Insights Network

- A Blockchain Data Exchange

29. December 2017 (v0.5)

Brian Gallagher
Insights Network
bg@insights.network

Darwin Lo
Insights Network
darwin@insights.network

Peter Frands Frandsen
Partisia
pff@partisia.com

Jesper Buus Nielsen
Partisia
jbn@partisia.com

Kurt Nielsen
Partisia
kn@partisia.com

Abstract

In the present day, data brokers collect data on individuals around the world from a wide array of sources, both online and offline. The data is packaged into profiles, which they sell to organizations who can use them to make decisions that impact the lives of ordinary people without them ever being aware of it. Recent technological advances, such as the Blockchain and Secure Multiparty Computation, are making it possible to build a superior platform for conducting market research while putting control, as well as monetization, of the data into the hands of the people who generate it.

Content

1 Disclosure

Nothing herein constitutes an offer to sell, or the solicitation of an offer to buy, any tokens, nor shall there be any offer, solicitation or sale of Insights Network INSTAR tokens in any jurisdiction in which such offer, solicitation or sale would be unlawful. You should carefully read and fully understand this whitepaper and any updates. Every potential token purchaser will be required to undergo an on-boarding process that includes identity verification and certain other documentation, which you should read carefully and understand fully because you will be legally bound. Please make sure to consult with appropriate advisors and others.

This white paper describes our current vision for the Insights Network platform. While we intend to attempt to realize this vision, please recognize that it is dependent on quite a number of factors and subject to quite a number of risks. It is entirely possible that the Insights Network platform will never be implemented or adopted, or that only a portion of our vision will be realized. We do not guarantee, represent or warrant any of the statements in this white paper, because they are based on our current beliefs, expectations and assumptions, about which there can be no assurance due to various anticipated and unanticipated events that may occur.

Please know that we plan to work hard in seeking to achieve the vision laid out in this white paper, but that you cannot rely on any of it coming true. Blockchain, cryptocurrencies and other aspects of our technology and these markets are in their infancy and will be subject to many challenges, competition and a changing environment. We will try to update our community as things grow and change, but undertake no obligation to do so.

2 Background

2.1 Problem

There are organizations, called data brokers, that collect data on people from various sources, both online and offline. Most notably, they purchase data from Internet services and mobile applications that collect information on their users and track their in-app behavior.

The data is used to create profiles for individual consumers. Acxiom, a top data broker, has on average 1,500 pieces of information on more than 200 million Americans. Acxiom and other data brokers are able to combine all the information they have on each individual to generate in-depth consumer behavior reports across a wide range of industries, which they sell.

Data brokers make a lot of money. In 2012, Acxiom was reported to have made $1.13 billion in sales, earning a profit of $77.26 million. Forbes reports that Big Data Analytics is a $200 billion per year industry and that by 2019 nearly all businesses will be customers of a data broker such as Acxiom.
But consumers, who are actually driving the industry, do not share in the profit.

Simultaneously, they experience many negative consequences as a result. Centrally managed databases allow hackers to steal large amounts of personally identifying information in just one attack, which allows for large-scale identity theft and fraud. Recently, hackers breached Equifax's systems and stole personally identifiable data on more than 140 million Americans. This is not an isolated incident. There have been several attacks in the past, including on Acxiom, and if unchecked there will be more in the future.

Consumers must demand a new standard for the storage of personally identifiable and sensitive information that is being used in market research. This is currently being implemented in the GDPR throughout Europe, and soon, the US will follow suit.

2.2 Solution

Recent advances in decentralized storage, digital currencies, and smart contracts enable us to create a decentralized, incentivized platform for conducting marketing research and securely storing consumer data. Organizations will be able to use our platform to make requests for data from precisely defined populations amongst members of the Insights Network.

Insights Network users, not data brokers, would sell their data. By using smart contracts, which execute transactions between anonymous parties, users would be able to sell their data without disclosing their identity, only broad demographic information. A unique combination of a blockchain and Secure Multiparty Computation (SMC) makes it possible to enforce the exchange of data and payment between the provider and the requester of data without third-party involvement. While a blockchain makes the exchange transparent, SMC keeps data truly secure until an agreement has been reached and paid for.

We believe that by allowing participants to profit from their participation, organizations will acquire data that is both more relevant and more actionable as a result. At the same time, profits currently earned by data brokers will instead go to the rightful owners of the data, the consumers.

3 Our product

3.1 Overview

Primarily, we are serving two types of users: those who request data, which we call requesters, and those who provide it, which we call providers. Requesters are typically organizations, but anyone can buy INSTAR tokens and use them to submit a request for data to the Insights Network. Providers are users who comply with a data request by providing data; the ones who fit a data requesters' target demographic are compensated for their data in INSTAR tokens.

Requesters want to be able to collect data that is:

- **Relevant**. They want to be able to gather data from specific populations, for example, only those between the ages of 20 to 35.
- **Trustworthy**. The data that is collected is free from fraud. That is, the data is collected from targeted providers who have provided their data honestly. For example, requesters expect their surveys to be answered truthfully by their target demographic, and not say, a bot.
- **Timely and convenient**. Requesters should be able to get answers to their questions quickly without worrying about the details of how to reach their target demographic.

Providers want assurance of the following.

- **Consent.** Their data is not collected without their explicit permission.
- **Privacy.** Sensitive information, such as who they are, is not provided, only broad demographic details.
- **Payment.** They are paid for the data they provide.
- **Security.** Their data is handled securely.

Traditional marketing research firms collect data over a fixed period of time and provide a single report to their clients. On our platform, requesters would never have to close their request for data. Their reports, which they would be able to view at any point during data collection, are updated as data from providers is received and forwarded by the smart contract.

As more data is available in our network, the data that the platform makes available to requesters becomes more comprehensive, which in turn gives the ability to design reports that are more comprehensive, including ones that combine data from more than one data request. Here are several examples on how the Insights Network could be used.

- A polling firm may want to run a poll to see who would win if the presidential election were re-run at this moment in time.

- McDonald’s may want to run a survey to solicit feedback on a new menu item.
- University classes can administer surveys throughout the quarter or semester to get feedback on the quality of instruction on an ongoing basis.
- A high-end auction can accept anonymous bids from verified individuals without having to perform KYC.

A related group of use cases addresses so-called asymmetric information by profiling of individuals and companies. A financial example could be a lender that knows less about the borrower’s ability to pay back the loan than the borrower. This may result in higher interest rates to an otherwise low risk borrower to compensate for the average expected risk. The solution to this problem is information that can profile as many individuals and companies as possible to separate e.g. low-risk from high-risk borrowers. Similar problems arise in the insurance business, where the insurance company knows less about the insured parties’ profile. Another example is product differentiation, where the supplier lacks information about the customers’ preferences in order to define the most suitable menu of products or contracts. One example of the latter could be the problem of designing the right menu of mobile subscriptions with respect to price and data, etc.

The Insights Network provides a platform that empowers the individual person or company to collaborate with e.g. banks and insurance companies to address this problem on equal terms. The Insights Network may become a unique source of information for profiling individuals and companies.

3.2 Functionality

The Insights Network is a secure infrastructure for decentralized data exchange, which has a number of different applications. In the beginning we will focus on surveys, but it will also be the foundation for other applications, such as royalty schemes, advertisements, and loyalty schemes.

The Insights Networks consists of the following components:

- Authentication
- User profiles
- Decentralized data brokerage
- Surveys
- Two-sided rating

A requester is a user who places a request on the Insights Network to get information from a specific population of users. Providers are users who fulfill such a request.

3.2.1 Authentication

A common practice among apps today is to allow their users to login using their Facebook account. We are building a similar service for apps to allow their users to log in using their Insights Network account. This would appeal to users who do not want to disclose their identity to their apps. As an added benefit, apps can use the Insights Network to give tokens to their users as part of a reward program and allow them to cash them in for rewards, such as airline miles.

The authentication is done confidentially using SMC.

3.2.2 User profiles

Users maintain an Insights Network profile. They can view their profile, which contains demographic information and other general, non-identifying information. They can make corrections, fill in missing details, and delete information. Furthermore, other than deleting information, they are minted INSTAR tokens to perform these actions.

For example, a 25-year-old woman living in Los Angeles may want to receive ads related to her political interests, so she may choose to keep the line that indicates she is a Republican. But she may not wish to receive surveys in relation to being a single mom, so she would delete that line. She can also see that her profession is unknown, and she may choose to fill it out in exchange for INSTAR tokens.

The user profiles are kept confidential either client-side or by using SMC.

3.2.3 Decentralized data brokerage

The basic purpose of Insights Network is to enable decentralized data exchange between providers and requesters, which consists of:

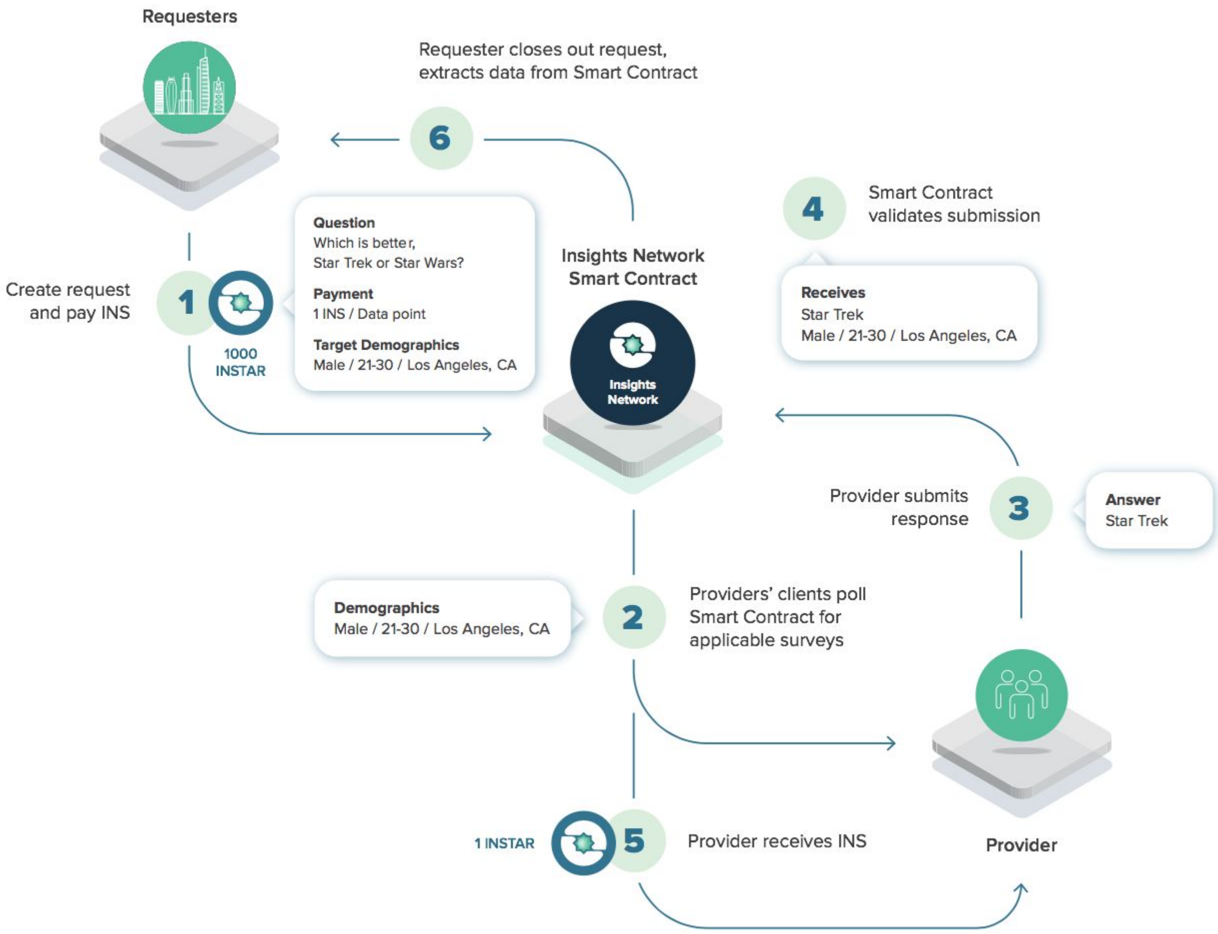
- Matching providers' profiles with the requester's demand
- Secure data transfer
- Secure payment

The data to be exchanged is initially survey data consisting of the relevant background information and user-supplied answers to the questionnaire. A unique combination of a blockchain and SMC secures the exchange of confidential information and payments without any third-party involvement.

3.2.4 Surveys

Anyone on the Insights Network, whether they are an individual or an organization, can publish a survey to the platform. The platform gives them the ability to specify a target demographic, as well as how many tokens users in the target demographic will receive when they submit a valid response. The process is as follows:

1. A requester publishes a survey, specifying a target demographic.
2. Users of our app who fit the target demographic get notified.
3. Users fill out the survey, submit it to the Insights Network, and funds are transferred to their account.


Requesters
Requester closes out request, extracts data from Smart Contract
6
Question
Which is better, Star Trek or Star Wars?
Payment
1 INS / Data point
Target Demographics
Male / 21-30 / Los Angeles, CA
Create request and pay INS
1
1000 INSTAR
Insights Network Smart Contract
Insights Network
4
Smart Contract validates submission
Receives
Star Trek
Male / 21-30 / Los Angeles, CA
Provider submits response
3
Answer
Star Trek
Demographics
Male / 21-30 / Los Angeles, CA
2
Providers' clients poll Smart Contract for applicable surveys
1 INSTAR
5
Provider receives INS
Provider

3.2.5 Two-sided rating

To handle unwanted behavior, the Insights Network implements a two-sided rating system, whereby a requester may rate providers and vice versa.

As a provider, it is possible to misuse the system and provide fake data. Some of this unwanted behavior can be detected by the requester while analyzing data. Giving the requester the opportunity to rate the providers will counteract such behavior.

As requester, data may be used outside of the agreed purpose. Part of such unwanted behavior can be detected by the provider as the final reports are published. Giving the provider the opportunity to rate the requesters will counteract such behavior.

Such decentralized regulation of unwanted behavior through two-sided rating is familiar from services like Uber and Airbnb.

3.2.6 Example: Gambeal

Gambeal is an iOS app that administers surveys to patrons of quick service restaurants, paying them a small cash reward for each survey they fill out. To prove their patronage, each user is required to attach a photo of their receipt along with each submission, which is validated prior to the rewards being issued. Gambeal has seen considerable growth without any deliberate marketing effort and processes thousands of transactions each week. It will be the first Insights Network partner app, and its success so far bodes well for other partner apps.

As part of the integration, Gambeal will swap out the authentication system with the Insights Network's SMC-based authentication system, which will allow users to log in without giving up their identity as, say, logging in with Facebook would. Only broad demographic details are released to the app, and by virtue of being part of the Insights Network, the user is able to monetize their data as they wish.

In addition, Gambeal will switch away from cash rewards to the INSTAR token, using the Mobius Universal Protocol API for the integration. Gambeal currently uses PayPal to make payments to users, which forces users to have to wait days for their rewards to appear in their PayPal account, as well as pay up to 3% in transaction fees. In addition to providing better privacy, using the Insights Network will allow Gambeal to process transactions quickly and at low cost to their users. Integrating with the Insights Network for authentication and rewards represents a clear improvement to the user experience.

3.3 Technical challenges

There are quite a few challenges that need to be solved. The following features of the Insights Network serve to enable as well as fulfill requests:

- **Identity verification**. Providers prove they are real people by authenticating with several verification partners -- such as a company that checks identity documentation, a service that performs background checks, or even their employer -- and getting a Digital Proof of Authenticity from the Insights Network, which they can use to prove their authenticity in transactions with requesters. Using techniques from distributed cryptography, the verification partners never relinquish their information to any other party, which preserves providers' privacy, and providers are compensated in tokens for going through this process.
- **Incentivized market research**. Requesters can publish a survey and receive survey responses using the Insights Network smart contract. The smart contract sends tokens to providers in the target population who have submitted a valid data point.
- **Blockchain-verifiable results**. At requesters' discretion, the data points of a data request could be recorded in the ledger for anyone to look at. Since the ledger is implemented using a blockchain, anyone looking at these records would have reasonable assurance that the data points have not been tampered with. This feature is important for certain kinds of surveys, such as polls and voting. If needed, the data can be encrypted using informationally-theoretically secure encryption on the blockchain.

- **Semantic validation of private data**. Without involving a third party, the provider will be able to withhold data until she receives payment from a requester and a requester will be able to withhold payment until the provider's data is proven to be valid.

4 Technical solutions

The two primary technical components of our solution are a blockchain and protocols for performing Secure Multiparty Computation (SMC). Both of these technologies use distributed cryptography to remove the need to entrust a third party with a critical role in order to execute a transaction; that is, they are both "trustless." At the same time, the two technologies provide complementary properties: A blockchain provides transparency, while SMC provides privacy.

SMC has been incubating in academia for more than three decades. Its use in industry has been limited due to its computational expense. But in 2008, Partisia, a company in Denmark, released a commercial deployment of SMC for the first time, replacing a traditional auctioneer in a double auction. Since then, performance has improved by orders of magnitude and at this time is beginning to gain adoption in industry.

Blockchains started with Bitcoin. Then Ethereum introduced a blockchain that supported smart contracts. We are betting that EOS, which uses a novel algorithm for block production to achieve a much higher transaction throughput than both Bitcoin and Ethereum, will be the next major blockchain, and it was what we will be building on.

We see blockchains and SMC growing in prominence together. Furthermore, we think they will frequently be combined for their complementary traits, as we are doing in our system. Along with our partner Partisia, who are assisting with the design and implementation of our system, we believe we are starting a new trend in trustless computing. In this section, we go into how we plan on using SMC and a blockchain to address the challenges outlined in section 3.3.

4.1 Blockchain

A blockchain is an implementation of a tamper-resistant public ledger. The state of the ledger is determined by consensus among independently operated servers, which form a network called a blockchain network. It is impossible for a single server or even a small group of servers acting in collusion to tamper with ledger entries without being detected by the other servers. As long as there are enough independent, protocol-abiding servers in the network, the ledger is tamper-proof.

In addition to keeping records of the movement of funds between accounts, blockchains are also capable of keeping records of changes to the state of a program. Programs that are hosted on a blockchain network are called smart contracts or decentralized apps. The Insights Network makes use of a blockchain and blockchain-based smart contracts for logging activities, assembling the output of a Secure Multiparty Computation, and transferring funds between accounts.

4.1.1 EOS Blockchain Platform

We are building on EOS, an upcoming blockchain operating system. Though it is unreleased, it is being developed rapidly, and as of this writing, we are currently building on the a local test node, as well as hedging our development by simultaneously building on Ethereum. This section goes into why we have chosen EOS.

Prior to EOS, Dan Larimer, the architect of EOS, architected two successful blockchain projects, Steemit and BitShares. Steemit is the only blockchain app that handles a realistic workload, 17,000 daily active users (dau). Graphene, the blockchain used in BitShares, has shown it can handle 20,000 transactions per second on a network. When released, EOS will have the greatest throughput of any blockchain network currently in existence, which will be needed to handle the level of activity we expect in the Insights Network.

In addition, EOS will have several features that make it suitable for operating an app such as the Insights Network.

1. EOS allocates resources, such as transaction bandwidth, to each account according to the number of EOS tokens held by that account. It also allows apps to pay for their users' usage, which means that users do not need to pay each time they use an app unlike with other blockchain networks, such as Ethereum.
2. Smart contracts and decentralized apps (dApps) can be upgraded to introduce new features and fix bugs, which will allow us to improve the Insights Network rapidly in response to real-world usage.
3. Many of our users are ordinary people, and there is no guarantee that their devices are secure; inevitably, some of their accounts will be compromised. Unlike other blockchain networks, EOS allows compromised accounts to be recovered with the help of a designated partner if they provide identity documentation and multi-factor authentication.

A portion of the proceeds from our token sale will be used to acquire and hold EOS tokens. Holders of EOS tokens are provided guaranteed transaction bandwidth across the network without disruptions caused by other activity occurring on the EOS blockchain. For example, in the event of an ICO or even a denial-of-service attack, users are still entitled to their share of the transaction bandwidth. EOS calls this "rate limiting."

Another consideration that is important to the Insights Network is security. Putting out a data request may involve paying out a large number of tokens. If someone were to make an unauthorized data request, an unwanted transaction would occur. EOS provides a couple of features that would help in this and other security-critical situations.

1. EOS allows our users to require that some operations are approved by multiple parties. In the case of a data request, our users could stipulate that making a request requires approval by several people within their organization. According to the EOS white paper, this feature, which it calls

"multi-user control," "is the single biggest contributor to security, and, when used properly, it can greatly eliminate the risk of theft due to hacking."

2. EOS allows apps to add a delay before sensitive operations are recorded in the blockchain, which is when they become irreversible. During the waiting period, users are notified by email or text that the operation is occurring, and they are given a chance to stop it if they did not authorize it. In the case of a data request, our users would be alerted to unauthorized data requests and be given the opportunity to cancel them.

We believe that EOS has a bright future. It is well-funded, having raised $185 million in just the first five days of its year-long token distribution, and we are well-supported by the EOS team. Given these and other considerations, we have determined EOS to be the closest fit for our needs.

4.2 Secure Multiparty Computation (SMC)

Secure Multiparty Computation belongs to a class of modern cryptographic solutions that enable computation on unknown data. This might seem impossible at first, but using the right cryptography, it is not, and in addition to SMC, this class of solutions includes techniques such as zkSNARKs and homomorphic encryption. SMC in particular achieves this goal by converting the computation into a distributed computation, in which no participant of the computation sees a full input but rather a derivation of it that does not on its own give information about the full input.

The seminal aspects of this concept can be traced back to Shamir (1979), with the theory being founded in the 1980s (Chaum 1988). Although SMC was shown in theory to be generally applicable in the mid-1980s, the computational complexity of SMC prevented its practical use for two decades. The first large scale and commercial use of SMC was done by the Partisia, a Denmark-based company, in 2008, when they used SMC to replace a traditional auctioneer on a double auction (Bogetoft et al. 2009).

Since 2008, the technology has matured both in terms of computational speed as well as the properties of the SMC protocols. The computational overhead has been reduced by approximately 1/1,000,000. The development of SMC can be traced by reading the following papers: Pinkas et al. (2009); Shelat and Shen (2011); Nielsen et al. (2012); Frederiksen and Nielsen (2013); Frederiksen and Nielsen (2014); Lindell and Riva (2015); and Nielsen et al. (2017).

For help with designing and implementing our custom SMC protocols, we have partnered with Partisia, who in 2008 developed and deployed the first production implementation of SMC. Since then, they have spun out a service[1] for encrypting data that uses SMC to avoid having to store a full copy of any key on a server, as well as a custom SMC-based dark pool for Tora[2], which processes $3 billion worth of transactions on a daily basis. Finally, they also help to maintain an open-source SMC framework called FRESCO. Arguably, no other organization has greater depth and experience at production-grade implementations of SMC.

[1] www.sepior.com

[2] www.tora.com

4.2.1 Custom SMC systems

SMC is applicable to a diverse set of applications. It is not a single protocol but a growing class of solutions, each with different characteristics. A number of SMC systems have been devised to meet the specific needs of different applications, such as key management and financial order matching.

Common to all SMC solutions are the following roles, which each person or organization holds one or more of:

1. The **Computing Parties** are responsible for carrying out the distributed computation.
2. The **Input Parties** have inputs for the computation that they would like to keep confidential. To this end, they use a technique called secret sharing to decompose each input into parts that are delivered to different Computing Parties. No Computing Party has more than one part for an input, and none of the parts alone give enough information to derive the original input.
3. The **Result Parties** are the parties whom the Computing Parties send their results to. The Result Parties assemble the data they get from the Computing Parties into the result of the overall computation.

Crucially, no party besides the Input Parties ever see the original inputs.

Custom SMC systems may differ along the following parameters:

- **Basic operations.** These are the operations that are used to define the computations. An SMC system will have either arithmetic or Boolean operations.
  - **Arithmetic operations** are more convenient for expressing statistical analyses
  - **Boolean operations** are more efficient at pattern matching.
- **Cryptographic primitives.** An SMC system will use one or more of the following cryptographic operations.
  - **Secret sharing:** a technique for splitting a piece of data into parts that by themselves do not give information about the original data. Secret sharing is very common in SMC systems.
  - **Oblivious transfer:** a class of protocols for data transfer in which the sender sends one of several pieces of data but does not know which one.
  - **Homomorphic encryption:** a class of schemes for producing ciphertexts that can be computed on.
- **Trust model**
  - **Self-trust:** A party assumes it can only trust itself.
  - **Honest majority:** A party must rely upon the majority of the parties being honest.

Different combinations of these parameters give rise to different properties:

- **Fault-tolerance:** Under self-trust, all parties are needed for the computation to proceed, and the system will fail if even one of the parties is unable or unwilling to participate. Whereas if a system merely relies on there being an honest majority, the system can proceed to completion even if some of the parties fail to carry out their duties.

- **Security**
    - **Passive security:** As long as all Computing Parties follow the protocol, none of the parties learn anything besides the output of the computation. Also known as semi-honest security.
    - **Active security:** None of the parties learn anything besides the output of the computation, even in the presence of malicious parties who are willfully trying to deviate from the protocol.
    - **Covert security:** The system is able to identify that a party has a 50% probability of being malicious -- which is suspiciously high -- and take punitive measures.
- **Performance.** Active security is often much less performant than passive security. Covert security provides similar guarantees to active security but is much more performant.

Due to the nature of the technology, custom systems are necessary to achieve acceptable levels of performance. Partisia has been developing custom SMC systems since 2008 and will be helping the Insights Network with the design and implementation of a custom SMC system with the required security and performance guarantees. Notably, it will be one of the first SMC systems to interact with a blockchain.

4.2.2 The Insights Network SMC solution

SMC will be used to solve the following technological challenges in the Insights Network solution: SMC-based authentication and SMC-based profile matching and data point validation. To achieve good performance, we have designed custom protocols for each of these use cases.

In our SMC-based authentication, the computation produces a Digital Proof of Authenticity, which can be thought of as a certificate, and is carried out by the provider, the Insights Network, and verification partners selected by the provider. The proof contains the provider's EOS account name, basic demographic details, and is signed by the Insights Network. The more certainty there is around the provider's authenticity -- that is, the more high-quality verification partners vouch for the provider and the more basic demographic details the provider provides -- the more she is paid.

The SMC receives the following inputs, which remain unknown to all participants of the computation besides the original owner.

- Insights supplies the private key used to sign the Digital Proof of Authenticity.
- The provider supplies her EOS account name, which will be included in the Digital Proof of Authenticity, as well as demographic details she wants to include in the certificate.
- Each verification partner, who are provided with the provider's EOS account name, supplies basic demographic details.

During the computation, the identity of the provider is verified and her basic demographic details are cross-checked. If successful, the requester gets a certificate she will be able to use to prove her authenticity in future transactions.

For SMC-based profile matching and data point validation, we use a two-party SMC protocol with covert security between the requester and the provider. Covert security is a relaxed version of active security, in which a party can be proven to have a 50% probability of having cheated. Cheating can be punished and deterred by publishing proofs of cheating to the blockchain. By using covert security instead of fully active security, we are able to provide safeguards against malicious behavior while at the same time providing a vastly more performant solution, which is important for achieving a good user experience.

A consideration we must take is that the Insights provider client may be running on a laptop or phone with limited network connectivity and computing power. The SMC protocol we have designed addresses these limitations with the following properties:

- Minimal trips: There are only two rounds of communication between the requester and the provider, which minimizes the effect of network latency and poor connectivity.
- Asymmetric computation: The requester will do the bulk of the work so that the provider's computer or mobile device is not overloaded.

Finally, using standard encryption schemes, it may be possible for someone in the future to take advantage of improved processing power to crack encrypted data stored on the blockchain using brute force. On the other hand, our system uses information-theoretically secure encryption, which cannot be cracked even if an attacker had unlimited time and computing resources. Hence, in our system, only the requester can decrypt the data they have purchased, both now and at any point in the future.

4.3 Solving the technical challenges

4.3.1 Proof of authenticity

Providers preserve their anonymity while participating in the Insights Network. But requesters need to know that, even though they are anonymous, the providers who are fulfilling their data requests are real people who are being truthful about their characteristics. This is the process by which a provider provides this reassurance.

Our system uses information from multiple parties in order to verify a provider's identity. For example, one party might be the provider's employer. Another party might take the provider's state-issued identity document and validate it.

We call these parties verification partners. If the provider passes verification, the system issues a digital proof of authenticity signed by the Insights Network and publishes it to the blockchain, where it is accessible to requesters. The digital proof of authenticity is a digital document containing the provider's blockchain account name, which is used as the endpoint identifier in the peer-to-peer protocol that requesters and providers communicate over; it is similar in function to a TLS certificate. Using SMC, our system is designed to allow each party to provide the information they hold as an input to the verification process without disclosing it to the other parties.

As regards identity verification, the inputs are the confidential information held by the verification partners, and the computation to be performed is verifying that information and generating a digital proof of authenticity. When the SMC computing parties are finished with their sub-computations, they send their intermediate results to the provider who assembles them into a digital proof of authenticity, signed by the Insights Network.

The Insights Network is not assigned a sub-computation but contributes its private key as an input, which is used in the computation to generate the signature for the digital proof of authenticity.

4.3.2 Submitting a data request

For a small fee, a requester can send a data request to the Insights smart contract. The request would include a survey for providers to fill out, a pattern describing the target population, and how much to pay for a valid data point. The request would also include tokens to be held in escrow for the smart contract to pay out to qualified providers who have submitted valid data points. For extra security, since a large number of tokens could be involved, the requester could stipulate that each request be approved by multiple parties from their organization using the blockchain operating system's permission system. Once a requester is satisfied with the data points they have received, the requester can close out the data request, and any tokens remaining in escrow are returned to the requester's account.

4.3.3 Fulfilling a data request

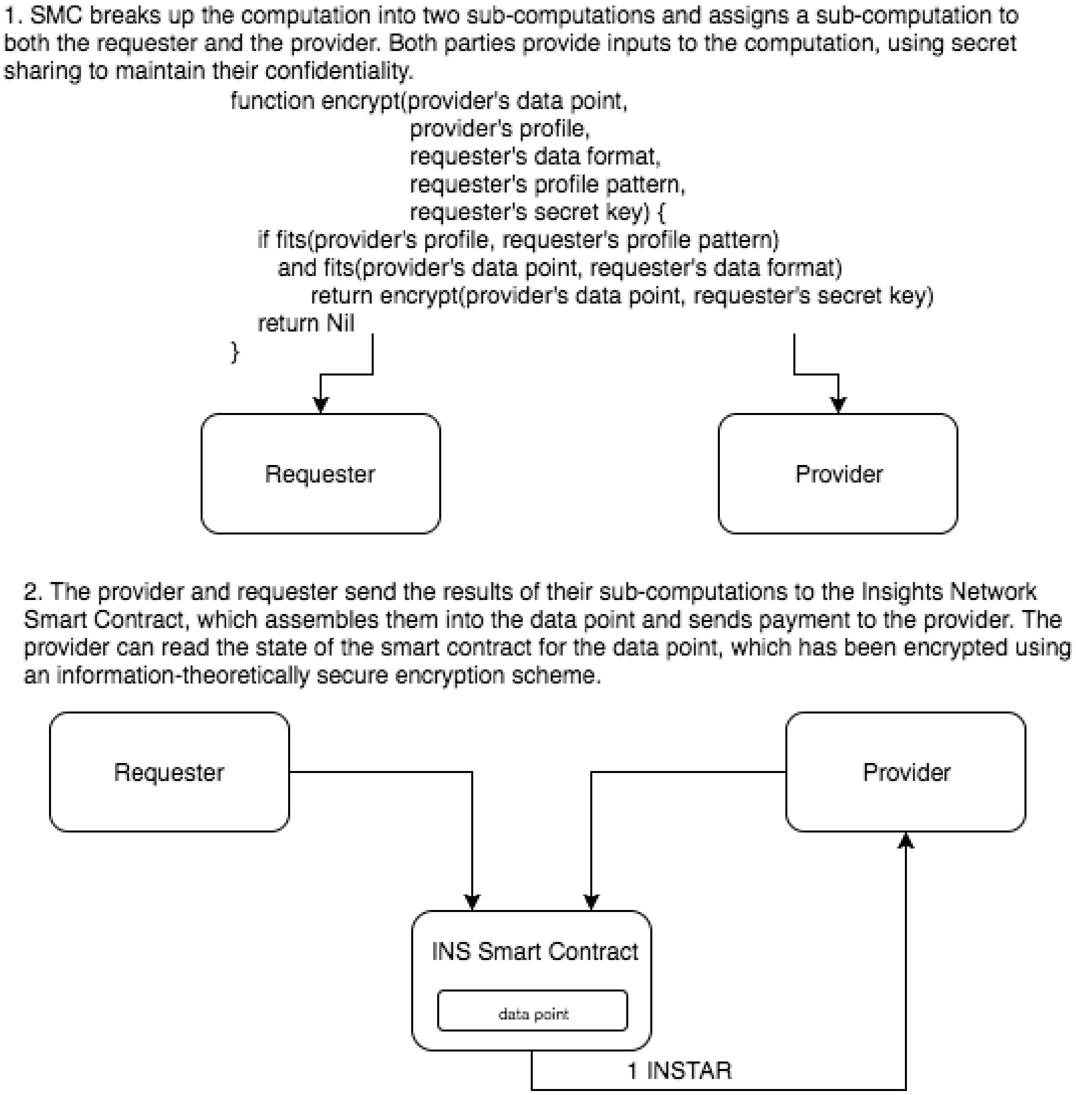
Periodically, the Insights Network client used by the provider will check with the Insights Network smart contract to see if there are open data requests. It downloads them locally and uses the provider's profile to select which ones to display to the provider. This section describes the process by which a provider submits a data point for a data request.

The provider chooses a data request and fills out the corresponding survey, which creates a data point. She then sends her certificate to the requester, along with the data request ID, to indicate interest in submitting a data point. The requester uses the Insights Network's public key to ensure that the certificate is valid and ensures that the name of the provider matches the account name in the certificate. If these check out, the requester generates a secret key that can be later used to encrypt (as well as decrypt) the data point and initiates a Secure Multiparty Computation with the provider.

Both parties provide inputs to the computation, which the protocol prevents from being disclosed to the other party. The provider provides her profile and the data point she created. The requester provides a profile pattern, a data format, and a secret key, which he generated for this particular submission and stored for later use.

The computation, which the two parties carry out cooperatively, consists of checking to see if the data point fits the data format and encrypting the data point using the secret key. Sub-computations are derived from this computation and assigned to the two parties who carry out them out and send the results to the smart contract for assembly into the final result. If the final result is an encrypted data point, as opposed to an empty value, then the smart contract sends the provider the tokens she is entitled to according to the terms of the data request. If not, then the data point was invalid, and her rating is reduced.

Notably, the encryption scheme used in our system is information-theoretically secure. That is, even if an attacker had unlimited computing power and time, he would not be able to crack the data point. This means that we can store the encrypted data point on the blockchain without having to worry that someone will be able to crack it ten years from now when computing resources have improved, which a naïve encryption scheme would be vulnerable to.



1. SMC breaks up the computation into two sub-computations and assigns a sub-computation to both the requester and the provider. Both parties provide inputs to the computation, using secret sharing to maintain their confidentiality.

```
function encrypt(provider's data point,
                 provider's profile,
                 requester's data format,
                 requester's profile pattern,
                 requester's secret key) {
  if fits(provider's profile, requester's profile pattern)
    and fits(provider's data point, requester's data format)
       return encrypt(provider's data point, requester's secret key)
  return Nil
}
```

2. The provider and requester send the results of their sub-computations to the Insights Network Smart Contract, which assembles them into the data point and sends payment to the provider. The provider can read the state of the smart contract for the data point, which has been encrypted using an information-theoretically secure encryption scheme.

4.3.4 Profile matching

A provider is only qualified to submit a data point if her profile matches a pattern provided by the requester. One way this could be implemented is for the provider to send her profile to the requester for vetting. But this could be a breach of privacy, since profiles may contain information that providers may prefer not to disclose. And in any case, the extent to which requesters need to know what is in a profile is whether it matches the pattern they are looking for; they don't actually need to know the profile's exact contents. For example, a requester may only need to know that a provider's age is from 21 to 30, not that the provider is 25 years old.

As described in Section 4.3.3, profile matching in our system is done inside a Secure Multiparty Computation. The provider provides her profile as an input to the SMC, which by virtue of the protocol is never exposed to the requester.

4.3.5 Secure exchange of validated data

Consider a requester and a provider. The requester is willing to buy data from the provider if the data meets certain requirements. But the provider is not willing to let the requester see the data prior to receiving payment, because if the requester is in possession of the data, the requester may choose to take it without paying. At the same time, the requester is not willing to pay for the provider's data unless they know the data is "good," which the requester defines as matching a pattern.

In our system, a data point is validated and encrypted inside a Secure Multiparty Computation. The provider and requester carry out sub-computations, but it is the smart contract that assembles the results together into the final result. Effectively, storing the validated and encrypted data point on the blockchain, which delivers the data to the requester, and sending payment to the provider occur in one transaction, which has the advantage of either completely failing or completely succeeding; that is, either a valid data point is delivered to the requester and tokens are sent to the provider or neither occur.

4.3.6 Storage

The major categories of data that pass through the Insights Network are providers' profiles, providers' data points, and the information used during the identity verification process for providers. Profiles will be simply stored on provider's local devices. Data points will be stored in encrypted form in the state of the Insights Network smart contract, which is recorded on the ledger. Profiles are used for population targeting, but SMC ensures they remain private.

4.3.7 Provider's demographic

Providers keep a profile, which includes demographic information, in their Insights Network client. Some of the fields in the profile are fixed, such as sex and year of birth. Others are mutable, such as household income and relationship status.

Providers may change mutable fields, but each change is logged in the blockchain. Requesters may use this information to avoid fraud; for example, a requester may want to avoid requesters who have changed their household income more than three times over the past three years. Also, if a requester detects suspicious activity, for a fee, he may flag the provider in our two-sided rating system.

5 Token Economics

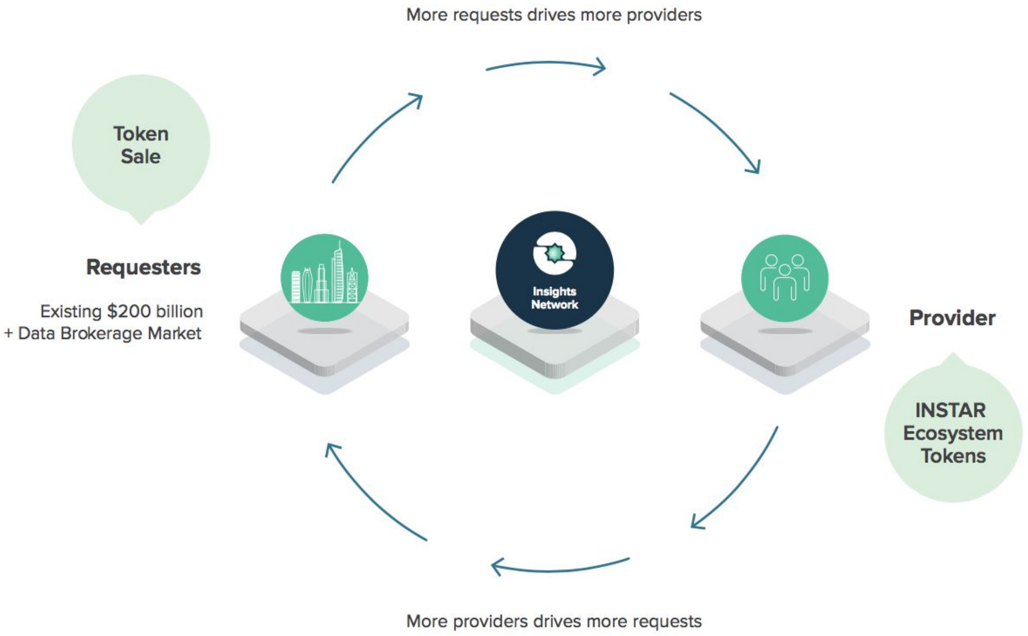
5.1 Two-sided Marketplace

The Insights Network is a platform that facilitates transactions between two distinct groups: requesters and providers. Requesters need information from providers and are willing to pay to get it. This dynamic is known as a two-sided marketplace.

Two-sided marketplaces are notoriously difficult to start up. The primary reason for a requester to place a request on the Insights Network is that it has providers who will fulfill it. At the same time, the primary reason for providers to join the Insights Network is that there are enough requests placed in order to make money. At the beginning, there are not enough of either group on the platform to attract the other.

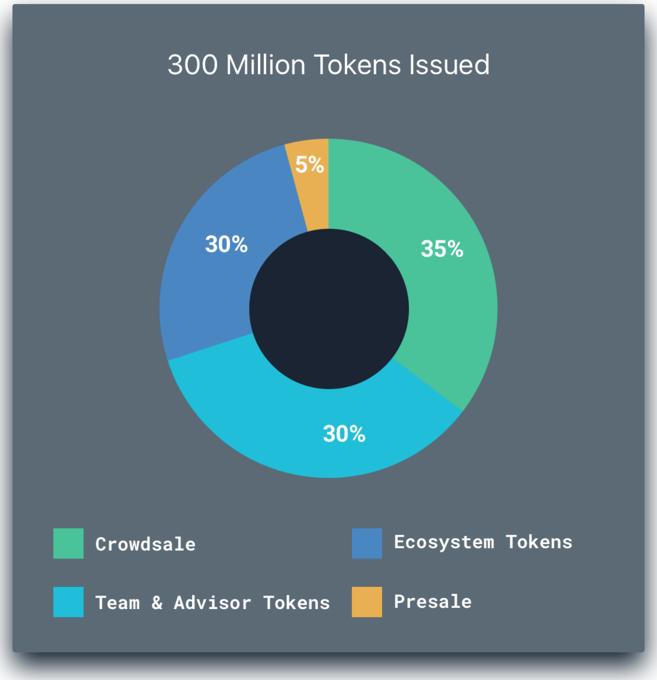
To help with starting up the marketplace, we are issuing a new ERC-20 token, which we call INSTAR. Users who place their data into the Insights Network and participate in market research will be rewarded with INSTAR tokens. Holders of INSTAR tokens will be able to place data requests or sell them to people who want to.

Providers will be able to sell the tokens paid out to them to future requesters, cash them out for things like airline miles or branded gift cards in the reward store, and send them to other users using the Insights Network wallet. As providers join the Insights Network, more requesters will be attracted to join and place data requests, which in turn would attract even more providers. This is known as a network effect, which could result in the network growing very large.



5.2 Distribution of Tokens

Tokens will be distributed in the following way:



Total Supply of INSTAR Tokens: 300MM

- 5% of tokens will be sold in a pre-sale at a discounted rate to the token sale
- 35% of tokens will be sold in the token sale
- 30% of tokens will be dedicated to the ecosystem
- 30% of tokens will be reserved by the company for team, advisors, operations, future engineering hires, R&D

90 million tokens will be issued to the ecosystem to be used as payment for early users to fill out their profiles, as well as participate in market research conducted by the Insights Network and early partners. By minting ecosystem coins to early adopters, we create hundred of millions of data points that create a viable data network for requesters to tap into. For example, if each early adopter receives a token for importing ten simple pieces of verified data into their client, these 90 million tokens generate 900 million data points for the Insights Network.

90 million INSTAR tokens will be set aside for the Company to be used as compensation for the Insights Network team, including the founders, employees, and advisors, which will incentivize them to increase the demand for the services offered by the Insights Network, and thus the demand for INSTAR tokens, by making it a compelling place to conduct market research.

It should be noted that during beta testing of the Insights Network Desktop Client, an Insights ERC-20 token will be used until the EOS platform is fully functioning and public, at which point there will be a one-time migration for token holders to swap their ERC-20 INSTAR tokens for the equivalent INSTAR EOS currency.

Token Sale Cap - 22,000 ETH

5.3 Use of Funds

Development - 50%
Operations - 25%
Marketing - 15%
Legal - 10%

6 Team

Brian Gallagher - W.P. Carey School of Business, Y-Combinator
Darwin Lo - Stanford Computer Science, Y-Combinator
Brandan Zaucha - W.P. Carey School of Business, Y-Combinator
Dylan Herman - Univ. Illinois, Engineering
Sebastian Leon - MIT Computer Science, Twitter, Data Science
Kurt Nielsen - CEO, Partisia, PhD Economics
Jesper Buus Nielsen - Research Scientist, PhD Cryptography, SMC, Co-founder Partisia
Peter Frands Frandsen - Computer Science, Aarhus University, Technical Director Partisia

If you have a passion for big data, are a computer scientist and love startups, contact us team@insights.network

6.1 Advisors

Jason Hamlin- A.C. Nielsen Data, Founder, Goldstockbull.com
Andrew Rosener - Founder, CEO, MediaOptions
David Gobaud - Co-Founder, Mobius.Network
Dino Amaral - Ph.D. Cryptography

6.2 Partners

Partisia is a pioneer in commercial implementations of SMC. The first large-scale and commercial use of SMC was done by Partisia in 2008 when it replaced a traditional auctioneer in a double auction for production contracts. Other key accomplishments include a spin-out called Sepior, which uses SMC to offer a pure cloud solution for trustless key management, and a spin-out called Secata, which provides off-exchange matching on financial markets for securities, as well as various privacy-preserving statistical analyses, such as collaborative credit rating.

7 Roadmap Draft

Insights: Q1-Q3 2017

- Proof of Concept
- Launch the Insights Network website
- Development work initiates on EOS test net

Insights: Q4 2017

- Updated white paper to include SMC
- Token Pre-sale
- Final platform specification with Blockchain and SMC

Insights: Q1 2018

- INSTAR Token CrowdSale
- INSTAR Wallet Integration Gambeal App Live on iOS
- INSTAR Platform online in alpha (with running implementations of SMC and blockchain)

Insights: Q2 2018

- INSTAR App Client Beta ERC-20 tokens redeemable in client
- INSTAR Platform in private beta with integrated blockchain rewards for SMC transactions

Insights: Q3 2018

- INSTAR Platform v1.0 with confidential requester criteria
- EOS Platform Beta Opens

Insights: Q4 2018

- INSTAR EOS blockchain Consumer Client Fully Functional

Insights: Q1 2019

- INSTAR is a fully decentralized platform

8 Conclusion

Data brokers collect intimate information on individuals without their permission and sell them to anyone who is willing to pay, even organizations that have considerable influence over the course of their lives, including universities, hospitals, and insurance companies. This is not just an invasion of privacy; it is surveillance. And there is little that consumers can do to stop this practice, even if they are among the few who know it is happening.

Fortunately, governments throughout the world are cracking down on data brokers. The European Union has enacted the General Data Protection Regulation, which regulates how organizations should handle what it calls “personal data.” Brazil prohibits transmitting data to another party that contains Personally Identifiable Information (PII). In the United States, Senator Edward Markey (D-MA) is sponsoring a bill called the Data Broker Accountability and Transparency Act of 2017.

We will work to promote regulations that demand a higher standard for the handling of sensitive, personal information. But even in the absence of regulations, we believe that our solution will be superior to existing data brokers and will win in a free market competition by providing superior information and putting consumers in control of their own data.

We predict that, in 10-20 years, due to the rise of decentralized technology, there will be no intermediaries. Organizations will use our platform to transact with consumers directly for their data. Organizations currently pay $200 billion per year for this data -- Forbes predicts that this will only grow. We think the Insights Network will grow to meet this demand and shift control and profit away from middle men such as Acxiom to the data's rightful owners, the consumers.

9 Sources

1. EOS.IO Technical White Paper
2. Multi Party Computation: From Theory to Practice
3. The secretive world of selling data about you (Newsweek)
4. 6 predictions for the $125 billion Big Data Analytics market in 2015 (Forbes)
5. Acxiom database hacked (Computerworld)
6. Equifax announces cybersecurity incident involving consumer information (Equifax)
7. Bogetoft P, Christensen DL, Damgaard IB, Geisler M, Jakobsen T, Kroejgaard M, Nielsen JD, Nielsen, JB, Nielsen K, Pagter J, Schwartzbach MI and Toft T (2009) Secure multiparty computation goes live, Lecture Notes in Computer Science, vol 5628, pp. 325–343.
8. Chaum D, Crepeau C, and Damgaard IB. (1988) Multiparty unconditionally secure protocols (extended abstract). In 20th ACM STOC, Chicago, Illinois, USA,May 24, 1988, ACM Press, pp. 11–19.
9. Shamir A (1979) How to share a secret, in Communications of the ACM 22, 11, pp. 612–613.
10. Pinkas B, Schneider T, Smart NP and Williams SC (2009) Secure Two-Party Computation Is Practical. Asiacrypt 2009.
11. Shelat A and Shen C (2011) Two-output Secure Computation With Malicious Adversaries. EUROCRYPT 2011.
12. Nielsen JB, Nordholt PS, Orlandi C and Burra SS (2012): A New Approach to Practical Active-Secure Two-Party Computation. CRYPTO 2012.
13. Frederiksen TK and Nielsen JB (2013) Fast and Maliciously Secure Two-Party Computation Using the GPU. ACNS 2013.
14. Frederiksen TK and Nielsen JB (2014) Faster Maliciously Secure Two-Party Computation Using the GPU. SCN 2014.
15. Lindell Y and Riva B (2015) Blazing Fast 2PC in the Offline/Online Setting with Security for Malicious Adversaries. CCS 2015.
16. Nielsen BN, Schneider T and Trifiletti R (2017) Constant Round Maliciously Secure 2PC with Function-independent Preprocessing using LEGO. NDSS 2017.