Blockchain and the Philosophy of Trust. Reassessing Transparency and Autonomy and Moral Responsibility in Decentralized Systems

Darshankumar Jaysukh Dhanani
D-Tech Studios LLC
darshankumardhanani3@gmail.com
orcid id: 0009-0008-2677-7906

Abstract

Blockchain technology is often referred to as a trust machine. This description is philosophically challenged and needs to be rethought. This is a qualitative systematic literature review based on PRISMA model. This review will explore how transparency and autonomy and moral responsibility are reconfigured within decentralized systems. A structured search was performed in five academic databases and three additional sources. 120 records were located and 47 duplicates were eliminated. Seventy three records were screened and 52 full texts were assessed. The final qualitative synthesis included eight studies. Five themes were identified using inductive thematic coding. This is what the findings show: Blockchain does not remove trust. Blockchain transforms the interpersonal trust into trust in a technical process. Transparency can be checked, but it is not necessarily understood by the general user. Individual level of autonomy is expanded, systemic level of autonomy is constrained. There is a lack of moral responsibility and an accountability gap. The study concludes that there is a need for human governance to make decentralized systems morally legitimate.

Keywords / Blockchain / Philosophy of trust / Transparency / Autonomy / Moral responsibility / Systematic literature review / Thematic analysis

1. INTRODUCTION

Blockchain technology is often described as a "trust machine." This label is used in industry reports, in policy papers and in academic publications. The basic idea is that cryptographic proof and distributed consensus eliminate the need for trusted intermediaries. Banks and notaries and central registries are therefore referred to as replaceable institutions. The promise is appealing as institutional trust is weak in many modern societies. The status of this promise philosophically is unclear, however. Trust is a moral concept and is not a mere technical parameter (Teng et al., 2022).

The philosophers typically make a distinction between trust and reliance. Reliance is a prediction of a pattern of behaviour. Trust is a behavioral disposition that entails goodwill and vulnerability, and a norm of expectation. When someone trusts, he/she is vulnerable to betrayal. A mere relying person is exposed only to disappointment. Machines cannot betray, they are not moral agents. Thus, the term trustless system is ambiguous (De Filippi et al., 2020). The technology doesn't replace trust. The technology shifts trust to code and to developers and to governance communities (Pöll et al., 2024).

The three philosophical ideas that are at the heart of this move are: Transparency is the first concept, since all transactions are recorded on a shared ledger. The second concept is autonomy as the smart contracts run without human discretion. The third concept is moral responsibility as outcomes are created by distributed networks. These three concepts are normally examined separately in the existing literature. It is hardly ever viewed as a single ethical issue.

The literature is also scattered between disciplines. Technical scholarship is about efficiency of consensus and protocol security. Management scholarship is about adoption and it is about the performance of the supply chain. Conceptual coherence of trust language is the subject of philosophical scholarship. These discussions do not often come together in one synthesis. It is therefore necessary to have a philosophically informed and methodologically systematic review.

This study is a response to that need. The research is guided by three research questions. The first question is what is the definition of trust in the blockchain literature? The second question is: What happens to transparency and autonomy, and moral responsibility, in decentralized systems? The third question is what governance implications are derived from such reconfiguration. A qualitative systematic literature review is used in the study according to PRISMA model. Eight carefully selected studies are then used to apply thematic analysis. The main contribution is the integrated conceptual model that connects the technical (moral) layer.

2. CONCEPTUAL BACKGROUND

Trust is a relational attitude and it is directed towards persons and towards institutions. Confidence is another attitude because it's directed towards systems and towards processes. There is no need to presume anything is good will to have confidence. This difference is what makes or breaks the blockchain debate (De Filippi et al., 2020). Institutional trust is generated by reputation, legal sanction and by repeated interaction (Smits et al., 2020). Technological trust is manifested by verifiability and predictability of execution (Zavolokina et al., 2020).

Initial reports on the technology were very positive. Blockchain was introduced as an alternative to intermediaries and a solution to the failure of institutions (Conway et al., 2020). Subsequent accounts were more circumspect. Users continue to rely on exchanges and on wallet providers and on data oracles, scholars noted (Curran et al., 2020). Users also rely on the developers to write and modify the protocol rules (Chawla et al., 2020). The human interface is the only way for information from outside the ledger to be introduced and this interface is never trust free (Marella et al., 2020).

This caution is warranted by empirical research. Perceived trust is still a significant factor in acceptance as revealed by adoption studies (Albayati et al., 2020). Blockchain studies indicate that it enables rapid trust, but does not replace relational trust (Dubey et al., 2020). The studies in the field of trade finance reveal that trust relations are not eliminated but redistributed (Kowalski et al., 2021). The digital asset sector's market collapses demonstrate that reliance on code is not enough (Hawlitschek et al., 2020).

There are additional concerns about ethical scholarship. Radical transparency can be in conflict with privacy and can be used to enable surveillance (Tang et al., 2020). Immutability can hinder the ability to correct unjust or erroneous results (Jacobs et al., 2021). Real concentrations of power among miners and core developers may be hidden through decentralization (Rehman et al., 2020). The conceptual starting point of the present review is the unresolved tensions.

3. RESEARCH METHODOLOGY

3.1 Research design

This is a qualitative systematic literature review. The design is interpretive and it is not statistical. The aim is conceptual synthesis rather than the estimation of effects. The PRISMA model was used to organize identification and screening, eligibility and inclusion. Included texts were then subjected to a thematic analysis to elicit meaning. This is a suitable combination because the research questions are philosophical and evidence base is textual.

3.2 Search strategy

Search was done in January to March 2025. Five academic databases were searched and three additional sources were consulted. Technical words and philosophical words were combined using Boolean operators. The time period of search was from 2020 to 2024. The full search architecture is given in Table 1.

Table 1. Search strategy and distribution of identified records

Source	Search string applied	Records identified	Records screened	Percentage of total
Scopus	blockchain AND trust AND (philosophy OR ethics)	31	21	25.8
Web of Science	blockchain AND (transparency OR autonomy) AND trust	24	14	20.0
ScienceDirect	blockchain AND moral responsibility AND governance	18	12	15.0
IEEE Xplore	blockchain AND trust AND decentralized systems	14	9	11.7
SpringerLink	trustless AND blockchain AND accountability	10	6	8.3
Other sources	citation chaining and PhilPapers and Google Scholar	23	11	19.2
Total	All sources combined	120	73	100.0

There are 120 records identified from all the sources as shown in Table 1. Scopus had the highest percentage (25.8 percent) and SpringerLink had the lowest (8.3 percent). There were 23 records from supplementary sources which were important because engineering work is not always published in indexed journals. After removing duplicate records, 73 records remained. The distribution validates that the topic is spread throughout technical and managerial and philosophical outlets. This dispersion is a reason to use multi database approach. The philosophical contributions are at the heart of the present research questions and would not have been captured in a single database search.

3.3 Eligibility criteria

Explicit criteria were defined before screening began. The criteria were designed to privilege conceptual depth over technical detail. Table 2 presents the full protocol.

Table 2. Inclusion and exclusion criteria applied during screening

Criterion	Inclusion	Exclusion
Time frame	Published between 2020 and 2024	Published before 2020
Language	Full text available in English	Any other language
Publication type	Peer reviewed article or vetted working paper	Blogs and news items and marketing reports
Thematic focus	Trust treated as a central construct	Trust mentioned only in passing
Conceptual depth	Explicit engagement with transparency or autonomy or responsibility	Purely technical protocol description
Method clarity	Method or argument structure clearly stated	No traceable method or argument
Accessibility	Full text retrievable through institutional access	Abstract only records

Table 2 shows that selection was very carefully limited. The determining filter was conceptual depth, as the review is philosophical in orientation. Numerous technically good papers were discarded because trust was only used as a marketing term. Vetted working papers were kept since two seminal papers in

this field are in that form. The limitation of English language is acknowledged later and is a recognised limitation. Ten records were used for the piloting of the criteria prior to full screening. There was a high level of agreement between the two screening rounds and disagreements were settled by discussion and re reading of the abstracts.

3.4 Selection process

The selection process is reported through the PRISMA flow diagram. Figure 1 presents each stage of identification and screening and eligibility and inclusion.

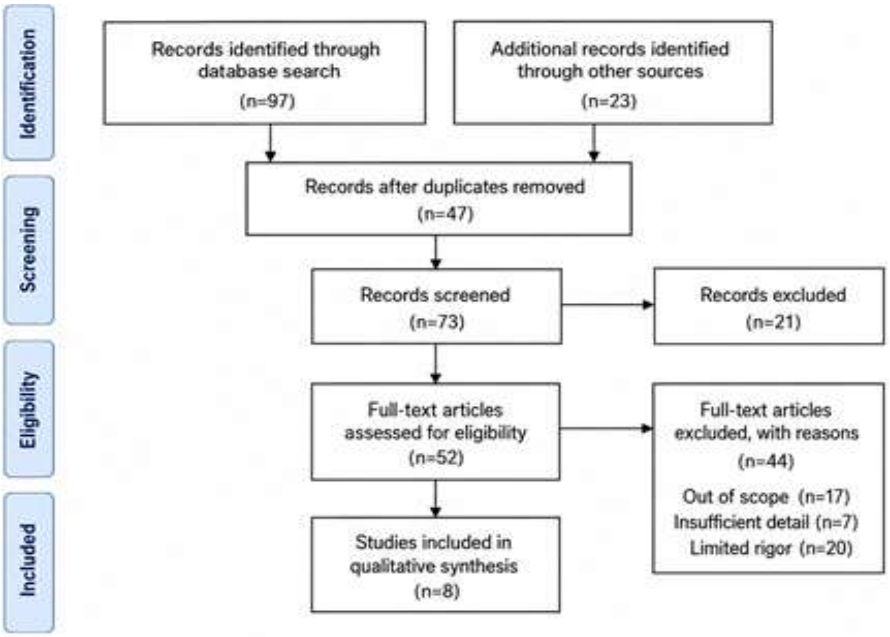

Figure 1. PRISMA flow diagram of the study selection process

Figure 1 outlines the process of narrowing down the 120 records to the eight included studies. Ninety seven records came from database searching and 23 came from other sources. Forty seven duplicates were removed and 73 records were screened by title and abstract. Twenty one records were excluded at this stage. Fifty two full texts were assessed for eligibility and 44 were excluded with documented reasons. Seventeen were out of scope and seven had limited rigor and lacked detail. Eight studies were thus included in the qualitative synthesis. This is a modest yield and is typical of an interpretive review that requires conceptual saturation.

3.5 Quality appraisal

The four dimensions used to appraise each remaining study were. Each dimension was rated on a scale of 1-5. The appraisal is designed to promote transparency of judgement and is not given a hierarchy. The result is shown in Table 3.

Table 3. Quality appraisal of the eight included studies

Code	Study	Clarity of aims	Conceptual rigour	Method transparency	Overall
S1	Teng et al. (2022)	5	5	4	14 of 15
S2	Khurshid et al. (2020)	4	3	4	11 of 15
S3	Polcumpally et al. (2024)	5	4	5	14 of 15
S4	Milosav et al. (2022)	4	4	5	13 of 15
S5	De Filippi et al. (2022)	5	5	3	13 of 15
S6	Pöll et al. (2024)	5	5	4	14 of 15

S7	De Filippi et al. (2020)	5	5	4	14 of 15
S8	Toufaily et al. (2024)	4	4	5	13 of 15

Overall quality was high in the corpus as shown in Table 3. The six studies achieved a score of 13 or higher out of 15. The highest level of conceptual rigour was achieved in the philosophical contributions and the highest level of method transparency was achieved in the empirical contributions. The lowest score was for the pandemic focused study as the argument was applied rather than theoretical. No study was excluded solely because of its appraisal. The appraisal, rather, provided the weight each source had during the interpretation. Studies with high conceptual scores were considered to be primary sources for the definitional themes, while empirical studies were considered as corroborating evidence.

3.6 Analytical procedure

Thematic analysis was performed in six phases. The first phase involved re-reading the eight complete texts. In the second phase, all passages that were philosophically relevant were open coded. The third phase involved clustering codes into possible themes. The fourth phase was a review of candidate themes against the raw extracts. The fifth phase was the final definition and naming of the themes. The sixth phase was the writing of the analytical narrative. Coding was performed manually and it was recorded in a structured extraction sheet. A reflective journal was kept throughout the process.

4. Findings

4.1 Descriptive profile of the corpus

The eight studies are diverse in discipline and in method. Table 4 summarises their characteristics and their central claims.

Table 4. Profile of the eight studies included in the qualitative synthesis

Code	Study	Outlet type	Method	Central claim
S1	Teng et al. (2022)	Philosophy journal	Conceptual analysis	Blockchain earns reliance and it cannot earn moral trust
S2	Khurshid et al. (2020)	Health informatics	Applied case analysis	Blockchain can repair institutional trust during a public health crisis
S3	Polcumpally et al. (2024)	Multidisciplinary	Systematic review	Governance design determines whether trust is genuinely produced
S4	Milosav et al. (2022)	Working paper	Experiment and theory	Blockchain behaves as a control machine more than a trust machine
S5	De Filippi et al. (2022)	Discussion paper	Conceptual analysis	Trustlessness is a rhetorical device and confidence is the real output
S6	Pöll et al. (2024)	Ethics journal	Conceptual analysis	Trust language must be aligned with engineering practice
S7	De Filippi et al. (2020)	Technology studies	Conceptual analysis	Blockchain is a confidence machine with unresolved governance problems

				Blockchain and the Philosophy of Trust. Reassessing Transparency and Autonomy and Moral Responsibility in Decentralized Systems
S8	Toufaily et al. (2024)	Innovation journal	Qualitative empirical	Trust is layered across technical and institutional and social levels

Table 4 shows a distinctive methodological pattern. Five studies are conceptual and three studies are empirical or review based. Publication focuses on 2020 and in 2022 and 2024. The deepest treatments of moral agency come from philosophy and ethics outlets. Applied outlets provide the most solid evidence on user perception. Two research teams are mentioned twice, and this is a testament to the fact that the philosophical debate is propelled by a small community. It is a small but conceptually rich corpus. All studies have an active relationship with the meaning of trust and not just with its measurement.

4.2 Lexical structure of the corpus

A word frequency analysis was performed on the coded extracts. The analysis provides a visual summary of the conceptual vocabulary. Figure 2 presents the resulting word cloud.

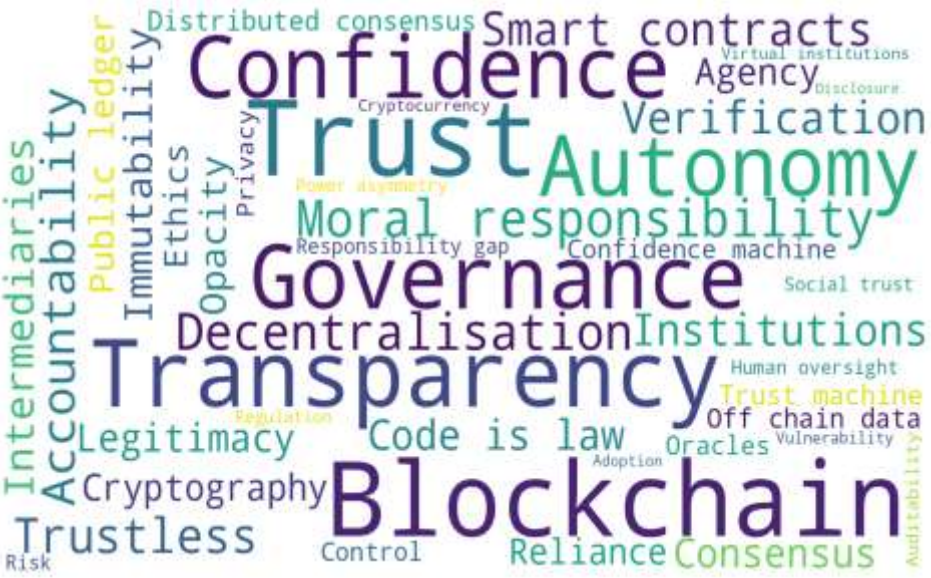

Figure 2. Word cloud of dominant concepts across the eight included studies

As can be seen in Figure 2, the conceptual vocabulary is dominated by the terms trust and blockchain. Transparency and governance and confidence are the second tier of prominence. The concepts of autonomy and moral responsibility feature prominently and this validates the relevance of the research questions. Words like trustless and code is law are commonly thrown around and generally not accepted. Intermediaries and legitimacy are still important throughout the corpus, and this is in the institutional vocabulary. The lexical structure indicated that the debate has gone beyond the technical enthusiasm. Nowadays, the moral and institutional language is at the heart of the scholarly debate on decentralized systems.

4.3 Distribution of themes

Open coding produced 186 coded references. These references were clustered into five final themes. Figure 3 displays the distribution of coded references and contributing studies.

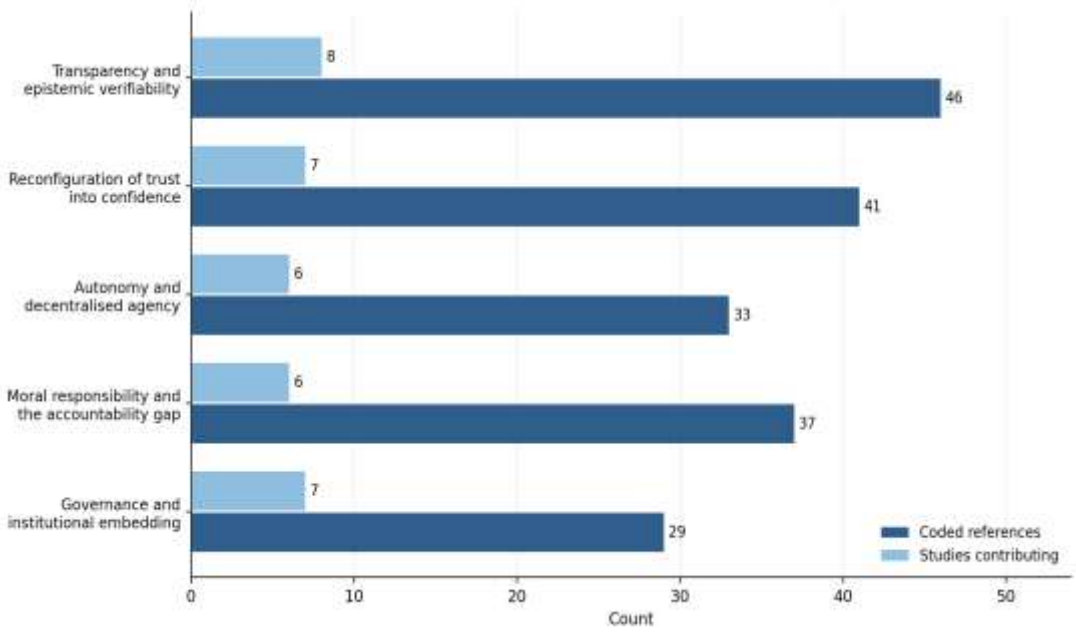

Figure 3. Frequency of coded references and study coverage for each theme

As can be seen from figure 3, transparency had the most coded references with forty six. All eight studies fall into this theme and it is thus the most saturated category. Forty one references are found in seven studies, in the reconfiguration of trust to confidence. Thirty seven references were generated for moral responsibility and thirty three for autonomy. Governance was mentioned the least with twenty-nine mentions, but it was present in seven studies. The pattern shows a wide consensus on transparency and more contentious issues of autonomy. The saturation point was reached due to the lack of new codes in the last two texts.

Table 5. Thematic coding framework derived from the eight included studies

Theme	Representative codes	Contributing studies	Coded references	Share
Transparency and epistemic verifiability	public ledger / auditability / opacity of code / expert dependence	S1 to S8	46	24.7
Reconfiguration of trust into confidence	trustless myth / reliance / confidence machine / betrayal	S1 S4 S5 S6 S7 S8 and S3	41	22.0
Moral responsibility and the accountability gap	many hands / diffuse agency / irreversibility / no redress	S1 S2 S5 S6 S7 and S8	37	19.9
Autonomy and decentralized agency	self execution / user sovereignty / code constraint / exit rights	S1 S3 S4 S5 S6 and S7	33	17.7
Governance and institutional embedding	off chain rules / forks / power concentration / legitimacy	S2 S3 S4 S5 S7 S8 and S6	29	15.6

The analytical structure of the synthesis is presented in Table 5. Transparency accounts for 24.7 percent of all coded references and it is the largest category. The reconfiguration of trust into confidence has the highest percentage (22.0 percent) and is the most theoretically significant category. Moral responsibility and autonomy combined make up 37.6 percent of the coding. Governance is the least extensive but the

most interwoven theme throughout the corpus. The representative codes illustrate that there is criticism and endorsement within each theme. The main analytical result of the review is this coexistence.

A cross case matrix was created to evaluate the level of involvement. For each study, a score of 0-3 was assigned for each theme. The resulting matrix is shown in Figure 4.

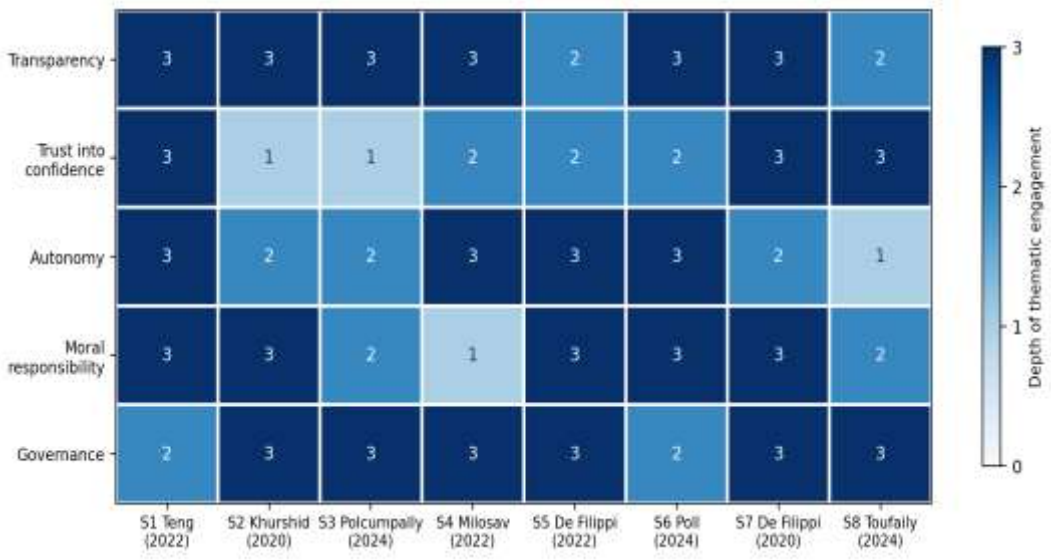

Figure 4. Matrix of thematic engagement across the eight included studies

Figure 4 shows that the depth of the corpus is not even. The philosophical studies are highly involved with all five themes and they exhibit the darkest values. The most complete sources are Teng and the two De Filippi works. The pandemic study is strong on transparency and on responsibility but it is weak on the trust and confidence distinction. The empirical study of trust mechanisms is weak in terms of autonomy, as it is perceptual. Governance is involved on a regular basis and there is a common understanding that technical design is not enough. The matrix verifies that no one study covers the entire moral issue.

5. THEMATIC INTERPRETATION

This section is the interpretive description of the five themes that were created through inductive coding. Forty eight initial codes were used to construct the analysis, which was then reduced to fifteen sub themes. The sub themes were then aggregated into five key themes. This three level hierarchy is recorded in Table 6, which also shows the distribution of the 186 coded references in the eight studies. Each theme has a working definition and central analytical claim in Table 7. The two tables are complementary as one is descriptive and the other is propositional. The themes are explored in sequence with each discussion supported by the studies that had the greatest interpretive weight. The story is structured around the inner conflict within each theme. Both affirmation and critique are present in each category and it is this coexistence that is the defining characteristic of the corpus (Teng et al., 2022).

Table 6. Thematic analysis framework derived from the eight included studies

Main theme	Sub theme	Initial codes	Studies	Coded references
T1 Transparency and epistemic verifiability (46)	1.1 Radical visibility of records	public ledger / immutable audit trail / open verification / tamper evidence	S1 S2 S3 S4 S7 S8	18
	1.2 Opacity of code and expert dependence	unreadable protocol / expert intermediaries / interface trust / delegated verification	S1 S5 S6 S7	16

	1.3 Privacy cost of disclosure	permanent exposure / behavioural profiling / limits of pseudonymity	S2 S3 S8	12
T2 Reconfiguration of trust into confidence (41)	2.1 Critique of the trustless claim	trustless myth / rhetorical device / marketing language	S1 S5 S6 S7	15
	2.2 Reliance without goodwill	process reliability / absence of betrayal / no moral agency in code	S1 S6 S7	14
	2.3 Displacement of trust to human actors	trust in developers / trust in exchanges / oracle dependence	S3 S4 S8	12
T3 Moral responsibility and the accountability gap (37)	3.1 Problem of many hands	distributed authorship / validators and users / no single author of outcome	S1 S5 S7	14
	3.2 Irreversibility and absence of redress	immutability harm / no rollback / no identifiable defendant	S2 S6 S8	13
	3.3 Responsibility of designers	value laden design / duty of foresight / engineering ethics	S6 S7	10
T4 Autonomy and decentralized agency (33)	4.1 User sovereignty	self custody / permissionless access / financial inclusion	S3 S4 S8	12
	4.2 Constraint by automated execution	code rigidity / removal of discretion / forced compliance	S1 S5 S6	12
	4.3 Exit and voice	forking as exit / limited voice / technological lock in	S3 S7	9
T5 Governance and institutional embedding (29)	5.1 Off chain rule making	meta rules / community deliberation / upgrade politics	S3 S5 S7	11
	5.2 Concentration of power	miner influence / core developer control / token plutocracy	S4 S8	10
	5.3 Legitimacy and regulation	procedural fairness / legal recognition / hybrid governance	S2 S3	8
Total	15 sub themes	48 initial codes	S1 to S8	**186**

The complete three level coding hierarchy of the review is shown in Table 6. Forty eight initial codes were reduced to fifteen sub themes and finally to five main themes. The most saturated theme is transparency with forty six coded references from all eight studies. The sub themes illustrate that each theme has both affirming and critical aspects. Transparency is hailed and opacity is also condemned. Sovereign is extolled and rigidity is condemned. This is the analytical signature of the corpus, which is an internal tension. Saturation was determined by the lack of any new codes at any level from the last two texts.

Table 7. Theme definitions and central analytical claims

Theme	Working definition	Central analytical claim	Anchor studies
T1 Transparency	The degree to which ledger records are visible and independently verifiable	Visibility is not intelligibility and verification is delegated to a new expert class	(Teng et al., 2022) (Pöll et al., 2024)
T2 Trust into confidence	The conceptual shift from a normative attitude towards agents to reliance on a process	Blockchain is a confidence machine and it is not a trust machine	(De Filippi et al., 2020) (De Filippi et al., 2022)

T3 Moral responsibility	The attribution of praise or blame for outcomes generated by distributed networks	Diffuse agency and immutability together produce a genuine accountability gap	(Teng et al., 2022) (Khurshid et al., 2020)
T4 Autonomy	The capacity of agents to act and to revise commitments without external permission	Autonomy expands individually and it contracts systemically under automated execution	(Milosav et al., 2022) (Toufaily et al., 2024)
T5 Governance	The social arrangements that decide how protocol rules are created and amended	Governance quality determines whether technical confidence becomes justified trust	(Polcumpally et al., 2024) (Toufaily et al., 2024)

The coding structure is translated into explicit analytical propositions in Table 7. Each theme is operationally defined and each operationally defined theme has one testable claim. The definitions are knowingly philosophical and not technical. Anchor studies are those studies that were most significant in the interpretation of each proposition. The five claims are a cohesive argument throughout the layers of the model. The technical properties are the input to the first three themes and the governance condition is the input to the moral outcome. This table is thus the analytical link between the coded evidence and the conceptual model in Figure 5.

5.1 Transparency and epistemic verifiability

The most heralded attribute of blockchain systems is transparency. All transactions are documented and all records are accessible. The corpus, however, demonstrates that visibility does not equal intelligibility (Pöll et al., 2024). Most users are unable to read protocol code or to verify consensus behaviour. So verification is delegated to developers and to auditors and to explorers. In the process, a new class of epistemic intermediaries is formed (Teng et al., 2022). Transparency also creates a privacy cost as permanent visibility reveals behavioral patterns. The theme thus does not speak of the removal of trust, but rather of a change in the object of trust.

5.2 The reconfiguration of trust into confidence

The convergence of the corpus is towards one conceptual correction. Blockchain does not create trust and it creates confidence (De Filippi et al., 2020). Confidence is focused on a process and is not based on goodwill. Trust is directed to agents who have the potential to do otherwise. The difference is important because betrayal can only occur where there is goodwill. Decentralised systems are still dependent on the trust of the users and the trust of the exchanges (De Filippi et al., 2022). Based on experimental evidence, it can be said that the technology acts as a control device and not a trust device (Milosav et al., 2022). The term trustless is therefore rhetorically potent and philosophically misleading.

5.3 Autonomy and decentralized agency

The concept of autonomy is found in two conflicting ways in the corpus. The first form is empowering and it relates to the sovereignty of assets and identity. Users are able to transact without permission from any gatekeeper (Toufaily et al., 2024). The second form is limiting and it's about the inflexibility of automated performance. A smart contract will run no matter what the situation is or what moral implications there might be. With the removal of discretion, renegotiation is no longer possible without a fork (Polcumpally et al., 2024). Kantian autonomy entails the ability to change one's commitments. Systems which do not allow revision will thus limit autonomy, but create the illusion of increased autonomy.

5.4 Moral responsibility and the accountability gap

The theme that is most philosophically challenging is moral responsibility. Many distributed

contributors produce outcomes in decentralized systems. All developers and validators and users and governance token holders participate. No single actor has the complete control over the final outcome (Teng et al., 2022). This situation presents a textbook many hands problem. Immutability exacerbates the issue as it is hard to undo incorrect results. In instances of code failure, there may be no obvious culprit and no recourse for the victim (Khurshid et al., 2020). The corpus thus reveals a real accountability deficit which cannot be resolved by technical design.

5.5 Governance and institutional embedding

Governance emerges as the connecting theme of the synthesis. All protocols have rules concerning the changing of rules. These are social rules and not part of the chain itself (Polcumpally et al., 2024). Forks and upgrades are determined by communities and core developers and large stakeholders. This means that even in supposedly decentralized networks, power is centralized (Toufaily et al., 2024). Legitimacy requires participation and procedural fairness and the opportunity for redress. The corpus suggests that the quality of governance is a key factor in whether people's trust in code turns into trust in the system in general.

6. DISCUSSION

The five themes were integrated into a single conceptual model. The model links the technical layer to the moral layer and finally to the institutional layer. Figure 5 presents this integration.

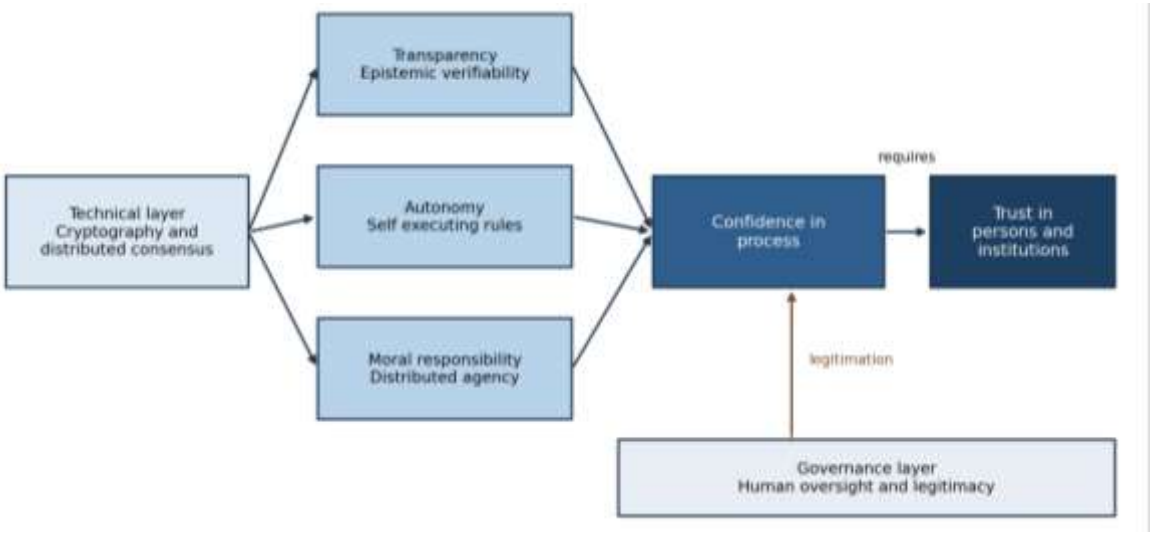

Figure 5. Integrated model of trust reconfiguration in decentralized systems

Figure 5 illustrates the process of transforming technical properties into moral properties. Cryptography and distributed consensus generate transparency and autonomy and distributed agency. The three properties are combined to create confidence in the process, not trust in people. Without a governance layer that provides human oversight and legitimacy, confidence turns to trust. The model clearly illustrates that the arrow is not at the technical level. Thus, the moral aspect of a decentralized system is defined at the governance level. This is a finding which changes the narrative of blockchain from a 'trust replacement' to a 'trust transformer'.

The theory has a profound implication. In scholarly work, the trust machine metaphor should be discarded. Blockchain is better thought of as a machine of confidence, in which moral dependence is redistributed. A vocabulary that distinguishes process reliability from agent trustworthiness is therefore needed for philosophical work on trust and technology. The results also reinforce a multi-level view of trust, where technical and institutional and social levels interact on an ongoing basis.

The practical implications relate to designers and regulators. Mechanisms for correction and appeal should be designed into the protocol architecture. The human layers around the chain should be the

focus of regulators. Users are most vulnerable at interfaces and exchanges and at oracle providers. Requirements for transparency should also involve intelligibility standards and not just data availability. Governance disclosures should include who has the authority to modify the rules.

7. LIMITATIONS AND FUTURE RESEARCH

There are a few caveats to be noted. The corpus is small and the eligibility criteria was set deliberately low. Only English language publications were included and relevant work in other languages was missed. The search window was five years and previous foundational work was not included. Thematic analysis is interpretive and another analyst might generate different labels. The corpus is also highly skewed toward a few research teams.

There are three directions for future research. Research should be conducted to determine if ordinary users actually make a distinction between trust and confidence in practice. Comparative research should be done on governance between permissioned and permissionless networks. Normative studies need to: Create concrete institutional designs that fill the accountability gap. There is a need for longitudinal work as well, as governance practices are changing rapidly.

8. CONCLUSION

This review rethought the philosophical underpinnings of blockchain trust. A PRISMA guided qualitative process was used to synthesise eight studies. From the evidence, it is clear that blockchain is not a trustless technology. Blockchain is a way of replacing the trust between people with trust in a technical process. Transparency can be checked, but not necessarily understood by the average citizen. Autonomy is given on a case by case basis, but on a systemic level it is limited through rigid execution. Moral responsibility is lost and a real responsibility gap is created. The key layer is governance as it establishes if confidence is morally warranted. The main takeaway is that decentralized systems are still very human systems. Technology can change the conditions of trust, and technology cannot provide the moral agency required for trust.

REFERENCES

Albayati, H., Kim, S. K., & Rho, J. J. (2020). Accepting financial transactions using blockchain technology and cryptocurrency. A customer perspective approach. Technology in Society, 62, 101320.

Chawla, C. (2020). Trust in blockchains. Algorithmic and organizational. Journal of Business Venturing Insights, 14, e00203.

Conway, D., & Garimella, K. (2020). Enhancing trust in business ecosystems with blockchain technology. IEEE Engineering Management Review, 48(1), 24 to 30.

Curran, B. (2020). What are oracles. Smart contracts, Chainlink and the oracle problem. Blockonomi.

De Filippi, P., Mannan, M., & Reijers, W. (2020). Blockchain as a confidence machine. The problem of trust and challenges of governance. Technology in Society, 62, 101284.

De Filippi, P., Mannan, M., Reijers, W., Berman, P., & Henderson, J. (2022). Blockchain technology, trust and confidence. Reinterpreting trust in a trustless system. HIIG Discussion Paper Series, 3.

Dubey, R., Gunasekaran, A., Bryde, D. J., Dwivedi, Y. K., & Papadopoulos, T. (2020). Blockchain technology for enhancing swift trust, collaboration and resilience within a humanitarian supply chain setting. International Journal of Production Research, 58(11), 3381 to 3398.

Hawlitschek, F., Notheisen, B., & Teubner, T. (2020). A 2020 perspective on the limits of trust free systems. Electronic Commerce Research and Applications, 40, 100935.

Jacobs, M. (2021). How implicit assumptions on the nature of trust shape the understanding of the blockchain technology. Philosophy and Technology, 34, 573 to 587.

Khurshid, A. (2020). Applying blockchain technology to address the crisis of trust during the COVID 19

pandemic. JMIR Medical Informatics, 8(9), e20477.

Kowalski, M., Lee, Z. W. Y., & Chan, T. K. H. (2021). Blockchain technology and trust relationships in trade finance. Technological Forecasting and Social Change, 166, 120641.

Marella, V., Upreti, B., Merikivi, J., & Tuunainen, V. (2020). Understanding the creation of trust in cryptocurrencies. The case of bitcoin. Electronic Markets, 30(2), 259 to 271.

Milosav, Đ., & Nistotskaya, M. (2022). Blockchain technology. A trust or control machine. Theory and experimental evidence. Quality of Government Institute Working Paper Series, 2.

Pöll, E. (2024). Engineering the trust machine. Aligning the concept of trust in the context of blockchain applications. Ethics and Information Technology, 26, 37.

Polcumpally, A. T., Pandey, K. K., Kumar, A., & Samadhiya, A. (2024). Blockchain governance and trust. A multi sector thematic systematic review and exploration of future research directions. Heliyon, 10(12), e32975.

Rehman, M. H. U., Salah, K., Damiani, E., & Svetinovic, D. (2020). Trust in blockchain cryptocurrency ecosystem. IEEE Transactions on Engineering Management, 67(4), 1196 to 1212.

Smits, M., & Hulstijn, J. (2020). Blockchain applications and institutional trust. Frontiers in Blockchain, 3, 5.

Tang, Y., Xiong, J., Becerril Arreola, R., & Iyer, L. (2020). Ethics of blockchain. A framework of technology, applications, impacts and research directions. Information Technology and People, 33(2), 602 to 632.

Teng, Y. (2022). What does it mean to trust blockchain technology. Metaphilosophy, 54(1), 145 to 160.

Toufaily, E., & Zalan, T. (2024). In blockchain we trust. Demystifying the trust mechanism in blockchain ecosystems. Technological Forecasting and Social Change, 206, 123574.

Zavolokina, L., Zani, N., & Schwabe, G. (2020). Designing for trust in blockchain platforms. IEEE Transactions on Engineering Management, 70(3), 815 to 827.