Digital Transformation and Transparency in Wine Supply Chain Using OCR and DLT

Conference Paper · February 2021

DOI: 10.1109/IT51528.2021.9390117

Stevan Cakic
University of Donja Gorica

Aida Ismailisufi
University of Donja Gorica

Tomo Popovic
University of Donja Gorica

Srdjan Krco
DunavNET

Digital Transformation and Transparency in Wine Supply Chain Using OCR and DLT

Stevan Cakic, Aida Ismailisufi, Tomo Popovic. *Senior Member IEEE,* Srdjan Krco, Nenad Gligoric, Srdjan Kupresanin, Vesna Maras

Abstract—**This paper describes an effort to utilize IoT, OCR, and blockchain technology to create wine track and trace system evaluated in a real-life environment. The research is focused on digital transformation in traditional wine supply chain, using computer vision to read the existing serial numbers labeled on bottles, so as to uniquely identifying individual bottles of wine and track the item life-cycle. The system provides mobile app to allow end consumers to scan each wine bottle and learn more about that particular product instance, its origin, authenticity, rating and potentially other characteristics. Status of each bottle is updated every time it has been scanned. To ensure the transparency of the recorded data and information immutability, the blockchain technology is used to record all relevant information into the ledger, e.g. all status updates for each item. The developed service enables tracking of a bottle throughout the supply chain, from a producer to a consumer, where traceability story is still not finalized. The consumed serial numbers are kept in the ledger, to avoid counterfeit scenario when one "bottle" (i.e. one serial number) is being sold multiple times. The life-cycle of bottle and its status changes are recorded, and the authenticity ensured facilitating verifiable identity of the authorized handlers and cutting edge cryptography, thus providing additional trust in the solution and transparency to all stakeholders.**

***Index Terms*—Blockchain, computer vision, digital transformation, track and trace, optical character recognition.**

I. INTRODUCTION

DIGITAL transformation relies on the use of various digital technologies in order to change how businesses connect and create value for their customers [1,2]. This transformation is taking place in all areas of industry and the agri-food sector is not left out [3]. In this paper, we are focusing on food track and trace application based on Internet of Things (IoT), computer vision, and blockchain digital ledger technology (DLT). Food safety and food supply chain applications of these technologies should improve end-to-end visibility and tracking and tracing food products through their lifecycle [4]. More specifically, we are focusing on tracking and tracing each individual bottle of wine using existing serial numbers on wine labels. The fact that there would be no need to introduce any changes to the existing scenario (no need to change expansive machinery for labeling of simple barcodes) will increase the number of potential users. Nevertheless, we have provided additional ability of generating and printing QR codes that could be labeled instead of together with existing barcodes. This work is a continuation of a pilot project implemented using smart tags that combines QR codes with functional ink and centralized data storage to keep product status information [5]. The pilot project based on QR codes with functional ink was implemented in collaboration with company "13. jul Plantaže" that is one of the largest wine producers in South Eastern Europe, with annual production of 17 million bottled products. Most of their popular wine sorts are bottled with labels that contain a unique serial number. Therefore, we considered the use of computer vision and optical character recognition (OCR) on mobile devices, enhanced by machine learning to improve the reliability of OCR, in order to read these serial numbers of individual wine bottles. In addition, the blockchain is used to keep track of each individual bottle in a decentralized peer-to-peer (P2P) data storage that is secure, immutable, trustless and distributed over all participating computer nodes [6].

This research focuses on wine track and trace use cases such as food safety, counterfeit detection and prevention, product rating, consumer awareness and others. We are currently focused on brand protection and counterfeit prevention, but the system may further be expanded to include various added value functions.

*Research supported by the Ministry of Science of Montenegro, grant agreement No. 01-788/2.

Stevan Čakić is with the Faculty for Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (e-mail: stevan.cakic@udg.edu.me).

Aida Ismailisufi with the Faculty for Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (e-mail: aida.ismailisufi@udg.edu.me).

Tomo Popović is with the Faculty for Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (e-mail: tomo.popovic@udg.edu.me).

Srdjan Krčo is with DunavNET, Bul. Oslobođenja 127/13, 21000 Novi Sad, Serbia (e-mail: srdjan.krco@dunavnet.eu).

Nenad Gligorić is with DunavNET, Bul. Oslobođenja 127/13, 21000 Novi Sad, Serbia (e-mail: nenad.gligoric@dunavnet.eu).

Srđan Kuprešanin is with Blockstruct GmbH, Zwoelfergasse 10/8/3, 1150 Vienna, Austria (e-mail: srdjan@0bsnetwork.com).

Vesna Maraš is with 13. jul Plantaže, Put Radomira Jovanovića br 3, 81000 Podgorica, Montenegro (e-mail: vesnam@t-com.me).

II. PROJECT BACKGROUND

This study is based on our previous research effort that developed "smart" QR codes use case for brand protection and counterfeit prevention [5,7]. In this project, we consider wine track and trace solution that can be used for implementing various further use cases for food and beverages, but also any other physical goods in very similar implementations. Motivation for the use of OCR is the fact that wine bottles already have unique serial numbers and as for the blockchain, the technology is used to provide transparent data across the value chain. The solution can be used for wine authentication, but also for promotion of national wine brands and sorts, raising awareness about product origin and quality, fully digital commodity trading platform, etc. In this research, we focused on wines coming from the company “13. jul Plantaže“. When each wine bottle is produced, more precisely a batch of wine of one type, information about that batch is stored in the central database. Information about each individual bottle cannot be retrieved automatically, which makes it hard to manage several million bottles produced annually.

In addition to the information provided at the time wine was produced, we can continue tracing each individual bottle throughout its lifespan. Every time a bottle is being scanned, we can collect information about its location and status such as unsold, sold, opened, consumed and rated. Interaction with users would provide product track and trace functionality, which can be utilized both by consumers and winemakers. The approach for this system is illustrated in Fig. 1. Winemaker labels each wine bottle with unique serial numbers (as it is already the case) or labels with QR codes. Production information about each bottle such as serial number, wine type, vintage year, geolocation, and region of origin, physical characteristics of the soil during the wine growing season, weather data (temperature, quantity of rain, hours of sunshine), any agri-technical measures etc. are stored in the database. Since centrally managed databases can easily be manipulated, for this project we propose to use blockchain as a fully immutable infrastructure layer, so that all stakeholders can trust that the data they operate with is accurate and not subject to alterations by any one party.

Optionally, distributors and retailers could integrate with the main app, or deploy their own versions of the mobile app and provide additional functionality and/or information about products through the supply chain. Every time a bottle is scanned, we collect more information about its location and change of state if any (sold, opened, consumed). This information could be used later for development of various applications with benefits for both wine makers and consumers.

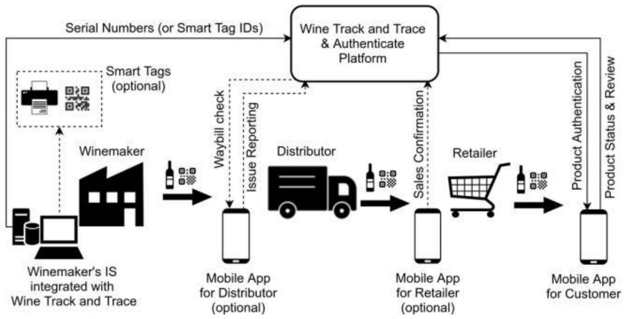


Fig. 1. Concept for the wine track and trace solution

In the brand protection and counterfeit use scenario, winemakers can use this wine track and trace platform to detect potential counterfeits, and users can use this app to get more in-depth information about the product, as well as to send feedback to winemakers [5]. In general, consumers would use an app on their mobile device to interact with a wine bottle that may be equipped with unique serial number, QR code, or NFC tag (Fig. 2).

The app will use mobile device sensors (camera, NFC reader, geolocation) to uniquely identify the product item and geolocate the bottle, while the rest of the information would be obtained from the user and/or pulled from the data stored. Finally, any change in the product state, i.e. from “in store” to “sold”, would be pushed back to the data store. In our experiment, we considered using a traditional database to keep information that is exchanged with the wine producer and blockchain based storage to keep information on product state and its changes throughout supply chain and each item‘s lifecycle.

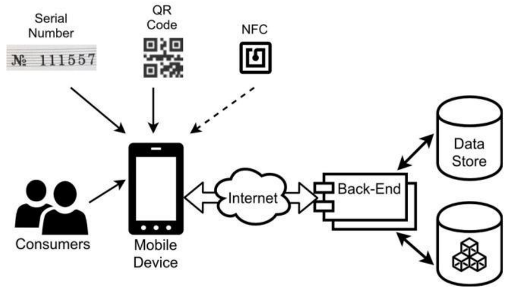


Fig. 2. Approach to the solution implementation with mobile app

III. Implementation and Discussion

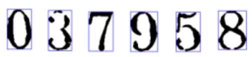
Main challenges with the implementation were the correct recognition of the serial numbers and blockchain data storage integration. Although various OCR tools for reading text/numbers are available and have been around for some time, it is not an easy task to read wine bottle serial numbers. There are several reasons for that, but most importantly different camera types and resolutions on various mobile devices, serial numbers are read from a label that is attached to a curved object (wine bottle), wildly varying lighting conditions and shadows in the real world, and in our case some serial number have horizontal lines across them [8]. As for the blockchain part, besides the learning curve, the challenge is the selection of platform and configuration and we want to investigate any potential performance issues for real world implementation of a project on this scale. The rest of the system was adopted from the previous project as described in [5].

A. Reading Serial Numbers

In order to read serial numbers from wine bottle labels we employed Tesseract OCR [8, 9]. In addition to Tesseract, we used jTEssBoxEditor tool in order to create a new trained data set, a new language in Tesseract terminology, in order to improve character detection for our application [10]. The default training for Tesseract OCR worked well, but were not good enough for this app [8].

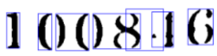
To create a new trained dataset that works with Tesseract, we needed to use jTessBoxEditor. Creating a new language requires both Tesseract and jTessBoxEditor installed, but also a new set of images that will be used for training process. For this process we manually collected around 150 images of wine labels and serial numbers using wines from the company “13. jul Plantaže”. The set contains several bottle labels including retakes of same labels. The quality of images varied in resolution, angle, shadows, wine type, etc. The jTessBoxEditor is run from terminal using Trainer option. While in Trainer option, we selected the folder where Tesseract is installed and folder where the images for training were stored. A large number of files is created in this process, but most importantly the new language will be stored in a subfolder called tessdata. It contains a newly generated file called language_name.traineddata, where language_name is the name we entered prior to the run. This file is later used by Tesseract when we want to use the result of this training to improve Tesseract OCR accuracy. However, our new language and Tesseract was not ready for use yet. Some text, here serial numbers, from the images was wrongly recognized. For each image, we can check what serial number was recognized by checking the files with .box extension.

An example where Tesseract and jTessBoxEditor successfully recognized a serial number is given in Fig. 3. The figure shows how boxes/rectangles are drawn around each character (a digit of the serial number), and provides details about each character taken from specific file (.box) generated for this image in the process of generating new language. It can be seen that each number is correctly fitted in a separate box, as shown in the table with box coordinates. However, there were situations where this process detected wrong characters and did not correctly placed the boxes around the serial number digits. An example of such situation is given in Fig. 4.



Char	X	Y	Width	Height
[illegible]	45	26	50	82
[illegible]	130	28	48	83
[illegible]	213	31	50	81
[illegible]	299	31	56	82
[illegible]	393	32	50	81
[illegible]	480	29	58	80

Fig. 3. An example of correctly recognized serial number



Char	X	Y	Width	Height
[illegible]	20	37	22	61
[illegible]	74	39	24	62
[illegible]	105	39	57	63
[illegible]	169	39	82	60
[illegible]	226	31	65	71
[illegible]	271	36	37	60
[illegible]	333	31	44	61

Fig. 4. An example when the serial number was not correctly detected

Because of some spacing between characters jTessBoxEditor missed proper creation of boxes/rectangles around characters. In situations like this, it was needed that we fix this by manually creating rectangles by hand in jTessBoxEditor around each character in order to help Tesseract to detect characters properly.

B. Mobile App for Consumers

This version of mobile app allows consumers to scan serial numbers from wine bottles in order to read more about wine and even get additional specifics related to this particular wine bottle (i.e. possible counterfeit). Technology that is used for application development is ReactNative [11]. The use scenario is illustrated by mobile app mockup design depicted in Fig. 5. It is assumed that the user is in front of the bottle that she or he wants to scan to learn more about the product.

When the user starts this app, he/she should click on button indicating the status of the particular instance of the product they are about to scan (Fig. 5). Then, the user selects the wine type. In our case, serial numbers are unique within a wine type, which means that two same serial numbers cannot occur once the wine type is pre-selected. The next step is taking a picture of the serial number. In left lower screen in Fig. 5, the blue rectangle guides the user to scan serial number for wine bottle. That image, only the part from blue rectangle, is passed to the server side of the system for processing.

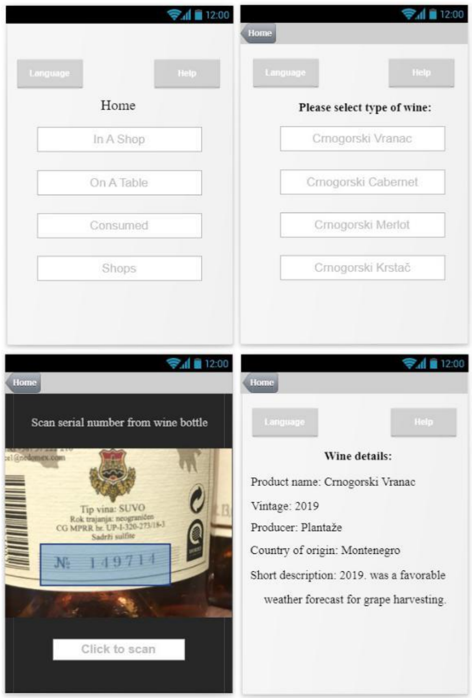


Fig. 5. Mobile app front-end design

Upon running the app, the user is presented with a screen that requires input about the current state of the product. This is how the app “kindly” requires user to act as a human sensor and provide us the current state of the bottle prior to the scan. The user selects the current state (i.e. “in a shop”) and continue to the next screen to select the type of the wine. This second step could be removed by using an automated label recognition, but this will be done as a part of our future work. After that, the user is prompted with a screen that uses built-in camera to scan the serial number. Finally, the user is provided with information on that particular bottle.

There are couple of possible outcome scenarios. First scenario, if serial number is not detected correctly, i.e. the server does not recognize the serial number and the user confirms that the number is not read correctly, we keep that image on the server together with the serial number of the device. The user is then offered to manually enter the serial number and, if the number is in the database, the user receives the information about that wine bottle. Second scenario is when the serial number is recognized correctly and the user is provided information about the bottle as shown in Fig. 3 earlier. Finally, there is a situation where the number is recognized correctly, but not found in the database. This last scenario may indicate the issue of authenticity.

C. Server-side of the System, Back-end

The back- end of this system is responsible for three main functions: a) image processing in order to correctly extract the serial number of the bottle; b) winemaker database used to store information about the wines, but also to maintain status of each individual bottle; c) heuristics used to detect possible issues with products, i.e. counterfeit detection; and d) blockchain data store here used for research purposes, but with possibility to later include other stakeholders and create new usage scenarios.

The image processing uses preprocessing to enhance the image prior OCR detection step. Mainly, this is done to clean the background, reduce the noise, and remove horizontal lines [8]. The OCR detection is implemented using Tesseract OCR improved with the new language trained using around 150 images of serial numbers obtained from wine bottles. The current rate of recognition is around 87%, but we have to take into account the fact that some of the images are of bad quality and taken without guidelines. Also, we need to expand the training set to include a larger number of images. Therefore, the system will be configured to keep unrecognized images and ask users to repeat the scan or enter the serial number manually. That way, a dataset needed for training of the OCR will be created.

The winemaker database contains information about each wine type, serial numbers, vintage year, product description, region of origin. It can also contain more specific information about each vintage and production batch. In addition, the database contains status of each individual bottle. This additional information allows creation of heuristics that can be used to detect issues (fake wine, reused bottles, “strange” locations, etc.) [5]. Finally, the interface to blockchain is introduced for research experimentation.

D. Blockchain Data Store

Blockchain technology was first introduced for use in financial sector and it represents data storage and processing infrastructure based on decentralized P2P network. Blockchain network is characterized by [12]:

- Decentralization – the data is spread over multiple participants, and the integrity of the data is verified and secured by many decentralized parties who are incentivized to “play by the rules” and disincentivized from cheating by various types of consensus mechanisms, rewards and punishments;
- Immutability - information integrity is created by immutable ledgers, blocks rely on each other and all stakeholders verify the integrity of data and transactions;
- Security – the elimination of single point of failure and usage of advanced cryptography based on fully self-sovereign private-public key infrastructure makes this architecture more secure;
- Efficiency - currently existing blockchain platforms have vastly improved performance compared to previous versions, their efficiency is reflected in terms of cost, settlement speed, and risk management, as well as dramatically reduced energy consumption;
- Transparency - transactions details are shared between all stakeholders involved in blockchain network.

With all these characteristics, blockchain is finding its use in various industries [13]. Several blockchain applications in agri-food sector have been identified [14,15].

Traditional databases represent storing data in one place, where security cannot be guaranteed, data can be lost and altered by any person having access to it. Traditional databases rely on CRUD paradigm (Create, Read, Update, Delete), while in blockchain networks we only have Create, and Read, with some newer blockchain platforms also supporting Update function, with a significant advantage over the traditional database model in that any Update also preserves the original state of the data, instead of overwriting it. Delete is not supported on a true blockchain platform, one can only mark a piece of data as “no longer valid/in use”, which simulates Delete, but again with the full history always available. The scenarios for using a traditional database are:

- Data needs to be under exclusive control of one party (centralized),
- Permissions on data accessibility can be changed,
- Data is of temporary nature and needs to be deleted,
- Data is not valuable enough to the business relationship, so storing it immutably on the blockchain would not result in meaningful benefits to the stakeholders.

On the other side, blockchain provides trusted data verification and there is no single party that can be designated as an administrator. The use of blockchain in supply chain has already shown promising results in multiple pilot and commercial implementations around the world [16]. It can be implemented as a system for storing information about products or storing records of transactions between partners in supply chain. It can be implemented as a private or public network, as well as be integrated with existing traditional ERP systems.

There are various blockchain platforms which can be used for storing data. In this research we considered Multichain [17,18] or 0bsNetwork [19]. The following code illustrates importing data into Multichain network:

```
publish stream1 key1 '{"json": {"bottle_id":"123456","prod_date":"2020-01-02", "city": "Podgorica", ”field”:”P1acPd3”}}'
```

End result from the transaction is hashed data written in the blockchain: *PNKtVb4Kdh0UILj5nD3Qq34rtfgwPVKW3r*. Nodes in the blockchain uses this hashed data to verify transactions and data.

The 0bsNetwork is a platform for business applications focused on powerful features, speed, scalability and light on resources. An example of data transaction in 0bsNetwork is shown in Fig. 6.

The idea in this setup is to use blockchain DLT as a transparent data storage where we will keep the data, i.e. bottle id, status update, and date/time stamp. All the other information about that bottle of wine will be kept in a standard database such as MySQL, using the Blockchain as a trustless intermediary. When the bottle is scanned, the Blockchain could confirm that the id corresponds to the data for the bottle with a certain serial number in our database, thus it is possible to trace the bottle, e.g. if the bottle was already purchased there is no way that the bottle can show up in market again. Such a scenario would indicate a possible fraud or counterfeit situation. Blockchain holds the tracking data for bottle of wine with its serial number and indicates when the bottle left manufacturing building and when it showed up in the market. In other words, the “ownership" of the bottle can be traced, and the history of usage. Such a setup would provide trust into the system and all the necessary information to all potential stakeholders, in this case consumers, winemakers, distributors, and retailers. Winemakers would stay responsible for providing additional details about that wine series, vintage, origin, and other information.

```
data = [{ 'type':’int’,
          'key': ‘bottle_id’,
         'value’:’123456’
        },
       { 'type':’date’,
         'key': ‘prof_date’,
         'value’: ‘2020-01-02’
       },
       { 'type':’string’,
         'key': ‘city’,
        'value’:’Podgorica’
       },
       { 'type':’string’,
         'key': ‘field’,
        'value’:’1P1acPd3’
       }]
```

Fig. 6. Code Snippet for Using 0bsNetwork Blockchain

Additionally, 0bsNetwork blockchain platform would enable us to store data generated by IoT sensors in the vineyards and the processing facilities on the blockchain itself, for added security.

0bs blockchain platform is interesting from the perspective of performance, but this system is commercial and comes with the licensing costs. On the other hand, Multichain platform may increase costs if there is a need for more server/node power and speed for mining and searching data through these nodes [20,21].

IV. Conclusion

This paper describes an effort to implement digital transformation in wine supply chain with the focus on product traceability. The computer vision is used to read the serial numbers for wine labels and blockchain is used to keep information on all status updates for each wine bottle, thus providing transparency to all possible stakeholders.

Tesseract OCR was selected for the solution, and the system is collecting additional images of serial numbers in order to create a training dataset that will be used to enhance the OCR results when reading serial number from wine bottles. The OCR feature is used by the mobile app that allows end consumers to scan each wine bottle and learn more about that particular product instance, its origin and authenticity. Blockchain technology is currently being tested as an additional data storage where all of the status changes are recorded.

Future research will include further functional and performance evaluation of blockchain network and exploring architecture of the system that will rely only on the blockchain based storage, as well as adding IoT-generated wine production data. We will also expand the solution to utilize QR codes and possibly NFC tags should the company decides to install them on each bottle.

Acknowledgment

The research is supported in full by the scientific research project “Digital transformation of agriculture production and food supply chain in Montenegro - DIPOL”, though the grant awarded by the Ministry of Science of Montenegro, grant agreement No. 01-788/2.

References

[1] D. L. Rogers, “The digital transformation playbook: Rethink your business for the digital age,” Columbia University Press, 2016.
[2] A. Maheshwari, “Digital Transformation, Building Intelligent Enterprises” , Wiley, 2019.
[3] N. M. Trendov, S. Varas, M. Zeng, “Digital Technologies in Agriculture and Rural Areas: Status Report”, Food and Agriculture Organization of the United Nations, Rome, 2019
[4] H. Sundmaeker at al., “Internet of Food and Farm 2020”, Chapter in “Digitising the Industry,” River Publishers, ISBN 9788793379817, pp 129 - 150., 2016
[5] S. Sandi at al., “Smart tags for brand protection and anti-counterfeiting in wine industry,” 23rd IEEE Information Technology Conference IT 2018, Žabljak, Montenegro, pp 1-5, 2018.
[6] M. Swan, “Blockchain, Blueprint for a New Economy,” O’Reilly, 2015.
[7] N. Gligoric et al., “SmartTags: IoT Product Passport for Circular Economy Based on Printed Sensors and Unique Item-Level Identifiers”, Sensors 19(3): 586, pp 1-26, 2019.
[8] S. Cakic et al., “The Use of Tesseract OCR Number Recognition for Food Tracking and Tracing”, 24rd IEEE Information Technology Conference IT 2020, Žabljak, Montenegro, pp 1-4, 2020.
[9] Tesseract Github repository with documentation. Accessed Mar 2020. [Online]. Available: https://github.com/tesseract-ocr/tesseract
[10] K. El Gajoui, F. Ataa Allah, M. Oumsis, “Training Tesseract Tool for Amazigh OCR”, in 15th International Conference on Applied Computer Science (ACS '15), 2015.
[11] Eisenman, Bonnie. Learning React Native: Building Native Mobile Apps with JavaScript. O'Reilly, 2018.
[12] F. Knirsch, A. Unteweger, D. Engel, “Implementing a blockchain from scratch: why, how, and what we learned”, EURASIP Journal on Information Security, 2019:2, pp 1-14, 2019.
[13] L. Pawczuk, R. Massey, J. Holdowsky, “Deloitte’s 2019 Global Blockchain Survey: Blockchain gets down to business”, Delloite, pp 1-52, 2019.
[14] M. Tripoli, J. Schmidhuber, “Emerging Opportunities for the Application of Blockchain in the Agri-food Industry,” FAO and ICTSD, 2018.
[15] L. Ge et al., “Blockchain for Agriculture and Food; Findings from the pilot study.” Wageningen, Wageningen Economic Research, Report 2017-112, pp 1-34, 2017.
[16] F. M. Benčić, P. Skočir, I. P. Žarko, "DL-Tags: DLT and Smart Tags for Decentralized, Privacy-Preserving, and Verifiable Supply Chain Management," in IEEE Access, vol. 7, pp. 46198-46209, 2019.
[17] A. Ismailisufi et al., “A Private Blockchain Implementation Using Multichain Open Source Platform”, 24rd IEEE Information Technology Conference IT 2020, Žabljak, Montenegro, pp 1-4, 2020.
[18] Multichain Documentation. Accessed on Mar 15, 2020. [Online]. Available: https://www.multichain.com
[19] 0bsnetwork Documentation. Accessed on Mar 15, 2020. [Online]. Available: https://www.0bsnetwork.com/
[20] K. Biswas et al., “Blockchain based wine supply chain traceabillity system”, Future Technologies Conference, Vancouver, BC, Canada, 2017.
[21] W. Zheng, et al., “NutsBaSS: A blockchain as a service platform”, in IEEE Access, vol. 7, pp. 134422-134433, 2019.