Bohnanza!

A Traceable Soybean Supply Chain using a Blockchain

Nikolaus Vertovec
vertoven@student.ethz.ch
Megan Morrow
mmorrow@student.ethz.ch
Sascha Stocker
sastocke@student.ethz.ch
Felix Richter
richterf@student.ethz.ch
Ingvar Groza
igroza@student.ethz.ch
Samuel Bernet
sbernet@student.ethz.ch

Introduction

Food product fraud is not a recent development, and often the frequency at which it occurs is driven by consumer trends. In fact, the organic food product market share in Switzerland grew by 4.4% between 2007 and 2017 ("*Demand for organic*", 2018). Unfortunately, fraudulent behavior in this realm can be difficult to detect even in cases where adulteration and counterfeiting may be concerned. In particular, though, increasing preference for organic over inorganically grown crops has shone the spotlight onto more than one occasion of a product being misrepresented as an organic label throughout the supply chain. While not always intentional, it does allow the seller to charge a premium. The commodity crop, soy, is worth highlighting as it is a highly versatile crop, and where human consumption is concerned it can be made into everything from cooking oil to tofu (Krull, 2018). With such value, it is not a surprise that this crop is susceptible to fraud. In 2017, a 36 million pound shipment of soy from the Ukraine to the United States somehow shifted from conventional to organic where "the addition of the "USDA Organic" designation boosted value by approximately $4 million" (Whoriskey, 2017). But sadly this is not the only instance of such a malicious transformation. Moving forward, efforts towards increasing the robustness of traceability measures are imperative especially as demand for organic soy grows. As such, *Bohnanza!*, a traceable soybean supply chain using a Blockchain will be laid out as a possible aid in this quest.

Problem Statement

As the world is moving towards a more sustainable organic food production, producers need to guarantee the integrity of their supply chain. Consumers are seeking more information on their products and the assurance of upholding the organic food production standards. Currently, consumers can only trust the producer not to violate any standards. Simultaneously there are many vulnerabilities when it comes to the production of soybeans. Every participant in the supply chain may act as a fraudulent actor. The farmer may add non-organic soybeans to his organic certified harvest to increase his profits. The trader is also a party that can defraud the consumer by mixing both non-organic and certified produce and sell the mixture as organic in order to increase his earnings. Each subcontractor is also an entity with an incentive to cheat the system by adding non-organic produce and then gaining from the certified soybean price. Consequently, each party has to be examined in order to uphold the trust of the consumer that his product is in fact organically produced.

With the help of Blockchain, *Bohnanza!* ensures full traceability of organic-certified soybeans produced in the Ukraine and transported to Switzerland. The system informs the producers of attempts of tricking the production standard by a party within the

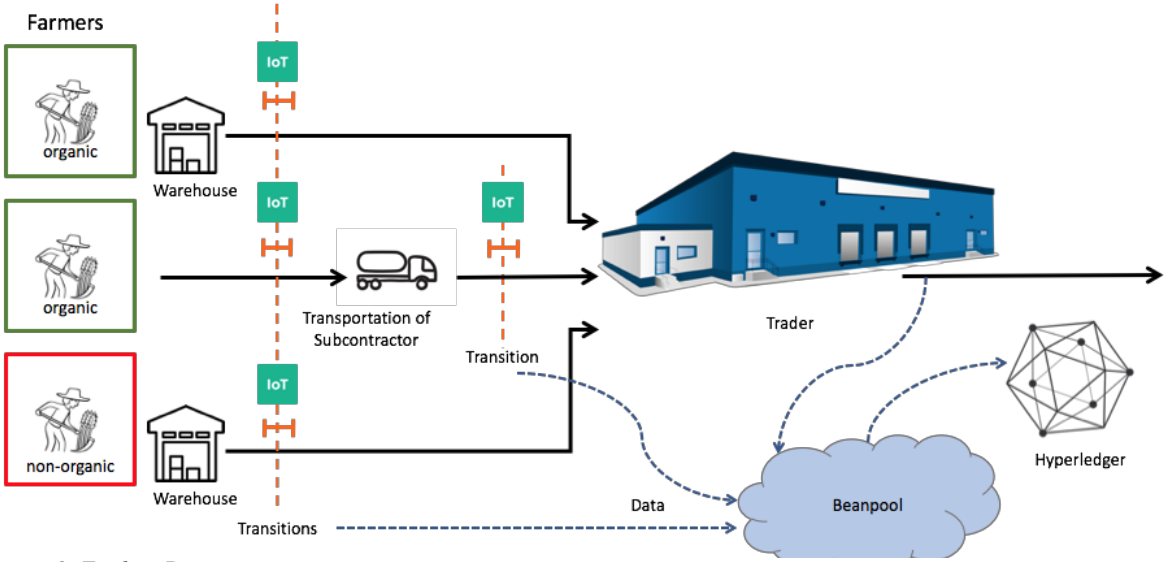

Figure 1: Trading Process

supply chain. This will ensure the trust of the consumer and allow them to track their product down to the field that produced it.

Our Solution

Initially, farmers in the Ukraine are registered in our system. They will be recorded with the location, number and size of their fields. It is also noted whether the farmer is a certified organic producer or not.

The farmers harvest their soybeans once a year. The harvesting is recorded with the help of IoT sensors. The duration of the harvest can be recorded as well as the movement of the tractor that does the harvesting. Next to the tractors, there will be trucks that are loaded with the soybeans. These trucks are also equipped with sensors that record the time and location of the truck. With the help of volume sensors, the volume of the harvest can be measured, leading to a precise evaluation of the harvest amount. All measurements have to match the expected data that is stored in our system. This means the yields of the field have to be in proportion with the size of the field and the expected harvest thereof. Furthermore, the location of both the tractors and the trucks have to align with the location of the field to a degree of measurement deviation. The evaluation of the data will be made in a *verification process* when a handover occurs.

After the harvest, the farmer stores the soybeans in their respective warehouse, if one is available. The soybeans require cleaning and a time period in order for them to dry. During this period, we are utilizing further IoT devices that can record the time and place the soybeans were handled since they have to be filled into large bags when they are stored. We can check with light sensor for instance whether a bag was opened before it was supposed to. Further we can have weight sensors monitoring if any soybeans are added or taken away. These records would be added to the data from the harvest to be verified in a similar manner.

The farmer may also choose to have his harvest directly picked up by a subcontractor. When such a handover occurs in the supply chain, it will trigger a *verification process*. It verifies the data of the transacted soybeans with the help of further sensors. The location of both parties and the amount of transacted soybeans is recorded. This data requires verification in order to uphold the integrity of the supply chain, ensuring both parties have little opportunity to cheat the system.

The subcontractor then delivers the product to the trader. His route to the warehouse of the trader is known to the system. The recordings of the subcontractor's movement with the harvested soybeans will be added to the data that requires verification. With the route and time of the transport recorded, the consumer can trust that the subcontractor had no opportunity to add non-organic soybeans to the harvest.

The subcontractor will hand over the soybeans to the trader, triggering another *verification process*. Again, the data will need to be verified in a similar process as from farmer to subcontractor, only now the data from the subcontractor and the trader is used.

At the trader's warehouse, IoT devices are used to safeguard against his ability to mix any of the delivered products, as he may be the recipient of both organic and non-organic soybeans. Before the trader's goods will be shipped off to Switzerland, they are once again double checked by the *verification process.* Each validated transaction will be added to the Hyperledger, creating a fully traceable supply chain for both the consumer and the producer.

The Verification Process

In order to ensure full traceability of the bean harvest from the farmer to the consumer, a publicly available transaction ledger must be created that can be easily verified by any user. At the same time, the *verification process* must ensure that no two parties can collaborate to outsmart the system. The presented method will not only ensure full traceability of the bean harvest and all its transactions but will also be able to find the party responsible if the produce is tampered with. As will become evident, a simple mempool as implemented with bitcoin would not suffice to verify the harvest transaction, as also corrupt transactions must enter the chain, unless the harvest is destroyed.

The first step to the *verification process* is the creation of a so-called *beanblock*. A *beanblock* is a data package that contains: a unique ID; the amount of harvest linked to *beanblock* (e.g. 300 tons); a type (e.g. organic bean); a source (either a field from which the produce was harvested or a link to a previous *beanblock*); an owner (e.g. Farmer or Trader ID); a data hash that can be used to uniquely link an attached database and finally a *verification bit*.

When a set of *beanblocks* are linked in a chain, they can be used to trace a harvest of beans from a buyer over multiple traders and subcontractors, to a farmer. Through the use of the *verification bit*, the first *beanblock* in the chain to have its *verification bit* set to false can be identified as the corrupt member in the chain. Before such a *beanblock* can enter the chain though, it must be peer-reviewed and verified.

The *verification process* of a *beanblock* before it can enter the chain could either be done by an external group who all members must trust or, through the use of incentive design, can be done by members of the trading process. Our solution opted for latter. In order to ensure that self-regulation works, a separate token system based on reputation was developed that shall be explained in a later chapter on incentive design. The key concept is that any member who takes part in the *verification process* has the opportunity to gain reputation yet can also lose reputation by cheating or by falsely verifying a *beanblock*.

The first step of the *verification process* is for the owner of a *beanblock* to encrypt his *beanblock* with a public key. The *verification bit* of the *beanblock* needs to be handled separately and is not encrypted with the rest of the *beanblock*. This ensures that later in the *verification process*, it remains possible to change the verification bit. An asymmetric key

Figure 2: Verification Package

encryption scheme is used in order to ensure that every step of the process must be certified and cannot be tampered with by a third party. The encrypted *beanblock* is then inserted into a *verification package*. A *verification package* contains an encrypted and compressed *beanblock*; a private key with which to unlock the *beanblock* and view the data; a *verification list* that starts out empty; a *verification value* and finally a transaction certificate (see figure 2). As previously mentioned, the *beanblock* is only created once a participant, such as a farmer, has finished handling a bean harvest. The transaction certificate is used to ensure that a second party, such as a transportation subcontractor or trader has taken control of the harvest and the harvest cannot be tampered with anymore by the previous owner. The *verification value* that is assigned to a *verification package* is dependent on the initial reputation of the *beanblock* owner. Once a *verification package* has been generated it can be submitted to the *beanpool*. The *beanpool* is implemented in a similar distributed fashion as a mempool and serves the same purpose. Every member who wishes to validate a *verification package* from the *beanpool* can do so by randomly picking a *verification package* from the pool. A member can only extract a copy of the *beanblock* by adding his name to the *verification list*. This ensures that at all time, it is possible to check who has had access to the *verification package*, and how that person voted on the *verification package*. If the verifiers decide to vote in favour of the *beanblock* (i.e. that the handling of the harvest was done in a correct fashion), the *verification value* of the *verification package* is increased, else it is decreased. The amount by which the *verification value* is changed depends on the reputation of the verifier. Therefore, the higher the reputation of a verifier, the more he is trusted by the system. Once a vote has been given on a *verification package*, it can be submitted back into the *beanpool*. Internally the *beanpool* checks the *verification value* and uses a lower and upper threshold to decide if the *verification bit* of the *beanblock* should be set or not. Once either threshold has been passed, the *verification package* is taken out of the *beanpool* and the *beanblock* is added to the chain.

Once a *verification package* has been dissolved and the *beanblock* has been added to the chain, the *beanpool* system checks the *verification*

list and updates the reputation of the various members accordingly. There are different possibilities to do this that will not be elaborate at this point.

As a result of the constantly falling cost of IoT devices, they have become increasingly more affordable and many options for tracking shipments and harvests have arisen in the past years. Therefore IoT devices could shape the majority of sources for the verification data that is linked to a *beanblock*. As previously mentioned, every *beanblock* has a hash that links to an external database. This database does not need to be decentralized, as tampering with data after a *beanblock* has been submitted is impossible due to the attached hash that links a *beanblock* with a database and a corresponding data set. For more information on how this is done please refer to standard uses of hashing in cryptography (OnlineHashCrack, "Hashing in Blockchain explained", 2019). For the verification, a verifier can gain access to the IoT data by copying and decrypting the *beanblock* and then using the hash to find and read the IoT data. If the IoT data supports the claim of the *beanblock*, the verifier will naturally vote in favour of the *beanblock*, else he will vote against the *beanblock*.

Filtering through IoT Data to see if tractor movement, crop yields, pesticide use etc. seems legitimate for a *beanblock* claim (i.e. organic bean harvest of x tons over y number of days etc.), can be tedious. But there is no reason a verifier cannot use software to help process the IoT data and look for outliers in datasets. Thus a verification should not take too long and the *verification process* should remain scalable for a large number of *beanblock* submissions.

Incentive Design

The goal of our system and the implemented incentive design was to minimize any gains that an entity could potentially obtain by cheating the system. We first set out to identify where and how the current systems were being cheated and identified the following main sources:

- Farmers or traders buy non-organic beans, and mix them with an organic bean harvest in order to sell a larger amount of beans as organic.
- Trader repackages old or out of date goods as new.

Interest in this method of fraud increases dramatically during periods where yields are negatively impacted, for example by natural disasters.

Goal of our system introduced incentives

To counteract the motivation to cheat, our system demands that every entity that seeks to sell or buy, in our example soybeans, has to be registered in the system. Here each transaction is checked and verified, ensuring the produce is actually linked to an organic field. It would not be of interest to sell outside our system as the price per kg would be lower for organic goods that offer less traceability. The security measures of the system should help to quickly identify fraudulent data and entities that are attempting to cheat the system, and punish them accordingly. This should encourage all participants, regardless of their economic size, to buy and sell honestly.

Every entity is given a reputation value that is based on their past transactions and whether or not their goods have checked out to be as labelled (i.e. match the *beanblock* claim). The reputation can be logged and tracked via a token system. We thus refer to tokens that represent a reputation value, as *repbeans.*
In order for a transaction to take place both parties have to agree to and sign off on the *verification package* that is submitted to the *beanpool*. Each entity now takes part in verifying an at random assigned *verification package* and casts a vote whether or not the *verification package* checks out. The reputation value of the entity casting the vote is then added to the current verification value of the *verification package*. Once a *verification package* has been verified the reputation of each entity that casted a vote on the *verification package*, or was the original owner, is adjusted. In order to be able to manage the reputation of various participants in a decentralized fashion, it is required to deposit *repbeans* to the *beanpool*. Once a participant has completed analyzing its assigned *verification package* and enough votes have been cast for the packet to clear the *beanpool* and enter the blockchain, *repbeans* are paid out again. The amount paid out depends on the final vote on the *verification*

package and how this aligns with the participants vote. If the participant's vote is the same as that of the final vote, he will receive his original *repbeans* that he deposited and get a little added bonus, if his vote differs from the general consensus of the *verification package,* he will only receive a partial amount of his deposited *repbeans.*

For *verification packages* submitted by entities that have a low reputation the verification value required for the *verification package* to be entered into the Blockchain is increased. Therefore, farmers, traders, and participants in the system with a high reputation will be preferred as their transactions will go through faster. The various parameters such as the thresholds for the reputation values and the amount *repbean* bonus one receives have not been closely studied yet and require further research and tuning. It is also possible to have these parameters be dynamic. This way, when the *beanpool* is quite full, the reward for evaluating a *verification package* could increase and thus provide an incentive to verify packets and decrease the load on the overall system. The incentive design is motivated by common concepts in game theory. It only requires a few anonymous and honest participants for the system to work, as *verification package* are always dealt out in a random fashion, making collaboration impossible to cheat the system near impossible. Even if over 50% of participants desire to cheat the system, there is no guarantee that they will succeed and there is a high chance that they will instead do irreparable damage to their own reputation.

Hyperledger

Today's consumers increasingly want to know where the products they buy are coming from and under which circumstances they have been produced. At the same time, globalisation leads to complex supply chains, making it harder for companies to proof the quality of their goods. In addition, fraud and food safety scandals diminished the trust between customers and suppliers. (Trienekens *et al*., 2012; Wognum *et al.*, 2010). This issue can be tackled by using blockchain technology (Kshetri, 2017; Wüst & Gervais, 2018).

In our opinion, the best suited blockchain system is Hyperledger. Hyperledger is a platform to build blockchain systems for companies. In the following paragraphs, we will explain why public ledgers - such as Bitcoin or Ethereum - are not practical for our challenge and highlight the main features making permissioned blockchains - like Hyperledger Fabric or R3 Corda - the system of choice (Valenta *et al.*, 2017).

First, public blockchains are permissionless, meaning that anyone can join the network anonymously having the same rights and actions at hand and access to the same data. In our case surely not everyone in the world should be able to submit *beanblocks* for verification, but only certain members. Using Hyperledger the company itself can manage who joins the network and with what kind of rights. For our challenge we came up with the following roles: Suppliers, submitting the *beanblocks* for verification; Validators, verifying the *beanblocks*; and Visitors, people who want to check the supply chain of products they bought. These different groups also allow for adjusted access rights to potentially sensitive data and therefore being in compliance with regulations on privacy protection laws. Note though, that participants can belong to more than one group, thus allowing, for example farmers, to both submit *beanblocks* and also verify packets in the *beanpool.*

Next, current public ledgers typically are not designed for scalability or efficiency. This includes storage limitations as well as unsustainable consensus algorithms e.g. proof of work. Especially the proof of work leads to long transaction times at high costs. The transaction costs of public ledgers with money like tokens are heavily influenced by speculation. Due to the lack of tokens in the Hyperledger concept this uncertainty is non-existent. As discussed in more detail in the previous sections our *verification process* resembles a proof of stake system. (Davies, n.d.)

Another advantage of Hyperledger is its flexibility. For example, one can add CouchDB to Hyperledger Fabric which could be used to store the IoT data in an easy to query fashion. The company also has full control over the Blockchain and therefore can make changes to the design as soon as problem appear and therefore can more easily adapt the

system as needed. In addition, the community around Hyperledger counts several prestigious companies such that more functionalities and further development can be expected.

All in all, a Hyperledger system combines the intrinsic trust establishing of Blockchain systems with the needs of a company to have a fast, scalable and well managed system.

Proof of Concept

Software alone does not allow us to track the origins of a soybean harvest. This is where IoT devices become useful. We developed a concept using simple hardware, which can be utilized everywhere (e.g. GPS and timers in smartphones and smartwatches). Keeping scalability in mind, we implemented a simple framework of our solution and present it as a proof of concept. We used a Raspberry Pi coupled with a GPS device to measure location and harvesting time. All necessary data gathered we then uploaded into a MySQL database and there this data gets hashed and returned, which ensures the data cannot be manipulated. This data hash is stored in the *beanblock*, which will enter the Blockchain implemented with Hyperledger Composer, after being confirmed in the procedures described in the *verification process*.

We implemented our data aggregation with a Raspberry Pi, but to make the process less energy intensive, one can consider using an esp8266 or an Arduino instead.

Harvesting

It is most crucial to define where the beans were harvested. As such, we opted for a Raspberry Pi with a GPS module allowing us to see where and how long the harvesting machine was operating. This data will then get uploaded into the MySQL database and its hash returned. Having saved GPS location and approximate size of the beanfield on a separate static MySQL database we can then collate whether the farmer was harvesting on one of his fields and if the yield of the field matches preexisting information on the farmer.

The truck where the harvest is loaded onto is can also be equipped with a similar GPS module and a volumeter. Here also the data gets uploaded into the MySQL database and its hash returned. A *beanblock* will be created and sent into the *beanpool* for the final inspection once all necessary steps have been recorded and the harvest it passed on to a trader.

Trading

Each further trading step will look similar. Every movement of a harvest (e.g. from truck to warehouse) will be recorded and finally uploaded into a MySQL database and a new *beanblock* containing database hash values will be created and passed into the *verification process*.

The proof of concept we present can be initially tested with just a couple of farmers and traders and will hopefully establish trust in the system. The implementation of recording data, uploading the data into a database (in our example a MySQL database) and after verification uploading data into a Blockchain has already been implemented and seems to work in a robust fashion.

Outlook for the Future

While the *Bohnanza!* system was designed with scalability in mind, there exist a number of opportunities and considerations that should be taken into account. To start, there are two components of *Bohnanza!* that must be completed in order to implement out in the field. Beyond this, there exist other technological considerations which offer both challenges and opportunities for development. First, due to time constraints, not all components of the *Bohnanza!* system were implemented. With this in mind, an immediate opportunity would involve completing the configuration of the Raspberry Pi GPS module. More specifically, during the hackathon,

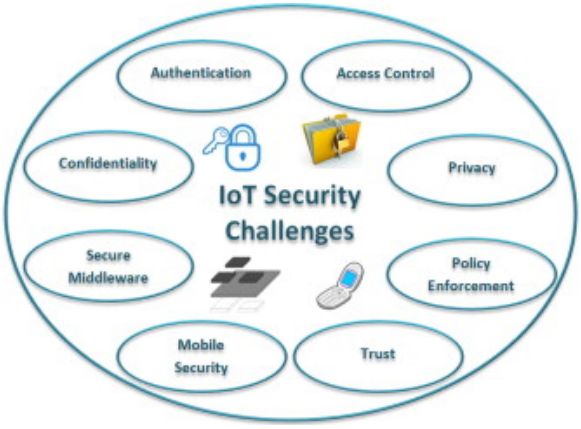

Figure 3: IoT Security Challenges (Source: Sicari et al., 2015)

connection to local public Wi-Fi failed, though this issue may be resolved through the creation of a static IP address on the device. Upon successful connection, real GPS data would be able to enter the MySQL database and linked with a *beanbloc*k.

However, one consideration is that the module will always require network connection in order to relay data from the field, so it is assumed that the soybean plantations will have the appropriate equipment and accessories to allow for this. In addition to resolving GPS connectivity issues, the *verification process* with the *beanpool* was not implemented due to the aforementioned time constraints.

Other Considerations

It is, of course, pertinent to realize that blockchain and IoT technologies are still in their infancy. As such, there remain many possibilities which have yet to be explored; however, challenges should be expected as well. For *Bohnanza!* specifically, the initial implementation costs may be high, creating a barrier to entry for smaller farms whose budgets are likely more constrained. To remedy this, these growers might consider partnering with larger farms to gain market access, thereby encouraging their own internal growth.

In a broader sense, one must also consider the general security of IoTs. In fact, the security of the *Bohnanza!* system relies tremendously on this factor. Security of IoTs has been a hot topic as of late, with a multitude of questions that researchers and industry professionals alike are attempting to address (Figure 3). Such points of vulnerability could allow for foreign entities too tamper with, or otherwise comprise, data integrity. IoTs communicating with public networks are especially susceptible to such malicious attacks (Khan & Salah, 2018). Additionally, IoTs can be fairly simplistic and as WIRED magazine wrote "they lack robust security solutions and encryption protocols" (2017) as a result. Current measures that can help protect against these security threats include the use of IoT gateways or integrated approaches at the company level that analyze each layer of data movement (*"IoT is coming...Here's what to do"*, 2017)

As a final note, one must reflect upon the IoT device itself (in this case, Raspberry Pi connected with a GPS module). Global navigation satellite system (GNSS) modules tend to consume a notable amount of power and also time in order to communicate with satellites. For remote farms, this can take around 20 minutes if the system initialized after a cold start in which the device was in a deep sleep (Williams, 2017). While a hot start would cut this time down significantly, it would also consume power accordingly. So power requirements are worth considering, though luckily recent advances are pushing towards improved batteries and less power intensive GNSS units.

In conclusion, there exist a number of opportunities both on the initial implementation side of the *Bohnanza!* system as well as on the development end. With scientific advances, including improvements in IoT security measures and battery technology, *Bohnanza!* will become a scalable control mechanism that will remain affordable. Furthermore, similar technological innovations will cause the costs of various system components to decrease allowing smaller farms access to the market without relying on existing partnerships. We thus hope to set up a system that will allow self-regulation at an affordable price, without external influences. We also see potential in adapting *Bohnanza!* to other trading situations, such as the oil and cotton trade and expanding the possibilities of collecting verification data with the help of drones and satellite surveillance systems. With new innovations in distributed ledger technology and more complex global trading networks, *Bohnanza!* is predicted to kick off and create a fairer, transparent economy that gives back control to the local farmers and traders.

References

Davies, A. (n.d). Pros and Cons of Hyperledger Fabric for Blockchain Networks. https://www.devteam.space/blog/pros-and-cons-of-hyperledger-fabric-for-blockchain-networks/

Demand for organic food grows strongly in Switzerland. (2018). Retrieved from https://www.swissinfo.ch/eng/business/agriculture_demand-for-organic-food-grows-strongly-in-switzerland/44088596

IoT is Coming Even if the Security Isn't Ready: Here's What to Do. (2017). Retrieved from https://www.wired.com/brandlab/2017/06/iot-is-coming-even-if-the-security-isnt-ready-heres-what-to-do/

Trienekens, J.H., Wognum, P.M., Beulens, A.J.M., van der Vorst, J.G.A.J. (2012). Transparency in complex dynamic food supply chains. Advanced Engineering Informatics, https://doi.org/10.1016/j.aei.2011.07.007.

Khan, M. A., & Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. Future Generation Computer Systems, 82, 395-411. doi:10.1016/j.future.2017.11.022

Krull, C. (2018). Uses for Soybeans. Retrieved from https://ussoy.org/uses-for-soybeans/

Kshetri, N. (2017). Blockchain's roles in meeting key supply chain management objectives,
International Journal of Information Management,
https://doi.org/10.1016/j.ijinfomgt.2017.12.005.
OnlineHashCrack. (n.d.). Hashing in Blockchain explained. Retrieved March 10, 2019, from https://www.onlinehashcrack.com/how-to-hashing-in-blockchain-explained.php

Sicari, S., Rizzardi, A., Grieco, L., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. Computer Networks, 76, 146-164. doi:10.1016/j.comnet.2014.11.008

Valenta, M., and Sandner, P. (2017). Comparison of Ethereum, Hyperledger Fabric and Corda. FSBC Working Paper June 2017. Frankfurt School Blockchain Center.

Whoriskey, P. (2017). The labels said 'organic.' But these massive imports of corn and soybeans weren't. Retrieved from https://www.washingtonpost.com/business/economy/the-labels-said-organic-but-these-massive-imports-of-corn-and-soybeans-werent/2017/05/12/6d165984-2b76-11e7-a616-d7c8a68c1a66_story.html?noredirect=on&utm_term=.dd57a2fa5122

Williams, M. (2017). Why does it take so long to get a GPS fix? Retrieved from http://ozzmaker.com/take-long-get-gps-fix/

Wognum, P.M., Bremmers,H., Trienekens, J.H. van der Vorst, J.G.A.J., Bloemhof, J.M. (2010). Systems for sustainability and transparency of food supply chains – Current status and challenges, Advanced Engineering Informatics, https://doi.org/10.1016/j.aei.2010.06.001.

Wüst, K. and Gervais, A., Do you Need a Blockchain?, 2018 Crypto Valley Conference on Blockchain Technology (CVCBT), doi: 10.1109/CVCBT.2018.00011

Zheng, Z., Xie, S., Dai, H., Chen, X. and Wang, H. (2017). "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," 2017 IEEE International Congress on Big Data (BigData Congress), doi: 10.1109/BigDataCongress.2017.85