



Paperchain: Distributed Way of Publishing Paper on the Blockchain

Vidhyasagar Babalsure¹; Alnaz Inamdar²; Aadesh Mirajkar³; Aman Roy⁴; Ms. Poonam Gupta⁵

¹Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Pune, India

²Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Pune, India

³Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Pune, India

⁴Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Pune, India

⁵Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Pune, India

¹ sagarbabalsure@gmail.com; ² alnazinamdar22@gmail.com; ³ mirajkaraadesh@gmail.com;

⁴ royaman8757@gmail.com; ⁵ poonam.gupta@raisoni.net

10.47760/ijcsmc.2021.v10i04.007

Abstract— *Decentralization is the process by which the activities are distributed in such a way that there is no central, authoritative location or any central server where all records are stored. A blockchain is one such decentralized, distributed, a public network comprising of records called blocks which keeps track of transactions across many devices so that any involved block cannot be changed or tampered, without the alteration of all blocks.[1] Paperchain is a platform that integrates both these concepts. Paperchain is a decentralized platform that leverages the power of blockchain technology in empowering authors to take full ownership of their work and finances. Our platform is built on a transparent Peer-to-Peer network powered by programmable smart contracts to enable fair remuneration for all research work and journals. The long-term vision is to develop an open ecosystem where every geek is allowed to read and publish their research work on top of the Paperchain platform.*

Keywords – *Decentralization, blockchain, smart contracts*

I. INTRODUCTION

Just as the internet provides a ton of information, blockchain is a technology that enables the digital exchange of assets. Any asset can be converted into tokens.

[1]Blockchain is immutable and because of its append-only function, and as every transaction can be seen openly, it provides complete transparency for all users in a system. As a result of that, an ecosystem gets created that does not need a trusted authority because tampering with data or any record in the network is very difficult. The decentralization allows people to construct their open blockchain network.

[4] Basically there are three primary types of blockchains. First being public blockchains like Bitcoin and Ethereum, are designed to be fully decentralized, with no person has control over which transactions are tracked in the blockchain network or how they are processed. Second is Private blockchains also known as permissioned blockchains like Hyperledger in which participants need a permit to enter the ecosystem. Also, transactions are private and are only available to system participants that have been permitted to enter the blockchain network.

The most important point to keep in mind is that private blockchains are more centralized than public blockchains. Lastly followed by Hybrid blockchains. Hybrid here means that it combines the advantages or merits of a permissioned and private blockchain with the security and transparency benefits of a public blockchain. That gives corporations and startups significant flexibility to choose what data they want to keep public and transparent and what data they want to keep private. Paperchain makes use of Public blockchains as it is highly censorship-resistant since anyone is open to join the network, irrespective of location, nationality, etc. This makes it extremely tough for authorities to take them down. Another reason is, public blockchains have a token associated with them that is typically designed to incentivize and reward participants (authors in our case) in the blockchain ecosystem. More precisely, Paperchain uses [5] Ethereum blockchain which stores the work of the author where we will use the concept of Ethereum Virtual Machine (EVM) which can be coded using a programming language called solidity.

[1] Application programming interfaces (APIs) plays a key role to connect off-chain or external hardware and software with the ecosystem. [9] It enables connection as well as the transfer of data between the systems in such a way, external web servers can combine the functionalities of an existing blockchain network for specific use cases. So, an API is an important feature of a blockchain that one should always provide to uplift the blockchain's potential.

Paperchain will introduce new features that never existed before by making use of blockchain technology. It completely removes the presence of a central authority that generally can be hacked by taking central control over the server of any business organisation. Blockchain term was first coined in 2008 and implemented successfully as the infrastructure of Bitcoin in the year 2009 by an unknown person or an anonymous group called Satoshi Nakamoto. [1] Blockchain basically is distributed ledger or a huge chunk of database where all the records of transactions are kept track of related to all the participants present in the blockchain ecosystem or network. The Blockchain-based technology Purely operates on a peer-to-peer network, where a centralized third party is not required for handling the transactions. [2]

Precisely, Paperchain will accept the paper when the author will upload it and undergoes some checks such as plagiarism check to make sure the content of the research paper is not copied from any other resources and is fully authentic work. Another check it will go through is citation error checking if any. And lastly, checks for the standard format to see if everything is in proper format like all heading constraints are the same, if all paragraphs in the paper are uniformly written using the same font size, colour, etc.

This concept which paperchain has used is somewhat resonant with [7] bitcoin where miners will try to solve the (PoW) proof-of-work algorithm within 600 seconds using high-end computers and electronic devices to make a transaction successful. In a similar manner, Paperchain will let the author decide whether to keep the asset paid or free of cost before finally publishing the research paper. If someone wishes to read or download the paid-paper in the paperchain then that person has to make a transaction that is to purchase the research paper by making use of tokens or [8] Initial Coin Offering (ICO). These tokens are papercoins in our system. Papercoin is like a currency (token) that is provided to both publishers and reviewers as a bonus. All the transaction-related process will be performed with papercoin only as these ICOs are only valid in paperchain network. Initially, when a user has to exchange their currency with papercoin they can use that currency to buy papercoin from anyone in paperchain network who is having it. Thus this ecosystem makes a great deal with all the researchers and reviewers around the globe because of the papercoin method that we have introduced.

However, with daily advancements in technology, the way we operate and use it also goes on varying. As of now, scalability is the issue that Paperchain might face in future as more users join the network. The solution to this can be making these records [9] off-blockchain and use blockchain to keep and access information. This project still needs significant development before it releases to the real world.

II. LITERATURE SURVEY

The politics of publication [10] article published on nature states the issues of paper publication in 2003. According to an article, in 2003 there was huge pressure in biomedical science to publish the research paper in a top journal. One student named Jim Jargon published a paper on 'Cell Paper' which illustrates that publishing a paper on a famous journal is more important than the message want to give scientist through the paper.

According to Jim's paper if you are publishing a paper in a top journal then you are a good scientist or researcher. It does not matter how good your idea.

And all this was happening because those who gave money and designations now they evaluate scientist based on a number of a paper published, journals impact factor. Besides, scientists should be evaluated based on the quality of work they have done.

Understanding modern banking ledgers through blockchain technologies: future of transaction processing and smart contracts on the internet of money[11] In more recent times, blockchain applications have emerged that go far beyond their first application domains in virtual currencies, for instance, they are now important in areas such as domain registration, crowdfunding, prediction markets and even gambling. Second-generation blockchain technologies enable the execution of simple transactions and the carrying out of computation on a network, where e.g. payments become conditional on the state of some internal or external variables (much the same way as financial derivatives have a payout that is a function of an underlying financial instrument). This is the basis for ‘smart contract’[12] technologies, which we shall see can be important building blocks for these new application areas. As a consequence of these second-generation technologies, a number of developments in this field have begun to appear which include third party data ledgers (Devanbu *et al.* [2001]), e-contracts/smart contracts and virtual contracts (Buterin [2014b]; Kosba *et al.* [2015]; Swan [2015]), e-assets or remote asset title transfers (Halevi *et al.* [2011]) and further applications, discussed in Czepluch *et al.* [2015]. It has then described a number of features that are vital from a financial application perspective, including permissions, data integrity, data security and data authenticity as well as important regulatory requirements relating to account provisioning for financial asset reporting, and the blockchain aspects that can help adhere to these

Blockchain, sometimes referred to as Distributed Ledger Technology (DLT), makes the history of any digital asset unalterable and transparent through the use of decentralization and cryptographic hashing. For more details[13]

Blockchain-Based E-Voting System[14] As we know the current system of voting is too faulty and controversial. After every election, we heard that there were some problems with the EVM[15] machines. And due to that votes are credited into the wrong candidate’s account. To solve that problem this paper gives a model to build an e-voting system using the full power of smart contract which secure voter’s identities. Each election process is instantiated using some set of rules called smart contracts. And these smart contracts are defined separately for each district. The steps involved in this process are following:

1. Election creation: Elections ballots are created districts-wise using smart contracts.
2. Voter registration: By using an election card given by the government voters has to register on this ballot.
3. Vote transaction: When election these votes are transferred to district level nodes and from there, votes are appended to the blockchain database.
4. Verify vote: After data is appended to blockchain voters receive transaction ID for verifying vote.
5. Tallying result: Each ballot by using smart contract counts the number of votes and the final result of a smart contract is published.

“Towards a Decentralized Process for Scientific Publication and Peer Review using Blockchain and IPFS”[16], According to this paper, the current process of publication has some problem quality, performance, cost, and accuracy. The proposed system of this paper tries to slide the infrastructure from publishers to the scientific community.

It gives Three essential features:

1. The peer reviewers are selected and recognized based on the system where review reports can be rated.
2. The IPFS[17] Peer-to-Peer network is used to distribute scientific papers by giving open access infrastructure.
3. It provides a transparent and decentralized platform using blockchain technology for the peer-review process.

And the main intention of using this decentralized platform is to reduce the control of publishers on their centralized system.

PubChain: A Decentralized Open-Access Publication Platform with Participants Incentivized by Blockchain Technology[18]. The main motive of this paper is to make papers free for reading to all communities by breaking the pay system of the current publishing platform. Instead of storing all papers in one place like servers, it uses blockchain technology to store papers globally so they are available to all peers in the network. For storing papers it uses a decentralized and distributed file system named IPFS. In order to store data, architecture consists of four layers named blockchain layer, virtual machine layer, routing layer, and storage layer. This decentralized system provides hassle-free service without failure at a single point in time. By making cryptocurrency like bitcoin[19] it provides an incentive to all participants to make the system more trustworthy.

III. PROPOSED SYSTEM

The working and the implementation of the whole project is quite complicated and made up of many tiny parts. All the tiny parts assemble together to make a very interesting tool which is decentralized as well as distributed, i.e. there is no authority of a single unit. Every node works in the same direction to achieve a singular goal and maintaining the identical version of the database across the network with the use of Blockchain.

The major challenge while implementing this project is that the person has to keep in mind that the data is not as tiny as a transaction which can be kept easily inside the Blockchain in few tiny bytes, whereas in the case of paperchain, the data is a scientific paper which can take few megabytes which may seem very tiny if you see it as an individual but in long run that data when accumulated with few thousands unit can seem very large and even can chock up the whole network.

Apart from that, there are few more challenges which should be tackled in the initial stage. Pornographic content, plagiarized document, self-plagiarism to name a few. We are going to deal with all of them one by one and implement the protocols wherever it is required to check for the unwanted traffic and deal with it on each node basis. Solving this problem on a centralized theme is very easy but it is not worth it compared to decentralized setting. Looking at the picture below, it can be easily visualized that the traditional methodology is quite clean and very easy to manage but in Blockchain it is very tough to do so.

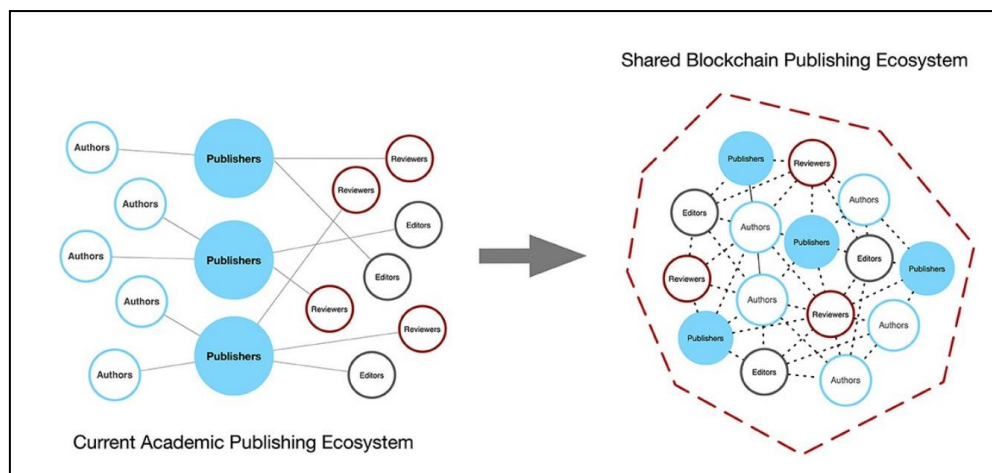


Fig. 1 Traditional vs Blockchain mode of paper publishing

I. Pornographic content detection.

Anything which is build for a noble cause can be used in a multiple of unwanted ways in which pornographic content peddling can surpass every other ways of illegal use. As we know the whole system is decentralized and no one can take decision on there own. So for that, we need some ground rules on which everyone agrees before.

Solving this issue is a two phase process. In the very first phase, first we need to extract the whole image from the document. The image can be of multiple types, such as – JPG, PNG, GIF, SVG, etc. We need to take care of all of them and extract them accordingly. After the first phase is done, we need to send image individually to the second phase to check if the image is a good image or a bad one. Thinking of both the phases as a separate entity is not correct. They both are correlated and works on the same go.

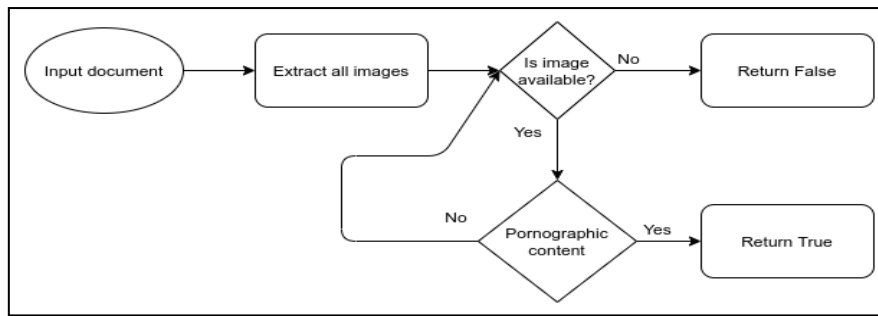


Fig. II Pornographic detection – Phase I

As seen in the figure above, the process is quite simple in phase I. At the very start of the process, we get a document which should be in PDF format. It is made compulsory to do so to maintain a uniformity. Processing different type is very hectic and also according to industry standard this decision is taken place. In that document we search all the images that is present. In the end, every image has to go through pornographic content detection which is basically phase II. If all the images passes the test successfully then there is no issue but if any image fails the test, the flag will be raised.

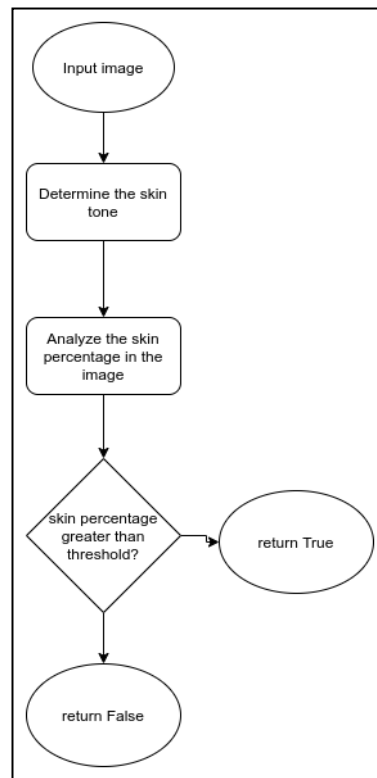


Fig III Pornographic detection – Phase II

In phase II, the image goes into several more steps to be checked accordingly. Every person’s skin color can be different, so first thing that needs to be done is to get to know the skin color of the person. After getting to know that, the next step is to know in the given image how much of skin percent is visible. This data can play a keyrole in getting to know the correct conclusion. At last, just check if the skin percentage is crossing the threshold which is per-decided. Following these steps can lead us to the final result.

II. Plagiarism Check

For checking the plagiarism, there are many tools and API available in the market. In our case scenario, offline mode of checking that is a necessity. Our implementation uses a tool called difflib library. using SHA1 hashlib function and SequenceMatcher function it gives the how much percent content is same or copied.

III. Final implementation

The final implementation uses the all the above implementations as subpart to work as a whole. The application works in two modes. First mode is about submitting the paper and the second mode focuses on how to access any of the document which has been pushed to the blockchain earlier.

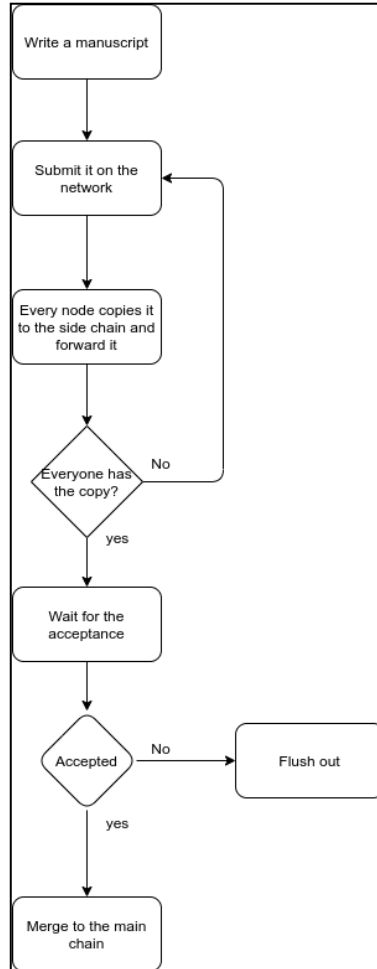


Fig IV Paper Submission

THE SUBMISSION OF THE PAPER IS DONE TO THE SYSTEM IN THE FORM OF FLOODING. THE PAPER GOES TO EVERY NODE AND GET APPENDED TO THE SIDE CHAIN. IT IS NOT MERGED TO THE MAIN CHAIN BECAUSE THE BLOCKCHAIN IS APPEND ONLY DATABASE. WE CAN'T REMOVE IT AFTERWARDS. ONLY AFTER GETTING THE ACCEPTANCE, IT IS MERGED TO THE MAIN CHAIN.

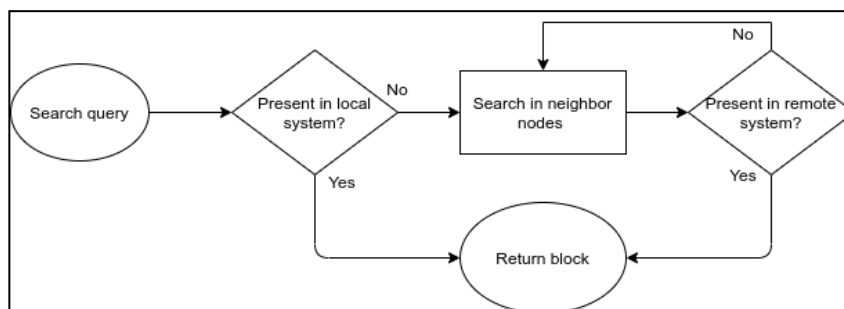


Fig V Paper Search

In Fig 05, The search is quite simple. As it is known already, everyone has a copy of the whole Blockchain on every node. In most cases, the paper will be found at the local system itself. The only exception is when some new node has been added to the network but your chain is not updated. In that case, the system tries to pull it from their neighbors and find the required paper.

IV. RESULT & DISCUSSION

Thus, by implementing the proposed Paperchain system, we have achieved an upper hand over the old broken system. Due to this, the research paper will get published in an easy and stipulated amount of time. Paperchain will also keep an eye on the irrelevant content and copied content if it is present in the research paper, and based on which it will decide whether to discard it or not. Paperchain is also responsible for providing equal benefits to the authors and the reviewers for their work. These benefits are rewards that are made available in terms of papercoins. The circulation of papercoins and the research paper in the system based on the demand and need highlights the functionalities of Paperchain and thus makes it more flexible and efficient. Table 01 represents the complete analysis of various attributes and functionalities in both the system. Paperchain thus results in bridging the gap of low factors of the former system with the help of blockchain technology.

Attributes	Factors	Architecture	
		Paperchain	Traditional methodology
Reliance	Accountability	High	Low
	Immutability	High	Medium
	P2P transaction	High	Low
Content	Data transparency	High	Low
	Detection of foul content	High	Medium
	Redundancy	Low	Medium
Performance	Quality retention	High	High
	Speed of process	High	Low
	Maintenance cost	Medium	High

Table 01: Analysis of attributes and factors of Paperchain versus traditional methodology.

1. As a Blockchain-based application or any system in which Blockchain serves as the key is responsible for extreme power consumption, this Paperchain system will also face some criticism. But when we put light on the previous methodologies we come to know that big names like IEEE, Springer, etc are multi-equipped and facilitated institutions thus, the power consumption is eventually high there also. So the disadvantage of power consumption in paperchain remains unsolved but, the decentralized mechanism breaks the tie on the centralized mechanism in the former system. Hence, we will notice a gradual transition in the intellectual mob who will prefer our Paperchain system.

2. IEEE and other reputed journals are known for maintaining the quality of a research paper, as their paper undergoes a lengthy and well-inspected process to get published. Considering this quality factor, our Paperchain will take a little time to gain attention but as soon as it does, Paperchain will experience an increase in the slope of the people acceptance graph. Paperchain also provides a thorough inspection process which is in the form of three checks, that ultimately add, to the quality maintenance of the paper.

V. CONCLUSION

The decentralized environment of the Paperchain will prove out to be the most promising part of all the features in the system. Paperchain will speed up the process of publishing a research paper therefore, it will increase the frequency of papers getting published every year. The security, privacy, and reliability of the Paperchain system outweigh all the hurdles and disadvantages faced by the research community. Our Paperchain system is composed of all the functionalities that were, required by the broken former system. On the other hand, the emergence of paper currency known as papercoins and its circulation in the system will turn out to be a more valuable asset in the paper publishing world. Thus, after evaluating the effectiveness of the proposed system, it is proved that Paperchain will be the most efficient publishing platform.

REFERENCES

- [1]. Leible S, Schlager S, Schubotz M and Gipp B(2019) A Review on Blockchain Technology and Blockchain Projects Fostering Open Science. *Front. Blockchain* 2:16. doi:10.3389/fbloc.2019.00016
- [2]. Zeeshan, J., Third, A., Bachler, M., and Domingue, J. (2018). Peer-reviews on the blockchain. In 1893 1st Workshop on Reframing Research (RefResh) (Cologne, Germany).
- [3]. Casino, F., Dasaklis, T. K., and Patsakis, C. (2018). A systematic literature review of blockchain-based 1419 applications: Current status, classification and open issues. *Telematics and Informatics* 36, 55–81. 1420 doi:10.1016/j.tele.2018.11.006
- [4]. Zheng, Z., Xie, S., Dai, H.-N., Chen, X., and Wang, H. (2017). An Overview of Blockchain Technology: 1898 Architecture, Consensus, and Future Trends. In *Proceedings of IEEE 6th International Congress on Big 1899 Data*. doi:10.1109/BigDataCongress.2017.85
- [5]. Buterin, V. (2014). A Next-Generation Smart Contract and Decentralized Application Platform. 1399 Available online at: http://blockchainlab.com/pdf/Ethereum_white_paper-a_1400_next_generation_smart_contract_and_decentralized_application_1401_platform-vitalik-buterin.pdf
- [6]. Kakavand, H., Kost De Sevres, N., and Chilton, B. (2017). The Blockchain Revolution: An Analysis 1594 of Regulation and Technology Related to Distributed Ledger Technologies. *Social Science Research 1595 Network* (SSRN). doi:10.2139/ssrn.2849251
- [7]. Gervais, A., Karame, G. O., Wust, K., Glykantzis, V., Ritzdorf, H., and Capkun, S. (2016). On 1528 the Security and Performance of Proof of Work Blockchains. In *Proceedings of the 2016 ACM 1529 SIGSAC Conference on Computer and Communications Security (CCS '16)* (Vienna, Austria), 3–16. 1530 doi:10.1145/2976749.2978341
- [8]. Li, J. and Mann, W. (2018). Initial Coin Offering and Platform Building. *Social Science Research Network 1620* (SSRN). doi:10.2139/ssrn.3088726
- [9]. Hepp, T., Sharinghousen, M., Ehret, P., Schoenhals, A., & Gipp, B. (2018). On-chain vs. off-chain storage for supply- and blockchain integration. *It - Information Technology*, 60(5-6), 283–291. doi:10.1515/itit-2018-0019
- [10]. Liang, X., Shetty, S., Toshi, D., Kamhoua, C., Kwiat, K., and Njilla, L. (2017). ProvChain: A Blockchain 1622 Based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability. In 1623 *Proceedings - 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing 1624 (CCGRID)* (Madrid, Spain: IEEE), 468–477. doi:10.1109/CCGRID.2017.8
- [11]. 20 march 2003 Lawrence, P. The politics of publication. *Nature* 422, 259–261 (2003). <https://doi.org/10.1038/422259a>
- [12]. G.W. Peters, E. Panayi Understanding modern banking ledgers through blockchain technologies: future of transaction processing and smart contracts on the internet of money *Banking Beyond Banks and Money*, 2016 ISBN : 978-3-319-42446-0 *New Economic Windows* (2016), pp. 239-278
- [13]. Szabo, N., 1994. Smart contracts.
- [14]. Nofer, M., Gomber, P., Hinz, O. et al. Blockchain. *Bus Inf Syst Eng* 59, 183–187 (2017). <https://doi.org/10.1007/s12599-017-0467-3> Nofer, M., Gomber, P., Hinz, O. et al. Blockchain. *Bus Inf Syst Eng* 59, 183–187 (2017). <https://doi.org/10.1007/s12599-017-0467-3>
- [15]. E-Voting 2018 Friðrik Þ. Hjálmarsson, Gunnlaugur K. Hreiðarsson “Blockchain-Based E-Voting System”, School of Computer Science Reykjavik University, Iceland {fridrik14, gunnlaugur15}@ru.is
- [16]. D. A. Kumar and T. U. S. Begum, "Electronic voting machine — A review," *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)*, Salem, India, 2012, pp. 41-48, doi: 10.1109/ICPRIME.2012.6208285.
- [17]. Tenorio-Fornés, Antonio, Jacynycz, Viktor, Llop-Vila, David, Sánchez-Ruiz, Antonio, Hassan, Samer, “Towards a Decentralized Process for Scientific Publication and Peer Review using Blockchain and IPFS”, 2019 *Attribution-NonCommercial-NoDerivatives 4.0 International* doi:10.24251/HICSS.2019.560, url:<http://hdl.handle.net/10125/59901>
- [18]. Q. Zheng, Y. Li, P. Chen and X. Dong, "An Innovative IPFS-Based Storage Model for Blockchain," 2018 *IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Santiago, Chile, 2018, pp. 704-708, doi: 10.1109/WI.2018.000-8.
- [19]. T. Wang, S. Chang Liew and S. Zhang, "PubChain: A Decentralized Open-Access Publication Platform with Participants Incentivized by Blockchain Technology," 2020 *International Symposium on Networks, Computers and Communications (ISNCC)*, Montreal, QC, Canada, 2020, pp. 1-8, doi: 10.1109/ISNCC49221.2020.9297213.
- [20]. S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008.