The Eastasouth Management and Business
Vol. 01, No. 03, May, pp. 128 - 136
ISSN: 2985-7120, DOI: 10.58812/esmb.v1.i03

Causality-Preserving Distributed Decision Fabrics for AI-Native Financial Enterprise Architectures

Nithesh Gudipuri
Raymond James & Associates, Independent Researcher, St. Petersburg, FL 33716, USA

Article Info

Article history:

Received May, 2023
Revised May, 2023
Accepted May, 2023

Keywords:

Auditability;
Blockchain;
Cap Theorem;
Causal Consistency;
Causal Inference;
Decision Provenance;
Distributed Systems;
Enterprise Architecture;
Explainable Artificial Intelligence;
Financial Technology;
Permissioned Ledgers;
Shapley Additive Explanations

ABSTRACT

Today, financial enterprise architectures are increasingly incorporating artificial intelligence (AI) capabilities into distributed transaction-processing systems, creating a tension between the need for rapid service availability and the need for a causally coherent audit trail of AI-made decisions. In this paper, we marshal research on causal consistency models, causal inference, explainable AI (XAI), and blockchain-based financial infrastructure to propose a conceptual framework called the causality-preserving distributed decision fabric. Causal consistency, which is formalized in causal-memory and session-guarantee models, is explored as a compromise between the high availability of eventual consistency and the coordination cost of linearizability, which is limited by the CAP theorem. Methods for causal inference and model-agnostic explainability, such as Shapley-value attribution and local surrogate modeling, are examined as ways to make automated financial decisions interpretable to those who manage the risks in financial institutions and regulators. Distributed ledger architectures are viewed as an audit trail that can track and order AI-backed decisions between organizational lines. The synthesis suggests that none of the single mechanisms alone meet the simultaneous requirements of availability, explainability, and auditability; thus, a layered architecture of causal-consistency protocols, an explainability layer, and a permissioned ledger is proposed. Comparisons on a conceptual level regarding consistency models and explainability methods, as well as architectural diagrams depicting the proposed layering, are presented. The findings suggest that maintaining causality is a key prerequisite for trustworthy AI-native financial systems, but that interoperability and computational load are major obstacles to implementation.

Corresponding Author:

Name: Nithesh Gudipuri
Institution: Raymond James & Associates, Independent Researcher, St. Petersburg, FL 33716, USA
Email: Nithesh.gudipuri@gmail.com

1. INTRODUCTION

Financial institutions have increasingly been implementing distributed, service-oriented, and cloud-native information technology (IT) environments, in which artificial intelligence (AI) processing elements are becoming commonplace in credit scoring, fraud detection, algorithmic trading, regulatory reporting, and beyond. These architectures rely on data being simultaneously replicated and processed across nodes that are geographically separated from each other, creating a conflict

between the need for the data to be available quickly enough for the operation to perform its function and the need for the data to be available in a way that is understandable and verifiable for governance purposes.

This tension has been expressed in theory in the CAP Theorem [3] , which states that in a distributed data store, consistency, availability, and partition tolerance are mutually exclusive, meaning that if a network partition occurs, you will have to compromise on at least one of the three properties. The boundary has been further clarified in subsequent work, where gradations of consistency are identified ranging from eventual to linearizability—availability versus strict ordering, with the latter having a higher coordination overhead [5]. However, between these extremes, causal consistency leaves ordering of causally related operations unchanged, but allows unrelated operations to be observed in different order at different replicas, which was first formalized with the causal-memory model [1] and subsequently expanded with session-guarantee protocols limiting staleness a client can observe [4].

In an AI-driven financial decision system, the lack of causal ordering isn't just a nuisance; it can result in decision records that make it seem like an automated action took place before the event that triggered it, compromising the accuracy of the data used downstream in models, as well as the ability to defend the decision under the regulatory lens. Concurrently, the rise of black-box machine learning for high-stakes economic decisions has driven a separate but related research field of causal inference and explainable AI that aims to explain the behavior of a model to a human reviewer, but not just on an out-of-sample test set [6], [8]. Another line of research focused on distributed ledger technology (DLT) as a way of capturing and storing evidence of transactions and decisions that cannot be changed or removed over time and across organizational boundaries [13, 14].

This paper brings together these three streams of research, which have grown almost independently of each other, into a conceptual integrated fabric known as a causality-preserving distributed decision fabric. The goal of the synthesis is to investigate if causal consistency protocols, explainability techniques, and distributed ledger architectures can be integrated to ensure that AI-native financial systems are also available, interpretable, and auditable. The paper proceeds as follows: Section 2 summarizes the literature for causal consistency, causal inference, and explainable AI, while Section 3 proposes a layered architecture for blockchain in the financial sector. Section 4 conducts a comparative discussion on both the advantages and disadvantages of the mechanisms discussed in Section 2. Section 5 concludes with implications and limitations.

2. LITERATURE REVIEW

2.1 Causal Consistency in Distributed Systems

Causal consistency is a particular consistency model in the range of consistency models for replicated data stores. Ahamad et al. formalised the concept of causal memory as a model in which all operations that are causally related must be observed in the same relative order by all processes, either through a sequence of program order relations or by a process that sees a previous operation that wrote the memory word, but not through a sequence of program order relations [1]. This definition enables causal consistency to be achieved without the need for global coordination of the operations that should be performed, as a replica only has to 'follow' the causal relationships of the operations it has witnessed, rather than enforce a single global order on all operations.

Another set of guarantees that a client session can ask a weakly consistent replicated store to provide, known as session guarantees, was operationalized by Terry et al., which include monotonic writes, writes-follow-reads, monotonic reads and read-your-writes [4]. These guarantees can be understood as an

approximation of causal consistency, a predictable behaviour for a single session, without the need for causal metadata to be maintained across the whole system. Viotti and Vukolić's survey places causal memory and session guarantee in a larger scale of consistency, which ranges from eventual consistency, different types of session and causal consistency, sequential consistency and linearizability, and notes that stronger guarantees come at a higher coordination cost and at the expense of lower availability in the event of a partition [5].

In contrast, Bailis et al. present a complementary argument, showing that a large class of transactional guarantees, called highly available transactions (HAT), can be achieved without compromising availability or introducing coordination into the normal operation as long as the guarantees don't require detecting a global order across all transactions [2]. Unlike linearizability, they show that causal consistency is still in the class of guarantees that can be achieved without coordination, making it a strong candidate for latency-sensitive financial workloads where a defensible causal view of decisions is still required. Together, this literature shows that causal consistency is achievable, is sufficient to ensure that an AI decision record will not include any event that could be considered to precede itself, and is not so strong as to cause the global coordination bottlenecks that are inherent in strict serializability across geographically distributed data centers.

Table 1. Comparative Overview of Distributed Consistency Models Reviewed in Section 2.1

Consistency Model	Ordering Guarantee	Availability Under Partition	Coordination Overhead
Eventual consistency	None across replicas; convergence only	High	None
Session guarantees	Per-session (read-your-writes, monotonic reads/writes)	High	Low (client-scoped)
Causal consistency	Preserves order of causally related operations	High	Low (per causal chain)
Sequential consistency	Single total order, not tied to real time	Reduced	Moderate–high
Linearizability	Single total order consistent with real time	Lowest under partition	High

Note: Availability and overhead ratings reflect qualitative characterizations reported across the cited sources rather than measured benchmarks.

2.2 *Causal Inference and Explainable AI*

A parallel line of research focuses on the interpretability of the AI models to be deployed on top of such a consistency layer. Pearl suggests the causal inference problem is more complicated than associational machine learning and requires the use of tools other than the models trained to predict outcomes from observations, such as structural causal models, do-calculus interventions, and counterfactual reasoning; in fact, models trained only to predict outcomes from observational data cannot alone support the counterfactual questions that regulators and risk managers ask when considering a financial decision [6]. Zhao and Hastie take this argument further, illustrating that post-hoc interpretation methods of black-box models can only be interpreted as causal under certain structural assumptions, and warning that feature-attribution scores do not necessarily correspond to causal effects [7].

In the model-agnostic explainability literature, Ribeiro, Singh and Guestrin present Local Interpretable Model-Agnostic Explanations (LIME), a surrogate model of the local decision

boundary for an arbitrary classifier trained on a small number of perturbed examples near the instance to explain [12]. Lundberg and Lee then present a family of additive feature-attribution methods, including LIME, under a common cooperative game-theoretical framework known as SHapley Additive exPlanations (SHAP), in which each feature gets a contribution based on Shapley values that satisfies consistency and local accuracy properties that cannot be guaranteed by LIME [11]. These and related approaches are classified by the survey of Carvalho, Pereira, and Cardoso, which focuses on the scope of the method (local or global), model dependency (model-specific or model-agnostic), and the type of results produced (and trade-offs between fidelity to the underlying model and computational cost) [9].

Specifically, in the field of financial risk management, Bussmann et al. show how SHAP-based attribution can be used to explain the credit risk classification results of machine learning models and how this can enable risk managers to trace back a model's output to the credit policy documentation required by the regulations [8]. Further, in Lessmann et al.'s benchmarking of credit scoring classification algorithms, more accurate ensemble models are often less interpretable than traditional scorecards, which further motivates the explainability layer proposed in Section 3 [10] for developing more accurate yet more interpretable classification models.

2.3 Distributed Ledger Architectures in Finance

The third relevant stream is the use of distributed ledger technology as an auditable record-keeping substrate used for financial services. Ali et al. conducted a literature review of blockchain applications in financial services, finding that the three common motivations for the adoption of blockchain technology are transparency, disintermediation, and the immutability of transactions [13]. They also list common barriers to blockchain adoption in the financial services sector such as the uncertainty of regulation, interoperability with legacy core-banking systems, and concerns regarding transaction throughput with permissionless consensus protocols. The review also separates permissionless ledgers that focus more on the decentralisation characteristics of blockchain networks, while compromising on throughput and latency, from permissioned ledgers that limit their access to participants, thereby being better suited to regulated financial institutions' latency and governance needs.

In Frizzo-Barker et al.'s systematic review on blockchain as a disruptive technology for business, financial services are one of the earliest and most active sectors of blockchain applications, although most of the existing published literature is conceptual or exploratory, and organisational and inter-firm governance issues, not technical constraints, often decide whether the blockchain pilot moves to production [14]. Collectively, this literature establishes that a permissioned ledger is a potential write path that could record a sequence of entries that has no causal ordering guarantees (as described in Section 2.1), but still be a way of recording the provenance of AI-driven financial decisions

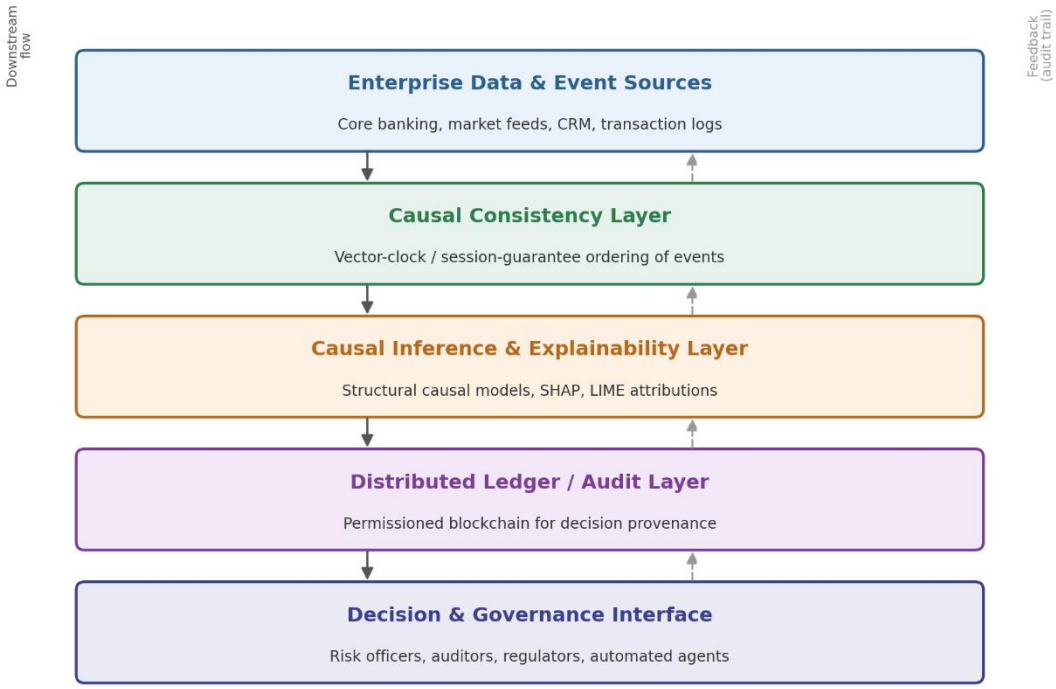

Figure 1. Layered Architecture of the Proposed Causality-Preserving Distributed Decision Fabric

3. PROPOSED CONCEPTUAL ARCHITECTURE: THE CAUSALITY-PRESERVING DECISION FABRIC

The literature reviewed in Section 2 suggests that causal consistency, explainability, and ledger-based provenance provide solutions for different failure modes: Causal consistency allows for the recording of an AI decision before its causal antecedents become visible, or for its generation to be acted upon before they are recorded, and thus, making the reasoning behind the decision legible to a human audience; ledger-based provenance, on the other hand, allows for an immutable, shared record of the decision sequence spanning organizational boundaries. None of these mechanisms, when used alone, targets failure modes of the other two, thus calling for a layered architecture where each mechanism is applied at the place in the decision pipeline where it is most useful.

The proposed causality-preserving distributed decision fabric is shown as five layers in Figure 1. The lowest layer is the collection of enterprise data and event sources, such as core banking transactions, market data feeds, customer relationship records, and so on, all with logical timestamps based on a vector-clock or hybrid logical clock scheme consistent with the causal-memory formalism [1]. The causal consistency layer consumes these tagged events and enforces that any downstream consumer, including an AI inference service, sees causally related events in a consistent order; and that causally unrelated events can be processed concurrently and independently without causing coordination costs, as Bailis et al. demonstrate that such a guarantee is not needed for this class [2].

The causal inference and explainability layer is fed causally ordered event streams and generates a decision and an explanation. The explanation is generated in a Shapley value-based attribution that is consistent with the SHAP framework, in which the contribution of the i-th input feature to the output of a model is calculated as a weighted average over all possible feature subsets:

$$\varphi\_i = \Sigma\_\{S \subseteq F \setminus \{i\}\}[|S|!\,(|F| - |S| - 1)! \,/\, |F|!] \times [f(S \cup \{i\}) - f(S)]$$

Specifically, each value of f(S) represents the model's prediction for the output function given all the remaining features, except feature i, S is a subset of F, and f(S) is the model's prediction for the output function, when feature i is not included in the set of features S. This formulation ensures that the attributions add up to the difference between the model's prediction for the instance and its prediction on the baseline example [11]. If a more rapid and locally consistent explanation is sufficient, the layer can be fitted to a local surrogate model in a similar fashion as described in the LIME procedure [12]; this choice is determined by the fidelity-versus-cost trade-off reported in the explainability literature [9] and argued further in Section 4.
The distributed ledger layer is updated with the causally ordered event identifiers and the associated decision output and summary explanation, and stores them on a permissioned ledger shared by the respective institutional participants, in line with the permissioned architecture pattern found to be more suitable for regulated financial institutions [13]. It is important to note that an audit trail created as such does not contain the property that a recorded decision always refers to as its cause, an event that comes after it in the ledger, but is instead added in the causal order below it. The highest level of decision and governance oversight presents the ordered, explained, and ledgered decisions to risk officers, internal auditors and, as appropriate, external regulators.

This layering is done with deliberate care with respect to the coordination requirement, as only a single point of coordination (ledger-commitment) is introduced, as opposed to the causal-consistency and explainability layers, which leverage the same per-session, per-participant, non-coordinated guarantees used in HAT systems [2] that emerged from the literature reviewed in Section 2.1, and which it was found that the cost of stronger guarantees increases dramatically when a system must coordinate across all replicas as opposed to across causal dependencies of a single session, as presented in the paper of [5].

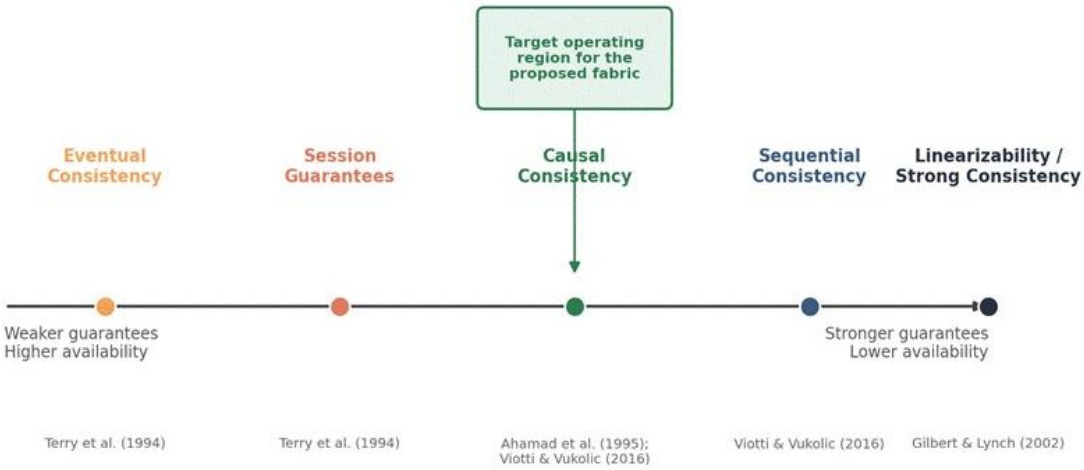

Figure 2. Positioning of Causal Consistency Along the Distributed Consistency Spectrum

Table 2. Comparison of Explainability Approaches Reviewed in Section 2.2

Method	Scope	Model Dependency	Theoretical Basis	Representative Source
LIME	Local (single prediction)	Model-agnostic	Local surrogate linear approximation	Ribeiro et al. (2016) [12]
SHAP	Local, aggregable to global	Model-agnostic	Shapley values (cooperative game	Lundberg & Lee (2017) [11]

Method	Scope	Model Dependency	Theoretical Basis	Representative Source
			theory)	
Interpretable scorecards	Global	Model-specific	Additive linear coefficients	Lessmann et al. (2015) [10]

4. DISCUSSION AND COMPARATIVE ANALYSIS

Table 1 classifies consistency models discussed in Section 2.1 along the dimensions of ordering guarantee, availability in presence of network partition, and coordination overhead and positions causal consistency as the model adopted by the lower layer of the fabric, due to its convenient position on both dimensions with respect to sequential consistency and linearizability. This selection is consistent with the argument presented in the causal-consistency literature that the strongest guarantee that can be achieved without inter-replica coordination during normal operation is causal ordering [1, 2, 5], an efficient allocation of consistency strength for a component that has the main requirement that AI decisions are not recorded out of causal order.

The approaches to explainability that have been reviewed in Section 2.2 are summarized in Table 2. The comparison suggests that while SHAP-based attribution provides stronger theoretical guarantees (in particular consistency and local accuracy based on cooperative game theory), it also comes with a higher computational cost of computing or approximating a sum over feature subsets, LIME is lower in computational cost, but sacrifices the stability guarantees that SHAP can provide [9], [11], [12]. A hybrid deployment is recommended, however, for a financial enterprise fabric that needs to explain a high volume of decisions with low latency, like real-time fraud scoring models, where SHAP should be used for decisions to be formally audited or challenged by regulators, while LIME or similar lightweight surrogates can be applied to their routine, high-volume scoring, and as shown by Lessmann et al. in credit-scoring models specifically [10], the accuracy-interpretability trade-off discussed in that article is also relevant here.

Table 3 attempts to consolidate the drivers and barriers to blockchain adoption from the financial-services literature and to align them with the ledger layer suggested in Section 3. It also shows that the key challenges, such as regulatory uncertainty and integration with legacy core-banking systems, are not just technical, but organizational and involves agreeing on the governance of the ledger, vetting participants, and dispute-resolution procedures, regardless of the underlying consistency and "explainability" design.

A second observation is that causal consistency and causal inference are very similar, yet very different problems. Causal consistency, as formalized by Ahamad et al., and further developed by later distributed-systems papers, is a concern of the observable order of operations across replicas [1, 5]. As described by Pearl, and in question by Zhao and Hastie, causal inference is the question of whether or not a statistical relationship among the inputs and outputs of a model is due to a causal mechanism [6, 7]. The proposed fabric relies on the former to ensure that decision records are well-ordered, while the latter, together with the explainability layer, relies on the latter to ensure that recorded explanations correspond to real causal mechanisms, a distinction that, as noted by Zhao and Hastie's cautionary analysis, should be made explicit in any production deployment [7].

Last but not least, the synthesis implies that the key challenge is not just the individual parts but the combination of composing coordination-free causal consistency, computationally demanding

explainability, and cross-institutional commitment to the ledger, all under a single latency budget. Since cross-participant coordination is unavoidable, the architecture focuses inevitable latency at the layer that has auditability as the critical constraint, not throughput as does the rest of the world [2], [5] where coordination is really needed.

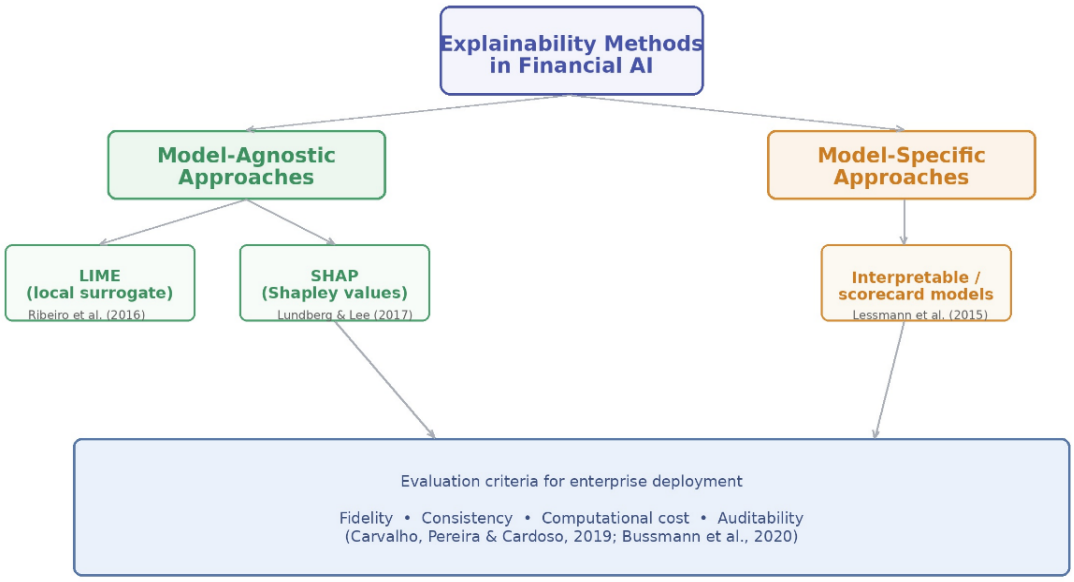

Figure 3. Taxonomy of Explainability Methods Applicable to Financial AI Models

Table 3. Drivers and Barriers to Blockchain Adoption in Financial Services

Theme	Description	Representative Source
Transparency	Shared, tamper-evident record visible to permissioned participants	Ali et al. (2020) [13]
Disintermediation	Reduced reliance on centralized intermediaries for settlement	Ali et al. (2020) [13]
Regulatory uncertainty	Ambiguous treatment of ledger-based records under existing financial regulation	Ali et al. (2020) [13]
Legacy integration	Difficulty interfacing permissioned ledgers with existing core-banking systems	Frizzo-Barker et al. (2020) [14]
Governance dependency	Pilot-to-production progression governed by inter-firm agreement more than technical readiness	Frizzo-Barker et al. (2020) [14]

5. CONCLUSION

This paper has synthesized the work in causal consistency, causal inference, and explainable AI, combined with the work in financial infrastructure based on blockchain, into a conceptual architecture: the causality-preserving distributed decision fabric, that will reconcile the availability requirements of financial systems that are native to AI with the auditability and interpretability requirements of financial governance. The synthesis suggests that causal consistency is a realistic and coordinated-light guarantee that an AI decision record upholds the cause and effect of the events it represents, that SHAP and LIME are complementary mechanisms for explaining, and that the selection of these mechanisms should depend on the audit stakes of a decision, and that PDLs offer a plausible, though organizationally rather than technically limited, means for recording the provenance of AI decisions across institutional boundaries. In this distributed-systems literature review, causality preservation can

be understood as a necessary but not sufficient precondition for a reliable financial architecture in the age of AI: If a decision is unexplainable or unauditable, a causally consistent record of the decision is a limited governance asset. Empirically testing the proposed layering under realistic financial workloads, especially on the latency it introduces at the ledger-commitment layer and how accurate explanations will be generated for the coordination-light regime described in Section 3, would benefit future work on this synthesis.

ACKNOWLEDGEMENTS

[Author(s) to insert acknowledgements of funding sources, institutional support, or individual contributions, as applicable.]

REFERENCES

[1] M. Ahamad, G. Neiger, J. E. Burns, P. Kohli, and P. W. Hutto, "Causal memory: Definitions, implementation, and programming," *Distrib. Comput.*, vol. 9, no. 1, pp. 37–49, 1995, doi: 10.1007/BF01784241.

[2] P. Bailis, A. Fekete, A. Ghodsi, J. M. Hellerstein, and I. Stoica, "HAT, not CAP: Towards highly available transactions," in *14th Workshop on Hot Topics in Operating Systems (HotOS XIV)*, Santa Ana Pueblo, NM: USENIX Association, May 2013, doi: 10.1145/2463676.2465279.

[3] S. Gilbert and N. Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services," *ACM SIGACT News,* vol. 33, no. 2, pp. 51–59, 2002, doi: 10.1145/564585.564601.

[4] D. B. Terry, A. J. Demers, K. Petersen, M. J. Spreitzer, M. M. Theimer, and B. B. Welch, "Session guarantees for weakly consistent replicated data," in *Proceedings of 3rd International Conference on Parallel and Distributed Information Systems*, 1994, pp. 140–149, doi: 10.1109/PDIS.1994.331722.

[5] P. Viotti and M. Vukolić, "Consistency in non-transactional distributed storage systems," *ACM Comput. Surv.*, vol. 49, no. 1, Art. no. 19, 2016, doi: 10.1145/2926965.

[6] J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Commun. ACM*, vol. 62, no. 3, pp. 54–60, 2019, doi: 10.1145/3241036.

[7] Q. Zhao and T. Hastie, "Causal interpretations of black-box models," *J. Bus. Econ. Stat.*, vol. 39, no. 1, pp. 272–281, Jan. 2021, doi: 10.1080/07350015.2019.1624293.

[8] N. Bussmann, P. Giudici, D. Marinelli, and J. Papenbrock, "Explainable AI in fintech risk management," *Front. Artif. Intell.*, vol. 3, Art. no. 26, 2020, doi: 10.3389/frai.2020.00026.

[9] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electron.*, vol. 8, no. 8, Art. no. 832, 2019, doi: 10.3390/electronics8080832.

[10] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *Eur. J. Oper. Res.*, vol. 247, no. 1, pp. 124–136, 2015, doi: 10.1016/j.ejor.2015.05.030.

[11] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017, pp. 4765–4774, doi: 10.5555/3295222.3295230.

[12] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.

[13] O. Ali, M. Ally, P. Clutterbuck, and Y. K. Dwivedi, "The state of play of blockchain technology in the financial services sector: A systematic literature review," *Int. J. Inf. Manage.*, vol. 54, Art. no. 102199, 2020, doi: 10.1016/j.ijinfomgt.2020.102199.

[14] J. Frizzo-Barker, P. A. Chow-White, P. R. Adams, J. Mentanko, D. Ha, and S. Green, "Blockchain as a disruptive technology for business: A systematic review," *Int. J. Inf. Manage.*, vol. 51, Art. no. 102029, 2020, doi: 10.1016/j.ijinfomgt.2019.10.014.