A Blockchain Framework in Compliance with Data Protection Law to Manage and Integrate Human Knowledge by Fuzzy Cognitive Maps: Small Business Loans

Swati Sachan
Management School
University of Liverpool
Liverpool, United Kingdom
swati.sachan@liverpool.ac.uk

Dale S. Fickett
Robins School of Business
University of Richmond
Richmond, United States
dfickett@richmond.edu

Nan Ei Ei Kyaw
Management School
University of Liverpool
Liverpool, United Kingdom
N.K.Kyaw@liverpool.ac.uk

Rituparna Shome Purkayastha
Management School
University of Liverpool
Liverpool, United Kingdom
r.shome@liverpool.ac.uk

Renimol S
Department Of Computer Science
University of Liverpool
Liverpool, United Kingdom
sgr.sibyr@liverpool.ac.uk

***Abstract*— New firms have minimum accessibility to data, unlike large data repositories of businesses operating for longer periods. They rely on the knowledge exchange and perceptions of domain experts outside the organisation to enhance the capability of the decision support systems. This paper presents a framework on blockchain technology to develop a secure knowledge management service which addresses the challenges imposed by European and US data protection laws and high on-chain computational costs. It presents a hybrid off-chain and on-chain smart contract computation methodology for the scalable application of blockchain technology. It requires low on-chain storage cost and low cost of frequent data imports for off-chain AI computations. The framework demonstrates a real-world application of knowledge management by blockchain to integrate the judgment of multiple financial experts to design loan eligibility policies.**

Keywords*— *blockchain, artificial intelligence, knowledge-base, loans, small business, finance

I. Introduction

A. Integrate Human Expertise by Blockchain

The advancement of artificial intelligence (AI) based decision support systems (DSS) has led to a growth in recognition of the importance of integrating human experts and data patterns to enhance automated decision-making [1]. To achieve this, organisations have implemented an augmented workflow environment that enables the exchange of judgements and perceptions among domain experts within or outside the organisation. Knowledge management is crucial in transforming experts' implicit knowledge into explicit and accessible formats [2]. This is particularly beneficial for new firms with limited access to data, as they often rely on their team members, advisors, and external stakeholders for knowledge sharing. Collaboration among multiple organisations can accelerate knowledge transfer and diversification, but a shared incentive for data-sharing commitment by experts remains a challenge.

Participatory modelling (PM) is a crucial step towards the development of DSS as it involves the input of domain experts. PM methods such as concept mapping [3], causal loop diagrams [4], Bayesian networks [5], and fuzzy cognitive mapping [6] enable the transfer of unstructured knowledge into plausible beliefs stored in a structured knowledge base (KB) [7][8]. However, knowledge elicitation from experts is time-consuming, and creating a comprehensive multi-stakeholder KB can be challenging due to competition between firms, insecure personal identity and access control concerns. A KB could be logic, semantic network, rules, and frame representation [9].

Blockchain technology facilitates secure data sharing and auditing among multiple organisations. Its decentralised architecture enables the secure submission, review, and synchronisation of data by experts from multiple organisations without third-party involvement. Each organisation can connect multiple devices to the network, called nodes or peers, responsible for verifying and adding data blocks to the chain. The immutability of blockchain technology enables automated evaluation and auditing of transactions, registering accurate and tamper-proof data access records critical for regulatory compliance and legal purposes. Yang et al. [10] proposed a decision model to identify a suitable decentralised and transparent platform for multiple experts for knowledge exchange and data accessibility management. Schniederjans et al. [11] presented the future blockchain trends in knowledge management to enhance supply chain digitisation. A review of the implications of blockchain technology in knowledge management is presented by Frozza et al. [12].

B. Blockchain Immutability and Data Protection Laws

The immutability characteristic of blockchain technology presents a challenge with regard to data protection laws. For instance, the EU's General Data Protection Regulation (GDPR) stipulates the "Right to be Forgotten," which mandates the permanent erasure of personal information [13]. In line with the GDPR and the US's state and federal data protection legislation, such as CCPA, data subjects' (human experts) consent is required to process, access, or withdraw their explicit knowledge in a dataset to ensure accountability, transparency, privacy, and security. Merlec et al. [14] presented the use of smart contract (SC), a program stored on a blockchain to dynamically manage the consents of data subjects, regulators, and data owners under the GDPR.

The paper introduces hybrid off-chain and on-chain smart contract methodologies to store access logs and data subject consent on the blockchain. In this approach, the elicited data,

979-8-3503-1019-1/23/$31.00 ©2023 IEEE

personal identity, and cryptographic keys are stored off-chain, while the access logs and consent states are stored on-chain. It ensures the accuracy of the consent states computed off-chain by all host organisations with a simple smart contract mitigating any mistrust in off-chain computations. An application of this framework to integrate the judgment of multiple experts in designing small business loan eligibility policies is demonstrated in this paper.

II. Blockchain Framework to Manage Knowledge-Base (KB) in Compliance with Data Protection Law

The architecture to manage KB by blockchain is shown in Figure 1. All aspects of this Figure are discussed in this section from subsection (A) to (F).

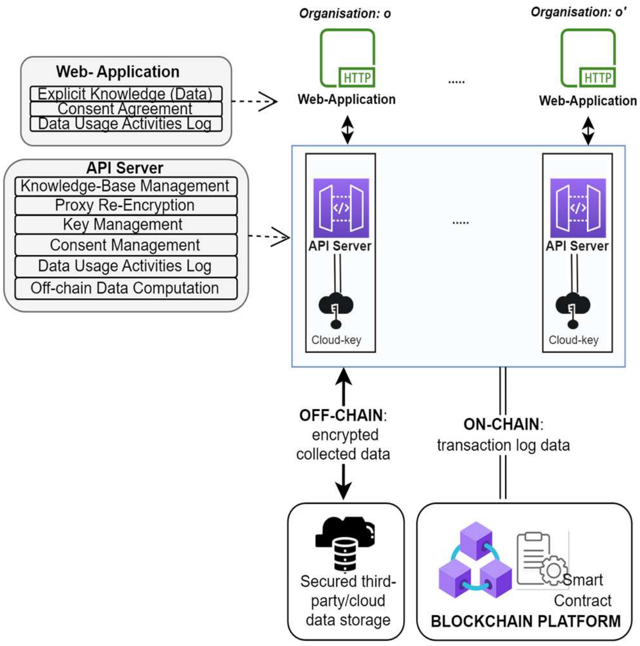

Fig. 1. Architecture to Manage Knowledge-Base by Blockchain

A. *Elicitation of expert knowledge in a web application*

The initial layer depicted in Figure 1 portrays a web application designed to collate data from multiple experts, considered as data subjects within this framework. The web application serves as the front-end interface for experts' knowledge elicitation.

Two modules: consent management and data usage activity log in a web application, allow the experts to self-evaluate their contribution and monitor their data usage activities and consents. The data elicited from the experts and their consent parameters are transferred between the web application and an API server of an organisation through pull and push mechanisms, respectively.

B. *Data management and interaction with blockchain and other services by web API*

The API server facilitates interoperability between multiple services, including the web application, blockchain platform, off-chain data storage, and cloud-key management. The knowledge elicited from multiple experts through the web application is stored off-chain in a third-party data storage server, and the data access transaction logs are stored on-chain in a blockchain platform. In later stages, auditors retrieve the transaction logs stored in the blockchain to perform audits on protected chronological data access records. The developers or data scientists retrieve off-chain and on-chain data to perform various off-chain artificial intelligence and data analytics computations.

C. *Off-chain and on-chain data storage*

According to data protection laws such as GDPR, permanently uploading personal data or sensitive information to the immutable blockchain is not permissible. The data elicited from human experts be denoted by D_h, and the transaction log data is denoted by D_T.

- Off-chain secured storage of personal identity and non-transactional data:

The data D_h mapped with KB is non-transactional data. It is too large for efficient storage in a blockchain network due to low storage speed, scalability issues, and the high cost of frequent data imports to perform off-chain AI computations. Therefore, encrypted files of non-transactional data are stored in a secured third-party off-chain database (or cloud). The KB in this database is contributed by human experts from multiple organisations and shared among them to infer collective knowledge. Information in the KB is categorised and stored in decrypted, mutually exclusive files for easy navigation and accessibility control. Access to these decrypted files requires the approval of data access consent by the data subjects.

The personal identity of data subjects is not locally stored in the server configuration files. Instead, it is stored in a cloud-based Key Management System that allows for the deletion of personal identity if requested by a data subject of an organisation.

- On-chain storage of transactional log

The blockchain stores transactional logs of data subjects' consents and the number of accesses by organisations for off-chain computations. The transactional records can neither be removed nor mutated, thus establishing the system as tamper-proof. The consent requests are sent from the consent management API to the web application to receive the data subjects' responses. Their response to various consents is stored in a private blockchain network.

The data subjects do not access the blockchain directly, and it does not manage data subjects. However, their data is shared across different organisations. APIs of these organisations are used to access the data for off-chain computations. Privacy of blockchain networks is enhanced by segregating the networks into channels, where each channel supports peer-to-peer communication.

D. *Proxy re-encryption (PRE)*

A human expert (or data subject) serving as a delegator is denoted by $h \in \{1, \dots, H\}$ and a user in an organisation serving as a delegate, such as data scientists and developers, is denoted by $o \in \{1, \dots, O\}$.

The proxy re-encryption scheme is used for the encryption and decryption of data (D_h) obtained from human experts. The pair of delegator's and delegate's public key and secret key is represented by (pk_h, sk_h) and (pk_o, sk_o), respectively. The data (D_h) is encrypted by utilising the delegator's public key; $\overline{D_h} = Enc(pk_h, D_h)$. The encrypted data is then stored off-chain in a third-party secured database. The delegate requests decryption permission from the delegator by notifying the public key pk_o. If a delegator approves the data access, then a re-encryption key $(rk_{h\to o})$ is generated specifically for a delegate. It is sent to a proxy system to re-encrypt the previously encrypted data stored off-chain; $\overline{\overline{D_h}} = ReEnc(rk_{h\rightarrow o}, \overline{D_h})$. After re-encryption, the delegate can use decrypt the data by using its secret key sk_o; $D_h = DeCr\left(sk_o, \overline{\overline{D_h}}\right)$.

E. Off-chain and on-chain smart-contract computation

The ability of a blockchain ledger to remain immutable enforces the trust between multiple data subjects and organisations due to the unalterable history of consents recorded as transactions in each block in a chain. Three types of consent are obtained from data subjects h to access elicited data stored off-chain: consent to the initialise data acquisition process (CAQ), consent to withdraw data access (CW), and consent to access the data (CA). The state of these consents and the timestamp of the events are stored in the blockchain. The data subjects can review and audit their data usage activity logs and withdraw access in case of consent violation. A violation is detected by comparison of off-chain and on-chain consent states through Smart contracts (SCs).

A complete data consent state of a given human expert h to an organisation o is represented by $C_{state}^{h,o}$. It is a set of three consent states: $CAQ_{state}^{h,o}$, $CW_{state}^{h,o}$, and $CA_{state}^{h,o}$ as shown in Expression (1).

$$C_{state}^{h,o} = \left\{CAQ_{state}^{h,o}, CW_{state}^{h,o}, CA_{state}^{h,o}\right\} \qquad (1)$$

These states are computed off-chain in an organisation's API by utilising on-chain transaction data. Suppose, o and o' are two different organisations. The on-chain SC compares the set of states $C_{state}^{h,o}$ and $C_{state}^{h,o'}$ for a given human expert h by organisation o and o', respectively. Expression (2) shows the condition of valid execution.

$$\left(\forall c: c \in C_{state}^{h,o} \; and \; c \in C_{state}^{h,o'}\right) \Rightarrow C_{state}^{h,o} = C_{state}^{h,o'} \qquad (2)$$

F. Off-chain Knowledge Integration

A fuzzy Cognitive Map (FCM) allows multiple experts to integrate uncertain and imprecise knowledge into a diagrammatic format. FCMs, are signed digraphs consisting of concepts denoted by $X_i, i \in \{1, \dots, I\}$ and causal relationships between nodes $w_{i,i'}$, where $i \neq, i'$[15]. A logical connection of concepts imposes causal relationships in an FCM. The weight of an edge between two nodes is the strength of the relationship between two concepts, such that $w_{i,i'} \in [-1,1]$. It represents the aggregated judgment of multiple human experts for a given question or rule in a KB. An adjacency matrix represents the weight between concepts. The casual weight between the edges is obtained by first defining the fuzzy membership function to get an expert's opinion on a given linguistic term. Second, fuzzy implication rules such as Mamdani and Larsen are applied to quantify the response proportion for each linguistic term for each edge between concepts [16]. Then, activated membership functions are aggregated by fuzzy aggregation functions. Finally, a crisp value of the weight $w_{i,i'}$ is calculated by defuzzification methods such as the centre of gravity and simulation by Petri-net [17].

III. A Study on Small Business Loans

A. Problem Statement

This framework is implemented to integrate the heuristic knowledge of multiple experts to understand small business lending rules. Both commercial lenders and charitable organisations utilise the combined knowledge of industry best practices and peer advice when evaluating small business loan applications. FCM represents integrated knowledge. The human experts (h) can include external advisors from financial organisations such as large retail banks and internal members of a charitable organisation's loan committee. Charitable organisations are the primary user of KB. However, both organisation types can access and own the integrated KB to perform off-chain statistical analysis and develop data.

B. Results on Implementation of Blockchain

The performance of Hyperledger Fabric 1.4 was evaluated for transaction latency and throughput in a network of four organisations [18][19]. Figure 2 shows a decrease in throughput with increasing load and number of nodes and a marginal increase in average commit time with more than six nodes. Decentralised networks prioritise resilience at the cost of performance, while centralised networks distribute work among nodes for optimal performance.

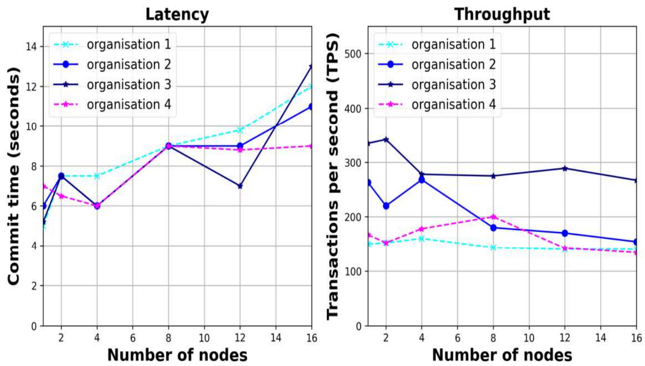

Fig. 2. Average latency and throughput of 800 transactions

C. Off-chain Computation for FCM

The integration of knowledge from twelve human experts by FCM as shown in Figure 3. The red node represents the decline criteria, and the green node represents the referral criteria of a loan application. The influential concepts towards loan eligibility are represented by the computed weight of the edge between two variables. For example, the three concepts for "default in the last t years" positively and negatively impact the target concept, loan eligibility. These concepts have the following weights: 0.92 for "0 default", 0.50 for "1+ default" and -0.92 for "2+ default" respectively.

IV. Conclusion

This paper proposed a framework to manage the knowledge aggregated from multiple domains from different organisations by blockchain technology which secures the data subjects' identity and controls the accessibility of data. It ensures counterparty trust as uploading too much information in the blockchain, especially personal identity, could result in too much transparency and volition of data protection law; in contrast, too little stipulates a lack of trust. A hybrid off-chain and on-chain smart contract computation methodology are proposed for the scalable application of blockchain technology which requires low on-chain storage cost and low cost of frequent data imports for off-chain AI computations. It demonstrates the results of a study to aggregate the heuristic knowledge of multiple experts from four organisations to understand small business lending rules.

ACKNOWLEDGEMENT

This work is funded by the University of Liverpool-Management School. The authors are grateful to RVA Works Enterprise for the knowledge exchange regarding their Open Trellis program for small business lending.

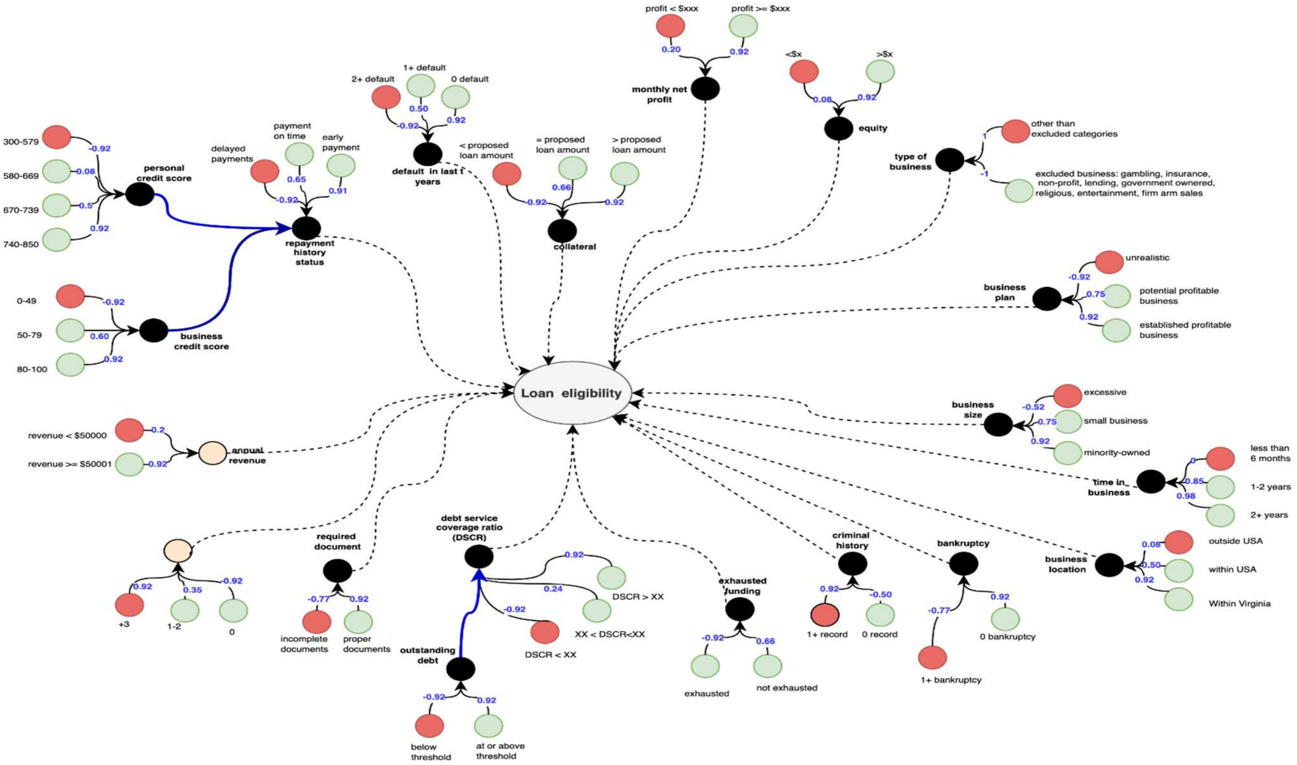

Fig. 3. Off-chain knowledge integration by FCM to assess loan eligibility dynamics by 12 small business loan experts in four organisations

REFERENCES

[1] S. Sachan, J. B. Yang, D. L. Xu, D. E. Benavides, and Y. Li. "An explainable AI decision-support-system to automate loan underwriting," Expert Systems with Applications, 144, 113100, Apr 2020.

[2] C. McInerney, "Knowledge management and the dynamic nature of knowledge." Journal of the American society for Information Science and Technology, Vol. 53(12), pp. 1009-1018, Oct 2002.

[3] W.M. Trochim, "An introduction to concept mapping for planning and evaluation." Evaluation and program planning, 12(1), pp.1-16, Jan 1989.

[4] T. Tip, "Guidelines for drawing causal loop diagrams." Systems Thinker, Vol. 22(1), pp. 5-7, Feb 2011.

[5] D. Heckerman, "A tutorial on learning with Bayesian networks." Innovations in Bayesian networks, pp. 33-82, 2008.

[6] E.I. Papageorgiou and J.L. Salmeron, "A review of fuzzy cognitive maps research during the last decade." IEEE transactions on fuzzy systems, Vol. 21(1), pp. 66-79, May 2012.

[7] S. Sachan, "Fintech Lending Decisions: An Interpretable Knowledge-Base System for Retail and Commercial Loans." In International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, pp. 128-140, 2022.

[8] S. Sachan, F. Almaghrabi, J. B. Yang, and D. L. Xu, "Evidential reasoning for preprocessing uncertain categorical data for trustworthy decisions: An application on healthcare and finance." Expert Systems with Applications, 185, pp. 115597, Dec 2021.

[9] S. Sachan, J. B. Yang, and D. L. Xu, "Global and local interpretability of belief rule base." Developments of Artificial Intelligence Technologies in Computation and Robotics: Proceedings of the 14th International FLINS Conference (FLINS 2020), pp. 68-75, 2020.

[10] W. Yang, S. Garg, Z. Huang, and B. Kang, "A decision model for blockchain applicability into knowledge-based conversation system." Knowledge-Based Systems, 220, p.106791, May 2021.

[11] D.G. Schniederjans, C. Curado, and M. Khalajhedayati, "Supply chain digitisation trends: An integration of knowledge management.", International Journal of Production Economics, Vol. 220, p.107439, Feb 2020.

[12] T. Frozza, F. Galli, E.P. de Lima, and E.R. da Silva, "Implications of Blockchain technology in Knowledge Management: a literature review." In 3rd International Symposium on Supply Chain 4.0: Challenges and Opportunities of Digital Transformation, Intelligent Manufacturing and Supply Chain Management 4.0, pp. 92-98, 2019.

[13] M. Finck, "Blockchain and the General Data Protection Regulation, Can distributed ledgers be squared with European data protection law?", European Parliamentary Research Service, PE 634.445 – July 2019.

[14] M.M. Merlec, Y.K. Lee, S.P. Hong, S.P. and H.P. In, "A Smart Contract-Based Dynamic Consent Management System for Personal Data Usage under GDPR.", Sensors, Vol. 21(23), p.7994, Nov 2021.

[15] R. Taber, "Knowledge processing with fuzzy cognitive maps." Expert systems with applications, Vol. 2(1), pp.83-87, Jan 1991.

[16] M. Mazandarani and X. Li, "Fractional fuzzy inference system: The new generation of fuzzy inference systems.", IEEE Access, Vol.8, pp.126066-126082, Jul 2020.

[17] S. Sachan and N. Donchak, "Generalised stochastic Petri-net algorithm with fuzzy parameters to evaluate infrastructure asset management policy.", In 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-8, Jul 2020.

[18] Q. Nasir, I.A.Qasse, M. Abu Talib, and A.B. Nassif, "Performance analysis of hyperledger fabric platforms.", Security and Communication Networks, Sep 2018.

[19] M. Kuzlu, M. Pipattanasomporn, L. Gurses, and S. Rahman, "Performance analysis of a hyperledger fabric blockchain framework: throughput, latency and scalability.", In 2019 IEEE international conference on blockchain (Blockchain), pp. 536-540, Jul 2019.