

Vertrauen im Kontext – Messung und Operationalisierung

Susanna Wolf, Jan Fiete Schütte, Marc Hauer, Christopher Koska

Einleitung

Vertrauen ist für den Erfolg von datengetriebenen Unternehmungen ein entscheidender Faktor. Vor allem mit Blick auf die radikalen Umbrüche, die durch die Einführung von neuen Technologien ausgelöst werden, gilt Vertrauen für viele Unternehmen als die wichtigste Ressource (Benrath 2019). Insbesondere wenn personenbezogene Daten erhoben und verarbeitet werden, müssen User z.B. darauf vertrauen können, dass ihre Daten sicher sind und alle Beteiligten verantwortungsbewusst mit ihren Informationen umgehen. Wenn ein Unternehmen das Vertrauen verliert, kann das schwerwiegende Auswirkungen haben, beispielsweise den Verlust von Mitarbeitenden oder Kundschaft. Dies kann zu negativen Reputationseffekten führen, aber auch zu rechtlichen Konsequenzen und großen finanziellen Schäden. Positiv gewendet kann Vertrauen ein wichtiger Wettbewerbsvorteil sein. Wenn User die Wahl zwischen verschiedenen IT-Unternehmen haben, gibt es Tendenzen, dass sie in der Regel demjenigen gegenüber am loyalsten sind, dem sie am meisten vertrauen (Capgemini 2017). Deshalb ist es für Unternehmen unerlässlich, Vertrauen in ihre digitalen Produkte und Dienstleistungen aufzubauen und aufrechtzuerhalten. Doch was ist eigentlich Vertrauen? Und wie lässt es sich messen und operationalisieren?

Ein Blick auf die Vertrauensforschung zeigt: Es gibt kein allgemeingültiges Erfolgsrezept, um Vertrauen in IT-Unternehmen aufzubauen oder zu festigen. Vertrauen ist graduell und entzieht sich einer eindimensionalen Messung. Der Akt des Vertrauensschenkens hängt u.a. von der subjektiven Bereitschaft ab, sowie von der konkreten Situation, in der sich diese Personen oder Personengruppen zu einem bestimmten Zeitpunkt befinden, aber auch von der spe-

zifischen technologischen Anwendungspraxis. Deshalb lassen sich die Messgrößen von Vertrauen rein als Variablen im Rahmen eines mehrdimensionalen Vertrauensmodells kategorisieren. Die Metriken, die für die Messung und Operationalisierung von Vertrauen in datengetriebenen Projekten verwendet werden, lassen sich nur über den jeweils konkreten Anwendungskontext bestimmen.

Dieser Beitrag stellt ein Konzept zur kontextspezifischen Operationalisierung und Messung von Vertrauen für datengetriebene Geschäftsmodelle vor. Das Konzept baut auf den Erkenntnissen und Erfahrungen des interdisziplinär agierenden *Datenethik-Hubs* auf, der initial als *Community of Practice* organisiert und von DATEV zur Umsetzung der unternehmensinternen *Directive Datenethik* (DATEV 2021) ins Leben gerufen wurde (PWC 2023).¹ Im ersten Schritt wird die begriffliche Grundlage für das hier vorgeschlagene Konzept entwickelt. Anschließend wird in einem zweiten Schritt ein Werkzeug vorgestellt, welches für vorhandene Datenbestände u.a. Verantwortlichkeiten und zulässige Verarbeitungszwecke katalogisieren kann und sich bei DATEV als eine zusätzliche Option für die verantwortungsbewusste Operationalisierung einer vertrauenswürdigen Datennutzungspraxis etabliert. Im dritten Schritt werden verschiedene Ansätze und Methoden zur Messung von Vertrauenswürdigkeit vorgestellt. Der vierte Schritt beinhaltet die konzeptionelle Erweiterung des Datenkatalog-Modells, um den Boden für die kontextspezifische Messung von Vertrauen vorzubereiten. Eine abschließende Reflexion fasst die Ergebnisse zusammen. Ziel des Artikels ist es, in den akademischen Diskurs zur Operationalisierung und Messung von Vertrauen einen praxisorientierten Beitrag einzubringen, indem er abstrakte Vertrauenskonzeptionen kritisch hinterfragt und anwendungsbezogen erweitert.

1 Mit dem Projekt »Digitale Verantwortung leben. Datenethik bei DATEV« hat das Softwareunternehmen 2021 den CDR Award (Corporate Digital Responsibility Award), verliehen von BVDW und Bayern Innovativ, in der Kategorie »CDR und Mitarbeitende« gewonnen. Ausgezeichnet wurde das Projekt, da die Directive Datenethik (»DATEV Datenethik-Leitlinie«) im Rahmen eines partizipativen Prozesses gemeinsam mit diversen Stakeholdern und Mitarbeitenden entwickelt und anschließend in einer initialen, bereichsübergreifenden Community of Practice, dem sogenannten Datenethik Hub, nachhaltig in der Organisationsstruktur verankert wurde.

Begriffliche Eingrenzung des Untersuchungsgegenstandes

Die sozialwissenschaftliche Organisationsforschung unterscheidet drei Formen von Vertrauen anhand eines Ebenenmodells (Wagenblaus 2018, 1803):

- Auf der Makroebene wird das Vertrauen in gesamtgesellschaftliche Veränderungsprozesse diskutiert, also z.B. ganz grundsätzlich das Vertrauen in digitale Transformation, in Künstliche Intelligenz oder in Technik. Auch wenn wir im politischen Kontext vom Vertrauen in Demokratie oder im wirtschaftlichen Kontext vom Vertrauen in den materiellen Wert von Geld sprechen, dann befinden wir uns, wie Luhmann (1968) gezeigt hat, auf der Ebene des generalisierten Vertrauens oder Systemvertrauens.
- Auf der Mesoebene wird das Vertrauen in konkrete Institutionen oder projektbezogene Kooperationen adressiert. Die Vertrauenswürdigkeit von bestimmten Unternehmen, Institutionen und organisationsübergreifenden Kooperationen beruht hier insbesondere auf institutionalisierten Prozessen und Strukturen. Auf dieser Ebene geht es folglich um ein sehr spezifisches Vertrauen.
- Auf der Mikroebene wird das Vertrauen als ein Moment diskutiert, der aus der Interaktion und Beziehung zwischen den Vertreter*innen einer bestimmten Institution und den Vertrauensgebenden entstehen, beispielsweise zwischen den Mitarbeitenden eines Unternehmens und der Kundschaft eines Unternehmens. Auf dieser Ebene geht es vor allem darum, persönliches Vertrauen aufzubauen.

Im Fokus dieses Beitrags steht das spezifische Vertrauen, beispielsweise in eine konkrete Institution, also die Mesoebene. Darüber hinaus möchten wir zeigen, inwiefern das Vorgehen hierbei (insbesondere durch die Integration in eine CDR-Strategie)² auch positiv auf das generalisierte Vertrauen bzw. Systemvertrauen (Makroebene) einzuwirkt und von persönlichem Vertrauen (Mikroebene) getragen wird. Für die weitere Eingrenzung des Untersuchungsgegenstandes sind ferner analytische Vertrauensmodelle hilfreich, die zwischen vertrauensgebender Instanz (A), Vertrauensobjekt (X) und vertrauensnehmender

2 Einen wissenschaftlichen Einstieg in das Konzept der Corporate Digital Responsibility bietet u.a. Filipović (2023).

Instanz (B) unterscheiden (Michalski 2019, 44).³ In unserem Fall ist die vertrauensnehmende Instanz (B) ein datenverarbeitendes Unternehmen. Für dieses Unternehmen ist es in einem nächsten Schritt hilfreich – mit Blick auf das Vertrauensobjekt (X) – zwischen der externen und internen Wert-Ebene von Vertrauen zu differenzieren (van de Poel 2013, 135ff):

- Intern: Aspekte wie Sicherheit, Verlässlichkeit, Robustheit, Wartungsfreundlichkeit, Kompatibilität, Qualität, Effektivität und Effizienz lassen sich beispielsweise mit Blick auf die Leistungsstärke als Eigenschaften (interne Werte) von technischen Artefakten (Vertrauensobjekten)⁴ verstehen; insofern als Bedingungswerte für Vertrauen. In der Praxis werden die internen Eigenschaften der Vertrauensobjekte häufig als terminale Werte betrachtet. Tatsächlich handelt es sich dabei aber zumeist um instrumentelle Werte, die sich als Mittel zum Erreichen eines terminalen Wertes einsetzen lassen (van de Poel 2013, 137).
- Extern: Da es ganz wesentlich von der Bereitschaft der vertrauensgebenden Instanz (A) abhängt, dem technischen Bezugsobjekt (X) Vertrauen zu schenken, ist Vertrauen oder die Vertrauenswürdigkeit von »X« kein interner, sondern ein externer Wert. Denn die Bereitschaft Vertrauen zu schenken ist u.a. von den individuellen Einstellungen und persönlichen Präferenzen der vertrauensgebenden Instanz, aber auch von der konkreten Situation bzw. dem spezifischen Anwendungskontext abhängig.

Weiterhin findet in Schritt 4, »Kontextspezifische Erweiterung des Datenkatalog«, das ABI-Modell nach Schoorman et al. (2007) Anwendung, um konkrete Operationalisierungsmaßnahmen einzelnen Vertrauensaspekten zuzuordnen. Nach dem ABI-Modell sind diese *Ability* (Fähigkeit), *Benevolence* (Wohlwollen) und *Integrity* (Integrität). Durch diese Einstufung wird gezeigt, wie und warum die einzelnen Maßnahmen das Vertrauen im Sinne des oben genannten Ebenenmodells stärken.

Die Herausforderung bei der Messung und Operationalisierung von Vertrauen wird vor dem Hintergrund der hier skizzierten Modelle deutlich sicht-

3 Zum Entscheidungsvertrauen der vertrauensnehmenden Instanz vgl. auch Kaminski 2017, 167–169

4 Mit Blick auf den IEEE Std 7000™-2021 (IEEE Standard Model Process for Addressing Ethical Concerns during System Design) lassen sich die technischen Artefakte (Vertrauensobjekte) u.a. als SOI (system of interest) definieren.

bar. Vertrauen ist ein mehrdimensionaler Relationsbegriff. Die Frage »Wer (A) vertraut wem (B) in Bezug auf was (X)?« grenzt den Untersuchungsgegenstand dabei kontextbezogen ein. Da Vertrauen stets in spezifischer Referenz auf einen bestimmten Zusammenhang entsteht, zielt dieser Beitrag darauf ab, die Vertrauenswürdigkeit von konkreten Datenverarbeitungsszenarien (X) zu messen. Wird Vertrauen zu unspezifisch erfasst, entstehen zahlreiche Fehlerdimensionen. Deshalb ist der jeweils konkrete Bezug auf das Vertrauensobjekt (über seine effiziente Nutzbarkeit hinaus) ebenso relevant wie der auf die vertrauensnehmende Instanz (auch unabhängig vom Vertrauensobjekt). Nur auf diese Weise lassen sich Fehlerquellen durch eine einseitige Referenz identifizieren und eliminieren.

Anwendungsszenario: Operationalisierung einer vertrauenswürdigen Datennutzungspraxis

Datenverfügbarkeit gehört zu den Schlüsselkonzepten, wenn es um datengetriebene Geschäftsmodelle und Prozessautomatisierung geht. Um dies im Dienst der Genossenschaft voranzutreiben, beschäftigt sich das Data Office als Organisationseinheit bei DATEV mit Data Governance, also mit dem kontrollierten Verfügbarmachen von Daten durch adäquate Systeme und Prozesse sowie Daten zu erheben, zu verwalten und zu nutzen (Al-Ruithe 2019, Janssen 2020). Data Governance unterstützt dabei, Daten rechtskonform, effizient und wertorientiert zu verarbeiten. Der Anspruch für den hier angeführten Use Case ist auch, eine verantwortungsvolle Datenverarbeitungspraxis auf Basis der Unternehmenswerte von DATEV aufrechtzuerhalten. Das hier behandelte Anwendungsszenario zeigt exemplarisch, wie Unternehmenswerte – mitunter Vertrauenswürdigkeit – über ihre Verankerung in interne Prozesse fortlaufend in ihrer Umsetzungspraxis gestärkt werden. Das Beispiel geht insbesondere auf zwei Aspekte ein: zum einen wie digitale Verantwortung (im Sinne von Corporate Digital Responsibility, CDR) bei DATEV ausgerichtet ist, zum anderen wie sie exemplarisch zum Tragen kommt. Der hier gewählte Fokus bezieht sich auf das konzeptionelle Entwickeln und Etablieren eines wertorientierten Datenkatalogs.

Mit Bezug auf die unternehmerische Verantwortung (Corporate Social Responsibility, CSR) und ihrer Vertiefungsrichtung CDR (Anzinger 2021, Seidl 2022) können Unternehmen als Vertrauensnehmer*innen sicherstellen, dass ihr Wertbezug – somit grundsätzlich auch ihr Bezug auf Vertrauens-

würdigkeit – beim Entwickeln und Betreiben von Technologie gegenüber ihren Stakeholdern als Vertrauensgeber*innen präsent bleibt und dadurch mittelbar wie unmittelbar in Produkten, Systemen und Services als Vertrauensobjekte zur Darstellung kommt. CSR bei DATEV ist nach der klassischen Triple-Bottom-Line gegliedert: in die ökonomische, ökologische und soziale Säule (Slaper und Hall 2011). Das bedeutet: DATEV agiert im Bewusstsein der Verantwortung für Wirtschaft, Umwelt und Gesellschaft (DATEV 2023a). Digitale Verantwortung verortet DATEV als Softwarehaus bei nachhaltiger Innovation und vor diesem Hintergrund in der ökonomischen Säule. Insofern bewegt sich das hier beleuchtete Beispiel auf der Mesoebene und adressiert insbesondere wertorientierte Data Governance als Voraussetzung für die Leistungsstärke von datengetriebenen Produkten (internes Vertrauen) sowie deren vertrauenswürdiges Potential aus Stakeholder-Perspektive (externes Vertrauen). Dies zeigt sich auch mit Blick auf die Operationalisierung der erwähnten Unternehmenswerte.

Sie strukturieren dabei die Fokuspunkte von Digitaler Verantwortung im Rahmen von DATEVs Datenverarbeitungspraxis, festgehalten in der Directive Datenethik. Diese Directive als interne Regulatorik ist direkt im DATEV-Verhaltenskodex – *Code of Business Conduct* (CoBC) (ebd.) verankert und zeigt durch den konsequenten Rückbezug: Digitale Verantwortung leitet sich aus den grundsätzlichen Unternehmenswerten von DATEV ab. *Vertrauenswürdig*, *Leistungsstark*, *Partnerschaftlich*, *Führend* und *Nachhaltig* sind die Unternehmenswerte, die hierbei zum Tragen kommen:

- Vertrauenswürdigkeit ist dabei insofern kontextgebunden, als sie darauf abzielt, gemeinsam mit den je beteiligten Stakeholdern Chancen und Herausforderungen der entsprechenden datenbasierten Lösungen für den individuellen Fall angemessen zu reflektieren, »um so ein gemeinsames Verständnis für langfristige Ziele und nachhaltiges Handeln zu entwickeln« (DATEV 2021).
- Dieser Anspruch findet sich im »ganzheitlichen Wertemanagement« (ebd.) bei Leistungsstark wieder. Hier geht es darum, Innovationen wertorientiert voranzutreiben und mit Blick auf Technologie-Entwicklung und -Einsatz sicherzustellen: »Im Zentrum steht dabei immer der Mensch mit seinen individuellen Bedürfnissen und Interessen« (ebd.). Das bedeutet auch, digitale Lösungen zu entwickeln, die Mitglieder und Nutzer*innen bei ihrer Arbeit bestmöglich unterstützen und gut zu ihnen passen. Wenn

es darum geht, mit datenbasierten Innovationen Stakeholdergruppen gerecht zu werden, impliziert dies auch Fairness.

- Der Wert Partnerschaftlich geht insbesondere auf Mitgestaltungsoptionen ein: über aktiven Stakeholder-Einbezug zur Entwicklung und Weiterentwicklung von DATEV-Produkten und -Services sowie Feedbackschleifen geht es darum, »Datensouveränität genossenschaftlich [zu] ermöglichen« (ebd.).
- Um »das Potential für einen zukunftsweisenden Umgang mit Daten zu erkennen und verantwortungsvoll auszuschöpfen« (ebd.), fördert DATEV im Sinne des Wertes Führend das Bewusstsein der Mitarbeitenden für wertorientierte Technologiegestaltung und beteiligt sich am gesellschaftlichen Diskurs.
- Dabei stellt die Directive Datenethik final klar: die Unternehmenswerte bleiben in ihrem Bezug auf eine verantwortungsvolle Datenverarbeitungspraxis nicht dadurch beständig und relevant, dass sie einmalig regulatorisch fixiert sind. Es gilt, sie fortlaufend diskursiv zu prüfen, umzusetzen und darüber zu aktualisieren. Das bedeutet, sie lebendig werden zu lassen und zu leben – unter anderem über fortlaufendes Use-Case-bezogenes Umsetzen. Diesen Aspekt hält die Directive mit Bezug auf einen fünften Wert fest: »Nachhaltig. Gemeinsam Standards leben« (ebd.).

Wie die Unternehmenswerte von DATEV bei der Datenverarbeitungspraxis verinnerlicht und gleichzeitig aktuell bleiben, daran hat verantwortungsvolle Data Governance grundsätzlich ihren Anteil. Data Governance bedeutet hier unter anderem: Datenverarbeitungsprozesse auch infrastrukturell und organisatorisch zu unterstützen, um die Daten-Wertschöpfung innerhalb eines definierten regulatorischen sowie wertbezogenen Rahmens effizient zu gestalten. DATEV zielt darauf ab, verantwortungsvolle Data Governance dabei als sogenannte *Good Data Governance* zu operationalisieren: Neben Effizienz und Wertschöpfung ist die Wertperspektive im Sinne Digitaler Verantwortung (CDR) hier ebenso im Fokus wie partnerschaftliche Vernetzung mit Bezug auf vertrauenswürdige Datenräume insbesondere im europäischen Kontext. Insofern lässt sich auch dieser Anwendungsfall mit Bezug auf den hier thematisierten wertorientierten fachlichen Datenkatalog als ein Werkzeug von *Good Data Governance* betrachten, das dem zuletzt genannten Umsetzungsanspruch der Directive Datenethik Rechnung trägt.

Klar ist, dass es sich hierbei zunächst um interne Prozesse handelt. Der Wertbezug mit Blick auf die vertrauensgebende Person liegt nun darin, dass

die fortwährende prozessuale Referenz auf Unternehmenswerte – auch hinsichtlich perspektivischer Datenverarbeitungsvorhaben – dem Anspruch Rechnung trägt, dass die Entwicklung und der Betrieb von Technologie nicht losgelöst von CDR-Aspekten gedacht und umgesetzt werden. So sind interne (für diesen Fall: leistungsstarke datengetriebene Produkte) wie externe Vertrauensdimension (für diesen Fall: Vertrauen externer Stakeholder) stets reflektiert. Im genannten Wert-Kontext zueinander stärken sie über interne Prozesse die Werteverbindung von (zukünftigen) Produkten, Systemen und Services. Der Wertbezug kommt schließlich während der Nutzung durch entsprechende Stakeholder faktisch zum Tragen.

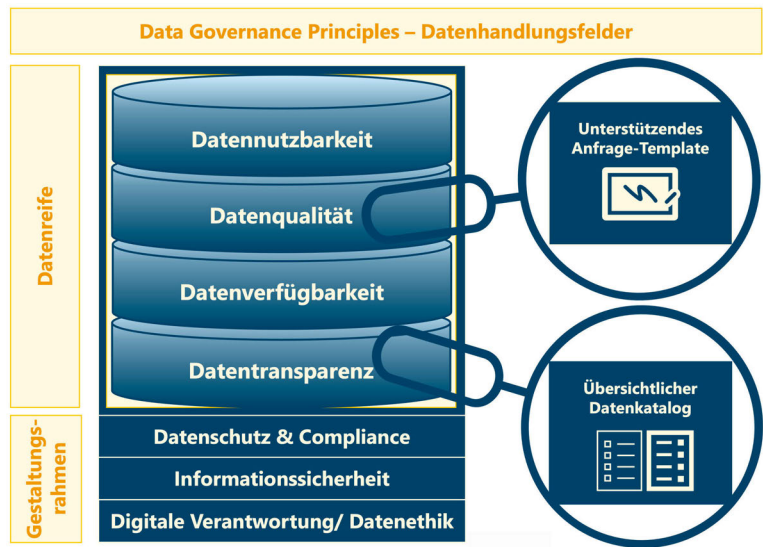


Abbildung 1: Data Governance Principles – Datenhandlungsfelder

Wie Mitarbeitende bei DATEV Datenkultur gemeinsam gestalten, halten die *Data Principles* fest. Sie beschreiben wesentliche Handlungsfelder im Rahmen der Datenverarbeitungspraxis. Dass sich das Nutzbarmachen von Daten an grundlegenden strategie- und wertorientierten Zielen nachhaltig ausrichtet, ist ein Grund dafür, dass sie zentral im *Data Office* festgehalten sind und dort – gemessen an je verschiedenen Kriterien – gemonitort werden. Einzelne Um-

setzungsmethoden und -werkzeuge, sogenannte *Wegbereiter*, operationalisieren die *Data Principles*. Die Abbildung zeigt, dass es zwei verschiedene Arten von *Data Principles* gibt:

- 1) Gestaltungsprinzipien als regulatorische und Werte-Basis legen die Leitplanken bei der Nutzbarmachung von Daten fest. Zu den Gestaltungsprinzipien gehören neben Datenschutz, Informationssicherheit und Compliance auch Digitale Verantwortung und Datenethik. Das bedeutet: die oben genannten Unternehmenswerte für eine verantwortungsvolle Datenverarbeitungspraxis sind über den Verweis auf die Directive Datenethik als Leitplanke direkt in die *Data Principles* integriert. Diese Verknüpfung bildet die normative Basis für wertorientierte *Data Governance*.
- 2) Datenreifepinzipien treiben innerhalb dieses Gestaltungsrahmens den Reifegrad des Datenbestands voran – mit Fokus auf Transparenz, Verfügbarkeit, Qualität und Nutzbarkeit.

Abbildung 1 zeigt zwei Wegbereiter, die in den Datenreifepinzipien Transparenz und Qualität verankert sind: den fachlichen Datenkatalog, sowie ein unterstützendes Anfrage-Template für Mitarbeitende, mit Datenverarbeitungsvorhaben, beispielsweise zur Produktentwicklung. Diese beiden Wegbereiter mit dem jeweiligen Prinzip, dem sie zur Umsetzung verhelfen, sollen im Rahmen des Anwendungskontextes in ihrer Wertorientierung näher betrachtet und daraufhin auf die unten dargestellte *Data Prosumer Journey* bezogen werden.

Die Journey dient in ihrer grundsätzlichen Form dazu, einen kontrollierten Lebenszyklus von Datenprodukten in Kombination mit verbundenen Bedarfs- und Umsetzungsprozessen der verantwortlichen und beteiligten Mitarbeitenden zu ermöglichen. Für datengetriebene Projekte sind das beispielsweise die verantwortlichen Mitarbeitenden, die eine bestimmte Datenanalyse umsetzen möchten, wie sogenannte *Data Scientists*. Ein Datenprodukt kann also beispielsweise eine Datenanalyse sein, die auswertet, wie effizient eine Software interne Prozesse von Kunden*innen automatisiert. Der genannte Lebenszyklus umfasst die verschiedenen Phasen des Datenprodukts – von den Konzeptions- bis zu den Löschvorgängen.

Abbildung 2 zeigt, inwieweit bei DATEV der *Data Scientist* als *Data Prosumer* im Fokus steht. *Prosumer* setzt sich dabei begrifflich aus *Consumer* und *Producer* zusammen, was bedeutet: *Data Scientists* als *Prosumer* fragen für die Analyse Daten an (*Consumer-Rolle*) und stellen in Folge über ihre Analy-

seergebnisse auch wieder Daten (als Datenprodukte) bereit (*Producer-Rolle*). Mit Fokus auf die *Data Principles* tritt an dieser Stelle Datentransparenz in den Vordergrund: für die Rolle *Data Consumer* ist wichtig, dass Daten für die geplante Analyse leicht im Bestand auffindbar sind, dass sie auf Metadatenbasis in ihren Eigenschaften (bspw. Vertraulichkeit oder zweckgebundene Verarbeitung) verständlich sind, sowie, dass leicht zuordenbar ist, wofür sie sinnvoll und wertorientiert genutzt werden können und wofür nicht. Diese Anforderungen spielen für *Producer* ebenso eine wichtige Rolle: nur wenn die Aspekte von Auffindbarkeit, Verständlichkeit und Zuordenbarkeit für sie transparent sind, gelingt es, entstehende Analysedaten wertschöpfend in den Bestand zu integrieren, um diese im Rahmen ihrer Verarbeitungszwecke wieder zur Verfügung zu stellen.

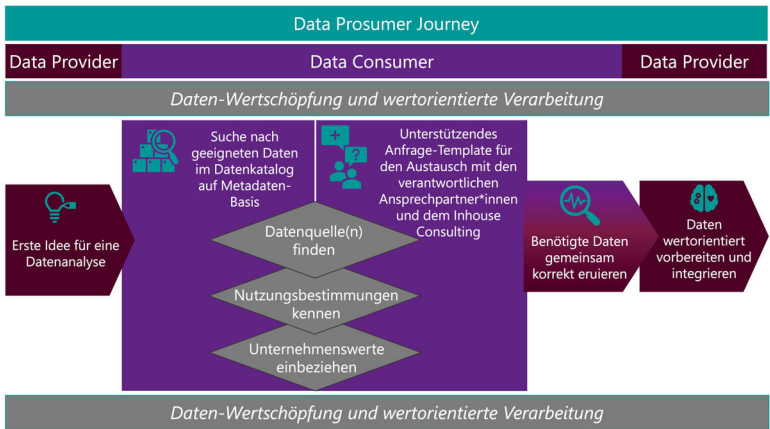


Abbildung 2: Ausschnitt aus der Data Prosumer Journey

Voraussetzung für diese gute Übersicht und damit das Datenreifeprinzip Transparenz bietet ein fachlicher Datenkatalog. Damit ist die fachliche Katalogisierung vorhandener Datenbestände auf Metadatenbasis gemeint. Verknüpft sind die Datensätze unter anderem mit entsprechenden Informationen über zulässige Verarbeitungszwecke sowie verantwortliche Ansprechpartner*innen. Dadurch ist ein zuverlässiges Wissens- und Informationsmanagement gegeben, um die Daten im Dienst der Genossenschaft effizient und vertrauenswürdig verfügbar zu machen. Hierbei zeigt sich,

dass der fachliche Datenkatalog als Werkzeug Unternehmen eine verantwortungsvolle Data Governance ermöglichen, CDR-Orientierung unterstützen und so Unternehmen in ihrer Rolle als vertrauensnehmende Instanz stärken kann. Er dient dazu, Leistungsstärke von Datenprodukten zu unterstützen (internes Vertrauen) und dadurch letztlich dazu, darauf einzuzahlen, wie dieses Potential für Vertrauenswürdigkeit wahrgenommen wird (externes Vertrauen).

Vor diesem Hintergrund benötigen *Data Prosumer* Fachkenntnisse zur unmissverständlichen Anwendung des Datenkatalogs.

Hierzu gehören:

- Verständnis der mit den Datensätzen verbundenen Logik für eine rechtskonforme Verarbeitung
- Kenntnis der Verantwortlichkeiten und Prozesse
- Transfer der Unternehmenswerte auf eine wertorientierte Datenverarbeitungspraxis
- Neben diesem Wissen ist das stete Bewusstsein für wertorientierte Verarbeitung relevant. Dazu gehört auch, Herausforderungen bei Datenanalysen wie einer ungewollten Verzerrung der Analyseergebnisse stets im Blick zu behalten. Über Dialog- und Weiterbildungsformate ist dabei relevant, folgende Aspekte zu adressieren:
 - Bewusstsein für relevante Verzerrungs-Arten (Bias)
 - Kenntnis von Optionen, diese Risiken zu reduzieren
 - Wissen um die Optionen, gewollten Bias (bspw. positive Diskriminierung, um Fairness-Potenziale auszuschöpfen) zu integrieren

Das Wissen um sowie die Sensibilität hinsichtlich der oben genannten Punkte zählt auf ein weiteres *Data Principle* ein: Datenqualität. Stichworte in diesem Zusammenhang sind unter anderem Korrektheit, Aktualität und Redundanzfreiheit der Daten. In Prüfung ist vor diesem Hintergrund auch ein digitales Template für Anfragen mit erster und zweiter Ebene. Das Template integriert auf der ersten Ebene grundsätzliche fachliche und organisatorische Informationen, wie zulässige Verarbeitungszwecke und verantwortliche Ansprechpartner*innen, eine Skizze der geplanten Datenverarbeitung, auf welche Unternehmenswerte das Vorhaben grundsätzlich einzahlt und welche Stakeholder insbesondere zu den betroffenen Zielgruppen gehören. Die zweite Ebene zielt darauf ab, Maßnahmen, Chancen und Herausforderungen der Da-

tenverarbeitung – direkt unter Berücksichtigung von Wertereferenzen – für das Unternehmen und die jeweilige Zielgruppe im Voraus zu bedenken.

Instrumente zur Messung von Vertrauenswürdigkeit

Um die voran beschriebene Operationalisierung durchzuführen und ihren Erfolg zu bewerten, ist es erforderlich, konkrete Messinstrumente für Vertrauen zu finden und anzuwenden. Eine Messung des Vertrauens in ein Produkt (wir folgen hier dem Normverständnis, dass auch Dienstleistungen Produkte sind) wäre prinzipiell über die Akzeptanzrate, oder die Akzeptanzgeschwindigkeit denkbar. Jedoch gibt es hierbei starke Einschränkungen der Aussagekraft. Zum einen gibt es keine offiziell genormten Richtwerte, ab denen man von einem hohen oder niedrigen Vertrauen sprechen könnte. Darüber hinaus können die Gründe für die (Nicht-)Akzeptanz in einer Vielzahl an möglichen Ursachen und deren Kombination miteinander – Abseits der Vertrauensmessung – begründet liegen. Es kann beispielsweise daran liegen, dass es keine oder nur schlechtere/teurere Alternativen gibt oder, dass das Produkt oder die Dienstleistung ein Novum darstellt, für das zunächst einmal ein Markt entstehen muss. Deshalb ist es, wie oben skizziert, sinnvoller, die Vertrauenswürdigkeit zu messen und diese Messung an die Stakeholder zu kommunizieren, um damit die Bildung von Vertrauen zu unterstützen.

Eine gängige Möglichkeit die Vertrauenswürdigkeit in ein Produkt oder eine Dienstleistung zu bewerten, ist das Durchführen und Auswerten von Tests. Der Testbegriff ist sehr umfassend und nicht einheitlich definiert. Gemäß ISO 29119-1:2022⁵ ist Testen ein »Satz von Aktivitäten, die durchgeführt werden, um die Entdeckung und/oder Bewertung von Eigenschaften eines oder mehrerer Testelemente zu erleichtern«, während ein Testelement »das Ergebnis einer Aktivität ist, wie z.B. Management, Entwicklung, Wartung, Tests selbst oder andere unterstützende Prozesse«. Für einen fachlichen Datenkatalog bietet es sich u.a. an, Tests bezüglich der Datenqualität (sind die Informationen bspw. vollständig, ausreichend aktuell, korrekt, redundanzfrei) durchzuführen. Für Anwendungen auf Basis einer initialen Metadaten-Recherche spielen Tests zu den Themen Zweckbindung, Vorhersagequalität (bspw. bei Analysen) und Nicht-Diskriminierung potenziell eine große Rolle. Derartige Tests basieren in der Regel auf der Berechnung von sogenannten Qualitäts- und Fairnessmaßen (Verma und

5 Software- und Systemengineering – Software-Test- Part 1: Allgemeine Konzepte.

Rubin 2018). Erreichen die gewählten Maße einen vorab definierten Mindestwert, gilt der entsprechende Test als bestanden. Sollte das Resultat eines Tests negativ ausfallen, wird das Produkt entsprechend den erwarteten Eigenschaften verbessert. Je besser also Produkte, bzw. die einzelnen Komponenten eines Produktes, getestet sind, desto höher fällt ihre Vertrauenswürdigkeit bezüglich ihrer Eigenschaften und Leistungen aus.

Tests allein sind dabei nicht ausreichend, um pauschal eine hohe Vertrauenswürdigkeit zu attestieren. Insbesondere, wenn Ziel- oder Interessenskonflikte vorliegen, zum Beispiel, weil in datengetriebenen Projekten keine klaren Rollen- und Funktionstrennungen existieren (wie beispielsweise Planung und Umsetzung durch die gleichen Personen), könnten Tests gezielt nur für Aspekte konzipiert werden, die unabhängig davon umgesetzt werden können, um dies als Qualitätssicherung zu kommunizieren und auf dieser Basis eine hohe Systemqualität zu behaupten. Insofern ist es empfehlenswert, eine Rollentrennung zu berücksichtigen zwischen denjenigen Expert*innen, die Produkte konzipieren und denjenigen, die Tests als kontrollierende Instanz in die Qualitätssicherung integrieren.

Dabei bestehen – wie grundsätzlich bei der Durchführung von Analysen und der Kommunikation ihrer Ergebnisse – Risiken: bspw. dass Testergebnisse nicht adäquat interpretiert oder missverständlich kommuniziert werden.⁶ Eine Option, diese Risiken zu reduzieren, kann in sogenannten Audits liegen. Mit Blick auf die Bedeutung von Audits durch unabhängige Prüfinstitutionen für Vertrauenswürdigkeit von Nutzenden ist erwähnenswert, dass diese neben der Selbstverpflichtung von Unternehmen im Rahmen von Ethikleitlinien eine bedeutende Rolle spielen.⁷

Auch für den Begriff Audit gibt es unterschiedliche Auffassungen. Zwei davon sind besonders relevant: Inspektionen auf der Grundlage von Normung (z.B. der Korpus der ISO-Normungsdokumente, der die Grundlage für akkreditierte Zertifizierungen bildet) und Bias-Audits von Plattformen und Institutionen (u.a. Ada Lovelace Institute in Europe (2023) oder die Initiative Zertifizierte

6 Klar ist, dass es auch Vorfälle gibt, bei denen Ergebnisse wissentlich falsch kommuniziert werden. Dies bewegt sich allerdings insbesondere im Bereich dessen, was rechtlich unzulässig ist (legale Ebene) und nicht mehr rein im datenethischen Bereich (legitime Ebene).

7 Gerade für Technologien wie Künstliche Intelligenz ist dies der Fall (ZVKI 2022; TÜV-Verband 2021).

KI des Fraunhofer IAIS (2023)). Nach ISO 19011:2018 ist ein Audit ein »systematischer, unabhängiger und dokumentierter Prozess zum Erlangen von objektiven Nachweisen [...] und deren objektiver Auswertung, um zu bestimmen, inwieweit die Auditkriterien [...] erfüllt sind« (ISO 19011: 2018).

Diese Definition bildet die Grundlage für alle Verweise auf Audits im gesamten Korpus der ISO-Normungsdokumente. Sie unterscheidet ausdrücklich zwischen Erstparteien-Audit (*1st party audit*), Zweitparteien-Audit (*2nd party audit*) und Drittparteien-Audit (*3rd party audit*). Erstparteien-Audits werden von der Organisation selbst oder in ihrem Namen durchgeführt (ISO 19011: 2018). Sie sind geeignet, um interne Vertrauenswürdigkeit in die Funktionalität interner Regelungen und Prozesse sowie deren Gültigkeit aufzubauen. Es geht also primär um die Vertrauenswürdigkeit der internen Stakeholder, bspw. in die Wirksamkeit des eigenen Managementsystems. Zweitparteien-Audits werden auf Initiative von externen Stakeholdern durchgeführt, die ein Interesse an den Qualitätsstandards einer bestimmten Organisation haben (z.B. Kunden, oder Institutionen in deren Auftrag). Drittparteien-Audits werden von unabhängigen Prüforganisationen durchgeführt, die bspw. Konformitätszertifikate oder -registrierungen ausstellen, oder von staatlichen Stellen (ISO 19011: 2018). Bei Zweit- und Drittparteien-Audits handelt es sich also um Audits durch Externe. Das bedeutet, dass Prüfpersonal Zugang zu den relevanten Informationen und Systemen erhält, um bspw. ein Produkt oder einen Prozess zu untersuchen. Dies schließt auch Tests ein. Die Aussagekraft eines Audits hängt damit auch von der Vertrauenswürdigkeit des auditierenden Prüfpersonals und des durchgeführten Auditprozesses ab. Diese Vertrauenswürdigkeit wiederum kann durch Zertifikate, die die Prüfinstitutionen selbst innehaben bemessen werden.

Ein Zertifikat ist eine Auszeichnung durch eine kompetente Instanz, die einer Person oder Sache eine bestimmte Eigenschaft bescheinigt und in dem hier behandelten Kontext auf die Vertrauenswürdigkeit eines Vertrauensnehmers oder eines Vertrauensobjekts abstellt. Bei Zertifikaten ist grundsätzlich zwischen allgemeinen und akkreditierten Zertifikaten zu unterscheiden. Im Grunde kann jeder ein Zertifikat als Bescheinigung ausstellen. Die Aussagekraft eines Zertifikats durch eine Instanz, die sich (bspw. über entsprechende Nachweise) noch kein Vertrauen verdient hat, ist jedoch sehr gering. Je vertrauenswürdiger eine Instanz ist, desto aussagekräftiger sind Zertifikate, die diese Instanz ausstellt. Am vertrauenswürdigsten sind die sogenannten akkreditierten Zertifikate, die nur durch akkreditierte Zertifizierungsstellen ausgestellt werden können. Diese wiederum sind durch die Akkreditierungsstel-

le eines Landes legitimiert (Bundesnetzagentur 2018; Datenschutzkonferenz 2020). Der europäische Rechtsrahmen stellt sicher, dass es in jedem Land der Europäischen Union genau eine Akkreditierungsstelle gibt (in Deutschland ist das die Deutsche Akkreditierungsstelle GmbH, DakkS), die (auf der Makroebene) die Kompetenz der Zertifizierungsstellen (auf der Mesoebene) bestätigt. Da akkreditierte Zertifikate auf Normen basieren müssen, sind die Mindestanforderungen zur Ausstellung akkreditierter Zertifikate für Zertifizierungsunternehmen vielfach reflektiert und eindeutig formuliert. Akkreditierte Zertifikate können sich auf verschiedene Vertrauensobjekte der jeweiligen Zertifizierungsstelle beziehen u.a.: ihre Managementsysteme (ISO/IEC 17021–1: 2015), ihre Mitarbeitenden im Sinne einer Personenzertifizierung wie Sachverständige (ISO/IEC 17024: 2021) und ihre Produkte und Prozesse (ISO/IEC 17065: 2013; Datenschutzkonferenz 2019).

Es gibt noch über Zertifizierungen hinausgehende Konzepte zur Bemessung der Vertrauenswürdigkeit. Besonders relevant sind dabei Kriterienbasierte Ansätze. Dafür wird eine Sammlung an Kriterien erstellt, die als Indikatoren für Vertrauenswürdigkeit fungieren. Das Maß an Vertrauenswürdigkeit kann dann zum Beispiel über die Anzahl erfüllter Kriterien ausgedrückt werden. Da je nach Produkt nicht alle Kriterien anwendbar sind, könnte auch das Verhältnis zwischen erfüllten und nichterfüllten Kriterien ein angemessenes Maß darstellen. Ein solches Maß ist nicht dafür geeignet die absolute Vertrauenswürdigkeit auszudrücken, jedoch um die relative Vertrauenswürdigkeit zwischen ähnlichen Produkten und Dienstleistungen (basierend auf dem gleichen Kriterienkatalog)⁸ zu ermitteln. Einen ähnlichen Ansatz verfolgen auch aktuelle Überlegungen zum Thema Corporate Digital Responsibility.⁹ Darüber hinaus können Checklisten auch alle bisher genannten Aspekte (also Tests, Auditierung und Zertifizierung) umfassen und ermöglichen damit eine holistische Bemessung der Vertrauenswürdigkeit. Gleichzeitig gibt es keine korrekte oder vollständige Sammlung an Aspekten, die in einer Checkliste erfragt werden sollten. Somit muss eine geeignete Checkliste anwendungsbezogen erstellt, sowie vergleich- und überprüfbar sein, um belastbare Aussagen daraus ableiten zu können.

8 Angemessene Kriterien für einen solchen Kriterienkatalog können mit Blick auf KI zum Beispiel auf Basis des VCIO Modells (VDE 2022, 7–47) gewonnen werden.

9 Es gilt dabei auch, die verschiedenen Konzepte von Digitaler Ethik zu berücksichtigen (Koska 2023).

Parallel zu den genannten Möglichkeiten kann die Vertrauenswürdigkeit eines Produktes auch im Kontext einer Risikobemessung bzw. Technikfolgenabschätzung betrachtet werden. Ein höheres Risiko verweist auf eine erhöhte Verletzbarkeit und erfordert deshalb ein höheres Maß an Vertrauen (Koska et al. 2024). Dadurch entsteht ein Bedarf, Mechanismen einzusetzen, die das Risiko reduzieren, bzw. Konzepte, mit den Risikofolgen kontrolliert umzugehen (Risikomanagement), um somit die Vertrauenswürdigkeit zu stärken. Diesen Weg geht zum Beispiel auch der Entwurf des AI-Acts (European Commission 2021), der eine Risikoklassifizierung auf Basis der Technologie und des Anwendungskontextes [AI-Act Annex II, Annex III und Art. 52] vorsieht. Je nach Klassifizierung werden unterschiedliche Anforderungen an ein Produkt gestellt, zum Beispiel in Bezug auf Tests, Auditierung, Zertifizierung und Transparenz. Die Kopplung der obligatorischen Etablierung von Schutzmechanismen an eine vorige Risikobemessung bzw. Technikfolgenabschätzung ist ein häufig diskutierter Lösungsansatz, um Vertrauenswürdigkeit herzustellen, ohne eine Überregulierung zu verwirklichen (Hallensleben et al. 2020).

Kontextspezifische Erweiterung des Datenkatalogs

Der folgende Abschnitt erläutert insbesondere am Beispiel des Datenkatalogs, mit welchen Methoden und Werkzeugen bei der Operationalisierung vorgegangen werden kann sowie was es dabei zu beachten gilt. Ziel ist es, eine ganzheitliche Basis für Vertrauenswürdigkeit zu schaffen – nicht nur im Rahmen des Datenkatalogs selbst, sondern mit vorausschauendem Blick auf die damit verknüpften Prozesse und Produkte.

Hierfür lässt sich wertorientierte Technologiegestaltung bei datengetriebenen Projekten grundsätzlich in drei Phasen unterteilen (IEEE SA 2021):

- 1) Begutachtung der Ist-Lage des *System of Interest* und Entwicklung von wertorientierten Zielen
- 2) Maßnahmenplanung und -durchführung – *Ethical Value Requirements*
- 3) Maßnahmenbewertung und Iteration – Qualitätssicherung der *Ethical Value Requirements*

Für die Kategorisierung einzelner Maßnahmen bietet sich das ABI-Modell an. Es ist ein Konzept für die Modellierung von Vertrauen, beschrieben durch

Schoorman et al. (2007) auf Grundlage von McAllister (1995). Im Fokus stehen dabei folgende Vertrauensaspekte:

- Ability (Fähigkeit): Die Fähigkeit des Vertrauensnehmers, die erwarteten/angebotenen Aufgaben innerhalb der Vertrauensbeziehung zu erfüllen.
- Benevolence (guter Wille): Der gute Wille gegenüber dem Vertrauensgeber, dessen Interessen zu wahren.
- Integrity (Integrität): Ein schlüssiges und konsistentes Verhalten des Vertrauensnehmers u. a. durch die fortwährende Einhaltung (selbst)gesetzter Grundsätze, welche für den Vertrauensgeber Relevanz haben.

Durch die Zuordnung von Maßnahmen zu diesen Schwerpunkten, wird deutlich, welches Ziel eine Maßnahme verfolgt. Ebenso kann die Leistung des Vertrauensnehmers für diese drei Aspekte gemessen werden, beispielsweise durch Umfragen, und es können daraus nötige Maßnahmen abgeleitet werden.

Auf Prozessebene kann hierfür bspw. der Standard IEEE 7000–2021 (IEEE SA 2021) hinsichtlich vorgesehener Prüfschritte und -kriterien berücksichtigt werden, insofern Umfang und Inhalt zum Zielprodukt passen. Es handelt sich um eine Prozessnorm für Produktentwicklung unter Berücksichtigung einer Werte-Reflexion (*Ethics by Design*), welche für das gesamte Phasenmodell anwendbar ist.

Phase 1: Begutachtung der Ist-Lage des *System of Interest* und Entwicklung von wertorientierten Zielen

Zunächst gilt es, den Anwendungskontext des Vertrauensobjekts (*System of Interest*, SOI [IEEE SA. 2021, 22]) zu skizzieren. Um Vertrauenswürdigkeit für ein bestimmtes Produkt (auf der Mesoebene) aufzubauen, ist unter anderem die Identifizierung und der Einbezug der relevanten Stakeholdergruppen (bei Bedarf über entsprechende repräsentative Personen) entscheidend. Der Datenkatalogs als SOI integriert sich weiterhin in eine bestehende Produktlandschaft. Der Katalog bietet aber nicht nur die Basis dafür, eine Vielzahl an unterschiedlichen Datennutzungsszenarien zu verwalten, sondern ermöglicht auch neue Anwendungsfelder wertorientiert zu erschließen. In diesem Sinne kann der Datenkatalog u. a. als ein Management-Tool betrachtet werden, das dazu dient, datengetriebene Projekte vertrauenswürdig zu verwalten

und zu gestalten, insbesondere in Bezug auf deren Beziehung zueinander, wie sie in Systemen mit gegenseitiger Referenz auftreten können (*System of Systems*, SoS [IEEE SA. 2021, 22]).

Bei der Entwicklung von Zielen gilt es, bestehende Richtlinien, Kodizes und Werte innerhalb des Unternehmens anzuwenden. Dafür setzt DATEV unter anderem die *Directive Datenethik* ein. Unter Einsatz der fünf Unternehmenswerte werden Handlungsbedarfe ermittelt und kontextspezifische Ziele gesetzt sowie Kriterienkataloge erstellt. Durch die fortwährende Umsetzung der *Directive* zielt DATEV darauf ab, auf die Integrität nach dem ABI-Modell einzuzahlen. Der Datenkatalog ist ein Werkzeug, welches insbesondere durch geprüfte Daten die Intention unterstützt, dass die ihm zugrundeliegenden Werte in datengetriebenen Projekten berücksichtigt werden. Dies kann einerseits dazu beitragen, die Integrität (*Integrity*) zu verbessern und andererseits das Zutrauen in die Fähigkeiten (*Ability*) des Unternehmens stärken, da hochwertige Datensätze die Leistung der Produkte erhöhen.¹⁰

Außerdem können auch Werkzeuge aus der Organisationsmethodik genutzt werden, wie *Canvas-Formate* (*Ethics Canvas* (Online Ethics Canvas n. d.), *Data Ethics Canvas* (Open Data Institute n. d.), *Digital Product Ethics Canvas* (Gerlach 2019)), welche ebenfalls dabei unterstützen, die Rahmenbedingungen und Implikationen der Produkte strukturiert zu reflektieren und zu überblicken. Ein *Canvas* ist eine visuelle Organisationsmethodik, die komplexe Projekte übersichtlich und nachvollziehbar darstellt. Das Canvas bietet den Rahmen für mehrere untergeordnete Kategorien, die als Schlüsselemente eines Projekts zu skizzieren sind. Im Rahmen des Datenkatalogs wird derzeit beispielsweise mit Canvas-Formaten gearbeitet, um den Anfrageprozess durch interne Stakeholder (bspw. o.g. Data Prosumer) zu strukturieren. Dies bietet Potenzial für einen effektiven Anfrage-, Beratungs- und Datenbereitstellungsprozess.

Für weitere Tätigkeiten, wie die der Technikfolgenabschätzung mit Fokus auf die Wertorientierung, gibt es eine große Auswahl von Werkzeugen. Als Einstieg bietet sich das Verzeichnis der OECD an, welches eine Vielzahl an Hilfsmitteln zum Thema »Vertrauenswürdige KI« umfasst und zudem umfangreiche Filter- und Suchoptionen bietet (OECD.AI Policy Observatory n.

10 Der erwähnte IEEE 7000 beginnt analog zu Phase 1 mit den Fragen nach dem Concept of Operations und dem Context Exploration Process (Kap. 7). Hierbei werden die Ziele der Organisation, sowie das ethische Umfeld ergründet. Im nächsten Schritt, dem Ethical Values Elicitation and Priorization Process (Kap. 8), werden die kontextspezifischen Werte ermittelt und priorisiert.

d.). Für die Verwendung in Projekten oder Unternehmen lässt sich empfehlen, eine eigene Werkzeugsammlung zu erstellen und zu pflegen, die speziell auf die individuellen Bedürfnisse zugeschnitten ist. Eine Auswahl ergänzend zum OECD-Katalog sei an dieser Stelle genannt:

Für Technikfolgenabschätzung in ethischer Dimension können Reflexionsmethoden wie das MEESTAR-Modell (Weber 2016, Ethical OS n. d.) oder Workshop-Formate, bspw. Ethical Explorer (Ethical Explorer n. d.), hilfreich sein. Workshops bieten die Möglichkeit, Stakeholder unter den relevanten Diversitätsgesichtspunkten miteinander ins Gespräch zu bringen und über die verschiedenen Projektphasen hinweg als multiperspektivischen Reflexionskreis, angepasst an das jeweilige Produktszenario, einzubinden. Andere Standards und Frameworks in Bezug auf beispielsweise Risikomanagement, Informationssicherheit (ISO/IEC 27001) und KI-Systeme (ISO/IEC 42001, ISO/IEC TR 24368, ISO/IEC 23053) können ergänzende Leitplanken für die Umfeldbeobachtung und Zielentwicklung stellen. Sollte die Anwendung von Prozessstandards nicht möglich sein, kann stattdessen auf *Playbooks* aufgesetzt werden, wie den *AlgoRules* (Algo.Rules n. d.) oder dem *etami open guidebook* (Etami open guidebook n. d.). Um Prozesse und Datenflüsse aufzuschlüsseln, können Programme, wie beispielsweise der *Data Process Modeler* (Bayern innovativ n. d.), eingesetzt werden. Die hier genannten Werkzeuge sind zum Teil umfangreich und erfordern mitunter ethische Expertise für einen effektiven Einsatz. Geeignete Werkzeuge sind deshalb kontextspezifisch (insbesondere nach Umfang und Aufwand entsprechend der Produkt- und Prozessgegebenheiten) auszuwählen oder anzupassen.

Ebenso gibt es für die wertorientierte Ausrichtung von Unternehmen eine Vielzahl an Kodizes und Leitbildern, welche als Orientierungspunkt für die Reflexion der unternehmenseigenen Haltung dienen können. Als Inspiration genannt seien hier zum Beispiel die *Digital Responsibility Goals* (Meier et al. 2022) und der CDR-Kodex (Corporate Digital Responsibility Initiative 2023). Bei der Formulierung und Reflexion des eigenen Leitbildes sind partizipative *bottom-up*-Prozesse essenziell, um bereits bestehende Unternehmenswerte zu integrieren und der eigenen Unternehmenskultur angemessen Rechnung zu tragen. Darauf aufbauend lässt sich die Operationalisierung und Messung der Vertrauenswürdigkeit weiter optimieren, bspw.: von der unternehmensinternen Datenethik-Leitlinie (*Directive*) über interdisziplinäre Dialogformate zu Data Governance und Datenethik (*Datenethik Hub*, *Enterprise Data Council*) bis hin zu den konkreten internen Beratungs- und Unterstützungsleistungen mit

Blick auf datengetriebene Projekte (Datenkatalog; PWC 2023; DATEV 2023b, S. 7).

Die erste Phase ist entscheidend für den weiteren Verlauf der Entwicklung und den Aufbau von Vertrauenswürdigkeit. Ohne eine geordnete, klare und nachvollziehbare Aufarbeitung der Rahmenbedingungen, sind Ziele nur schwer zu formulieren und eine vertrauenswürdige Entwicklung schwer zu kommunizieren. Im gleichen Zug ermöglicht diese tendenziell explorative Phase eine Konkretisierung der Produktleistungen (Hauer et al. 2023). In Bezug auf das Vertrauen wird erkundet, wer die potenziell vertrauensgebende Instanz ist. Außerdem wird festgestellt, mit welchen Maßnahmen die vertrauensnehmende Instanz die Vertrauenswürdigkeit der Produkte stärken kann.

Phase 2: Maßnahmenplanung und -durchführung – *Ethical Value Requirements*

Auf Basis der festgelegten Ziele und Bedingungen gilt es im nächsten Schritt konkrete Maßnahmen zu planen und durchzuführen. Funktionalitäten und Systemeigenschaften (*Ethical Value Requirements*, EVR [IEEE SA. 2021, 18]) werden unter Berücksichtigung der Werte aus Phase 1 bestimmt und tragen diese in die praktische Umsetzung. Bei bestehenden Produkten werden an dieser Stelle nötige systemische Anpassungen gefunden, um fortlaufend auf ein einheitliches und vertrauenswürdiges Produktportfolio einzuzahlen. Um die Umsetzung zu gewährleisten, sollten für diese Phase ebenfalls geeignete Dokumentationswerkzeuge und Kommunikationsmaßnahmen bestimmt werden. In Bezug auf wertebasierte Produktentwicklung können ergänzende *Ethical Checkboxes* in bereits etablierte Checklisten und Projektprozesse integriert werden. Aus dem Umfeld der *Data Science* sei hier beispielhaft das Werkzeug *deon* (deon n. d.) genannt, womit sogenannte *Ethical Checkpoints* (ECP) in datengetriebenen Projekten erstellt werden können. Es kann hilfreich sein, auch in dieser Phase Workshop-Formate zu nutzen, um die Suffizienz

der verabschiedeten Maßnahmen aus Sicht verschiedener Stakeholder beurteilen zu können.¹¹

Für die Maßnahmenplanung und -durchführung greift DATEV bei Bedarf auf Dialogformate zu wertorientierter Data Governance, wie den *Enterprise Data Council* oder den Datenethik-Hub, zurück. In regelmäßigen Terminen und einzelnen Workshops werden die priorisierten Werte aus Phase 1 in konkrete Anforderungen und Maßnahmen übersetzt. Durch die Dokumentation in gemeinsamen Arbeitsumgebungen ist es möglich, diese zu reflektieren und regelmäßig zu überprüfen. Der Datenkatalog trägt dazu bei, dass sich Anfrageprozess und Datennutzung nachvollziehen lassen. Dies ist eine wichtige strukturelle Voraussetzung für *first party audits* oder Tests. In diesem Sinne findet ein Beitrag zu den Parts *Integrity* und *Ability* des ABI-Modells statt.

Phase 3: Maßnahmenbewertung und Iteration – Qualitätssicherung der *Ethical Value Requirements*

Zur Bewertung der erfolgten Maßnahmen, wird der Erfüllungsgrad von wer-tebasierten Zielen geprüft. Hierzu können Aktivitäten aus Phase 1 herangezogen werden, wie die Ergebnisse der Technikfolgenabschätzung. Des Weiteren können externe Instanzen für Audits oder Zertifikate hinzugezogen werden. Erreicht das geschaffene Produkt nachweislich und für die Stakeholder nachvollziehbar die gesteckten Ziele, dann stärkt dies die Vertrauenswürdigkeit des Vertrauensobjekts und in der Regel auch die der vertrauensnehmenden Instanz, insbesondere wenn die Technikgestaltung in Orientierung an den gesetzten Werten erfolgt ist (Integrität). Wurde ein Kriterienkatalog angelegt, kann dieser auf Erfüllung – intern und/oder extern – geprüft werden. Sollte die Erfüllungsrate geringer ausfallen als gewünscht, gibt er Orientierung für die Fokuspunkte der weiteren normativen Produktentwicklung. Bei ausreichender Erfüllungsrate kann der Kriterienkatalog selbst und der Grad der Erfüllung geprüft oder zertifiziert und als Vertrauensgrundlage an die vertrauensgebende Instanz kommuniziert werden.

11 Ein weiteres mögliches Ergebnis dieser Phase ist ein anwendungsbezogener Kriterienkatalog, welcher konkrete Anforderungen an das Produkt stellt. Dieser kann als Basis für externe Prüfungen genutzt werden, um die Qualität des Produkts von außen zu bestätigen. Um vage Metriken zu vermeiden, ist eine Finalisierung der Kriterien in Phase 2, statt schon in Phase 1, vorzuziehen.

Im Sinne des ABI-Modells kann durch Tests unter Berücksichtigung der Produktleistung die Ability der vertrauensnehmenden Instanz untersucht werden. Eine hohe Leistung lässt sich im Kontext des geprüften Produkts intersubjektiv nachvollziehbar extern kommunizieren. Gleichzeitig kann zusätzliche wertorientierte Weiterentwicklung die Integrität des Vertrauensnehmers fördern. Über Szenarienkataloge lassen sich gebündelte Kriteriensets prüfen und Aussagen über die Vertrauenswürdigkeit für spezifische Szenarien treffen.¹²

Die Directive Datenethik lässt sich im Kontext der Maßnahmenbewertung für den genannten Use Case als Möglichkeit heranziehen, wertorientierte Anforderungen an datengetriebene Produkte interdisziplinär zu überprüfen. Die o.g. Dialogformate kommen u.a. dann zum Einsatz. Hierbei referenziert der Prozess auf die Phase 1, insofern Potenziale mit Blick auf operationalisierte Werte identifiziert werden. Dieses Vorgehen bietet eine Basis, um eine perspektivische Auditierung strukturiert vorzubereiten.

Zusammenfassung und Fazit

Der hier vorgeschlagene Weg, Vertrauen in datengetriebene Projekte aufzubauen und zu festigen, verknüpft konzeptionelle Überlegungen mit einem praxistauglichen Lösungsansatz. Ausgehend von analytischen Vertrauensmodellen und Ansätzen aus der sozialwissenschaftlichen Organisationsforschung skizziert dieser Beitrag ein bedarfs- und anwendungsorientiertes Werkzeug, um Vertrauen zu messen und zu operationalisieren. Mit Blick auf den inflationären und zumeist vagen Gebrauch des Begriffs Vertrauen grenzt dieser Beitrag den Untersuchungsgegenstand zunächst auf das Konzept der Vertrauenswürdigkeit von konkreten Datennutzungsszenarien ein. Im anschließenden Anwendungsszenario dient ein wertorientierter fachlicher Datenkatalog zur Operationalisierung einer vertrauenswürdigen Datennutzungspraxis. Der nächste Schritt erörtert problemorientiert die Möglichkeiten von Vertrauensmessung. Abschließend zeigt der Beitrag, welche weiteren Werkzeuge sich zur Sicherstellung einer vertrauenswürdigen

12 Um die Vertrauenswürdigkeit des Produkts für verschiedene Nutzungsprofile differenziert zu betrachten, werden die jeweils relevanten Kriterien in Kriteriensets zusammengefasst. Beispielsweise kann zwischen B2B- und B2C-Nutzung unterschieden werden, da sich die beiden in Nutzungsgrundlage und -umständen unterscheiden.

Datennutzungspraxis anbieten und wie diese zu einer kontextspezifischen Erweiterung des Datenkatalogs führen können.

Zusammenfassend lässt sich festhalten, dass sich die Operationalisierung und Messung von Vertrauen nicht allein über Akzeptanzraten oder -geschwindigkeiten steuern lässt. Die Kombination von internen und externen Tests, Audits und Zertifizierungen mit kriterienbasierten Ansätzen bietet eine umfassendere Bewertung der Vertrauenswürdigkeit. Dabei bleibt der Aspekt »Kontext« entscheidend: So geht es im Rahmen dieses Beitrags nicht darum, ein basales Urvertrauen zu begründen oder eben genau unbegründet zu lassen. Sondern es geht darum, Vertrauen in seiner Dimension als Entscheidungsvertrauen zu verorten. Aus wirtschaftlicher Sicht betrachtet: Kunden entscheiden sich, unter Bezug auf ihre aktuelle Einsicht in den jeweiligen Zusammenhang, für oder gegen einen bestimmten Service, ein Produkt oder eine Dienstleistung. Persönliche Vertrauensbeziehungen auf der Mikroebene stärken das Vertrauen in bestimmte Unternehmen oder Institutionen auf der Mesoebene. Dies trägt wiederum dazu bei, das allgemeine Systemvertrauen auf der Makroebene aufrechtzuerhalten. Dieser Zusammenhang erklärt die oben ausgeführte hohe Bedeutung von Stakeholder-Einbezügen über unterschiedliche Formate. Für die IT-Branche hat dies zur Folge, dass insbesondere eine vertrauenswürdige Datennutzungspraxis im Rahmen der aktuellen Technikgestaltung essenziell für das Vertrauen in die Zukunft datengetriebener Geschäftsmodelle ist. Es ist die Voraussetzung dafür, Mehrwert aus Daten gemeinsam auf einem ethischen Fundament zu gestalten. Zukünftige Forschung und Entwicklungen sollten darauf abzielen, anwendungsorientierte und kontextspezifische Lösungen für die Vertrauensbewertung zu entwickeln und gesellschaftliche wie auch unternehmerische Belange miteinbeziehen. Mit Blick auf die Wertorientierung gilt es eine angemessene Balance zwischen innovativer Technikgestaltung und angemessener Regulierung zu finden.

Literatur

- Ada Lovelace Institute. 2023. »Ada in Europe.« Accessed July 18, 2023. <https://www.adalovelaceinstitute.org/our-work/europe/>.
- Al-Ruithe et al. 2019. »A systematic literature review of data governance and cloud data governance.« In: personal and ubiquitous computing, 2019, Vol.23 (5–6), p.839-859. <https://doi.org/10.1007/s00779-017-1104-3>.

- Anzinger, Heribert M. 2021. Corporate Digital Responsibility. In: Nietsch, Michael (Hg.): Corporate Social Responsibility Compliance. München, S. 611–633.
- Algo.Rules. N. d. »Regeln für die Gestaltung algorithmischer Systeme.« Accessed July 18, 2023. <https://algorules.org/de/startseite>.
- Bayern Innovativ. N. d. »Data Process Modeler. Wie aus Transparenz Wettbewerbsvorteile werden.« Accessed July 18, 2023. <https://www.bayern-innovativ.de/de/seite/data-process-modeler>.
- Benrath, Bastian. 2019. »Die wichtigste Ressource für Tech-Unternehmen. Vertrauen der Nutzer.« In Frankfurter Allgemeine Zeitung. Last modified September 19, 2019. <https://www.faz.net/aktuell/wirtschaft/diginomics/vertrauen-der-nutzer-wichtigste-ressource-fuer-tech-unternehmen-16386248.html?GEPIC=s5>.
- Bundesnetzagentur. 2018. »Verordnung (EG) Nr. 765/2008«. Accessed July 18, 2023. https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Technik/D_MUEF/VO_EG_765_2008.html.
- Capgemini Digital Transformation Institute survey. 2017. »The key to Loyalty.« Online unter: https://www.capgemini.com/wp-content/uploads/2017/11/dti_loyalty-deciphered_29nov17_final.pdf.
- Corporate Digital Responsibility Initiative. 2023. »Kodex.« Accessed July 18, 2023. <https://cdr-initiative.de/kodex>.
- Datenschutzkonferenz (DSK). 2019. »Akkreditierungsprozess für den Bereich »Datenschutz« gemäß Art. 42, 43 DS-GVO.« https://www.datenschutzkonferenz-online.de/media/oh/20190315_oh_akk_c.pdf.
- Datenschutzkonferenz (DSK). 2020. »Anforderungen zur Akkreditierung gemäß Art. 43 Abs. 3 DS-GVO i. V. m. DIN EN ISO/IEC 17065«. Last modified October 8, 2020. https://www.datenschutzkonferenz-online.de/media/ah/20201008_din17065_Ergaenzungen_deutsch_nach_opinion.pdf.
- DATEV eG. 2021. DATEV Directive Datenethik. Online zukünftig unter: <https://www.datev.de/web/de/ueber-datev/das-unternehmen/corporate-responsibility/>.
- DATEV eG. 2023a. DATEV-Verhaltenskodex – Code of Business Conduct. Online unter: <https://www.datev.de/web/de/m/ueber-datev/das-unternehmen/compliance/verhaltenskodex/>.
- DATEV eG 2023b. Whitepaper »Datenschutz und Unternehmenssicherheit bei DATEV. Online unter: <https://www.datev.de/web/de/m/ueber-datev/daten-schutz/>.

- deon. N. d. »An ethics checklist for data scientists.« Accessed July 18, 2023. <https://deon.drivendata.org/>.
- Deutscher Ethikrat (Hg.). 2023. »Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz.« Stellungnahme. Accessed July 18, 2023. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>.
- etami open guidebook. N. d. »The open guidebook on legal, trustworthy, and ethical Artificial Intelligence.« Accessed July 18, 2023. <https://guidebook.etami.org/>.
- Ethical Explorer. N. d. »Ethical Explorer.« Accessed July, 2023. <https://ethicalexplorer.org/>.
- Ethical OS. N. d. »Ethical OS Toolkit.« Accessed July 18, 2023. <https://ethicalos.org/>.
- European Commission. 2021. »Proposal for a Regulation of the European Parliament and of the Council laying down perationa rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.« Accessed July 18, 2023. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>.
- Filipović, Alexander. 2023. »Corporate Digital Responsibility« (Preprint). In *The Routledge Handbook of Responsibility*, edited by Maximilian Kiener. New York, London.
- Fraunhofer IAIS. 2023. »Künstliche Intelligenz sicher und vertrauenswürdig gestalten.« Accessed July 18, 2023. <https://www.iais.fraunhofer.de/de/forschung/kuenstliche-intelligenz/ki-zertifizierung.html>.
- Gerlach, Robert. 2019. »The Digital Product Ethics Canvas.« Accessed July 18, 2023. <https://www.threebility.com/post/the-digital-product-ethics-canvas>.
- Hallensleben, Sebastian, Carla Hustedt, Lajla Fetic, Torsten Fleischer, Paul Grünke, Thilo Hagendorff et al. 2020. »From Principles to Practice. An interdisciplinary framework to operationalize AI ethics.« Edited by VDE and Bertelsmann Stiftung. AI Ethics Impact Group. <https://www.ai-ethics-impact.org/en>.
- Hauer, Marc, Lena Müller-Kress, Gertraud Leimüller, Katharina Zweig. 2023. »Using Assurance Cases to assure the fulfillment of non-functional requirements of AI-based systems – Lessons learned.« In *IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. DOI: 10.1109/ICSTW58534.2023.00040.

- IEEE SA. 2021. »IEEE Standard Model Process for Addressing Ethical Concerns during System Design« Accessed July 18, 2023. <https://standards.ieee.org/ieee/7000/6781/>.
- ISO/IEC 17021–1:2015 Konformitätsbewertung – Anforderungen an Stellen, die Managementsysteme auditieren und zertifizieren.
- ISO/IEC 17024:2021 Konformitätsbewertung – Allgemeine Anforderungen an Stellen, die Personen zertifizieren.
- Janssen, Marijn et al. 2020. »Data governance: Organizing data for trustworthy Artificial Intelligence.« In: Government information quarterly. Vol.37 (3), p.101493, Article 101493. <https://doi.org/10.1016/j.giq.2020.101493>.
- Kaminski, Andreas. 2017. »Hat Vertrauen Gründe oder ist Vertrauen ein Grund? – Eine (dialektische) Tugendtheorie von Vertrauen und Vertrauenswürdigkeit.« In Praxis und »zweite Natur«, edited by Jens Kertscher and Jan Müller, S. 167–188. Brill | mentis. https://doi.org/10.30965/9783957438249_017.
- Koska, Christopher. 2023. Ethik der Algorithmen. Auf der Suche nach Zahlen und Werten. (1. Auflage 2023.) Berlin: Springer Berlin; J.B. Metzler (Techno:Phil – Aktuelle Herausforderungen der Technikphilosophie, 6). <https://doi.org/10.1007/978-3-662-66795-8>.
- Koska, Christopher, Julian Prugger, Sophie Jörg, Michael Reder. 2024. Die Verlagerung von Vertrauen vom Mensch zur Maschine. Eine Erweiterung des zwischenmenschlichen Vertrauensparadigmas im Kontext Künstlicher Intelligenz. Zeitschrift für Praktische Philosophie (ZfPP).
- Leitfaden zur Auditierung von Managementsystemen. (ISO 19011:2018); Deutsche und Englische Fassung EN ISO 19011:2018. Note 3.1.
- Luhmann, Niklas. 1968. Vertrauen. Ein Mechanismus der Reduktion sozialer Komplexität. Stuttgart: Enke.
- McAllister, D. J. 1995. »Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations.« In Academy of Management Journal 38 (1), S. 24–59. DOI: 10.2307/256727.
- Meier, Jutta, Hermsen, Kai, Bauer, Jochen und Bjoern Eskofier. 2022. Digital Responsibility Goals – A Framework for a Human-Centered Sustainable Digital Economy with a Focus on Trusted Digital Solutions. DOI: 10.3233/SHTI220377.
- Michalski, Niels. 2019. Normatives und rationales Vertrauen in Europa. Eine ländervergleichende Untersuchung gesellschaftlicher Vertrauensniveaus. Wiesbaden: Springer Fachmedien Wiesbaden.

- OECD.AI Policy Observatory. N. d. »Catalogue of Tools & Metrics for Trustworthy AI.« Accessed July 18, 2023. <https://oecd.ai/en/catalogue/tools?terms=&page=1>.
- Online Ethics Canvas. N. d. »Online Ethics Canvas.« Accessed July 18, 2023. <https://www.ethicscanvas.org/>.
- Open Data Institute. N. d. »What is the Data Ethics Canvas?«. Accessed July 18, 2023. <https://theodi.org/article/the-data-ethics-canvas-2021/>.
- PWC (Hg.). 2023. Digitale Ethik & Verantwortung. Wie eine langfristige wirkungsvolle Verankerung in ihrem Unternehmen gelingt.
- Schoorman, F. David, Roger C. Mayer, James H. Davis. 2007. »An Integrative Model of Organizational Trust: Past, Present, and Future.« In *The Academy of Management Review* 32 (2), S. 344–354. DOI: 10.5465/amr.2007.24348410.
- Seidl, Martina. 2022. »Corporate Digital Responsibility: Stimulating Human Centric Innovation and Building Trust in the Digital World.« In: Jacob, Kai et al. (Hg.): *Liquid Legal – Humanization and the Law*. Cham, S. 55–81.
- Slaper, Timothy F., Tanya J. Hall. 2011. »The Triple Bottom Line: What Is It and How Does It Work?« In *Indiana Business Review* 86 (1), S. 4–8. <https://www.ibrc.indiana.edu/ibr/2011/spring/pdfs/article2.pdf>.
- TÜV-Verband. 2021. »Studie Künstliche Intelligenz.« Online unter: https://www.tuev-verband.de/?tx_epxelo_file%5bid%5d=856780&cHash=2a8f26dde3cbe286f5d30efd9759dd.
- van de Poel, Ibo. 2013. »Werthaltigkeit der Technik.« In *Handbuch Technikethik*, edited by Armin Grunwald und Melanie Simonidis-Puschmann. Stuttgart: J.B. Metzler, S. 133–137.
- VDE. 2022. »VCIO based description of systems for AI trustworthiness characterisation. VDE SPEC 90012 V1.0 (en).« Accessed July 18, 2023. <https://www.vde.com/resource/blob/2242194/a24b13db01773747e6b7bba4ce20ea60/vcio-based-description-of-systems-for-ai-trustworthiness-characterisationvde-spec-90012-v1-0--en--data.pdf>.
- Verma, Sahil, Julia Rubin. 2018. »Fairness definitions explained.« In *Proceedings of the International Workshop on Software Fairness – FairWare '18*. the International Workshop, edited by Yuriy Brun, Brittany Johnson und Alexandra Meliou, S. 1–7. Gothenburg, Sweden, 29.05.2018 – 29.05.2018. New York, New York, USA: ACM Press.
- Wagenblass, Sabine. 2018. »Vertrauen.« In *Handbuch Soziale Arbeit. Grundlagen der Sozialarbeit und Sozialpädagogik* (6., überarbeitete Auflage), edi-

ted by Hans-Uwe Otto, Hans Thiersch, Rainer Treptow und Holger Ziegler. München: Ernst Reinhardt Verlag.

Weber, Karsten. 2016. »MEESTAR² – Ein erweitertes Modell zur ethischen Evaluierung soziotechnischer Arrangements.« Conference Paper: Zweite transdisziplinäre Konferenz zum Thema »Technische Unterstützungssysteme, die die Menschen wirklich wollen« 2016. https://www.researchgate.net/publication/311699459_MEESTAR_-_Ein_erweitertes_Modell_zur_ethischen_Evaluierung_sozio-technischer_Arrangements.

ZVKI. 2022. ZVKI-Online-Befragung: »Wissen, Nachvollziehbarkeit und bewertbare Erfahrungen – Zutaten für vertrauenswürdige Künstliche Intelligenz (KI).« Online unter: <https://www.zvki.de/zvki-exklusiv/fachinformationen/online-befragung>