Applications

This sections outlines 3 types of application that realistically could be built on top of the Contribution Graph. The first 2 are straightforward examples in the Collaboration space & will be developed as open source sample apps to illustrate the Contribution Graph. The third one is inspired by previous work done in the Health space by the SAN authors and serves as an example of how existing apps can use the SAN network to store mission-critical verified data. The examples are the following: The Virtual startup (san-startup), Time Tracking (san-timetrack), Proof-Of-Fitness (san-fit).

Startup (float-based SAN roles)

This scenario is applicable when people are working on a project full-time, do not immediately need to generate personal income for the project, have not incorporated yet but would like to already assign an ownership split to project founders.

Often when startup founders initially start working together there is no legal framework for their collaboration. Incorporation & assignment of IP usually only happens when a startup starts generating revenue or accepts a seed-stage investment. At this point the founders will often already have invested significant amounts of time in the startup & discussions about equity splits often turn into fights or worst case startup death.

The Contribution Graph can be used to settle on an equity split for a startup project earlier in its lifecycle than is usual. The founders create a project in SAN & assign percentage-based roles to each founder. E.g. Jane, John & Joseph decide to work together on a “Unicorn Haircare” startup. Since Joseph is not ready to work full-time on the project yet, the future equity split is set at $0.4 + 0.4 + 0.2$. This is done by assigning the corresponding floating-point values to the roles connecting the founders with the project.

The main benefit in this scenario is transparency if the project raises money later on: each founder’s role is clearly reflected in the Contribution Graph & can be inspected by potential investors to answer common questions like: “Who really worked full-time on this? Who joined when? What is the commitment of this team?”

Time Tracking (sum-based SAN roles)

This scenario is applicable when people decide to work together on a project but would like to generate income early from the project as contractors.

Each contributor tracks the time they spend on a project similar to how a contractor would invoice a client by generating timesheets. In this case the contributor simply logs their time in the SAN Graph. Project members approve each other's "Timesheets". To keep the amount of data stored in the blockchain small it is recommended to only store timesheet summaries. i.e. monthly like a freelancer's timesheets, not per-minute. The SAN Graph can be queried to calculate the total sum of time each member has spent on the project.

This sum can then be used by a third-party contract to automatically enumerate contributors based on the time they have spent on the project. An example of this could be an open-source project where funding is assigned to a third-party contract that interacts with the SAN Graph to assign payouts to contributors each week.

Proof-of-fitness (event-based SAN roles)

This scenario is useful if the project contribution is event-based rather than time-based, i.e. the value for the common project is generated when certain events occur. This example is based on a concept developed for the SAN creators' own startup FitChat in the Health space called "Proof-of-Fitness".

At FitChat we think people owe it to their colleagues & society to live a healthy active lifestyle. Inactivity is killing us: diabetes & obesity are on the rise simply because we live a far too sedentary lifestyle. Illnesses related to inactivity are a burden on society & can derail projects due to unexpected health-related downtime.

What if you could *prove* you are living a healthy lifestyle by e.g. publishing your daily step-count gathered by a wearable in the SAN Graph. We call this "Proof-of-Fitness". One common example in corporate wellness is a stepcount-based rally for charity where a certain amount of money will be donated to charity based on how many steps a team have taken collectively.

In this example each team member would install an app on their phone that automatically uploads a daily stepcount to the SAN Graph. A third-party contract would hold the funds to be allocated to the charity & pay out according to each team's performance.

Technical Considerations

Ecosystem & Tech Stack

We plan to build on top of *Ethereum* & will contract with an independent blockchain expert to assure the correctness & security of the SAN system. We aim to integrate

with nascent distributed identity verification (IDV) services like Civic. *We do not desire to provide IDV ourselves long-term.* Initially we will however contract with existing IDV providers to bootstrap the SAN community until the nascent distributed IDV industry is mature.

Our interest is ONLY in building a system to manage the “Contribution Graph”. In this we follow the UNIX philosophy: “Write programs that do one thing and do it well. Write programs to work together. Write programs to handle text streams, because that is a universal interface.”

The core technology stack will be kept minimalistic & consist of a series of well-documented open source *command-line programs to interact with the public “Contribution Graph”* running on Ethereum. On top of this core stack a user-friendly web interface will be built which can be used from the main SAN website. Sophisticated users of course can fall back to the command-line tools if they desire to interact with SAN in a truly distributed manner.

Components

There are 3 main types of objects in the SAN stack: People, Roles & Projects. The 3 objects have similar data structures but instances of Roles & Projects will be created by consensus through voting by project participants. Initially instances of Person objects will be verified through third-party IDV services payable by each SAN user in SanCoin tokens. Long-term IDV should be done by projects like Civic.

Data Storage for each object is kept minimalistic & works by referring to outside url’s. Ultimately we are more interested in storing the graph between objects than detailed information about what each object represents. Each object is identified uniquely by a its blockchain id & consensus on the state of an object is reached through voting by all project participants. Both People & Project objects will store most of their metadata in outside systems. An exception is made for the Role object since that is at the core of the SAN graph. There we will allow storage of numerical metadata to represent the role, e.g. “X worked Y% on project Z” or “X worked Y hours on project Z”. Storing large objects, e.g. a Person’s avatar picture is out of scope for the SAN Graph & will be handled by external storage systems like Swarm, Storj or IPFS.

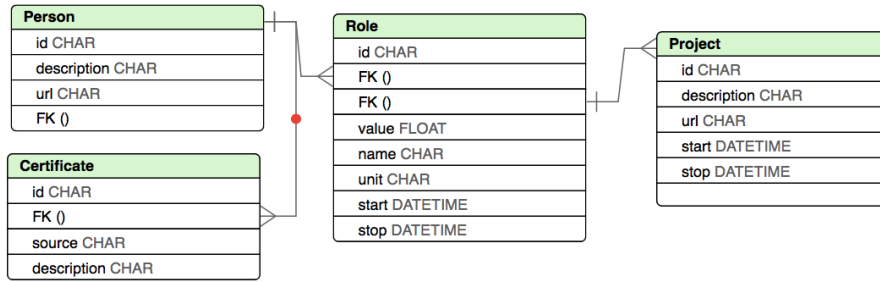
People (Verified Identities) form the core of the system. A strong validated identity is pre-requisite for establishing the basic trust in the system. The Person object is the answer to the question “Who?”. A people object consists of its unique id, a short description, an external url & relationships to its Certificates, Roles & Projects.

Certificates describe how a Person has been verified & will initially be issued only by the SAN network itself after third-party IDV. One person can have multiple certifications, e.g. id-card based through a service like Trulioo or account-based through e.g. a Github login. A weighted sum of all certificates represents the trust a Person has in the SAN network. Each certificate is associated with a single Person object.

Roles The Role object is the answer to the question “What?”. Besides id & external url it also contains three extra fields: Tag, value & unit. The tag field is a short text-based unique description of the role. (e.g. “CTO”) In the value field a numerical value can be stored to quantify e.g. %-based or hour-based roles. This field makes it possible for external contracts to perform actions based on the SAN Graph. The unit described by the value (e.g. hours, %, etc ...) is stored in a small text field named “unit”. Any change in a role field needs to be voted on by a project’s members.

Project The Project object is the answer to the question “Where?”. In a more generic form projects can be thought of as namespaces where Roles & People interact. A project just consists of a unique id, a description & an url. Any change in a project field needs to be voted on by a project’s members.

Data Structure



The above diagram shows the data structure modelled as a traditional relational database. In order to keep data storage cost low the number of fields has been limited. All fields will be written away in the blockchain in the data sections of either contracts or transactions. Editing will be done through a voting mechanism & modifications to the data will be written away in the data section of the voting transaction.

Rules

Some basic rules will be enforced by the contract to enforce data integrity.

- start <= stop for both Role & Project objects
- role start <= project start
- role stop <= project stop
- sum(all roles for person) <= 1.0 if unit type is sum

Business Considerations

SanCoin Token holder benefits

- Token holders will be able to create contributors, projects & roles in the platform.
- Token holders will be able to pay for making vote-based modifications to projects & roles they participate in.
- Token holders will be able to pay for initial manual verification of their account.
- Token holders will see the value of their tokens increase due to a gradual reduction of the supply.
- Top 100 weekly token holders will be rewarded with VIP status. VIP status will be visible on their account & can be used for expediting identity verification (their own or friends). Additionally the top-10 weekly token holders will be written into the SanCoin contract
- Summary: SanCoin will be a utility token that can be used on the SAN platform

Seeding

Select people from the startup & open source world will be invited to register & verify their SAN profile at no cost to seed the initial user base. In case of an overselling > 10M USD we will subsidise IDV for selected core members of the open source & startup community. In order to qualify for the IDV subsidy SAN account holders will need to provide proof that one of the following applies to them:

- Proof that they have participated in a world-class startup accelerator program such as Ycombinator.
- Proof that they have made major contributions to open source projects on e.g. Github.

Spending

We aim to raise approximately 1M USD & expect development to take up the majority of costs related to the initial release of SAN. The amount we wish to raise is in line with a typical first seed round for a startup & should give the core development team a runway of 1 to 2 years. We expect the initial expenses to be roughly split according to the following formula:

- Research & Development: 50%

- Marketing: 25%
- Legal & Operations: 25%

Research & Development Initial development will cover the following components of the SAN system: 1) SAN platform & open source command-line tools. 2) Sample app running on top of the Contribution Graph. 3) Web-based UI for interacting with the Contribution Graph.

Branding & Marketing The marketing budget has been kept modest on purpose since the Contribution Graph itself has inherent viral properties. We expect 2 main marketing pushes at the following events: 1) Branding at time of ICO 2) Marketing in open source community at launch.

Legal & Operations We expect to conduct a legal & security review of the SAN contract code. Operations includes all other costs related to running the project.

Overselling In case of an overselling > 10M USD we will subsidise IDV for select-ed core members of the open source community.

Team, Community Feedback, Further Research, References

Team

Project initiator & core committer is Francis Dierick, a software engineer with 20 years of experience building systems for the web & mobile. SAN is based on issues related to privacy & accountability while developing the blockchain backend for an app in the health space (FitChat).

Vidar Andersen is a startup maven: he regularly keynotes startup events & has an exceptional network in the Western startup world. As co-founder of the Plone project he is one of the earliest Open Source contributors. Francis & Vidar have worked together on several project before.

Puja Abbassi has built up an exceptional network in the Chinese startup world through his mentoring at Chinaccelerator (中国加速) He is connected to Open Source through his membership of both the Linux Foundation and the Cloud Native Computing Foundation.

Community Feedback Process

Even though all founding team members' contact info has been made publicly available we prefer to interact with the community through GitHub for reasons of transparency:

- Github: a git repo has been made available at <http://fdierick.github.com/san>

We welcome community feedback. We are especially interested in proposals for projects that can be built on top of the “Contribution Graph” & solicit early discussion to make sure our system is as flexible as possible.

Further Research

The following scenarios are out of scope for the initial ICO & simply serve as examples of the kinds of systems we think can be built on top of a strong “Contribution Graph”.

- Universal Basic Income (UBI): we believe the “Contribution Graph” may serve as a proven history of doing creative work regardless of income. This could be used as a weighting component in wealth distribution programs like UBI to make those fair & meritocratic. We believe that the combination of verified identity & the contribution graph can be used to create a fair & balanced UBI scheme.
- Credit Assignment: Mathematically the UBI scheme described above is similar to finding a solution for the so-called Credit Assignment problem & we believe a strong Contribution Graph will be a sub-requirement for solving it.
- Contribution Badges: allow SAN members to publish a “Contribution Score” on their own blogs & websites. One example could be a Contribution Badge on a Github profile.
- Reputation Systems: since all changes to the “Contribution Graph” are consensus-based reputation systems may be built based on a user’s contribution history.
- Strong Person-to-Person References: as the strength of their “Contribution Graph” grows users are disincentivized to give weak references to people they barely interacted with. See LinkedIn for a prime example of weak references.

Inspiration

SAN grew out of the founders’ frustration with sites like LinkedIn. Open source projects, side-projects & experiments are core to how modern creative professionals like designers, coders & entrepreneurs work. A modern creative’s worth is more closely linked to the projects they worked on than the companies they worked for. Yet these side-projects are poorly represented in current systems.

The SAN founders participated in Ycombinator's Startup School. Upon graduation they received a digital certificate & a spot to showcase their work on startupschool.org. This website was kept simple on purpose & answered just the three questions investors will ask first: "What is this? Who is working on it? Who is doing what?"

What if every open-source project & startup you contributed to had a simple page like that? What if we link all those pages together in a graph? What if we introduce strong identity & let contributors vote on how a project answers these simple 3 questions: "Who contributed what, where?". What if we call it a "Contribution Graph"?

References

- Yclist: List of YC-funded companies and their exit status (Dead / Alive / Exit) <http://www.yclist.com>
- Fuckedcompany: 90's website listing dead startups : https://en.wikipedia.org/wiki/Fucked_Company
- Civic : Secure & Protect Identities, Giving businesses and individuals the tools to control and protect identities: <http://www.civic.com>
- Trulioo: Manual, centralised IDV provider : <http://www.trulioo.com>
- Startupschool: <http://www.startupschool.org>
- The UNIX Philosophy: https://en.wikipedia.org/wiki/Unix_philosophy
- FitChat : Slack for Coaching: <http://www.myfitchat.com>
- YC Blog post : IPFS, CoinList, and the Filecoin ICO with Juan Benet and Dalton Caldwell : <http://blog.ycombinator.com/ipfs-coinlist-and-the-filecoin-ico-with-juan-benet-and-dalton-caldwell/>
- Storj : Blockchain-based, end-to-end encrypted, distributed object storage, where only you have access to your data. <https://storj.io>
- IPFS: A peer-to-peer hypermedia protocol

to make the web faster, safer, and more open. <https://ipfs.io>

FAQ

- *What's to prevent token printing?* The initial supply of tokens is fixed and the token contract will define a supply reduction function which periodically burns tokens.
- *When will SAN be launched?* The launch is expected in Q1 2018
- *Is there a token pre-sale? When will it start? How do I get in? When will it end?* A discounted SanCoin token pre-sale will start at an undisclosed time in July 2017. Register on the main website & ask about how to get

invited to the pre-sale. Invitees will receive access to a closed chat channel where the pre-sale will take place. We plan to hold the pre-sale over a roughly 24h time period to give participants from all timezones a chance.

- *When will the public ICO start?* The public ICO will start on July 24 2017. At this point anyone from the public will be able to buy SanCoin.
- *When will the public ICO end?* The public ICO will end one week after it starts.
- *How many coins will be available?* The initial supply is 2 Million SanCoin tokens.
- *How many coins will be allocated in the pre-sale?* A maximum of 400 thousand
- *How many coins will be allocated to the core team?* 1.2 million tokens
- *How many coins will be allocated in the ICO?* A total of 800 thousand tokens will be allocated (pre-sale + ICO)
- *Is there a supply limiting mechanism?* At least 200 thousand tokens will be burned while SAN is under development.
- *How can I buy, sell, hold or transfer tokens?* SanCoin is an ERC20-standard compliant token and can be traded as such.
- *What's different about this token?* SanCoin is the first token where in case of overselling a significant amount of value will flow back to the community in the form of IDV subsidies for core open source contributors.
- *Where do the funds collected in the ICO go?* All funds initially collected will go to the ICO contract. From there they will be transferred to other wallets according to project & founder needs.
- *How do I get a refund for the tokens I bought?* There are no refunds.
- *What can I do with SanCoin?* SanCoin is a standard ERC20 token, so you can hold it and transfer it. Once SAN launched you will be able to use SanCoin to request IDV of your account and to vote.
- *Will SanCoin be traded on any exchanges?* After the pre-sales and ICO we plan to work with exchanges to get SanCoin listed.
- *Is this a scam?* While we cannot guarantee the success of the SAN project we are very serious about building the Contribution Graph & expect to work full-time on it for at least one year after a successful ICO. We think the Contribution Graph can be of great benefit to creative professionals.
- *Why are you doing this?* The initial inspiration is related to our project FitChat: we were looking for a transparent way to store "Proof-of-Fitness" in the blockchain. While we think this particular use case has great value a more generic approach of storing proof-of-contribution might be of more immediate value to the creative community. Simply put: we think a system like this needs to be built.