

Trust and Its Redundancy

**A Note on the Sociocybernetic Term "Trust"
and Its Possible Meaning in the Trustless Web3**

v1.0

Originally published as **Vertrauen und seine Redundanz**

Translated by the author

Gerald Hock

April 3rd, 2025

Contents

I – Introduction: Trust in <i>trustlessness</i>	1
II – The Function of Trust	2
III – The Expectation Horizon "Trust"	3
IV – Orientation via Trust	5
V – Contingent States in Terms of Their Reference to Time: Do <i>public ledgers</i> Function as Trustworthy Consensual Domains due to <i>trustlessness</i> ?	6
VI – Conclusion: Non-Linearity and Self-Reference of Trust Relations	8

I – Introduction: Trust in *trustlessness*

In the context of Web3, the word *trustlessness* is commonly used to describe a fundamental quality of this digital infrastructure. Because trust is a fairly relatable everyday phenomenon, but is used as a specific term in the context of Web3, a more precise definition of its meaning seems worthwhile. And since in a more extensive theoretical study of the functionality of *public ledgers* one often comes across the question of the trustworthiness of *trustless systems*, it seems sensible to determine more precisely where trust does (and does not) occur in conjunction with blockchain technology and Web3.

Already in the first sentence of the introduction of the Bitcoin whitepaper one reads about "trusted third parties"¹ in the traditional banking system. But *wherein* lies the trust in them? Aren't they primarily functionally necessary intermediaries and a certain degree of trust in them is obligatory in order to participate in the system? Replacing the intermediary with (less fallible) automation finally also eliminates the prerequisite for possible trust relations – that *wherein* trust is placed is absent.

Buterin writes in his essay released in 2020 "Trust Models": "One of the most valuable properties of many blockchain applications is trustlessness..."ⁱⁱ According to the morphology of this term, one could assume the absence of trust. However, something else is meant. I would like to go into this in more detail with the help of Luhmann's sociocybernetic treatise "Vertrauen" ("Trust") with the subtitle "Ein Mechanismus der Reduktion sozialer Komplexität" ("A Mechanism of the Reduction of Social Complexity"¹); and explain why not the absence of trust, but its redundancy is a desirable state for a *public ledger*. Also, in order to subsequently establish it as an epistemic order, a "consensual domain".

Since blockchains (and *public ledgers*) aim to automate the bindingness of themselves and (partially) replace the intermediaries with a peer-to-peer system, there *can* be no relation of trust with them. They are simply not there (anymore) and therefore do not form part of the system.

So we're interested in the question: how can something be trusted that is attempting to dissolve the premise of trust between its users?

1 Translated by the author

II – The Function of Trust

Trust plays a crucial role in our everyday lives. Fleeting as well as intimate, friendly as well as purely business-related relationships are based on a more or less pronounced degree of trust. And especially where trust relations do not have to be negotiated and explicitly defined, trust seems to be sufficiently present, although it remains largely invisible in these contexts. Trust is thus not only formed explicitly under certain conditions, but also seems to resonate latently in a large number of interactions. Thus, trust is often an elusive predisposition for social interaction per se. Perhaps, it is indeed most effective where it remains unmentioned and the interaction yet persists.

A sociocybernetic view allows trust to be understood *functionally*, id est in cases where trust-(inter)dependent interactions persist. Because blockchains are progressive reformations of system states that are inherently limited by their architecture, *public ledgers* can largely be treated as closed systems.

But why should you trust a blockchain and/or *public ledger*, or rather: What does trust in them require? And what does the technology guarantee in terms of trust relations *from users towards the technology* - but above all - *between* users of the technology? We are interested in the ways of social interactions within Web3 being relevant in regard to functionally understood trust. Furthermore, we will deal with the question of why trust relations are non-linear. Since: giving more trust does not necessarily correlate with more trustworthiness of the recipient. The buzzword *trustlessness* in Web3 jargon even advertises its trustless technology as particularly trustworthy.

Now, let me bet on something: since you seem to have read this far, I expect you'll expect the rest of the text to be related to the title and to proceed roughly in the direction it thus far progressed. Based on your personality, your experiences and preferences in regard to media, your language skills and much more, you will develop possible scenarios for the future and thus also the continuation of texts you read; and this one, too. Your everyday life, your culture, your habits, your physical and psychological condition and so on establish patterns that make you make assumptions about consequences; for example, the assumption that a text is about what it says in its title, and that it doesn't get bogged down in an endless discussion about what second-order problems fundamentally exist with texts.

Likewise, this is the case here: trust me.

III – The Expectation Horizon "Trust"

First, we do not focus on the successful relations of trust, but on the failing one. Since failure determines limits and impossible conditions, trust can be negatively defined in this way, or its constitution can be determined more precisely, or: an outside of the horizon of a trust based relations can be described. Since you are apparently still reading, I do not – once again – want to disappoint your expectations and assert: If you have trusted so far, I ensure you have also expected something more or less specific. So, what about the interdependence of trust and the expectation it generates or is generated by?

Expectations, assumptions about the future, can be disappointed, but they can also be fulfilled. The quantity of all possible future states is distributed completely asymmetrically among these two poles. There are a multitude of options more of not meeting a specific expectation than meeting it. The looser the expectation is, the more likely it is that it will (somehow) be fulfilled. Consequently, one inevitably runs into social problems if one has generally too strict expectations regarding the behavior of those around him. (In special cases, however, strict expectation patterns make more sense, *exempli gratia* in bartering or in financial transactions – one expects a specific sum and not a rough amount.) If you have too loose expectations or no expectations at all towards your fellow human beings, you become disoriented due to a lack of anticipatory fixation points; also, one causes disturbances in the relation of trust with others, since those to a certain extent rely on reciprocity; one who never behaves according to common standards might be perceived as an indifferent beast. Such constellations ultimately end in a reflexive cycle: In a double-contingent social environment, one makes oneself mistrustful to others if one does not trust either. (It should be noted here that some social relations function in the reverse reflexive cycle of trust relations. Some (maybe material) values are trusted because others trust them as well. The fungibility of currencies is ensured, among other things, because they are used as largely familiar mediums of exchange. Talcott Parsons: "The rational ground for confidence in money is that others have confidence in money..."ⁱⁱⁱ)

Trust must therefore lie between these extremes of the spectrum of possible expectations in order to be functional.

Since expectations generate a projected frame for future behavior, those in the present can behave in anticipation. To quote Vitalik Buterin: "First, my simple one-sentence definition of trust: trust is the use of any assumptions about the behavior of other people."^{iv} According to Buterin's definition, the addition of expectation of behavior is the application (or "use") of the respective expectation

substantial for trust. Only with incorporated anticipation (meaning expectation) of present behavior trust can be established. Only when people "bet" on the future, only when a personal stake is placed on the anticipated scenario, does trust come about. Luhmann: "Those who show trust anticipate the future. One acts as if he is certain of the future. One could think that he overcomes time, at least time differences."^v (And since the present is already influenced by anticipating the anticipated future, recursivity and non-linearity can already be identified in the logic of trust.)

So, some commitment to the anticipated scenario is necessary, and one might assume that more commitment necessarily makes the scenario more likely to occur.

But the opposite also seems to be the case that an overly specific expectation denies the possibility of said expectation to be met. Trust is rendered impossible as soon as expectations are over- *or* under-determined. Excessive restriction of the degrees of variance in trust based relations is fatal if trust is tied to the rigid fulfillment of expectations. However, this does not seem to be possible in most cases and would generally be prone to failure. Consequently, two conditions are to necessarily be fulfilled in order for trust relations to persist:

- 1) From the truster emanating reliability guaranteed by a certain minimum of expectations and
- 2) from the truster emanating maximum of tolerance in regard to the expectable behaviour of alter.

Within this horizon of expectations, trust acts, metaphorically speaking, like a leash of social interaction, which enables the "straight walking" of intact social interactions by *and because* it constraints „straight walking“. Thus, trust is a restricting framework of a succeeding social interaction.

Once established, trust can only be maintained where the behavior of alter does not lie outside the horizon of expectations. In the context of Web3, this could mean that once trust has been established, it continues to exist; — *in particular because of the expected future rigidity of the system*. The advantage of a *public ledger* is therefore not only due to the automation of the bindingness of itself, but also, in second order, due to the expected future persistence of this very mechanism: *immutability*.

In the next section, this second-order trust-enhancing mechanism of rigidity will be furthermore discussed. Because, although a technology is potentially particularly trustworthy due to its sophisticated, possibly even infallible, functionality, this alone does not necessarily have a strong impact circumstance whether it is perceived as such (at this point in time). Trust also requires the dimension of time.

IV – Orientation via Trust

Trust, in a broader sense, represents the preference of a referential object relative to its absence, irrespective of a lack of verification. Despite a recognized difference between the known known and the known unknown, trust can exist in relation to the reference, that is, it can be operated with the confident assumption and predictability in regard to its future condition and/or behavior. The unconfirmed possibility that the opposite is true of a particular reference, id est it is not trusted, can be due to a greater amount of reasons compared to being trusted. Therefore – and this is how Luhmann explains the chaotic case of the ubiquitous absence of trust in his work *Trust* – the improbable case of relations of trust must first be established and, if necessary, restabilized.

"Trust in the broadest sense of trusting one's own expectations is an elementary fact of social life. In many situations, people have a choice as to whether or not they want to place their trust in certain respects. Without any trust, however, he could not get out of bed in the morning. Indefinite fear, paralyzing terror would seize him. He could not even formulate a definite distrust and make it the basis of defensive precautions; for that would require trusting in other respects. Anything would be possible. Nobody can stand such a sudden confrontation with the extreme complexity of the world."^{vi}

It requires (cognitive) contingency management. A reduction in uncertainty and thus the (cognitive) relief is necessary so that people can orientate themselves in their environment. Trusting people need time to be able to adjust their expectations in order to be able to establish functional relations of trust, including those in their everyday environment. If one follows Luhmann, then one must understand mistrust as a cognitively much more demanding form of perception than trust. As soon as an environment can be found that not only appears trustworthy, but also rewards long-term trust, the person who finds himself in it will feel relieved psychologically and cognitively. And he then can "invest" the capacities thus freed up in other matters. Trustworthy money, for example, is not superior to untrustworthy money simply because it is more suitable for the function of money, but because everyone using it would have the cognitive capacity free to take care of other things.² Hence: If you solve problems of trust (in money), you always solve more than those very problems. You create the conditions for people and, consequently, society to be able to orient themselves when the disorientation of existential fears decreases. A consensual trust is a fundamental technology not

2 Compare with this: "Trust strengthens, to allude to a well-known psychological theory, the "tolerance for ambiguity". This performance is therefore not to be confused with instrumental event control. Where such control can be ensured (id est "made present"), trust is unnecessary. Trust is needed to reduce a future of more or less indeterminate complexity." (Luhmann, Niklas. *Vertrauen*. p. 19, translated by the author)

only structures current conditions, but also determines the horizon of future ones. And therein lies the gain for social interaction, which includes financial transactions: Expectations can be met more reliably and therefore also be formed, even initiated.

The following chapter deals with how a consensual domain, in our case a "trusted public ledger", systematically allocates the future and determines the present.

V – Contingent States in Terms of Their Reference to Time:

Do *public ledgers* Function as Trustworthy Consensual Domains due to *trustlessness*?

The school of Radical Constructivism deals with the basic cognitive constitution of humans and with related epistemological questions. In short, one can say that within this intersection of branches of science, the observer is depicted as a condition of the possibility of the perceived and thus (in a specific sense) real world. A perceived object does not fall into perception; it is generated by a given cognitive structure. But then the question arises as to how coordination and orientation among people is possible at all, because if you follow this paradigm, there would be a substantial likelihood that everyone has a completely idiosyncratic and incongruent perception of "the world" – whatever that may then be. And although such would be possible, or even probable, there is immense agreement among people as to what is happening in the common world. We see – or at least can communicate about it – the same moon for example, we have a similar concept of a table, we all use the same internet, we all use the same *public ledger* and so on. This unity, "consensual domain", can not be derived solely from an identity of an all-outer "objective reality"; but because it is constructed for the most part in identical ways, similarly at least. A consensual domain thus results due to the equal forms of the interaction patterns of the respective observers. Such a consensual domain not only enables coordination among each other, it makes future coordination dependent on it.³ It has an effect of retention and, expressed in an economic metaphor, forms an anchor. And

3 Compare with this: "The observation of inter-systemic coordinated interactions, which may appear to an observer [...] as mutual orienting behavior of interaction partners, is based on the formation of a consensual domain among the individual participating organisms. [...] If the behavior of interacting organisms is of such a nature that it can be described as behavior directed toward each other and mutually conditioned [...] then a consensual domain can be assumed for the organisms concerned. Consensuality here does not mean unanimity or agreement in the sense of a common accord; a consensual domain [...] is initially merely a domain of behaviors of individual organisms in which (seen from the perspective of an observer) they coordinate their behavior with one another by mutually orienting themselves. Consensual domains of this kind thus arise/exist between all organisms when they - in whatever manner - enter into direct interaction with one another." (Rusch, Gebhard. *Erkenntnis, Wissenschaft, Geschichte*. Frankfurt am Main: Suhrkamp Verlag, 1987, p. 141, translated by author)

therefore this also inevitably occurs as effects of a *public ledger*. Hence the *public ledger* is the contingency for its own future states distributed over time.

The recursive logic of system states and their temporal inter-dependency can be better explained, “[...] if we distinguish between the *present future* and *future presents*. Every present has its own future as an open horizon of its future possibilities. One envisions a future of which only a selection can become present in the future. In advancing into the future, by selecting from these possibilities, it produces new presents and at the same time new future horizons for these presents.”^{vii}

In the horizon of a differently documented and then also referenced past (in the *public ledger*), the possibility of future connection operations and further overall systemic states is fostered. The documentation of the past(s) itself selects a possible future because the future then has to position itself as a derivation and has to fit in, so to speak, genealogically. Current (system) states impose future ones.

And precisely because the architecture of blockchain is structured in such a way that past system states are **necessary** premises of the present, the comparison to the concept of structural determinism from neurophysiology and radical constructivism seems plain. Traditionally, system states of software can be replaced by others largely without leaving any residue. The automated self-verification of data sets of the blockchain determines their respective past system states and integrates a verification of the past into the present. One could say that the blockchain is, by verifying its present, constantly pushing itself into the future.

Because the reference to past system states is *trustless* and automated, a problem of trust relations that would otherwise exist is eliminated. With a *public ledger* with a functioning consensus mechanism, there is no need to “bet” on the trustworthiness of the current state of the system, which would otherwise usually be the case with orientations regarding a future state of the system. (Ironically, the verification of trustworthiness is performed by the verifiers betting among themselves. Hashpower competes with each other; and for example miners bet against each other in the Proof of Work algorithm; trustworthiness-creating consensus about the state of the system is based on the activity of monetarily incentivized mistrusters.) Luhmann: "All planning and projections of future present, all indirect, long-term mediated, detour-conceived orientations remain problematic from the point of view of trust and require a reference to the present, in which they must be anchored."^{viii}

Because the introduction of *trustless public ledger's* referable presence creates a point of reference for any user, planning for a particular user can be done with less consideration of external factors. If everyone reliably has the same point of reference, there is no longer any need to exchange

information among each other constantly. This creates (ironically) a centralizing network effect towards this one informative node: the *public ledger* itself. The redundancy of trust, and thus trustworthiness, among users culminates in the (absolute) trust of that which makes trust redundant: trust in the *public ledger*. In such a scenario, the consensual domain of the *public ledger* enjoys extensive trust. And assuming that it can create adequate trustworthiness for the purpose of automating its own bindingness, this results in cognitive relief for the user on the one hand and, on the other hand, in mediated efficiency of all trust-related operations influenced by it.

Luhmann's definition of trust as a mechanism of the reduction of social complexity also applies, even though the technology, by its condition, minimizes trust. While *trustlessness* can exist and be considered a fundamental characteristic of the innovation itself, the fact of trusting in something that seeks to maximize trustworthiness though still persists. Because in practical application it is impossible for all users to understand, what creates the trustworthiness of that in which they trust. The point is: trust is necessarily gradually blind. And herein lies Luhmann's (auto-logical) argument: because mistrusting inspection will fail for practical reasons, social fabrics tend to remain remains. Also, to protect against cognitive overload, social systems operate based on trust in something or someone without ongoing verification and/or affirmative endeavors of why that is the case. The assumption that the why is answered is mostly inherent in the case of use.

Ultimately, the application of *public ledgers* relies upon the constitution as "socially enforced networks" and one must bear in mind that growth phases increase exponentially. For a network of users that depends on trust, this means: the highest rate of growth in total users correlates with the reflexive effect of trusting in trusting others; and not necessarily with a trustworthiness that emanates from the matter.

Thus, trust in technology does not arise solely because of its inherent trustworthiness, but rather comes increasingly from trust that is already being trusted. And insofar the use of the technology is interwoven with social processes and established patterns of behavior trustworthiness is created by that.

VI - Conclusion: Non-Linearity and Self-Reference of Trust Relations

In their fundamental architectural design, *public ledgers* appear as closed systems; and this not solely in consequence of a genesis of functional differentiation, but because their purpose is to document the record of progressively re-allocated (most often financial) values. The *public ledger* is designed as a closed system. All of the possible states of the *public ledger* are arrangements of the

elements that it exists of, including the memory logging that contingents each state of the system. Guaranteeing and preserving trustworthiness and thus social suitability as a "consensual domain" for (financial) applications consists in the fact that elements are neither added nor removed in a quasi-arbitrary manner, but that the "rules of the game" are preserved and remain largely "immutable". Users can trust because they can expect future system states to be a re-allocation of the elements of the current state. Although not every user understands but trusts how these states are generated, they can trust the *public ledger* because it is subject to rigid rules and not just because, and that would be a difference to traditional ledger systems, all users are in recursive trust relations with other users and thus operate with and on the dispositive "consensual domain" constituted by the respective *public ledger*. A crucial basis of trust in a *public ledger* is the mistrustfulness among the users, which does not go hand in hand with a restriction of its functionality, but with an increase in its suitability and usefulness. Paradoxically, the use of a *public ledger* is especially recommendable, if trading parties *do not* trust each other.

The *public ledger* therefore gets by with a reduction in trust between users in social terms; but not completely in relation to the technology as such. Where other financial transaction systems are structured in a trust-based manner in both dimensions, Web3 attempts to introduce *trustlessness*, id est to approach the dependency on trust relations between users asymptotically towards its redundancy, and thus: to reduce social complexity.

- i Nakamoto, Satoshi. *Bitcoin: A Peer-to-Peer Electronic Cash System*. bitcoin.org/bitcoin.pdf
- ii Buterin, Vitalik. *Proof of Stake*. New York: Seven Stories Press, 2022, p. 289
- iii Luhmann, Niklas. *Vertrauen*. Konstanz: UVK Verlagsgesellschaft GmbH, 2014, p. 90, translated by the author
- iv Buterin, Vitalik. *Proof of Stake*. New York: Seven Stories Press, 2022, p. 289
- v Luhmann, Niklas. *Vertrauen*. Konstanz: UVK Verlagsgesellschaft GmbH, 2014, p. 9, translated by the author
- vi Ibidem p. 1, translated by the author
- vii Ibidem p. 14, translated by the author
- viii Ibidem p. 15, translated by the author