Zeitschrift für Praktische Philosophie
Band 11, Heft 1, 2024, S. 475–510
www.praktische-philosophie.org
https://doi.org/10.22613/zfpp/11.1.19

„Trustless trust?“ – Zum Begriff des Vertrauens im Rahmen von Blockchainanwendungen

„Trustless Trust?“ – On the Concept of Trust in the Context of Blockchain Applications

Eva Pöll, Münster & Katja Stoppenbrink, München

Zusammenfassung: Gegen die in der Literatur zu Blockchainanwendungen verbreitete These, diese erforderten kein nutzerseitiges Vertrauen oder generierten ein neuartiges „vertrauenloses Vertrauen“, wird in diesem Beitrag zunächst gezeigt, dass auch in der Nutzung von Blockchainanwendungen Vertrauensbeziehungen eine Rolle spielen. Es wird dafür argumentiert, Vertrauen im Rahmen von Blockchainanwendungen als ‚Institutionenvertrauen‘ zu verstehen. Das klassische bilaterale interpersonale Vertrauensverständnis von Vertrauenssubjekt und Vertrauensobjekt bleibt dabei strukturell erhalten, die Zuschreibung von Vertrauenswürdigkeit durch das Vertrauenssubjekt erfolgt in einem default-and-challenge-Modell. Dabei zeigt sich bereits aus begrifflichen Gründen: Je ‚sicherer‘ das System, desto weniger bedarf es nutzerseitigen Vertrauens. Vertrauen setzt Vulnerabilität voraus. Diese ist mit Blick auf Blockchainanwendungen bei den meisten Nutzer:innen in hohem Maße vorhanden.

Schlagwörter: Vertrauen, Blockchain, Vertrauenswürdigkeit, Institutionenvertrauen, Technikethik

Abstract: Against the thesis sometimes asserted in the literature on blockchain applications that they do not require user trust or generate a new type of “trustless trust”, it is shown here that trust relationships also play a role in the use of blockchain applications. It is argued that trust in the context of blockchain applications is best understood as ‘institutional trust’. The classical bilateral interpersonal understanding of trust between trustor (trust subject) and trustee (trust object) remains structurally intact, the attribution of trustworthiness by the trust subject takes place

in a default and challenge model. This is already evident for conceptual reasons: The ‘safer’ the system, the less it requires trust on the part of the user. Trust presupposes vulnerability. With regard to blockchain applications, most users have a high degree of vulnerability.

Keywords: Trust, Blockchain, Trustworthiness, Institutional Trust, Ethics of Technology

1 Einleitung

Wissenschaftlich-technische Entwicklungen fordern unsere Begriffe und unsere ethischen Wertungen und normativen Einschätzungen heraus. Das ist eine übliche Diagnose in der Technikethik und zeigt sich aktuell auch mit Blick auf die Blockchaintechnologie und den Begriff des Vertrauens. Die einschlägige Forschungsliteratur dazu ist bislang vornehmlich technisch oder ökonomisch geprägt. Beratungsunternehmen eruieren in kommerzieller oder bestenfalls ‚grauer Literatur‘ die wirtschaftlichen Potenziale der Nutzung von Blockchain für unterschiedliche Anwendungen. Eine genuin ethische Perspektive wird bei der Beurteilung von Blockchain selten[1] eingenommen, sieht man von der pauschalen Zurückweisung etwa des Marktes für sogenannte Kryptowährungen auf der Grundlage von Blockchaintechnologie (à la ‚Das ist doch nur etwas für Kriminelle, die sich staatlicher Kontrolle und staatlichem Zugriff entziehen wollen‘) einmal ab. Dabei geht es im Rahmen der vorhandenen Literatur oftmals um eine neue Art oder gar die Möglichkeit eines Verzichts auf ‚Vertrauen‘[2] im Rahmen von Blockchainanwendungen. Beispielsweise wird das auf den ersten Blick paradox anmuten-

1 Zu ethischen Aspekten von Blockchainanwendungen vgl. etwa Lapointe und Fishbane (2019), Dierksmeier und Seele (2020) sowie Kirchschläger (2021). Philosophische, darunter auch ontologische Fragen adressiert das von Swan und De Filippi (2017) zusammengestellte Symposium in der Zeitschrift *Metaphilosophy*.

2 Wir referieren hier zunächst auf einen unspezifischen, in Alltagssprache und Literatur vorgefundenen Vertrauensbegriff, auf den wir uns als Begriff beziehen. Dies geschieht in einfachen Anführungszeichen. Im Zuge der vorliegenden Auseinandersetzung nehmen wir eine unseren Zwecken entsprechende Bestimmung des Vertrauensbegriffs vor. Auf diesen von uns eingeführten Vertrauensbegriff beziehen wir uns dann ohne Anführungszeichen. Aus Gründen besserer Lesbarkeit ist bei wiederholten Bezugnahmen in einem Absatz zumeist auf einfache Anführungszeichen verzichtet worden. Sollte dieser Leit-

de Schlagwort von „trustless trust“[3] als Vorteil von Blockchainanwendungen angepriesen. Andere verkünden oder fordern, dass Vertrauen fortan keine Rolle mehr spielen und durch Blockchainnutzung ‚ersetzt‘ werden könne bzw. müsse.[4] Diese vollmundigen Äußerungen oder Versprechungen wecken das Interesse der Philosophin oder des Philosophen: Es scheint uns lohnend, den Verheißungen in puncto Vertrauen, die mit der Blockchaintechnologie in Verbindung gebracht werden, genauer nachzuspüren. Dabei steht nicht so sehr die technische Dimension, die Funktionsweise einer Blockchain, im Fokus, sondern die begrifflich-konzeptuelle Dimension von ‚Vertrauen‘ im Zusammenhang mit Blockchainanwendungen. Die Forschungsfrage, die uns bewegt, lautet daher, ob sich tatsächlich, wie in der Literatur behauptet, von einem ‚Ersatz‘ von Vertrauen durch Blockchain sprechen lässt und – sofern dies nicht zutreffen sollte – was Vertrauen denn im Rahmen von Blockchainanwendungen bedeuten mag. Wir weisen *erstens* die These, Blockchainanwendungen kämen ohne ‚Vertrauen‘ aus („trustless trust“) zurück und argumentieren *zweitens* dafür, dass sich ‚Vertrauen‘ im Rahmen von Blockchainanwendungen adäquat als ‚Institutionenvertrauen‘ verstehen lässt. Zur Begründung werden wir zunächst auf das klassische Verständnis von ‚Vertrauen‘ in der – zumeist philosophischen – Literatur eingehen (2) und uns dann ausführlicher mit den Positionen befassen, die zum Zusammenhang von Blockchain und Vertrauen gegenwärtig vertreten werden (3). Dabei lassen sich neben der unseres Erachtens zurückzuweisenden Extremposition (‚kein Vertrauen erforderlich‘) unterschiedliche Verständnisse des im Rahmen von Blockchainanwendungen erforderlichen Vertrauens herausarbeiten. Diese Verständnisweisen teilen die Annahme, dass auch im Rahmen der Nutzung von Blockchainanwendungen Vertrauen erforderlich ist. Wir schließen uns dieser Annahme im Grundsatz an und plädieren in der Folge

idee der Verwendung von Anführungszeichen nicht durchgehend entsprochen werden, so bitten wir um Nachsicht.

3 Geprägt durch Hoffman (2014).

4 So spricht die vermutete – bis heute anonyme – unter dem Pseudonym Satoshi Nakamoto firmierende Autorengruppe des sogenannten Bitcoin White Paper, die als ‚Erfinder‘ der Kryptowährung Bitcoin gilt, von „a system for electronic transactions without relying on trust“ (Nakamoto 2008, 8). Vgl. auch Tang et al. (2020, 602f.): „[…] blockchain is a trustless computing infrastructure for large-scale applications that have the potentials to solve fundamental trust issues and allows financial transactions without intermediaries.“

für ein Verständnis von Blockchainvertrauen als ‚Institutionenvertrauen‘. Ein kurzes Fazit mit Blick auf die Perspektive von Nutzer:innen von Blockchainanwendungen – insbesondere in der Rolle von Verbraucher:innen – beschließt diesen Beitrag (4).

2 Begriffliche Grundlagen

Zunächst gehen wir auf den Begriff des ‚Vertrauens‘ in seinem traditionell vorherrschenden Verständnis in der philosophischen und zum Teil auch soziologischen Literatur ein und unterscheiden im Anschluss daran ‚Vertrauen‘ und ‚Vertrauenswürdigkeit‘ (2.1). Im Weiteren soll ausgehend von dem ‚klassischen‘ Verständnis von ‚Vertrauen‘ kurz untersucht werden, wie Vertrauen im Umgang mit oder in Bezug auf Technik verstanden wurde und wird (2.2). Die Fundstücke dieser kurzen Literaturschau können bereits als technikphilosophische, begriffsanalytische oder begriffsgeschichtliche Antworten auf die eingangs angesprochene Herausforderung (vormals) vorherrschender Begriffsverständnisse durch Technikentwicklung angesehen werden.

2.1 Vertrauen und Vertrauenswürdigkeit

‚Vertrauen‘ wird traditionell als Tugend, als Haltung, als mentaler Zustand bzw. genauer als mentale Einstellung verstanden, daneben aber auch als Bezeichnung für die Beziehung zwischen zwei oder mehr Personen, wenn zumindest bei einem der Relata, dem Vertrauens*subjekt*, die entsprechende mentale Einstellung (‚vertrauend‘) vorliegt. Ausgangspunkt für die (nicht nur philosophische) Diskussion um den Vertrauensbegriff ist eine *interpersonale* Beziehung, die zumindest auf einer Seite durch eine charakteristische Haltung gekennzeichnet ist, während die andere Seite, das Gegenüber (oder in einem Mehrpersonenverhältnis: die Gegenüber) als bloßes Vertrauens*objekt* diese Haltung nicht notwendigerweise selbst auch haben muss. Ein reziprokes Vertrauen ist selbstverständlich denkbar, möglicherweise sogar der phänomenologische Normalfall, wenn beide Relata sowohl Vertrauenssubjekt wie -objekt sind. Die Redeweise, dass zwei Personen *einander* vertrauen oder der Appell, man möge *einander* vertrauen, ist üblich und geläufig.

In der Literatur werden unterschiedliche Bezugsdimensionen, eine schwache und eine starke Lesart, der Vertrauensrelation unterschieden. Bei einem starken Verständnis vertraut man jemandem umfassend als Person. Formalisiert kann man von Vertrauen als einer zweistelligen Relation sprechen, welche die vertrauensschenkende Person (A) und die Person, der ver-

traut wird (B), als Relata verbindet. Die schwächere Lesart hingegen bindet Vertrauen explizit in einen Kontext ein: Statt jemandem grundsätzlich zu vertrauen, vertraut man der Person, in einem gegebenen Kontext etwas ganz Bestimmtes zu tun (oder zu unterlassen). Zu der Formalisierung kommt ein weiteres Relat, nämlich φ, als Beschreibung etwas (nicht) zu tun, hinzu: ‚A vertraut B zu φen' (McLeod 2021). Inwiefern diese Arten des Vertrauens zusammenhängen, ist Gegenstand aktueller Debatten. Die Überlegungen von Kelp und Simion zur Vertrauenswürdigkeit deuten darauf hin, dass die starke Lesart von Vertrauen in jemanden auf hinreichend viele Fälle von dünnem Vertrauen in dieselbe Person, hinreichend viele bestimmte Dinge zu tun, reduziert werden könnte (Kelp und Simion 2023).

Für McLeod (2021, einleitend) stellt Vertrauen als eine interpersonale Relation das „dominant paradigm of trust" dar. Andere Formen von Vertrauen (dazu sogleich und im nachfolgenden Abschnitt) variieren ihr zufolge das Grundmuster nur insofern, als unterschiedliche Relata eingesetzt werden, die mögliche Extension des Vertrauensobjekts über den Kreis menschlicher Personen hinaus erweitert oder durch andere mögliche Bezugsobjekte ersetzt wird. Andere Vertrauensrelationen, andere Formen von Vertrauen, seien „coherent only if they share important features of (i.e., can be modeled on) interpersonal trust" (ibid.).

Interpersonales Vertrauen erfordert nach der Standarddefinition (vgl. McLeod 2021) auf der Seite des Vertrauenssubjekts die folgenden notwendigen und zusammen hinreichenden Merkmale: (i) Vulnerabilität des Vertrauenssubjekts, (ii) Erwartung, dass das Vertrauensobjekt eine bestimmte Handlung durchführen kann – bzw. in der komplexeren Fassung von McLeod (sinngemäß) ‚sich darauf verlassen, dass das Vertrauensobjekt in der Lage sein wird, dasjenige zu tun, worauf wir vertrauen zu dürfen wünschen' („rely on others to be competent to do what we wish to trust them to do"; unsere Übers., d. Verf.) – sowie (iii) Sich-darauf-verlassen, dass das Vertrauensobjekt bereit sein wird, das auch zu tun. Vereinfacht lässt sich von einem Vertrauenssubjekt sprechen, das (i) darauf *angewiesen* ist, dass ein Vertrauensobjekt φ-en kann und das (ii) eine (optimistische) *Erwartungshaltung* hat, dass ein Vertrauensobjekt *φ-en kann* und das sich (iii) darauf *verlässt, dass* ein Vertrauensobjekt *φ-en wird*.

Die Vulnerabilitätsbedingung (i) ist besonders von Baier in ihrem zentralen Aufsatz zum Vertrauensbegriff (1986) herausgestellt worden. Bereits begrifflich setzt ‚Vertrauen' danach voraus, dass das Vertrauenssubjekt das Agieren des Vertrauensobjekts nicht kontrolliert oder kontrollieren

kann. Je mehr Kontrolle von Seiten des Vertrauenssubjekts ausgeübt wird, desto geringer wird der Raum, in dem das Subjekt dem Vertrauensobjekt vertrauen kann. Übernimmt das Subjekt die Steuerung, so vertraut es dem Objekt nicht, sondern lenkt dieses. Vertrauen setzt Handlungsspielräume auf Seiten des Vertrauensobjekts voraus. Entsprechend kann das Vertrauen des Vertrauenssubjekts auch durch nicht erwartungsgemäßes Handeln des Vertrauensobjekts oder durch den erwarteten Verlauf der Ereignisse störende, dazwischentretende dritte Ursachen enttäuscht werden. Nach Baier lässt sich Vertrauen daher als „accepted vulnerability in another's possible but not expected ill will (or lack of good will) towards one“ auffassen (Baier 1986, 235). Ähnlich geht auch Hartmann (2011) in seiner Vertrauensdefinition von drei notwendigen Merkmalen bzw. „wesentlichen Elementen einer plausiblen Beschreibung“ (2011, 82) eines Handelns als vertrauensvoll aus: ‚Vertrauen' lässt sich als (a) „relationale Einstellung“ (2011, 82), als (b) Handeln unter möglichen Handlungsalternativen („Verfügen über Optionen“, 2011, 86) und als (c) „Akzeptanz der durch Vertrauen ermöglichten Verletzungen“ bzw. „[a]kzeptierte Verletzbarkeit“ (2011, 99) verstehen.[5]

Die Merkmale (ii) und (iii) können neben der Vulnerabilitätsbedingung (i) als eine *Erwartungshaltung*, dass das Vertrauensobjekt eine Handlungsweise instanziieren *kann* (ii) und instanziieren *wird* (iii) verstanden werden. Mit anderen Worten: Das Subjekt nimmt an, dass das Objekt *φ-en kann* und *φ-en wird*. Es handelt sich also um eine *kompetenz-* und eine *performanz*bezogene Erwartungshaltung des Vertrauenssubjekts vis-à-vis dem Vertrauensobjekt. McLeod (2021) etwa bekennt sich zur schwachen Lesart von Vertrauen als der grundsätzlichen: „‚trust is generally a three-part relation: A trusts B to do X' […] —or ‚A trusts B with valued item C' […] or A trusts B in domain D“ [sic!, was die variierende Setzung von Anführungszeichen angeht]. Diese beiden letztgenannten Varianten mögen unseren alltäglichen Redeweisen entsprechen – dem Erzieher vertrauen wir in Bezug auf unser Kind, der Landschaftsbauerin in Bezug auf die Gartengestaltung usw. – doch lassen sich diese Formen von Vertrauen unseres Erachtens handlungstheo-

5 Hartmanns Bestimmung ist kompatibel mit der hier vorgeschlagenen, wenngleich er andere Schwerpunkte setzt. Logisch setzt die im folgenden zugrundegelegte Standarddefinition nach McLeod (2021) die drei Elemente der Beschreibung nach Hartmann voraus. Darauf kann hier nicht weiter eingegangen werden, da wir nur den interpersonalen Charakter der Standarddefinition hervorheben, keine vergleichend-exegetische Studie einzelner Forschungsbeiträge vorlegen möchten.

retisch auf die erstgenannte Form zurückführen: Wir vertrauen stets, indem wir erwarten, dass jemand etwas Bestimmtes (nicht) tun wird. Das Handlungssubjekt vertraut darauf, dass das Handlungsobjekt (*nicht) φ-en wird*. Die Kompetenzzuschreibung mag dabei zwar nicht fehlen, aber graduell unterschiedlich ausfallen, die Handlungserwartung hingegen ist grundlegend für – auch andere Formen von – Vertrauen.

Von der *begrifflichen* Frage nach den notwendigen und hinreichenden Merkmalen von ‚Vertrauen' ist die *normativ-evaluative* Frage zu unterscheiden, wann Vertrauen gerechtfertigt ist bzw. wann ein:e Akteur:in für das Entgegenbringen von Vertrauen, für die Vertrauen konstituierende Erwartungshaltung Dritten gegenüber, gute Gründe hat. Dies lässt sich deuten als die Frage nach der Vertrauenswürdigkeit (vgl. einleitend McLeod 2021): „that it [das zum Ausdruck gebrachte Vertrauen] successfully targets a trustworthy person".

Vertrauens*würdigkeit ist ein notwendiges (und je nach Konzeptualisierung*[6] auch hinreichendes) Merkmal für die Rechtfertigung des Vertrauens vertrauender Personen. In der 1:1-Vertrauensrelation kommt es auf die Vertrauenswürdigkeit des Gegenübers einer oder eines Vertrauenden an. Vertrauenswürdigkeit (engl.: *trustworthiness)* ist nach der hier eingeführten Terminologie eine Eigenschaft eines Vertrauensobjekts *(trustee)*, welche für das Vertrauenssubjekt *(trustor)* insofern zentral ist, als das Vorliegen dieses Merkmals die Vertrauenshaltung auf Seiten des Vertrauenssubjekts rechtfertigt. Ist das Vertrauensobjekt *vertrauenswürdig,* hat das Vertrauenssubjekt *gute Gründe* zu vertrauen.

Während Vertrauenswürdigkeit als ‚dünner Begriff' verstanden werden kann, für den es schlicht ausreicht, dass es gute Gründe für ein Vertrauen des Vertrauenssubjekts gibt, ohne dass eine inhaltlich reichhaltigere Bestimmung als erforderlich angesehen wird, gibt es in der Literatur auch divergierende Auffassungen, die substanzielle Bedingungen an Vertrauenswürdigkeit formulieren und diesen Begriff anspruchsvoller gestalten. So kann für Vertrauenswürdigkeit eine bestimmte Kompetenz auf Seiten des Vertrauensobjekts verlangt werden, ein Motiv oder eine Haltung entweder gegenüber dem Vertrauenssubjekt oder allgemeiner gegenüber der Allgemeinheit

6 McLeod (2021, einleitend) unterscheidet etwa, ob Vertrauen „warranted" (bei vorhandener Vertrauenswürdigkeit) oder nur „justified" ist (das kann nach dieser Auffassung der Fall sein, wenn eine epistemische Rechtfertigung auf Seiten der Vertrauenden vorliegt, aber tatsächlich keine Vertrauenswürdigkeit auf Seiten des Gegenübers).

(*erga omnes*) oder einer bestimmten Gruppe von Personen. Lehrer:innen etwa kann grundsätzlich von Seiten der Schüler:innen vertraut werden, weil es deren gesetzliche Aufgabe ist, sich auf bestimmte Arten und Weisen um die Schüler:innen zu kümmern (Erziehungs- und Bildungsauftrag). Soll nur ein Wohlwollen des Vertrauensobjekts vis-à-vis dem Vertrauenssubjekt für Vertrauenswürdigkeit des Vertrauensobjekts erforderlich sein, ergibt sich – ungeachtet des bei motivbezogenen Auffassungen stets auftretenden epistemischen Problems ((wie) kann das Subjekt Kenntnis über die Motivlage des Objekts haben?) – eine Unterinklusivität: Zu viele Phänomene, die wir in unserer Praxis als Vertrauensrelation ausweisen würden, könnten diese Bedingung an Vertrauenswürdigkeit des Vertrauensobjekts nicht erfüllen. Oftmals sind uns die Motive schlicht ‚egal‘, da wir lediglich eine bestimmte *Handlung* von Seiten des Vertrauensobjekts erwarten. Gegen diese in ihren inhaltlichen Ansprüchen an Vertrauenswürdigkeit abgesenkte Position wird oftmals vorgebracht, sie unterscheide analytisch nicht angemessen zwischen bloßer Zuverlässigkeit und echter Vertrauenswürdigkeit des Vertrauensobjekts. Dies ist ein Einwand, der nicht von der Hand zu weisen, aber u. E. unschädlich ist. Ein Vorteil dieser Sichtweise liegt in ihrer Inklusivität, was die Natur des Vertrauensobjekts angeht. Je weniger anspruchsvoll die Vertrauenswürdigkeitseigenschaft verstanden wird, desto mehr kann die Extension des Vertrauensobjekts erfassen, desto mehr mögliche Relata können in die Vertrauensrelation einbezogen werden. Der Verzicht auf die Motivlage des Vertrauensobjekts als eine notwendige Bedingung für dessen Vertrauenswürdigkeit ermöglicht erst die Redeweise von ‚Vertrauen in Institutionen‘ oder ‚Vertrauen in Technik‘ als Derivate der paradigmatischen interpersonalen Vertrauensbeziehung.

Eine Familie möglicher Einwände gegen das hier skizzierte Verständnis von ‚Vertrauen‘ lässt sich als ‚Moralitätsauffassung‘ zusammenfassen. Die Moralitätsauffassung kann unterschiedliche Aspekte hervorheben: Vertrauen als Haltung eines Vertrauenssubjekts kann nicht einfach nur fehlgehen und enttäuscht, sondern hintergangen und das Vertrauenssubjekt betrogen werden. Diese Möglichkeit betont Baier in ihren Arbeiten zum Vertrauensbegriff (1986 und 1991). Die Moralitätsauffassung versteht die interpersonale Vertrauensbeziehung damit als zumindest in manchen Hinsichten reziprok, d. h. auch auf Seiten des Vertrauensobjekts wird eine bestimmte Intentionalität unterstellt oder gefordert (welche als ein Sich-bemühen-um-Vertrauen, Sich-sorgen-um-das-Vertrauenssubjekt usw. interpretiert werden kann); auch das Vertrauensobjekt muss für manche Varianten der Mora-

litätsauffassung eine bestimmte Haltung vis-à-vis dem Vertrauenssubjekt manifestieren. *Vertrauenswürdigkeit* wird z.T. auch als eine charakterliche Disposition, mit anderen Worten: eine *Tugend*, aufgefasst. In Summe wird die Zuschreibung von Vertrauenswürdigkeit in der Moralitätsposition damit anspruchsvoller als in der hier vorausgesetzten ‚dünnen' Konzeptualisierung von Vertrauenswürdigkeit. Viele Phänomene und grundsätzlich taugliche Vertrauensobjekte fallen damit aus der Extension von ‚Vertrauen', ‚Vertrauensobjekt' und ‚Vertrauenswürdigkeit' heraus. Dieser letztgenannte Gesichtspunkt ist zugleich der wichtigste Grund, die Moralitätsauffassung zurückzuweisen. Sie baut zu sehr auf der Grundlage der idealtypischen interpersonalen Beziehung als Standardmodell der Vertrauensrelation auf und erschwert die Rede von ‚Technikvertrauen', ‚Vertrauen in Technik' oder auch – allgemeiner – ‚Vertrauen in Institutionen'. Da wir in unserer Vertrauenspraxis diese Phänomene vorfinden und mit fortschreitenden wissenschaftlich-technischen Entwicklungen eher noch mehr, jedenfalls nicht weniger von ‚Vertrauen in Technik' sprechen werden und sprechen können sollten, spricht dies für Konzeptualisierungsvorschläge, denen es gelingt, diese Sichtweisen extensional einzubeziehen – unter der Voraussetzung, dass intensional plausible und nicht völlig inkompatible oder verzerrende Vorschläge gemacht werden. Wir behaupten, dass wir mit dem Rückgang auf das Gerüst der drei notwendigen und zusammen hinreichenden begrifflichen Merkmale von ‚Vertrauen' und eine *default-and-challenge*-Auffassung von Vertrauenswürdigkeit ein zwar ‚dünnes', aber adäquates, sowohl unsere traditionell-interpersonalen Vertrauenspraxen als auch neuere Vertrauensextensionen erfassendes Vertrauensverständnis vorschlagen.

Die soziologische Literatur zur Analyse der gesellschaftlichen Funktion von Vertrauen unterstreicht diese Sichtweise, dass es nicht auf die Motivlage des Vertrauensobjekts ankommt. Vertrauen garantiert gerade in Situationen epistemischer Opazität die Handlungsfähigkeit des Vertrauenssubjekts. Misztal (1996) beispielsweise interpretiert Vertrauen als verinnerlichter Habitus, der in alltäglichen Routinen Stabilität und Schutz schaffe und funktional soziale Ordnungen aufrechtzuerhalten helfe, da er bei den Individuen als Vertrauenssubjekten eine Vorhersehbarkeitsannahme hervorrufe. Während individuelle Vorhersehbarkeitsannahmen der Vertrauenssubjekte sich als Irrtum herausstellen können, wird auf soziale Reputation und implizite kollektive Erinnerungen als Prädiktoren abgestellt, die individuelle Erwartungshaltungen rechtfertigen können und zugleich den Individuen eigene epistemische Überprüfungen der Vertrauensobjekte ersparen.

Damit werden qua Vertrauen soziale Interaktionen vereinfacht und Routinehandlungen ermöglicht.[7]

2.2 Vertrauen und Technik

Wie ist Technikvertrauen möglich, was impliziert es? Zunächst ist allgemeiner von Vertrauen in Technik oder Systeme zu sprechen; ein Vertrauenssubjekt schreibt einem Vertrauensobjekt eine bestimmte Funktionalität zu und erwartet eine bestimmte Ausführung oder Instanziierung dieser Funktionalität. Dabei wird die Ausführung als regelhaft verstanden: wenn x – dann y. Der Technikbegriff[8] orientiert sich hier weniger am Artefakt, denn an der Regelhaftigkeit. Passend dazu muss Vertrauen hier in der dünnen Lesart, als dreistellige Relation, verstanden werden. Die umfassendere, starke Lesart (Vertrauen in eine Person als Person) kann nicht auf das Vertrauen in Technik ausgeweitet werden, da Technik nicht intentional handeln und ihr keine Handlungsfähigkeit (*agency*) zugesprochen werden kann. Sie kann weder guten Willen zeigen (wie Baier (1986) es voraussetzt), noch Verpflichtungen (engl.: *commitment*) gegenüber der vertrauenden Person eingehen (wie Hawley (2014) und darauf aufbauend Lipman (2023) fordern). Durch das charakterisierend regelhafte Verhalten sind dennoch Erwartungen gegenüber der Funktionalität von Technik gerechtfertigt (Nickel 2013).

Dabei ist Vertrauen in Technik mehr als reines ‚Verlassen auf Technik‘ (*reliability*). Dass Vertrauen eine normativ reichhaltige Einstellung beschreibt, zeigt die Reaktion des Vertrauensgebers, wenn ein Artefakt nicht die erwartete Funktionalität aufweist: Man reagiert nicht nur mit Neugier und Überraschung, wenn der Drucker ein Dokument nicht ausdruckt oder wenn ein Textverarbeitungsprogramm den Fortschritt der letzten zwei Stunden verliert, sondern eher mit Frustration und Wut – ein Beweis für eine „richer normative attitude“ (Nickel 2013, 5) und mehr als bloß ein Urteil der Zuverlässigkeit (Nickel 2013). Dabei handelt es sich nicht nur um eine epistemische Haltung (wie es bei Verlassen auf eine Technik der Fall wäre), sondern vielmehr um eine emotional und normativ aufgeladene Haltung – ähnlich interpersonalem Vertrauen. Das bedeutet wiederum nicht, dass jede Interaktion zwischen Mensch und Technik auf Vertrauen beruht. Erst wenn

7 Vgl. z. B. Misztal (1996, 103): „Trust as habitus is a protective mechanism relying on everyday routines, stable reputations and tacit memories.“

8 Zum Technikbegriff und zur Regelhaftigkeit vgl. Grunwald und Juillard (2005).

die entscheidenden Voraussetzungen zusammenkommen, kann Vertrauen entstehen. McLeod (2021) folgend muss die vertrauensgebende Person sich in einer vulnerablen Position gegenüber der Technik befinden und nicht sicher voraussagen können, ob die erwartete Funktionalität tatsächlich vom Artefakt gezeigt wird (oder werden kann) (siehe (i) oben). Gleichzeitig muss es aber hinreichend sichere Hinweise geben, dass das Artefakt generell in der Lage ist, die Funktionalität zu zeigen (ii). Das ist insbesondere gegeben, wenn der durch die Entwickler:innen intendierte Verwendungszweck des Artefakts klar ist. Und schließlich muss sich der:die Nutzer:in auch tatsächlich auf das Artefakt verlassen (iii). In der Literatur lassen sich typisierend mindestens drei unterschiedliche Positionen zum ‚Technikvertrauen' unterscheiden.

Eine erste Position spricht affirmativ und ohne Umschweife von Vertrauen in Technik – mal verstanden als Artefakte, mal verstanden als Algorithmen und damit bestimmte (regelhafte) Prozesse (a). Eine weitere Position ist grundsätzlich skeptischer. Ihr zufolge ist – angesichts der als interpersonal verstandenen Vertrauensrelation – Technikvertrauen grundsätzlich kritisch zu betrachten und allenfalls im übertragenen oder analogen Sinne zu verstehen (b). Vielversprechender scheint auch mit Bezug auf Technik die institutionelle Sichtweise zu sein (c).

(a) Technikvertrauen als Vertrauen in Algorithmen

Technik als Gegenüber sorgt sich nicht um unser Wohl. Die Regelhaftigkeit von Technik könnte zwar in struktureller Analogie zu dispositionalen Eigenschaften wie etwa Tugenden verstanden werden; doch lässt sich das Modell der interpersonalen Relationen mit mentalen Zuständen auf beiden Seiten nicht in einem inhaltlich reichhaltigen Verständnis auf die Mensch-Technik-Interaktion übertragen. Technik bzw. technischen Systemen lassen sich grundsätzlich keine mentalen Eigenschaften wie Motive zuschreiben; ein motiv-basierter Ansatz des Vertrauensverständnisses stößt hier an Grenzen. Allenfalls kann auf Entwickler:innen- oder Hersteller:innenmotive abgestellt werden. Eine Ausnahme stellt möglicherweise eine als ‚intentional' beschreibbare, in einem sinnvollen, substanziellen Sinne ‚autonom' agierende Künstliche Intelligenz (KI) dar, die sich nach dem gängigen Verständnis als ‚starke KI' beschreiben lässt. Im Rahmen allgemeiner Technikethik und mit Blick auf die Frage nach einer angemessenen Konzeptualisierung von Technikvertrauen ist die Zuschreibung mentaler Eigenschaften im Zusammenhang mit KI von nicht nur theoretischem Interesse, doch sind diese Punkte

hier mit Blick auf Blockchainanwendungen (siehe sogleich in Abschnitt 3) nicht zu erörtern.

Relevante Motive auf Seiten des Vertrauensobjekts können in unserem Kontext allenfalls diejenigen der Blockchainbetreiber:innen oder -anbieter:innen sein. Hier lassen sich anhand der Unterscheidungen *public permissionless* vs. *permissioned* und *private* (siehe näher unter 3.1) unterschiedliche Idealtypisierungen vornehmen. Es könnte auch eine ‚staatliche Blockchain‘ geben, die im Rahmen von Leistungen der Daseinsvorsorge eingesetzt wird und die Interessen der Nutzer:innen (Bürger:innen, Einwohner:innen, Dienstleistungsadressat:innen) befördern soll. Eine solche Variante stellt gegenwärtig allerdings in einem nahezu vollständig von kommerziellen Akteur:innen beherrschten Umfeld eine Ausnahmeerscheinung dar.

Der Blockchain kann kein ‚guter Wille‘ zugeschrieben werden, auch keine sonstigen Motive, keine Benevolenzannahme vis-a-vis den Nutzer:innen. Sollen diese Anforderungen an den Vertrauensbegriff gestellt werden, so wird sich mit Blick auf Blockchain nicht von ‚Vertrauen‘ sprechen lassen. Blockchain könnte nicht nur nicht auf Vertrauen verzichten, sondern böte auch keine vertrauenslose Alternative, wie dies z.T. in der Literatur insinuiert wird (siehe sogleich).

(b) Sicherheit schafft kein Technikvertrauen

Schalten zuverlässig funktionierende Systeme die Risiken für das Vertrauenssubjekt weitgehend oder gar ganz aus, lässt sich begrifflich nicht mehr von Vertrauen sprechen, weil bereits die Vulnerabilitätsbedingung nicht erfüllt ist. Die Frage der Vertrauenswürdigkeit stellt sich dann gar nicht mehr. Schlagwortartig gilt daher, dass ‚Sicherheit‘ durch Technik somit ‚Vertrauen in Technik‘ ausschaltet – d. h. schon begrifflich unmöglich macht (vgl. Nissenbaum 1999; O’Neill 2020).

Es lässt sich im Anschluss an die klassische Idee Luhmanns (2014, 83f. und *passim)* von Vertrauen als Komplexitätsreduktion argumentieren, dass Vertrauen in Technik mehr noch als interpersonales Vertrauen eine Reaktion auf epistemische Opazität und damit einhergehende Risiken ist. Wo das Vertrauenssubjekt nicht sichergehen kann, wo es des für eine risikominimierende oder -ausschließende Beurteilung erforderlichen Wissens entbehrt, da ist es auf ‚Vertrauen‘ als Anerkennung der eigenen Vulnerabilität aufgrund epistemischer Defizite angewiesen. Wo die Schaffung der erforderlichen Wissensbasis prohibitiv hohe Transaktionskosten hervorriefe oder

sonst praktisch unmöglich ist, kann daher auch Technikvertrauen als Komplexitäts- (und damit: Kosten-)reduktion aufgefasst werden. Der Einfachheit halber bedient sich das Vertrauenssubjekt der Technik, macht die Technik zum Vertrauensobjekt und nutzt die Technik als – weiterhin risikobehaftetes – Mittel zum eigenen Zweck. Dabei wird oft übersehen, dass in den meisten Fällen Alternativen zum Technikeinsatz bestehen, das Vertrauenssubjekt auch auf den Technikeinsatz verzichten könnte und hinsichtlich des Kosten-Nutzen-Kalküls in der Entscheiderrolle bleibt. Das gilt freilich nur, solange es sich nicht um netzgebundene, (quasi-) monopolistische Strukturen auf Seiten potenzieller Vertrauensobjekte handelt. Im paradigmatischen Bereich von Social Media bleibt dem Vertrauenssubjekt oft nichts anderes übrig, als sich im vollen Bewusstsein der eigenen Vulnerabilität den vom Vertrauensobjekt erzeugten (Datenschutz-) Risiken auszusetzen.

Für Luhmann ist Vertrauen kein bloßes Zutrauen (*confidence*) oder Sich-verlassen-auf (*reliance*), sondern die bewusste Entscheidung, die eigene Vulnerabilität anzuerkennen, das Risiko in Kauf zu nehmen und zu ‚vertrauen', d. h. eine bestimmte Handlungsalternative gegenüber einer anderen, die eine andere Risikobeurteilung impliziert, zu bevorzugen. Phänomenologisch „ist Vertrauen zunächst personales" (2014, 27), welches sich mit steigender Komplexität der Lebenszusammenhänge in ein „Systemvertrauen" wandelt. Dies ist zugleich der Oberbegriff: „Systemvertrauen läßt sich nicht nur auf soziale Systeme, sondern auch auf andere Menschen als personale Systeme anwenden" (2014, 27), es verweist aber in seiner Rechtfertigungsdimension wiederum auf personales Vertrauen, denn nach Luhmann „liegt die rationale Basis des Systemvertrauens im Vertrauen in das Vertrauen anderer" (2014, 92).

Giddens (1990, 31) fasst Luhmann sehr treffend zusammen:

> Where trust is involved, in Luhmann's view, alternatives are consciously borne in mind by the individual in deciding to follow a particular course of action. Someone who buys a used car, instead of a new one, risks purchasing a dud. He or she places trust in the salesperson or the reputation of the firm to try to avoid this occurrence. Thus, an individual who does not consider alternatives is in a situation of confidence, whereas someone who does recognise those alternatives and tries to counter the risks thus acknowledged, engages in trust.

Der Gebrauchtwagenkauf als sprichwörtlicher „Market for Lemons" (Akerlof 1970) ist ein pointiertes Beispiel für eine der Standarddefinition entsprechen-

de Vertrauensrelation, welches zugleich die seitens des Vertrauenssubjekts eingegangenen Risiken unter Bestehen unterschiedlicher Handlungsoptionen verdeutlicht. Ohne ein ‚Auch-anders-handeln-können‘ lässt sich danach nicht von Vertrauen sprechen. Giddens selbst schließt sich dieser Analyse explizit (1990, 32: „unhelpful“; vgl. auch 1990, 90) nicht an, sondern nimmt eine eigene Bestimmung in zehn Punkten vor (1990, 33–36), von denen hier der dritte und der vierte Punkte besonders interessieren:

> 3. Trust is not the same as faith in the reliability of a person or system; it is what derives from that faith. […] All trust is in a certain sense blind trust! 4. We can speak of trust in symbolic tokens or expert systems, but this rests upon faith in the correctness of principles of which one is ignorant, not upon faith in the ‚moral uprightness‘ (good intentions) of others. Of course, trust in persons is always to some degree relevant to faith in systems, but concerns their proper working rather than their operation as such.“ (1990, 33f.)

Vertrauen lässt sich danach nicht auf Sich-verlassen-auf reduzieren, sondern setzt es voraus (vgl. die Bedingung (iii) der unter 2.1 eingeführten Standarddefinition). Systemvertrauen wird von Giddens als ‚Expertenvertrauen‘ rekonstruiert und die Moralitätsauffassung (s.o. unter 2.1) ebenso wie hier zurückgewiesen. Systemvertrauen oder Vertrauen in „symbolic tokens“ wie etwa Geld verweisen auf die weitere Option, Technikvertrauen zu verstehen, nämlich als ‚Vertrauen in Institutionen‘.

(c) Technikvertrauen als Vertrauen in Institutionen

Im Anschluss an Luhmann und Giddens lässt sich Technikvertrauen auch als ‚Vertrauen in Institutionen‘ verstehen. Damit wird die interpersonale Vorstellung inhaltlich-substanziell, nicht aber strukturell aufgeben. Während die beiden genannten Soziologen Systemvertrauen im Ergebnis auf Expertenvertrauen zurückführen, lässt sich ‚Institutionenvertrauen‘ zunächst auch direkt als ‚Vertrauen in Institutionen‘ auffassen. Vertrauensobjekt eines Institutionenvertrauens ist nicht ein regelhaft operierender, als Gegenüber oder Artefakt gar nicht (be)greifbarer Algorithmus (vgl. unter a), sondern eine Einrichtung, die ihrerseits zwar auch bestimmten (Funktions-) Regeln folgt, die aber nach einem bestimmten (normativen) Prinzip konstituiert ist (und z. B. ein Gründungsdokument, ein bestimmtes Gesetz oder sonstige Regulatorik als Referenz- oder Ausgangspunkt aufweist), deren Ratio bzw. Zweckhaftigkeit normativ festgelegt und deren Funktionsweise grundsätz-

lich nachvollziehbar und epistemisch zugänglich ist. Teng (2021, 389) stellt dies deutlich heraus: Institutionenvertrauen ist anders als das klassische interpersonale Modell „non-partner-relative", bezieht sich auch nicht auf ein Artefakt („non-thing-specific", ibid.), sondern lässt sich vielmehr als eine Erwartungshaltung konzeptualisieren, dass eine durch die Institution offerierte Dienstleistung akzeptabel und zuverlässig ausgeführt wird. Es geht um normative Erwartungen, die ein Vertrauenssubjekt einem Vertrauensobjekt ‚Technik' zuschreiben ‚darf'. Ist das Vertrauenssubjekt prädiktiv erfolgreich, wird das Vertrauensobjekt die erwartete Dienstleistung ausführen, so lässt sich *ex post* von gerechtfertigtem, anderenfalls von enttäuschtem oder fehlerhaftem Institutionenvertrauen sprechen. Teng spricht in diesem Sinne von „predictive and normative expectations" vis-à-vis Einrichtungen, die verstanden werden können als „entities carrying predefined normative qualities, such as moral, social, and legal norms" (2021, 390). Aufgrund der normativen Kraft, Dienstleistungsaussichten zu verbindlichen Zusagen zu machen, können die institutionellen Normen analog zu Versprechen, die gehalten werden müssen, aufgefasst werden (2021, 390). So wie wir etwa auch die Erwartung hegen, dass Bankangestellte sich den externen und internen Normen eines Unternehmens entsprechend verhalten, und dass das System ‚Bank' funktioniert, kann analog auch Technik (und damit auch eine Blockchainanwendung) normative Werte instanziieren und entsprechende Erwartungen erzeugen: „[…] technologies resemble institutions in their design capacity for carrying normative values and inviting relevant expectations about what they are supposed to do" (2021, 392). Diese Perspektive auf Technikvertrauen als auf Institutionen gerichtete normative Erwartung setzt keine interpersonale Beziehung voraus, das Vertrauensobjekt muss keine Kenntnis vom Vertrauenssubjekt haben, die Vertrauensbeziehung kann instantan, zeitlich distinkt und einmalig sein. Zugleich impliziert die normative Zuschreibung durch das Vertrauenssubjekt aber, dass es um mehr als bloße Verlässlichkeit in Bezug auf die Funktionstüchtigkeit der Technik geht.

Im Rahmen der Interaktion mit einer Institution (z. B. der Nutzung einer von ihr angebotenen Dienstleistung) hegen Vertrauenssubjekte prädiktive und normative Erwartungen, die nach der aktuellen Literatur auf drei Aspekten aufbauen: *erstens* frühere Erfahrungen mit der Institution (Lahno 2001; Möllering 2006), *zweitens* die wahrgenommenen strukturellen Absicherungen (wie Regeln, Richtlinien und Anreize), die bestimmte Verhaltensweisen fördern (Möllering 2006; Wingreen und Baglione 2005), und *drittens,* verbunden mit diesen Zusicherungen, die wahrgenommenen Wer-

te und Ziele, die von der Institution verfolgt werden (Teng 2021; Townley und Garfield 2013). Jede Interaktion mit einer Institution trägt zur Bildung des Rufs der Institution bei, der ein Indiz für ihre Vertrauenswürdigkeit ist (Alfano und Huijts 2020; Smits und Hulstijn 2020). Diese Interaktionen geschehen vermittelt durch Vertreter:innen der Institution als Intermediäre, die innerhalb interner Regeln, Rollen und Prozesse agieren, welche darauf ausgelegt sind, bestimmte Ergebnisse und Verhaltensweisen sicherzustellen (Lahno 2001; Wingreen und Baglione 2005). Institutionen integrieren dabei auch moralische Werte: „institutions can be understood as entities carrying predefined normative qualities, such as moral, social, and legal norms“ (Teng 2021, 390). Dementsprechend sind die Erwartungen, die ein Vertrauenssubjekt einer Institution als Vertrauensobjekt entgegenbringt, nicht nur prädiktiv, sondern auch normativ und dienen als Kriterien für die Einschätzung, ob das Vertrauen in die Institution gerechtfertigt ist (Teng 2023).

Informativ in diesem Zusammenhang ist auch der technikphilosophische Aufriss von Kaminski, welcher Technik als „Funktionierbarkeitserwartung“ (2010, 282f.) rekonstruiert. Damit ist der Weg eröffnet, unter Absehen von der Notwendigkeit der Zuschreibung mentaler Zustände, Handlungsintentionen, anderer normativer Erwartungen usw. Technik als Vertrauensobjekt zu begreifen. Die von uns zugrundegelegte dreigeteilte Standarddefinition ‚passt‘ in einem solchen Rahmen grundsätzlich auch auf Technik. Kaminski stellt auf die faktische ‚Nützlichkeit‘ von Vertrauen ab, denn für ihn ist

> [d]iese pragmatische Bedeutung von Vertrauen […] dem Unterschied Person oder Technik vorgelagert. Vertrauen in Personen wird vor diesem Hintergrund zu einer Spezifikation von Vertrauen. Man kann dann zwar noch immer zwischen Vertrauen in Personen und sich Verlassen auf Technik unterscheiden – das verwischt jedoch die pragmatische, grundlegende Bedeutung von Vertrauen.

Sein Fazit lautet: „Vertrauen in Technik ist daher kaum erklärungsbedürftiger als Vertrauen in Personen beziehungsweise: Beidem liegt dieselbe Erklärung zugrunde.“ (2010, 256) Es geht dabei um den Umgang eines Vertrauenssubjekts mit Risiko und Nichtwissen. In Kaminskis Worten: „Nur wenn beides zusammenkommt, das Vorhandensein von Risiko und Nichtwissen und ihr ‚außer Geltung sein‘, besteht Vertrauen“ (2010, 243). In einer pragmatischen Herangehensweise werden hier interpersonales Vertrauen sowie Technikvertrauen als zwei Ausprägungen von ‚Vertrauen‘ *tel quel*

ausgewiesen. Vertrauen liegt danach vor, wenn ein Vertrauenssubjekt eine bestimmte Erwartung ohne entsprechende Sicherheiten aufrechterhält und das Risiko, enttäuscht zu werden einzugehen bereit ist. Diese Ambivalenz besteht ungeachtet der Beschaffenheit des Vertrauensobjekts (Person oder Technik), weshalb sich ‚Vertrauen' schlechthin nach Kaminskis Auffassung als Oberbegriff eignet. Dieser pragmatische Ansatz Kaminskis steht insofern ‚orthogonal' zu der hier rekonstruktierten extensionalen Erweiterung des interpersonalen Vertrauensbegriffs auf Technikvertrauen als analog zu Institutionenvertrauen.

3 Vertrauen in der Blockchain

Bevor wir nun auf Vertrauen im Zusammenhang mit Blockchainanwendungen eingehen, sollen zunächst zumindest in nuce einige Grundlagen der Blockchain erläutert werden, die für das Verständnis der Funktionsweise einer Blockchain erforderlich sind und ohne deren Kenntnis das Thema ‚Vertrauen in der Blockchain' nicht sinnvoll erörtert werden kann (3.1). Wir untersuchen in einem zweiten Schritt verschiedene Begriffsverständnisse von Vertrauen in der Blockchain, die wir jedoch zurückweisen (3.2), und schlagen in einem dritten Schritt eine eigene Beschreibung vor, die die Stärken der bisherigen Begriffe aufnimmt und die Schwächen ausbessert (3.3).

3.1 Grundlagen der Blockchain in nuce

Die Blockchaintechnologie lässt sich grundsätzlich als eine dezentrale Möglichkeit der Speicherung von Daten verstehen (Bundesnetzagentur 2021). Während bei einer klassischen Datenbank die Daten zentral bei einer autorisierten Stelle gespeichert werden, werden sie in der Blockchain redundant an mehreren Stellen des Netzwerks abgelegt. Daten werden dabei in sogenannten Blöcken zusammengefasst, die chronologisch ihrer Erstellungszeit nach in eine Kette aneinander gehängt werden. Ein weiterer Unterschied zur zentralen Speicherung ist der Autorisierungsprozess neuer Daten: Anstatt dass eine zentrale Stelle neue Daten in die Datenbank einpflegt, wird das Netzwerk informiert, dass neue Daten angefügt werden sollen, indem Nutzer:innen eine Transaktion veranlassen. Außerdem unterscheidet sich, dass die Ersteller:innen von Datensätzen nicht durch eine zentrale Stelle zum Erstellen der Daten autorisiert werden, sondern die Notwendigkeit und Korrektheit der Daten durch kryptographische Verfahren beweisen, was durch andere Teilnehmende verifiziert wird. Kryptografische Verfah-

ren ermöglichen außerdem, durch pseudonyme Signaturen die Herkunft der Daten nachzuweisen (Bundesnetzagentur 2021). Das Netzwerk, das die Blockchain verwaltet, besteht aus verschiedenen Teilnehmenden (die Liste ist nicht trennscharf und Rollen können sich überlagern): Es gibt aktive Knoten, sogenannte Miners, die kompetitiv neue Blöcke erstellen und der Blockchain hinzufügen; passive Knoten, die neue Blöcke kryptografisch verifizieren und ggf. die Blockchain (redundant mit anderen) speichern sowie Wallet-Inhaber:innen, die Transaktionen in der Blockchain veranlassen und damit typischerweise die größte Gruppe von Teilnehmer:innen sind (Bundesnetzagentur 2021; Schlatt et al. 2016). Wie eine Transaktion schematisch abläuft, lässt sich wie folgt skizzieren:

Ein:e Nutzer:in gibt eine Transaktion auf (z. B. die Übertragung eines Kryptowährungsbetrags). Dieses Datum wird dann an das Netzwerk von aktiven Knoten gesendet, die dieses mit anderen Daten zu einem Block zusammenfassen. Jeder aktive Knoten versucht als erster seinen Block durch Lösen eines kryptografischen Rätsels zu verifizieren. Ist das gelungen, sendet er seinen Block zur Verifizierung an das Netzwerk aller Knoten. Stimmen diese der Korrektheit des Blocks zu, wird dieser an den je aktuellsten Block angehängt und bildet die neue Spitze der Blockchain. Damit ist die Transaktion öffentlich einsehbar[9] in der Blockchain gespeichert und gilt als abgeschlossen (Nakamoto 2008; Schlatt et al. 2016; Yaga et al. 2018). Blockchainimplementierungen lassen sich anhand zweier Dimensionen charakterisieren: erstens, wie begrenzt der Zugriff auf die darin enthaltenen Daten ist, und zweitens, ob die Teilnahme am Verwaltungsprozess eingeschränkt ist. In öffentlichen (engl.: *public*) Blockchains kann jede:r veranlassen, dass neue Daten hinzugefügt werden (d. h. neue Transaktionen durchführen), und auf bestehende Daten zugreifen. Private Blockchains (engl.: *private)* sind nicht frei zugänglich, sondern können nur durch Teilnehmende innerhalb einer Organisation oder eines Konsortiums genutzt werden. Typischerweise sind private Blockchains genehmigungsbasiert (engl.: *permissioned*), d. h. ein Konsortium oder eine Organisation bestimmt, wer einen Knoten im Netzwerk betreiben darf und somit eine Kopie der Blockchain speichern oder neue Blöcke erzeugen kann. In genehmigungsfreien (engl.: *permissionless*) Blockchains kann jede:r sich an der Verwaltung der Blockchain beteiligen.

9 Im Beispiel wird eine öffentliche Blockchain ohne Zugangsbeschränkung beschrieben. Dabei handelt es sich um den typischen Fall einer Blockchainanwendung.

Um eine Blockchain möglichst transparent zu gestalten, werden genehmigungsfreie Blockchains i. d. R. Open-Source entwickelt. Das bedeutet, dass der Code der Blockchain öffentlich einsehbar ist. Die Software wird zwar von Einzelpersonen entwickelt, die oft nur über die Plattform verbunden und in einer dezentralen Struktur organisiert sind, durch die Öffentlichkeit des Codes regulieren sich die Entwickelnden aber gegenseitig: Sie kontrollieren und verbessern untereinander Codeabschnitte, um sicherzustellen, dass einerseits die Korrektheit der Funktionalität gewährleistet ist und andererseits keine Hintergedanken, Vorurteile oder Hintertüren im Code enthalten sind. Auch wenn die Individuen *per se* nicht vertrauenswürdig sind, werden sie durch die internen Mechanismen der Blockchain so miteinander verbunden, dass sie dem Netzwerk als Ganzem dienen, indem sie ihre individuellen Interessen verfolgen (Lipman 2023). Während öffentliche, genehmigungsfreie Blockchains ein Maximum an Dezentralisierung erreichen, werden private, genehmigungsbasierte Blockchains zentral kontrolliert, da sie um die kontrollierende Organisation bzw. das kontrollierende Konsortium ‚herum' organisiert sind (Schlatt et al., 2016; Yaga et al., 2018).

Die Transparenz einer Blockchain hängt somit auch von der Art der Implementierung ab und ist bei öffentlichen, genehmigungsfreien Blockchains am höchsten. Durch die redundante, dezentrale Speicherung der Daten ist zudem deren Verfügbarkeit gewährleistet und das System resilient gegenüber Ausfällen einzelner Knoten (Bundesnetzagentur 2021). Dies führt auch zu einer hohen Datenintegrität: Sobald Daten in der Blockchain gespeichert sind, können sie praktisch nicht mehr geändert werden. Jeder neue Datenblock verweist durch einen Hashwert auf dessen Vorgänger und damit auf die ganze, vorausgegangene Kette an Blöcken. Daher können weder Angreifer noch autorisierte Personen einmal gespeicherte Daten (also Blöcke innerhalb der Kette) unbemerkt ersetzen oder verändern, denn beides würde den Hashwert des Blocks ändern. Um eine ungebrochene Kette zu erzeugen, müssten alle nachfolgenden Blöcke ab dem geänderten Block neu instanziiert werden, was unrealistisch rechenaufwändig ist, da ständig neue Blöcke zur ursprünglichen Kette hinzukommen (Nakamoto 2008). Diese Attribute von Blockchainsystemen, Transparenz, Verfügbarkeit und Datenintegrität, sind Faktoren, die als Grundlage für Vertrauen in diese Technologie gelten (Bundesnetzagentur 2021; Marella et al. 2020; Tang et al. 2020; Teng 2023; Völter et al. 2023).

3.2 Stand der Forschung: Vertrauen in der Blockchain

Es lassen sich mit Blick auf Blockchainanwendungen ganz unterschiedliche Positionen zur Frage von Vertrauen in der Blockchain in der Literatur ausmachen. Um diese klar beschreiben zu können, werden zuerst die oben eingeführten Bedingungen für Vertrauen auf den Kontext von Blockchainanwendungen übertragen und die Vertrauensbeziehungen beispielhaft erläutert. Darauf aufbauend werden wir nachfolgend verschiedene Ansätze der aktuellen Forschungsliteratur ansprechen, sie analytisch-systematisch kategorisieren und sodann zu entkräften versuchen. Im Anschluss unterbreiten wir einen Vorschlag für ein konzeptuelles Verständnis von Vertrauen in der Blockchain, das einen der anderen Ansätze aufgreift und nuanciert und so eine angemessenere Beschreibung der Form von Vertrauen in der Blockchain bietet.

Als Beispiel für eine konkrete Interaktion in einer Blockchainanwendung soll die Überweisung eines Betrages in einer Kryptowährung (wie etwa Bitcoin) dienen. Die Überweisung lässt sich generell als Transaktion von Daten verstehen. Bei einer Transaktion in der Blockchain ist das Vertrauenssubjekt die veranlassende Person, i. d. R. ein:e durchschnittlich informierte:r und fähige:r Nutzer:in. Diese vertraut dem Blockchainsystem Daten (im Beispiel: den Betrag der Kryptowährung) an, und sie vertraut darauf, dass das System die Daten den Erwartungen gemäß verarbeitet (im Beispiel: den Wert in die Wallet der empfangenden Entität überträgt und die Transaktion öffentlich sichtbar in der Blockchain speichert). Baier (1986) und Hardin (2002) folgend könnte man die Vertrauensbeziehungen bei einer Blockchaintransaktion so beschreiben: Ein:e Nutzer:in vertraut dem Blockchainsystem in Bezug auf den Kryptowährungsbetrag, dass dieser sicher und nachvollziehbar an die Empfänger-Entität weitergeleitet wird.

Die erste Bedingung für Vertrauen, d.i. (i) die Vulnerabilität des Vertrauenssubjekts (s. 2.1), lässt sich hier in dem Sinne verstehen, dass ein:e Nutzer:in (Vertrauenssubjekt) darauf angewiesen ist, dass das Blockchainsystem (Vertrauensobjekt) die Transaktion korrekt ausführen (φ-en) kann (s.a. Tab. 1). Bedingung (ii), die (optimistische) Erwartungshaltung des Vertrauenssubjekts, drückt sich in der Erwartung des:der Nutzer:in aus, dass das Blockchainsystem die von ihm:ihr angestoßene Transaktion korrekt ausführen kann. Im Beispiel: Der:die Nutzer:in erwartet, dass der veranlasste Betrag der Kryptowährung in die Wallet der empfangenden Entität übertragen und die Transaktion öffentlich sichtbar in der Blockchain gespeichert werden kann. Und schließlich drückt sich Bedingung (iii) darin aus, dass

φ-en	Korrekte und sichere Ausführung der Transaktion (z. B. sichere und nachvollziehbare Überweisung des Kryptowährungsbetrags)
Vertrauenssubjekt	Nutzer:in, der:die die Transaktion veranlasst
Vertrauensobjekt	das Blockchainsystem (im weitesten Sinne)
Bedingung (i) Vulnerabilität	Nutzer:in ist darauf *angewiesen*, dass das Blockchainsystem die Transaktion korrekt ausführen kann.
Bedingung (ii) (optimistische) Erwartungshaltung	Nutzer:in erwartet, dass das Blockchainsystem die Transaktion korrekt ausführen *kann*.
Bedingung (iii) Sich-auf-etwas-verlassen	Nutzer:in verlässt sich darauf, dass das Blockchainsystem die Transaktion korrekt ausführen *wird*.

Tab. 1: Vertrauensbedingungen im Rahmen von Blockchainanwendungen

sich das Vertrauenssubjekt darauf verlässt, dass die Transaktion vom Blockchainsystem auch entsprechend den Erwartungen verarbeitet wird.

(a) „Trustless trust"

Eine erste Position (‚Extremposition') lässt sich als „Vertrauen ohne Vertrauen" wiedergeben. Schon im Grundlagendokument „Bitcoin: A Peer-to-Peer Electronic Cash System" postuliert Blockchaingründer:in Satoshi Nakamoto die Ersetzung des interpersonalen Vertrauens durch Algorithmen: „cryptographic proof instead of trust" (Nakamoto 2008, 1). Reid Hoffman prägt, von der Idee Nakamotos inspiriert, den Begriff „trustless trust" (Hoffman 2014). Die Blockchaintechnologie soll Nutzenden ermöglichen, in direkten, aber abgesicherten Austausch treten zu können, ohne dass sie einer dritten Partei (etwa den nach der Finanzkrise angeschlagenen Banken) oder gar dem Gegenüber vertrauen müssten (Hoffman 2014; Nakamoto 2008). Vertrauen, das inhärent riskant ist, soll überwunden (Buterin 2015) und, so das Versprechen, automatisiert werden: Kryptographische Signaturen stellen die

Identität des Gegenübers sowie den Ursprung der Daten sicher, während die öffentliche Einsehbarkeit aller Transaktionen deren Korrektheit garantieren soll (Tang et al. 2020; Nakamoto 2008). Nissenbaum (1999) folgend würde diese absolute Absicherung bereits begrifflich verhindern, dass ‚Vertrauen‘ entsteht. Bei absoluter Sicherheit fehlt die Vulnerabilität des Vertrauenssubjekts (Bedingung (i)) und Vertrauen wird verhindert (Teng 2021). Nach der *trustless trust*-Auffassung sind Nutzende darauf angewiesen, dass das Blockchainsystem ihre Transaktion korrekt verarbeiten kann. Nach dieser Auffassung werden die Mechanismen dazu als so zuverlässig angesehen, dass diese sich nicht nur den Erwartungen gemäß verhalten *können*, sondern sich deterministisch so verhalten müssen (Nakamoto 2008; Tang et al. 2020).

Grundlegend für die Überzeugung, es bedürfe keines Vertrauens mehr bei der Nutzung von Blockchainanwendungen ist allerdings die Annahme, dass die Nutzenden kompetent genug sind, die Funktionsweise der Blockchain und die Implikationen der Sicherheitsmechanismen vollständig zu verstehen (Jacobs 2020; Teng 2022). Diese Annahme ist unrealistisch für durchschnittliche Verbraucher:innen[10], die i. d. R. weder die Zeit noch die Kompetenz haben, die Funktionsweise von Blockchainsystemen im Einzelnen zu verstehen. Sie sind deshalb nicht in der Lage, ihre eigene Sicherheit in der Verwendung gewährleisten zu können (Teng 2021). Jacobs hebt dabei die unterschiedlichen Betrachtungswinkel in der Literatur hervor: solche, die der *trustless trust*-Annahme folgen, fokussieren den Kontext der Betrachtung meist auf die Technologie an sich und haben eben nicht die praktische Verbraucherperspektive im Blick (Jacobs 2020).

10 Besonders aussagekräftig ist hierbei die Differenzierung, die Micklitz et al. als Dreiklang-Modell vorschlagen: Es unterscheidet zwischen verletzlichen, vertrauenden und verantwortungsvollen Verbraucher:innen (Micklitz et al. 2010). In einem bestimmten Kontext (z. B. Verwendung eines Blockchainsystems) sind verantwortungsvolle Verbraucher:innen weitgehend in der Lage und daran interessiert, sich in ein komplexes Thema einzuarbeiten. Die meisten Nutzer:innen gehören jedoch i.d.R. zu den vertrauenden Verbraucher:innen (Kenning und Wobker 2013), die sich darauf verlassen, dass andere dafür sorgen, dass Technologie korrekt funktioniert und sie für den Fall, dass es zu Funktionsdefiziten oder -ausfällen kommt, geschützt und ggf. entschädigt werden. Bei dieser Gruppe kann nicht erwartet werden, dass sie im Detail versteht, wie eine Blockchainanwendung funktioniert und wie ihre Sicherheit gewährleistet wird.

(b) Verteiltes Vertrauen

Manche Positionen, die Vertrauen in der Blockchain ausschließen, sehen dies nicht etwa als Vorteil der neuen Technologie (wie z. B. Nakamoto (2008)), sondern als Konsequenz eines engen Begriffsverständnisses von Vertrauen, das nur im direkten Austausch zwischen rationalen und emotionalen Agenten (Menschen) möglich ist (Jacobs 2020). Technologie könne sich nicht im eigentlichen Sinne verhalten und damit auch nicht die notwendigen Erwartungshaltungen erfüllen oder enttäuschen (Fries 2022).

Trotzdem ist Vertrauen in der Blockchain nötig: Es wird nicht einfach aufgehoben, sondern verschoben. Statt in direktem Kontakt einem Gegenüber zu vertrauen, verlagert sich das Vertrauen bei der Nutzung von Blockchainsystemen auf die Personen hinter dem System, so behauptet eine zweite Position in Abgrenzung zur *trustless trust*-Annahme (Walch 2019). Um eine Transaktion in der Blockchain abzuschließen, muss sie Teil eines neuen Blocks werden, der durch etliche Knoten des Netzwerks verifiziert wird (s. 3.1). Die Authentifizierung von Daten, die Verifizierung von Transaktionen und generell die korrekte Funktionsweise des Blockchainsystems hängt also an einer Vielzahl, genauer: an einem Netzwerk von Individuen.[11] Es besteht damit weiterhin ein *interpersonales* Vertrauensverhältnis zwischen einer Person, die die Transaktion aufgibt (Vertrauenssubjekt) und den Personen, die für die Funktion des Blockchainsystems verantwortlich sind (Vertrauensobjekt). Dass man sich gerechtfertigt auf die korrekte Funktionsweise des Blockchainsystems verlassen kann und den Erwartungen entsprochen wird (Vertrauensbedingungen ii und iii), hängt also an einem Kollektiv von Entwickler:innen und einem Netzwerk an Blockchainknoten (Wang et al. 2022, 34; De Filippi et al. 2020; Finck 2018, 13). Insofern sich so der Vertrauensbegriff von einer Person im direkten Kontakt auf ein Netzwerk aus Individuen weitet, lässt sich von *verteiltem Vertrauen („distributed trust")* sprechen (Mallard et al. 2014; Jacobs 2020).

In der Literatur sind dabei unterschiedlich starke Positionen zu finden, inwieweit dem Kollektiv Vertrauen entgegengebracht werden sollte. Die

11 Die korrekte Funktionsweise wird insbesondere bei öffentlichen, unbeschränkten Blockchains, die mit Open-Source-Software arbeiten, durch ein Netzwerk an Entwickler:innen sichergestellt. Bei privaten, zugangsbeschränkten Blockchains ist hingegen eine zentrale Stelle für die Programmierung und Instandhaltung zuständig, die nicht unbedingt durch andere Entitäten kontrolliert werden kann.

eine Seite gibt zu bedenken, dass Entwickler:innen und Stakeholder intransparente Interessen haben können (Fries 2022) und wichtige Entscheidungen oft nicht öffentlich nachvollziehbar getroffen werden (Teng 2021; 2023). Auf der anderen Seite finden sich zuversichtliche Stimmen, die sich auf die eingebauten Mechanismen der Blockchain verlassen: Transparenz und dezentrale Speicherung der Daten als zentrale Eigenschaft der Blockchain sind hier die wesentlichen Gründe für die Zuversicht. Außerdem ist nicht nur die Datenspeicherung, sondern auch die Machtverteilung dezentral und jede:r einzelne hat nur begrenzte Möglichkeiten, eigene Interessen durchzusetzen. Entscheidungen im Quellcode werden – insbesondere bei Open-Source-Software – auch öffentlich gemacht (Buterin 2015; De Filippi et al. 2020).

Genau diese Argumentation auf Grundlage der eingebauten Mechanismen führt jedoch zu Problemen: Um ihre Vulnerabilität (Bedingung (i), s.o.) informiert einzuschätzen und abzuwägen, müssten Nutzer:innen ermessen können, welche Netzwerke und Entwicklungskollektive vertrauenswürdig sind. Diese Zuschreibung von Vertrauenswürdigkeit durch das Vertrauenssubjekt wird jedoch auch bei dieser Position vernachlässigt und oft durch die Blockchaintechnologie und Verwendung von Open-Source-Software als automatisch und einseitig-induzierbar angenommen.

De Filippi et al. vertreten die These, ein Blockchainsystem sei nicht als „trust machine“ oder „trustless trust“-System zu verstehen, sondern als „confidence machine“ (2020, 2). Sie stellen dazu begrifflich ‚Vertrauen‘ (trust) und ‚Zutrauen‘ (confidence) gegenüber. Doch unterscheiden sich De Filippi et al. sowie unsere Analyse im Ergebnis (!) nur in der Annahme der Vulnerabilität der Nutzer:innen von Blockchainsystemen. Zutrauen lässt sich danach als Faktor von Vertrauen (vgl. etwa die soziologische Analyse bei Misztal oben unter 2.1) verstehen. De Filippi et al. gehen für die Blockchain phänomenologisch davon aus, dass eine Erwartungshaltung („predictability“, 2020, 4) und – implizit – ein Sich-verlassen der Blockchainnutzer:innen vorliegen (müssen). Dies entspricht den beiden Bedingungen (ii) und (iii) der hier zugrunde gelegten Standarddefinition von ‚Vertrauen‘ nach McLeod (2021). Zugleich stellen De Filippi et al. nicht in Abrede, Blockchainsysteme beruhten auf Vertrauen. Sie rekonstruieren Vertrauen in der Blockchain im Rahmen einer Zwei-Ebenen-Analyse als verteiltes Vertrauen: Im Ergebnis muss Vertrauen in die Governancestruktur bzw. die diese tragenden Akteur:innen (zweite Ebene) gegeben sein, damit in der Nutzungspraxis (erste Ebene) Zutrauen in ihrem Sinne bestehen kann:

> [...] when it comes to the standard operations of a blockchain-based system, it is nonetheless necessary to trust the actors securing and maintaining the underlying blockchain network, in order to guarantee a sufficient level of confidence in any of the blockchain-based applications operating on top of that network. Confidence in a procedural system such as a blockchain ultimately depends on the proper governance of that system. (2020, 11)

Zusammenfassend wird die Blockchain technisch als „confidence protocol that builds upon an external layer of trust" bestimmt (2020, 12). Fraglich ist, wie ‚extern' diese Ebene nach der Analyse von De Filippi et al. überhaupt ist, denn das von ihnen diagnostizierte Zutrauen

> depends on a variety of factors, including the collective management of the network by a large number of distributed actors (e.g. miners, validators) who—although they do not have the power to unilaterally influence the network—*nonetheless need to be trusted not to collude* in order to further their own interests, at the expense of the overall network. (2020, 2; unsere Hervorhebung)

Unseres Erachtens nähern sie sich damit stark der These, auch ‚im' Blockchainsystem sei Vertrauen erforderlich, denn die Erwartungshaltung, dass eine Technik bei Nutzung regelhaft Ergebnisse produziert und man sich auf sie in einer bestimmten Hinsicht verlassen kann, muss von einem Vertrauen in die Governance des Systems begleitet werden. Damit sind aber unsere Bedingungen (ii), (iii) und schließlich auch (i) erfüllt, denn die Nutzer:innen sind der Governance-Struktur und ihren Akteur:innen gegenüber – siehe den vorangegangenen Absatz – in einer Position der Vulnerabilität. Unseres Erachtens lässt sich die Einstellung der Nutzer:innen in der Blockchain nicht als ‚Funktionalitätszutrauen plus Governancevertrauen' rekonstruieren, sondern insgesamt besser als (bzw.: analog zu) Institutionenvertrauen beschreiben. (Dazu sogleich unter 3.3.)

(c) Vertrauen durch Code

Während sich manche auf einen sehr engen Vertrauensbegriff festlegen, der Vertrauen als ein Wagnis ohne Absicherung oder Kontrolle zwischen menschlichen Akteuren sieht (Fries 2022), erlauben andere eine Weitung des Begriffs: Teng beispielsweise rekonstruiert den Begriff dahingehend, dass Vertrauen vor allem auf Verlässlichkeit und Erwartungserfüllung be-

ruht: „[…] trust is applicable not just to humans but also to things, since trust does not require the trustee to have any condition other than reliability for developing a trust relation“ (Teng 2022). Die Möglichkeit Vertrauen technikseitig automatisch zu erzeugen, wie in den vorgenannten Positionen angesprochen, macht eine dritte Position explizit: Vertrauen wird nach dieser erzeugt, indem die Verhandlungsbeziehungen zwischen Akteuren im Code festgehalten und algorithmisch abgesichert werden (Finck 2018, 12), was auch durch den griffigen Slogan „in proof we trust“ (Werbach 2018, 29) beschrieben wird. Die Schaffung von Vertrauen wird in die technische Funktionalität verlagert. Insbesondere Werbach (2018) stellt dieses Merkmal der Blockchain als Novum und idiosynkratische Eigenschaft heraus. Auch andere sehen in der Klarheit, die der Code schafft, eine Vereinfachung der Komplexität in der Nutzung von Blockchainanwendungen einerseits und in den implementierten Werten Transparenz, Kryptographie und Konsens andererseits und somit eindeutige Signale, die Vertrauen induzieren (Teng 2022; Finck 2018, 12; Werbach 2018; Bundesnetzagentur 2021).

Unseres Erachtens mag diese Position zwar einen Teil der medialen Aufmerksamkeit, ja des ‚Hype‘ um Blockchainanwendungen als ‚besonders vertrauenswürdige‘ Technologie, erklären und plausibel machen, impliziert sie doch, dass sich die Vertrauenswürdigkeit ‚von selbst‘ aus der bloßen Struktur und Funktionalität, aufgrund der Dezentralität und Transparenz der Blockchain, ergebe. Doch halten wir dem entgegen, dass diese Position ein inadäquates Verständnis von ‚Vertrauenswürdigkeit‘ als ‚automatische, technikgenerierte Eigenschaft des Vertrauensobjekts‘ voraussetzt. Vertrauenswürdigkeit ergibt sich nach unserer Auffassung nicht ‚durch Code‘, sondern setzt die – im besten Falle: berechtigte – Zuschreibung von Vertrauenswürdigkeit in der Relation zwischen Vertrauenssubjekt und Vertrauensobjekt voraus. (Dazu sogleich.)

Auch innerhalb in der Ansicht, Vertrauen sei induzierbar, greift Vertrauen durch Code zu kurz: Es kann nicht die personalen Aspekte von Vertrauen aufgreifen und geht kaum über ein Verlassen auf Technik hinaus: „[…] feeling betrayed in case of failed trust cannot be appropriately directed at technological artifacts.“ (Jacobs 2020, 583; vgl. De Filippi et al. 2020). Sehr treffend drückt Butler (2021, 218) den Unterschied aus: „[…] reliance can exist without the attitude of trusting: one would then, should things not work out, be disappointed, yet not betrayed.“ Auf Code kann sich ein:e Nutzer:in allenfalls verlassen (*to rely on),* aber nicht nach dem hier rekonstruierten dreistelligen Vertrauensbegriff vertrauen *(to trust).*

Auch Fries (2022) lehnt diese Position ab, geht aber von einer anderen, affirmativen Ausgangsposition aus. Sie lässt sich darauf ein, dass ‚Code Vertrauen generiere', nur um diese Behauptung in Bausch und Bogen zu kritisieren – ohne freilich eine alternative Sichtweise vorzuschlagen. Diese Invektive kommt einem undifferenzierten ‚Blockchain Bashing' gleich. Unsere Position ist ähnlich kritisch gegenüber der Redeweise von ‚Vertrauen in Code', setzt dieser aber als positive These entgegen, dass Blockchainvertrauen am besten analog zu ‚Vertrauen in Institutionen' verstanden werden kann. Damit begeben wir uns zugleich in ein klassisches technikphilosophisches Fahrwasser, indem wir Technikvertrauen strukturell als Institutionenvertrauen konzeptualisieren.

(d) Vertrauen in Institutionen

Nach dem von uns favorisierten Verständnis lässt sich Vertrauen in Blockchainsysteme als *Vertrauen in Institutionen* (s. 2.2.c) verstehen (Teng 2021; Ostern 2018). Verbraucher:innen sind im Regelfall nicht kompetent, Blockchain als Technologie gut genug zu verstehen, um ihre Nutzung abzusichern (vgl. auch Lustig und Nardi 2015), sondern machen einen „leap of faith" (Teng 2021, 391) und vertrauen auf die korrekte Funktionsweise, die durch die Blockchain als Institution erwartet werden kann. Vertrauen wird so auch hier zur Reduzierung der Komplexität genutzt (Teng 2021; vgl. oben unter 2.2. (b) das klassische Diktum Luhmanns).

In der unserer These nahekommenden Literatur bleibt offen, was Blockchainvertrauen, verstanden als Institutionenvertrauen, umfasst. Teng argumentiert zwar, dass bei der Nutzung von Blockchainsystemen das Gegenüber nicht konkret klar sein muss („non-partner-relative") und sich durchaus normative und prognostische Erwartungshaltungen ergeben (Teng 2021), jedoch scheint auch hier angenommen zu werden, dass dadurch Vertrauen schon hinreichend bestimmt ist. Vertrauen wird so als technikseitig-induziert verstanden. Unseres Erachtens hingegen ist Vertrauenswürdigkeit keine technikgenerierte Eigenschaft, sondern Resultat einer Zuschreibung, die auch fehlgehen kann (dazu sogleich unter 3.3).

3.3 Vertrauen in der Blockchain als Zuschreibung von Vertrauenswürdigkeit

Vertrauenswürdigkeit ergibt sich u. E. aus der Zuschreibung durch ein Vertrauenssubjekt auf ein Vertrauensobjekt. Es handelt sich dabei um eine relational generierte Eigenschaft des Vertrauensobjekts, welche dieses nicht

selbst hervorbringen kann, sondern die sich nur in der Relation zu einem Vertrauenssubjekt durch dessen – durchaus irrtumsanfällige – Haltung vis-à-vis dem Vertrauensobjekt ergibt.

Die Blockchain beseitigt nicht die Vulnerabilität des Vertrauenssubjekts, hier: der Blockchainnutzer:innen („User“). Die beschriebene dreiteilige Vertrauensdefinition (Vulnerabilität – Erwartungshaltung des Vertrauenssubjekts bezogen auf die Kompetenz des Vertrauensobjekts – Erwartungshaltung des Vertrauenssubjekts bezogen auf die Performanz des Vertrauensobjekts) bleibt maßgeblich. Vertrauenswürdigkeit ergibt sich nicht durch die Eigenschaften der Blockchain, sondern in der Zuschreibung durch die Nutzer:innen in ihrer Rolle als Vertrauenssubjekte. Vertrauenswürdigkeit ist keine technikgenerierte Eigenschaft, sondern Resultat einer Zuschreibung, bei der die Vertrauenssubjekte auch irren können. Liegen die drei notwendigen Merkmale von ‚Vertrauen‘ vor, so ergibt sich im Normalfall für das Vertrauenssubjekt die Möglichkeit, dem System der Blockchain zu vertrauen bzw. – besser, da weniger zirkulär und informativer – Vertrauenswürdigkeit zuzuschreiben.

Dabei müssen drei Bedingungen gegeben sein, die es ermöglichen (rechtfertigen, gute Gründe dafür abgeben), Vertrauenswürdigkeit zuzuschreiben:

Bedingung 1 [B1]. Das Vertrauenssubjekt ist vulnerabel.

Wie bereits unter (2.1) erörtert, ist ohne Vulnerabilität die Rede von Vertrauen nicht sinnvoll bzw. bereits aus begrifflichen Gründen nicht möglich. (Siehe oben: Vulnerabilität ist das erste notwendige, nicht hinreichende Merkmal von Vertrauen.) Bei völliger Sicherheit eines Systems ist kein Vertrauen(müssen) erforderlich; die Systemnutzung führt dann ohne Risiken oder Unwägbarkeiten zum erwünschten, angestrebten Handlungsziel. Bestehen hingegen Risiken und damit Unsicherheiten, ergibt sich Vulnerabilität, derer sich das Vertrauenssubjekt *idealiter* bewusst ist und die sich ggf. sogar quantifizieren lässt.

Das ‚Vertrauen in Algorithmen‘ (nicht nur) bei Blockchainanwendungen entspricht dem soeben skizzierten Unsicherheitsmodell: Da es hier für außenstehende ‚gewöhnliche‘ Nutzer:innen praktisch nicht möglich ist, die Korrektheit des Codes zu vorab zu prüfen, müssen sie sich auf die Funktionsfähigkeit des Systems verlassen. Sie müssen ‚vertrauen‘, dass das System in ihrem Sinne funktionieren kann und wird. Das entspricht den beiden weiteren notwendigen Merkmalen des Vertrauensbegriffs (Kompetenz- und Performanzerwartung).

Wichtig ist, zu betonen, dass hier für die Rolle des Vertrauenssubjekts eine gewöhnliche Nutzer:innenperspektive eingenommen wird, nicht etwa die der sogenannten ‚Miner'. Grundsätzlich darf wohl auch eine Mehrheit benevolenter, wohlmeinender Miner angenommen werden; dies lässt sich aber nur als zufällige, nicht als strukturell notwendige Variable der Erwartungshaltung des Vertrauenssubjekts bei Blockchainanwendungen voraussetzen. Generell ergeben sich aus den Merkmalen einer Blockchain, der dezentralen Struktur, der Unveränderlichkeit und grundsätzlichen Transparenz der Eintragungen, gute Gründe für die Zuschreibung von Vertrauenswürdigkeit – diese ergeben sich aber nicht ipso facto, sondern lassen sich phänomenologisch am treffendsten als eine widerlegbare Vermutung beschreiben. Im Sinne eines ‚default and challenge'-Modells kann die Vertrauenswürdigkeit in Frage gestellt, die Zuschreibung von Vertrauenswürdigkeit durch das Vertrauenssubjekt nicht gerechtfertigt sein.

Die weiteren Bedingungen für die Zuschreibbarkeit von Vertrauenswürdigkeit durch das Vertrauenssubjekt ergeben sich daraus im nachfolgenden Sinne:

> [B2]. Das Vertrauenssubjekt hat eine Erwartungshaltung an Funktionsfähigkeit und Wahrscheinlichkeit der korrekten Ausführung des ‚Codes'. (=> entspricht den Vertrauensbedingungen (ii) und (iii))
>
> [B3] Das Vertrauenssubjekt ist, sofern es keine Anhaltspunkte für eine andere Einschätzung der Sachlage gibt, gerechtfertigt, dem Vertrauensobjekt (der Blockchain) Vertrauenswürdigkeit zuzuschreiben.

Die Vertrauenswürdigkeit generiert die Blockchaintechnologie nicht ‚automatisch' selbst. Vielmehr ist das Vertrauenssubjekt im Sinne eines epistemischen *‚default and challenge'* im Normalfall darin gerechtfertigt, Vertrauenswürdigkeit zu unterstellen und zuzuschreiben – bis sich Anzeichen für eine andere Auffassung, etwa betrügerische Absichten auf Seiten der Blockchainbetreiber:innen, Wallet-Anbieter:innen usw. zeigen. Die Kombination prädiktiver und normativer Erwartungen an den Prozess und das Ergebnis der Interaktion macht das Spezifikum des Vertrauens in eine Institution aus. Teng argumentiert mit Blick auf das Verständnis von Technikvertrauen als Institutionenvertrauen, dass Benutzer:innen Technik dabei wie Institutionen behandeln, Technik insofern *Ähnlichkeit* mit Institutionen aufweist: „Technologies resemble institutions in their design capacity for carrying normative values and inviting relevant expectations about what they are supposed to do" (Teng 2021, 392). Auch wenn Technik keine Institution ‚ist', weil sie – im Unterschied zu Kollektivakteurinnen, als die sich Institutionen

rekonstruieren lassen – selbst keine Intentionalität und Agency aufweist (s.o. unter 2.2 (c)), enthält sie dennoch eingebettete normative Werte und weckt Erwartungen an ihre Funktionalität. Institutionelles Vertrauen verbindet Vertrauen in strukturelle Absicherungen, also Vertrauen in Regeln und Systeme, mit interpersonalem Vertrauen (Lahno 2001; s.o.) – für das Verständnis von Technikvertrauen als Institutionenvertrauen gilt dies *analog*. Auch Blockchainvertrauen lässt sich in diesem Sinne am besten als Vertrauen in Technik in struktureller Analogie zum Vertrauen in Institutionen verstehen.

4 Fazit

Ausgehend von der Herausforderung des paradox anmutenden Geltungsanspruchs, Blockchainanwendungen böten Beispiele für „vertrauenloses Vertrauen“, kommen wir auf der Grundlage eines Literaturüberblicks zu einer typisierenden Unterscheidung (siehe Abschnitt 3.2) von vier unterschiedlichen Positionen zum Zusammenhang von ‚Blockchain‘ und ‚Vertrauen‘. Diese Liste ist dabei als offene Liste zu verstehen; andere Auffassungen sind denkbar. Neben der Position, ‚Vertrauen‘ sei in der Blockchain verzichtbar (a), lassen sich mehrere Positionen ausmachen, die auch in der Blockchain Raum für Vertrauensrelationen sehen, entweder als sog. „verteiltes Vertrauen“ (b), als „Vertrauen durch Code“ (c) oder als „Vertrauen in Institutionen“ (d). Indem wir Blockchainvertrauen als Institutionenvertrauen verstehen, legen wir uns – diesem Missverständnis sei hier vorgebeugt – nicht auf die Annahme fest, eine Blockchain ‚sei‘ eine Institution. Wir interpretieren das im Rahmen von Blockchainanwendungen erforderliche Vertrauen zwischen Nutzer:in als Vertrauenssubjekt und Blockchain als Vertrauensobjekt lediglich in dem Sinne, dass es strukturell mit Vertrauen in Institutionen vergleichbar ist und in diesem Sinne ‚als‘ Institutionenvertrauen verstanden werden kann. Damit behaupten wir, dass wir es in einer Blockchain weder mit bloßem, prima facie paradoxem technikgeneriertem Vertrauen („trustless trust“) zu tun haben noch mit Vertrauen in eine Gruppe (von Entwickler:innen, Verantwortlichen usw.) ‚hinter‘ den Mechanismen der Blockchain („verteiltes Vertrauen“) noch mit „Vertrauen durch Code“. Letztere Position geht wie „trustless trust“ von *eo ipso* technikgeneriertem Vertrauen aus, gesteht aber von vornherein zu, dass es sich um ‚Vertrauen‘ handelt.

Unseres Erachtens ist die Auffassung plausibler, da phänomenologisch adäquater, dass es sich im Zusammenhang mit Blockchainanwen-

dungen um ein Vertrauen als prädiktive und normative Erwartung an das Blockchainsystem analog zu möglichem Vertrauen in eine Institution, die bestimmten Grundsätzen und normativen Vorgaben folgt, handelt. Dieses Vertrauensverständnis ist begrifflich-analytisch sowohl von Vertrauen in das technisch-informatische Funktionieren als auch von Vertrauen in die ‚Akteur:innen hinter der Technik' zu unterscheiden. Vielmehr sorgt das Zusammenspiel von Technik und Entwickler:innen oder Verantwortlichen in Summe dafür, dass Kompetenz- und Performanzerwartung eines Vertrauenssubjektes berechtigt sein können – bis zum Beweis des Gegenteils bzw. bis zu Indizien, die diese Unterstellung unterminieren. Dies entspricht einem epistemischen default-and-challenge-Modell.

Aus der Anerkennung der fortbestehenden Vulnerabilität der Blockchainnutzer:innen (User, nicht: Miner) ergeben sich bestimmte normative Forderungen, die sich am besten als rechtspolitische Empfehlungen für einen angemessenen Schutz von Nutzer:innen in ihrer Rolle als Wirtschaftsbürger:innen oder Verbraucher:innen formulieren lassen. Entsprechend der Typisierung von Verbraucher:innen sollten hier Überlegungen zu unterschiedlichen Schutzniveaus und darauf abgestimmter Regulierung von Blockchainanwendungen erfolgen. Die Devise könnte dabei entsprechend der Standarddefinition von ‚Vertrauen' mit seiner Vulnerabilitätsbedingung lauten: So viel Sicherheit wie nötig, so viel Vertrauen wie möglich. Das bleibt aber weiteren Arbeiten vorbehalten. Vorliegend genügt es uns, wenn es uns gelungen sein sollte, das Verständnis von ‚Vertrauen' in der Blockchain ein wenig zu erhellen.[12]

12 Die Verfasserinnen danken den anonymen Gutachter:innen der *Zeitschrift für Praktische Philosophie* für ihre konstruktiven Rückmeldungen. Alle verbleibenden Mängel sind den Verfasser:innen zuzuschreiben. Dank gilt ebenso dem Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) für die Förderung des Verbundprojekts „Potenziale der Blockchain-Technologie für die Digitale Verbraucherteilhabe. Ethische, rechtliche und technische Implikationen der Entwicklung verbraucherfreundlicher Blockchain-Anwendungen (BlockTechDiVer) – Teilprojekt B", in dessen Rahmen der vorliegende Beitrag entstanden ist (Förderkennzeichen: 28V1401B20; https://app.dimensions.ai/details/grant/grant.9651721).

Literatur

Akerlof, George A. 1970. „The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism. “ *The Quarterly Journal of Economics* 84 (3): 488–500. https://doi.org/10.2307/1879431, zuletzt abgerufen am 24.05.2024.

Alfano, Mark und Nicole Huijts. 2020. „Trust and distrust in institutions and governance“. In *The Routledge Handbook of Trust and Philosophy*, herausgegeben von Judith Simon. New York: Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315542294-20/trust-institutions-governance-mark-alfano-nicole-huijts?context=ubx&refId=2537d680-3943-4851-a93d-3ca264f7ea82, zuletzt abgerufen am 28.04.2024.

Baier, Annette. 1986. „Trust and Antitrust“. *Ethics* 96 (2): 231–260. https://doi.org/10.1086/292745, zuletzt abgerufen am 30.03.2024.

Baier, Annette. 1991. „Trust“. In *The Tanner Lectures on Human Values*, Vol. 13, 109–174. Salt Lake City, UT: University of Utah Press. https://tannerlectures.utah.edu/_resources/documents/a-to-z/b/baier92.pdf, zuletzt abgerufen am 30.03.2024.

Bundesnetzagentur. 2021. „Die Blockchain-Technologie. Grundlagen, Potenziale und Herausforderungen“. Bonn: Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen. https://www.bundesnetzagentur.de/DE/Fachthemen/Digitalisierung/Technologien/Blockchain/Links_Dokumente/einfuehrung_bc.pdf?__blob=publicationFile&v=12, zuletzt abgerufen am 30.03.2024.

Buterin, Vitalik. 2015. „Visions, Part 2: The Problem of Trust“. *Ethereum Foundation Blog* (Blog). 27. April 2015. https://blog.ethereum.org/2015/04/27/visions-part-2-the-problem-of-trust, zuletzt abgerufen am 30.03.2024.

Butler, Aaron J. 2021. „Preliminary Reflections on the Ontological Significance of Blockchain Technology for Trust and Trustworthiness“. In *Digitalisierung aus theologischer und ethischer Perspektive. Konzeptionen – Anfragen – Impulse,* herausgegeben von Gotlind Ulshöfer, Peter G. Kirchschläger und Markus Huppenbauer, 211–225. Baden-Baden: Nomos.

De Filippi, Primavera, Morshed Mannan und Wessel Reijers. 2020. „Blockchain as a Confidence Machine: The Problem of Trust & Challenges of Governance“. *Technology in Society* 62: 101284. https://doi.org/10.1016/j.techsoc.2020.101284, zuletzt abgerufen am 30.03.2024.

Dierksmeier, Claus, und Peter Seele. 2020. „Blockchain and Business Ethics“. *Business Ethics: A European Rev*iew 29 (2): 348–359. https://doi.org/10.1111/beer.12259, zuletzt abgerufen am 30.03.2024.

Finck, Michèle. 2018. *Blockchain Regulation and Governance in Europe*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108609708, zuletzt abgerufen am 30.03.2024.

Fries, Isabelle. 2022. „‚In Code We Trust‘?: Zur Vertrauens-Verheißung der Blockchain-Technologie“. *Zeitschrift Für Evangelische Ethik* 66 (4): 264–276. https://doi.org/10.14315/zee-2022-660405, zuletzt abgerufen am 30.03.2024.

Giddens, Anthony. 1990. *The Consequences of Modernity*. Cambridge: Polity Press.

Grunwald, Armin und Yannick Julliard. 2005. „Technik als Reflexionsbegriff – Überlegungen zur semantischen Struktur des Redens über Technik". *Philosophia naturalis* 42: 127–157.

Hardin, Russell. 2002. Trust and Trustworthiness. Russell Sage Foundation series on trust. New York: Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610442718, zuletzt abgerufen am 30.03.2024.

Hartmann, Martin. 2011. *Die Praxis des Vertrauens*. Berlin: Suhrkamp.

Hawley, Katherine. 2014. „Trust, Distrust and Commitment". *Noûs* 48: 1–20. https://doi.org/10.1111/nous.12000, zuletzt abgerufen am 30.03.2024.

Hoffman, Reid. 2014. „The Future of the Bitcoin Ecosystem and ‚Trustless Trust' – Why I Invested in Blockstream". 17. November 2014. https://www.linkedin.com/pulse/20141117154558-1213-the-future-of-the-bitcoin-ecosystem-and-trustless-trust-why-i-invested-in-blockstream, zuletzt abgerufen am 30.03.2024.

Jacobs, Mattis. 2020. „How Implicit Assumptions on the Nature of Trust Shape the Understanding of the Blockchain Technology". *Philosophy and Technology* 34 (3): 573–587. https://doi.org/10.1007/s13347-020-00410-x, zuletzt abgerufen am 30.03.2024.

Kaminski, Andreas. 2010. *Technik als Erwartung. Grundzüge einer allgemeinen Technikphilosophie*. Bielefeld: transcript.

Kelp, Christoph, und Mona Simion. 2023. „What is trustworthiness?". *Noûs* 57: 667–683. https://doi.org/10.1111/nous.12448, zuletzt abgerufen am 30.03.2024.

Kenning, Peter, und Inga Wobker. 2013. „Ist der ‚mündige Verbraucher' eine Fiktion? Ein kritischer Beitrag zum aktuellen Stand der Diskussion um das Verbraucherleitbild in den Wirtschaftswissenschaften und der Wirtschaftspolitik". *Zeitschrift für Wirtschafts- und Unternehmensethik* 14 (2): 282–300. https://doi.org/10.5771/1439-880X-2013-2-282, zuletzt abgerufen am 30.03.2024.

Kirchschläger, Peter G. 2021. „Ethics of Blockchain Technology". In *Digitalisierung aus theologischer und ethischer Perspektive. Konzeptionen – Anfragen – Impulse*, herausgegeben von Gotlind Ulshöfer, Peter G. Kirchschläger und Markus Huppenbauer, 185–209. Baden-Baden: Nomos.

Lahno, Bernd. 2001. „Institutional Trust: A Less Demanding Form of Trust?". *Revista Latinoamericana de Estudios Avanzados* 15: 19–58. https://philpapers.org/archive/LAHITA.pdf, zuletzt abgerufen am 30.03.2024.

Lapointe, Cara und Lara Fishbane. 2019. „The Blockchain Ethical Design Framework". *Innovations: Technology, Governance, Globalization* 12 (3–4): 50–71. https://doi.org/10.1162/inov_a_00275, zuletzt abgerufen am 30.03.2024.

Lipman, Martin A. 2023. „On Bitcoin: A Study in Applied Metaphysics". *The Philosophical Quarterly* 73 (3): 783–802. https://doi.org/10.1093/pq/pqad030, zuletzt abgerufen am 30.03.2024.

Luhmann, Niklas. 2014. *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*. 5. Aufl. Konstanz und München: UVK Verlagsgesellschaft [Erstauflage 1968].

Lustig, Caitlin, und Bonnie Nardi. 2015. „Algorithmic Authority: The Case of Bitcoin“. In *2015 48th Hawaii International Conference on System Sciences*, 743–752. https://doi.org/10.1109/HICSS.2015.95, zuletzt abgerufen am 30.03.2024.

Mallard, Alexandre, Cécile Méadel und Francesca Musiani. 2014. „The Paradoxes of Distributed Trust: Peer-to-Peer Architecture and User Confidence in Bitcoin“, *Journal of Peer Production*. http://peerproduction.net/issues/issue-4-value-and-currency/peer-reviewed-articles/the-paradoxes-of-distributed-trust/, zuletzt abgerufen am 30.03.2024.

Marella, Venkata, Bikesh Upreti, Jani Merikivi und Virpi Kristiina Tuunainen. 2020. „Understanding the creation of trust in cryptocurrencies: the case of Bitcoin“. *Electron Markets* 30: 259–271. https://doi.org/10.1007/s12525-019-00392-5, zuletzt abgerufen am 30.03.2024.

McLeod, Carolyn. 2021. „Trust“. In *The Stanford Encyclopedia of Philosophy,* herausgegeben von Edward N. Zalta, Fall 2021. Stanford: Stanford University. https://plato.stanford.edu/archives/fall2021/entries/trust/, zuletzt abgerufen am 30.03.2024.

Micklitz, Hans-W., Andreas Oehler, Michael-Burkhard Piorkowsky, Lucia Reisch und Christoph Strünck. 2010. *Der vertrauende, der verletzliche oder der verantwortungsvolle Verbraucher? Plädoyer für eine differenzierte Strategie in der Verbraucherpolitik: Stellungnahme des Wissenschaftlichen Beirats für Verbraucher- und Ernährungspolitik beim BMELV vom Dezember 2010*. Berlin: Bundesministerum für Ernährung, Landwirtschaft und Verbraucherschutz.

Misztal, Barbara A. 1996. Trust in Modern Societies: The Search for the Bases of Social Order: Significance, Scope and Limits of the Drive Towards Global Uniformity. Cambridge: Polity Press.

Möllering, Guido. 2006. „Trust, Institutions, Agency: Towards a Neoinstitutional Theory of Trust“. In *Handbook of Trust Research,* herausgegeben von Reinhard Bachmann und Akbar Zaheer, 355–376. Cheltenham und Northampton, Mass.: Edward Elgar Publishing. https://doi.org/10.4337/9781847202819.00029, zuletzt abgerufen am 30.03.2024.

Nakamoto, Satoshi. 2008. „Bitcoin: A Peer-to-Peer Electronic Cash System“. Whitepaper. https://bitcoin.org/bitcoin.pdf, zuletzt abgerufen am 30.03.2024.

Nickel, Philip J. 2013. „Trust in Technological Systems“. In: *Norms in Technology*. [Philosophy of Engineering and Technology 9], herausgegeben von Marc J. Vries, Sven Ove Hansson und Anthonie W.M. Meijers, 223–237. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-5243-6_14, zuletzt abgerufen am 30.03.2024.

Nissenbaum, Helen. 1999. „Can Trust Be Secured Online? A Theoretical Perspective“. *Etica E Politica* 1 (2). https://www.openstarts.units.it/server/api/core/bitstreams/3d6b498a-d92a-46f0-9924-abc07f8b1d50/content, zuletzt abgerufen am 28.04.2024.

O'Neill, Onora. 2020. „Questioning Trust". In *The Routledge Handbook of Trust and Philosophy*, herausgegeben von Judith Simon, 17–27. New York: Routledge. https://doi.org/10.4324/9781315542294, zuletzt abgerufen am 30.03.2024.

Ostern, Nadine. 2018. „Do You Trust a Trust-Free Transaction? Toward a Trust Framework Model for Blockchain Technology". *International Conference on Information Systems ICIS 2018 Proceedings,* https://aisel.aisnet.org/icis2018/crypto/Presentations/3, zuletzt abgerufen am 30.03.2024.

Schlatt, Vincent, André Schweizer, Nils Urbach und Gilbert Fridgen. 2016. „Blockchain: Grundlagen, Anwendungen und Potenziale. White Paper". Bayreuth: Projektgruppe Wirtschaftsinformatik des Fraunhofer-Instituts für Angewandte Informationstechnik FIT. https://fim-rc.de/wp-content/uploads/2020/02/Blockchain_WhitePaper_Fraunhofer_FIT_2016.pdf, zuletzt abgerufen am 24.05.2024.

Smits, Martin, und Joris Hulstijn. 2020. „Blockchain Applications and Institutional Trust". *Frontiers in Blockchain,* https://doi.org/10.3389/fbloc.2020.00005, zuletzt abgerufen am 30.03.2024.

Swan, Melanie, und Primavera De Filippi. 2017. „Toward A Philosophy of Blockchain: A Symposium Introduction." *Metaphilosophy* 48 (5): 603–619. https://doi.org/10.1111/meta.12270, zuletzt abgerufen am 30.03.2024.

Tang, Yong, Jason Xiong, Rafael Becerril-Arreola, und Lakshmi Iyer. 2020. „Ethics of blockchain: A framework of technology, applications, impacts, and research directions". *Information Technology & People* 33 (2): 602–632. https://doi.org/10.1108/ITP-10-2018-0491, zuletzt abgerufen am 30.03.2024.

Teng, Yan. 2021. „Towards Trustworthy Blockchains: Normative Reflections on Blockchain-Enabled Virtual Institutions". *Ethics and Information Technology* 23 (3): 385–397. https://doi.org/10.1007/s10676-021-09581-3, zuletzt abgerufen am 30.03.2024.

Teng, Yan. 2023. „What Does It Mean to Trust Blockchain Technology?". *Metaphilosophy* 54: 145–160. https://doi.org/10.1111/meta.12596, zuletzt abgerufen am 30.03.2024.

Townley, Cynthia, und Jay L. Garfield. 2013. „Public Trust". In *Trust. Analytic and Applied Perspectives*, herausgegeben von Pekka Mäkelä und Cynthia Townley, 95–107. Leiden: Brill. https://doi.org/10.1163/9789401209410_007, zuletzt abgerufen am 30.03.2024.

Völter, Fabiane, Nils Urbach und Julian Padget. 2023. „Trusting the trust machine: Evaluating trust signals of blockchain applications". *International Journal of Information Management* 68: 102429. https://doi.org/10.1016/j.ijinfomgt.2021.102429, zuletzt abgerufen am 30.03.2024.

Walch, Angela. 2019. „In Code(rs) We Trust: Software Developers as Fiduciaries in Public Blockchains". In *Regulating Blockchain: Techno-Social and Legal Challenges*, herausgegeben von Philipp Hacker, Ioannis Lianos, Georgios Dimitropoulos und Stefan Eich, 58–82. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198842187.003.0004, zuletzt abgerufen am 30.03.2024.

Wang, Wenqian, Fabrice Lumineau und Oliver Schilke. 2022. *Blockchains: Strategic Implications for Contracting, Trust, and Organizational Design*. Elements in Business Strategy. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009057707, zuletzt abgerufen am 30.03.2024.

Werbach, Kevin. 2018. *The Blockchain and the New Architecture of Trust*. Information Policy Series. Cambridge, Massachusetts: The MIT Press.

Wingreen, Stephen C., und Stephen L. Baglione. 2005. „Untangling the Antecedents and Covariates of E-Commerce Trust: Institutional Trust vs. Knowledge-Based Trust“. *Electronic Markets* 15(3): 246–260. http://dx.doi.org/10.1080/10196780500209010, zuletzt abgerufen am 30.03.2024.

Yaga, Dylan, Peter Mell, Nik Roby und Karen Scarfone. 2018. „Blockchain technology overview“. *National Institute of Standards and Technology Internal Report* 8202. Gaithersburg: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8202, zuletzt abgerufen am 30.03.2024.