Von Datenknappheit zur Datenhoheit[1]

Endkonsument*innen als aktive Gestalter*innen für mehr Fairness in der Datensammlung

K. Valerie Carl

1. Einleitung

Daten spielen für die Entwicklung Künstlicher Intelligenz (KI)-basierter Systeme eine zentrale Rolle. Insbesondere im Kontext fairer KI wird die Bedeutung von Daten zunehmend diskutiert, da Daten eine zentrale Quelle für Unfairness darstellen können (z. B. Feuerriegel et al. 2020). Dabei spielt beispielsweise das Set an Trainings- und Testdaten eine große Rolle, um möglichst diskriminierungsfreie Systeme zu entwickeln. Auch die Art des Trainings, die Wahl der Modelle und wie diese Systeme implementiert werden, sowie die Wahl entsprechender Algorithmen (z. B. Wahl von Pre- oder Post-Processing Algorithmen) werden im Sinne fairer KI vielfach diskutiert (z. B. von Zahn et al. 2022).

1 Am Projekt COMET arbeiten das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) und die Goethe Universität Frankfurt/Main (GU) zusammen. An der Entwicklung der Mechanismen haben Prof. Dr. Oliver Hinz (GU), Prof. Dr. Oliver Thomas (DFKI), Constantin Brîncoveanu (GU), Dr. K. Valerie Carl (GU), Léon Dankert (DFKI), Dr. Tobias Dreesbach (DFKI), Rebecca Heigl (GU) und Aaron Witzki (GU) zusammengearbeitet. Das Projekt COMET wird vom Bundesministerium für Wirtschaft und Energie (BMWE) im Rahmen des Technologieprogramms „SmartLivingNEXT – Künstliche Intelligenz für nachhaltige Lebens- und Wohnumgebungen“ gefördert. Mit SmartLivingNEXT entsteht ein universelles, KI-basiertes Ökosystem für die einfache und kostengünstige Entwicklung intelligenter und an Nachhaltigkeit ausgerichteter Smart Living-Dienste und -Anwendungen. Insbesondere der unkomplizierte und sichere Datenaustausch zwischen allen Beteiligten wird damit ermöglicht.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Fairness kann in Bezug auf Daten noch weitergedacht werden. So kann faire KI bereits einen Schritt davor beginnen und auch eine faire Akquise von Daten von Endkonsument*innen, also den Datenerzeuger*innen, beinhalten (siehe Abbildung 1). Dabei gehen Gedanken hinsichtlich Fairness über diverse und diskriminierungsfreie Daten hinaus und beinhalten auch die Art und Weise wie Daten für das Training KI-basierter Systeme gesammelt werden. Das heißt beispielsweise, dass sich Endkonsument*innen tatsächlich bewusst sind, dass gerade Daten von ihnen gesammelt werden, an wen diese weitergegeben werden und wofür (bspw. für das Training KI-basierter Systeme) diese Daten eingesetzt werden. Oftmals ist dies in der Praxis nicht der Fall (vgl. Acquisti et al. 2018).

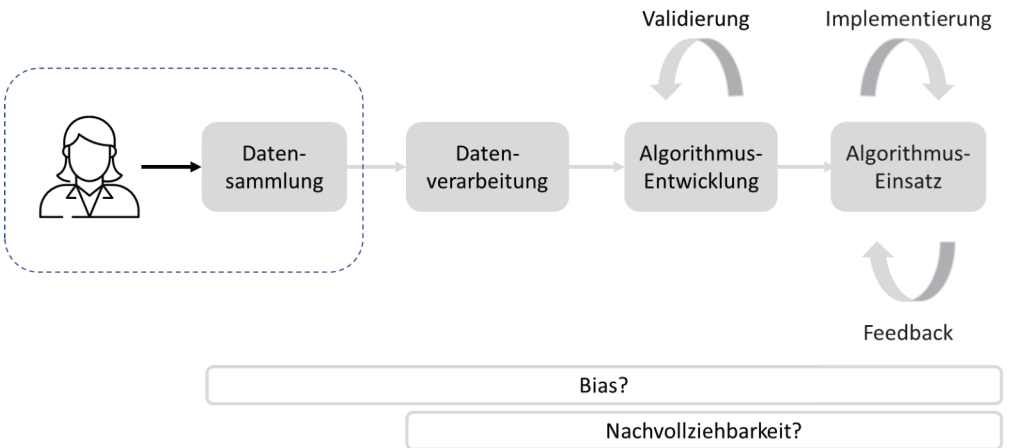


ABBILDUNG 1: BEISPIELHAFTE ENTSCHEIDUNGEN FÜR FAIRE KI (QUELLE: EIGENE DARSTELLUNG IN ANLEHUNG AN VAN HOUTEN 2020)

Endkonsument*innen-Daten werden nicht nur in den Unternehmen genutzt, mit denen Endkonsument*innen eine direkte Interaktion haben, sondern teilweise auch über Unternehmensgrenzen weitergegeben. So werden sie beispielsweise zwischen Töchtergesellschaften und der Muttergesellschaft geteilt. Aber auch unabhängige Unternehmen können ein Interesse an diesen Endkonsument*innen-Daten haben. Insbesondere kleine und mittelständische Unternehmen (KMUs) und Startups stehen oftmals vor einem Henne-Ei-Problem: Für das Training KI-basierter Systeme werden zumeist (größere) Datenmengen benötigt und stellen damit eine Voraussetzung für das

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Angebot innovativer digitaler Dienste und Produkte dar. Gleichzeitig sind derartige Daten vor Angebot solcher Produkte und Dienstleistungen in KMUs und Startups häufig nicht vorhanden (vgl. Witzki et al. 2025a). Entsprechend stellt sich die Frage, wie Daten für KI-basierte Services (extern) akquiriert werden können. Oftmals werden diese nicht über Unternehmensgrenzen hinweg zur Verfügung gestellt. Dafür können Datenschutzbedenken, Wettbewerbsvorteile, Interoperabilitätsprobleme oder Vertrauen in fremde Organisationen Gründe sein. Zudem ist die Weitergabe von Daten über Unternehmensgrenzen hinweg auch aus Fairness-Perspektive kritisch zu evaluieren. Ohne eine ausreichende Datengrundlage ist allerdings das Training KI-basierter Systeme nicht möglich. Gerade im Sinne eines fairen Wettbewerbs und Chancen für KMUs und Startups sowie das Aufbrechen sogenannter Datensilos (vgl. Tallon 2013) ist die Verfügbarkeit von Daten über Unternehmensgrenzen hinweg in einer fairen Art und Weise ein Schlüssel für innovative, verantwortungsvolle Systeme der Zukunft, und damit die Verfügbarkeit von Daten außerhalb einiger weniger Unternehmenskonglomerate.

Digitale Ökosysteme, speziell Datenökosysteme, bieten eine sich stetig entwickelnde Plattform, Daten zu akquirieren (vgl. Oliveira et al. 2019; Scheider et al. 2023). Sie können einen rechtlichen sowie Werte-Rahmen für den Austausch von Daten bieten sowie spezifische Mechanismen bereitstellen, die den Datenaustausch über Unternehmensgrenzen hinweg gesichert ermöglichen. Derartige Datenökosysteme entwickeln sich zunehmend in der Realität, beispielsweise in den Domänen Smart Living und Mobilität.

Bisher teilen im Rahmen von Datenökosystemen insbesondere Unternehmen untereinander Daten, die oftmals von Endkonsument*innen erzeugt werden (vgl. Scheider et al. 2023). Im Sinne von mehr Fairness, ist die Einbindung von Endkonsument*innen als aktive Akteure in diesen Austausch möglich. Für eine faire Akquise können in diesen Ökosystemen zusätzliche Mechanismen implementiert werden, die Endkonsument*innen ermächtigen, aktiv zu handeln und damit Daten fair und informiert bereitzustellen (vgl. Dreesbach et al. 2025). So behalten Endkonsument*innen die Hoheit über ihre eigenen Daten, werden in ihrer Souveränität gestärkt und gleichzeitig werden so gegebenenfalls mehr Daten über Unternehmensgrenzen hinweg fair und verantwortungsvoll verfügbar.

Im Folgenden werden zunächst Faktoren und Einflüsse auf eine faire Datensammlung über Unternehmensgrenzen hinweg im Kontext KI-basierte Systeme aufgezeigt, um ein Verständnis

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

zu ermöglichen, welche Aspekte zu einer fairen Datensammlung für KI-basierte Systeme beitragen können. Darauf aufbauend werden spezifische Mechanismen sowie Design-Entscheidungen in einem digitalen Datenökosystemen beleuchtet, die eine solche faire Datenbereitstellung in der Praxis unterstützen und flächendeckend möglich machen können. So wird für die Domäne Smart Living exemplarisch gezeigt, wie eine fairere Datensammlung sichergestellt werden kann. Die entwickelten Mechanismen sind domänenunabhängig anwendbar und sollen exemplarisch notwendige Entscheidungen und mögliche Designvorschläge aufzeigen. Zuletzt wird ein Ausblick auf mögliche weitere Entwicklungen gegeben.

2. Faire Datensammlung für KI-basierte Systeme

Im Sinne fairer KI und der freiwilligen Übernahme zusätzlicher Verantwortung, können Organisationen über gesetzliche Anforderungen hinausdenken und stärker im Sinne fairer KI handeln, als sie verpflichtet sind (vgl. Mihale-Wilson et al. 2021). Soll also die Datensammlung für KI-basierte Systeme besonders fair und verantwortungsvoll ausgestaltet werden, bieten rechtlich bindende Anforderungen einen guten Ausgangspunkt. Die Deutsche Datenschutz-Grundverordnung (DSGVO) hat für informiertes Einverständnis von Endkonsument*innen den rechtlichen Grundstein gelegt. In der Praxis geben Endkonsument*innen jedoch zumeist zu (längeren, teilweise sehr komplizierten) Datenschutzerklärungen per Checkmark die Zustimmung – oftmals ohne diese wirklich zu lesen.[2] Damit erfolgt zwar eine informierte Einwilligung, die Frage ist allerdings, wie bewusst und fair diese Zustimmung ist und ob Endkonsument*innen immer die Tragweite ihrer Entscheidungen bewusst ist. Unternehmen können zudem Interfaces so gestalten, dass Endkonsument*innen zur Zustimmung zusätzlich incentiviert („nudging“) werden (vgl. Acquisti et al. 2018). Insbesondere im Fall digitaler (Daten-)Ökosysteme erfolgt diese Zustimmung gegebenenfalls nicht nur für das Unternehmen, mit dem direkt eine Geschäftsbeziehung besteht, sondern im Rahmen der Datenschutzerklärung wird gegebenenfalls der Weitergabe an Unternehmen im größeren Unternehmensverband oder sogar für das ganze Ökosystem zugestimmt. Im Sinne einer möglichst fairen Datensammlung ist also die Frage, ob die Zustimmung zu regulären

[2] Einige Unternehmen ergreifen hier bereits zusätzliche Maßnahmen, um eine informierte Einwilligung zu unterstützen (z. B. One-Pager, FAQs, einfache Erläuterungen). Jedoch sind solche zusätzlichen Informationen und unterstützenden Angebote noch nicht flächendeckend verfügbar.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Datenschutzerklärungen (und AGBs) wirklich die volle Souveränität von Endkonsument*innen gewährleistet und eine faire Datenweitergabe über Unternehmensgrenzen hinweg sicherstellt.

Selbst wenn diese Erklärungen gelesen werden, ist die Frage, ob Endkonsument*innen immer verstehen, wie und wofür ihre Daten an wen weitergegeben werden. Aufgrund der teils komplexen und juristischen Formulierungen kann diese für Endkonsument*innen teilweise schwer nachvollziehbar sein. Entsprechend ist im Sinne einer fairen Datensammlung fraglich, ob in jedem Falle eine tatsächlich informierte Entscheidung getroffen werden kann. Im Sinne einer möglichst fairen Datenverwendung sollte Endkonsument*innen klar und verständlich erläutert werden, wofür ihre Daten durch wen genutzt werden (vgl. Carl/Hinz 2024), um eine tatsächlich informierte Zustimmung zu gewährleisten.

Zudem können oftmals keine Anpassungen der Datenweitergabe vorgenommen werden. Vielmehr müssen Endkonsument*innen, sofern sie ein bestimmtes Produkt oder einen bestimmten Service nutzen wollen, der Datenschutzerklärung und damit der hierin dargelegten Datenverwendung zustimmen. Oftmals können nicht für einzelne Zwecke Anpassungen vorgenommen werden. Sind solche Einstellungen anpassbar, steuern trotz allem manche Unternehmen die Datennutzung über die sogenannten Default-Settings, also Voreinstellungen, die festgelegt werden und die mit zusätzlichem Aufwand angepasst werden müssen. Endkonsument*innen müssen in diesem Fall aktiv Änderungen vornehmen, was von vielen Endkonsument*innen nicht genutzt wird (vgl. Acquisti et al. 2018). Eine Personalisierung ist nur durch Initiative von Endkonsument*innen möglich, oftmals nicht bekannt oder wird nicht genutzt. Zudem steuern Unternehmen über den damit verbundenen (zeitlichen) Aufwand, wie viele Anpassungen tatsächlich vorgenommen werden. Je länger eine entsprechende Anpassung dauert, desto weniger wahrscheinlich ist, dass viele Endkonsument*innen ihre Möglichkeit der Anpassung nutzen.

Endkonsument*innen haben in der Praxis oftmals nicht das Gefühl, die Datenverarbeitung tatsächlich kontrollieren zu können (vgl. Scheider et al. 2023). Diese Kontrolle ist für viele Endkonsument*innen allerdings eine zentrale Voraussetzung für das Teilen von Daten (vgl. Aitken et al. 2016). Zudem ist die Frage, wie einfach diese Kontrolle ausgeübt werden kann (vgl. Hanneke et al. 2023). Je niederschwelliger Lösungen sind, desto eher werden Endkonsument*innen ihre Rechte nutzen.

Datenakquise und deren Verwendung für KI ist nicht per se unverantwortlich, vielmehr können Maßnahmen ergriffen und Mechanismen implementiert werden, die diese fairer und damit

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

verantwortungsvoller machen. Wichtig ist das Bewusstsein von Organisationen, dass faire KI bereits bei der Akquise von Daten für das Training beginnt. Oftmals ist eine Fremdakquise über Unternehmensgrenzen hinweg notwendig, weil entsprechende Daten im Unternehmen nicht verfügbar sind. Dabei ist ebenfalls zu beachten, dass Verantwortung nicht an der eigenen Unternehmensgrenze aufhört, sondern das akquirierende Unternehmen auch bewusste Entscheidungen hinsichtlich aller beteiligten Akteure in Betracht ziehen sollte. So sollte sich das akquirierende Unternehmen fragen, ob die Daten von Endkonsument*innen als Datenerzeuger*innen verantwortungsvoll erhoben wurden und eine verantwortungsvolle und faire Weitergabe stattfindet. Somit ergeben sich im Sinne einer fairen Datenakquise und Verwendung als Baustein fairer KI folgende Schlüsselfragen:[3]

- SF1: Sind sich Endkonsument*innen der Datenverwendung bewusst?
- SF2: Erfolgt die Datenverwendung für Endkonsument*innen unerwartet über Unternehmensgrenzen hinweg?
- SF3: Ist die Datenverwendung für Endkonsument*innen verständlich?
- SF4: Kann eine Personalisierung der Datenverwendung durch Endkonsument*innen vorgenommen werden?
- SF5: Haben Endkonsument*innen die Kontrolle über das Teilen der Daten?
- SF6: Erfolgt die Datensammlung, -akquise und -verwendung bei allen beteiligten Akteuren verantwortungsvoll?

3. Datenökosysteme als Treiber für faire Datenakquise bei gleichzeitig höherer Verfügbarkeit

Digitale Datenökosysteme stellen eine Plattform für den Austausch von Daten dar (vgl. Oliveira et al. 2019). Je nach Governance und Struktur kann ein solcher Datenaustausch mehr oder minder verantwortungsvoll vonstattengehen. Bisher werden dabei vor allem zwei Akteure betrachtet: organisationale Datenbereitsteller (z. B. App-Anbieter) und organisationale Datenkonsumenten (z. B. KI-Service-Anbieter) (siehe Abbildung 2) (vgl. Scheider et al. 2023). Zukünftig

[3] Die vorgestellten Schlüsselfragen sollen als Diskussionsanstoß dienen und stellen keine abschließende Aufzählung dar. Auf Basis bisheriger Forschung und Beobachtungen in der Praxis wurden besonders präsente Schlüsselfragen ausgewählt, die eine zentrale Rolle für faire Datenakquise spielen.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

könnten Endkonsument*innen allerdings aktiver in diesen Datenaustausch eingebunden werden und souverän handeln (siehe Abbildung 2) (vgl. Oliveira et al. 2019).

ABBILDUNG 2: DATENAKQUISE VON ENDKONSUMENT*INNEN-DATEN (QUELLE: EIGENE DARSTELLUNG)

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Wenn Endkonsument*innen als souveräne Entitäten in Datenökosystemen agieren und ihre Daten informiert und bewusst für organisationale Datenkonsumenten bereitstellen können, werden einerseits Endkonsument*innen in ihren Rechten gestärkt. Dafür sollten sie ihre Rechte aber auch möglichst einfach ausüben können, um die Hürde für ihre Ausübung möglichst gering zu halten. So wird faire KI über Training und Implementierung hinausgedacht, indem auch eine faire Datenakquise verfolgt wird. So kann gegebenenfalls auch die Datenverfügbarkeit über Unternehmensgrenzen hinweg unterstützt werden. Wenn Endkonsument*innen ihre Daten freiwillig und informiert an andere Unternehmen übertragen, fallen für das bisher verwaltende Unternehmen (z. B. organisationaler App-Anbieter) rechtliche Bedenken weg, die diese Übertragung bisher behindert haben. Zudem können Endkonsument*innen zusätzlich incentiviert sein, ihre Daten freiwillig und informiert zur Verfügung zu stellen. Ähnlich wie beim Crowdfunding können Daten für das Training innovativer KI-basierter Services bereitgestellt werden mit der Motivation, diesen Service zukünftig selbst nutzen zu können. Beispielsweise in der Medizinforschung ist dieser Effekt bereits zu beobachten. Patient*innen stellen ihre Daten freiwillig zur Verfügung, damit eigene Krankheiten besser erforscht werden können oder um der Gesellschaft als Ganzes zu helfen (z. B. Aitken et al. 2016). Die meisten KI-basierten Systeme fallen nicht in die Kategorie der Medizinforschung. Trotz allem kann beispielsweise die Aussicht auf bessere Energieprognosen und Kosteneinsparpotenziale oder mehr Komfort Endkonsument*innen motivieren, ihre Daten freiwillig für das Training von KI-basierten Systemen bereitzustellen.

Um diese Souveränität zu ermöglichen und eine fairere Datenakquise zu verwirklichen, müssen verschiedene Mechanismen und Governances in digitalen Datenökosystemen implementiert werden, die Endkonsument*innen die direkte und einfache Kontrolle über ihre Daten geben und gleichzeitig die Möglichkeit, ihre Daten souverän freizugeben.

In der Domäne Smart Living werden solche Mechanismen im SmartLivingNEXT Ökosystem[4] durch das Projekt COMET[5] pilotiert (siehe Abbildung 3). Diese Mechanismen stellen einen möglichen Weg dar, Endkonsument*innen zu souveränen Akteuren in digitalen (Daten-)Ökosystemen zu ermächtigen und ein bewusstes und faires Datenteilen zu unterstützen. Endkonsument*innen werden über bestehende Apps aus dem Bereich Smart Living (z. B. Energieverbrauchsmonitoring und Assisted Living) an das digitale Ökosystem angebunden. Die pilotierten Mechanismen sind sowohl in

4 Für eine ausführliche Einführung des Ökosystems vgl. SmartLivingNext (2025).

5 Für eine ausführliche Einführung des Projektes vgl. SmartLivingNext in COMET (2025).

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

bestehenden Apps anwendbar wie im Projekt gezeigt wird, als auch in eigens entwickelten Apps umsetzbar. So können bestehende Unternehmen und Datenökosysteme aber auch sich neu entwickelnde Ökosysteme und Anwendungen auf die pilotierten Mechanismen zurückgreifen und sie domänenunabhängig implementieren. Anhand des Projektes COMET und der Domäne Smart Living soll exemplarisch gezeigt werden, wie Datensouveränität bei Endkonsument*innen einerseits zu faireren KI-basierten Systemen und andererseits zu besserer Datenverfügbarkeit führen kann.

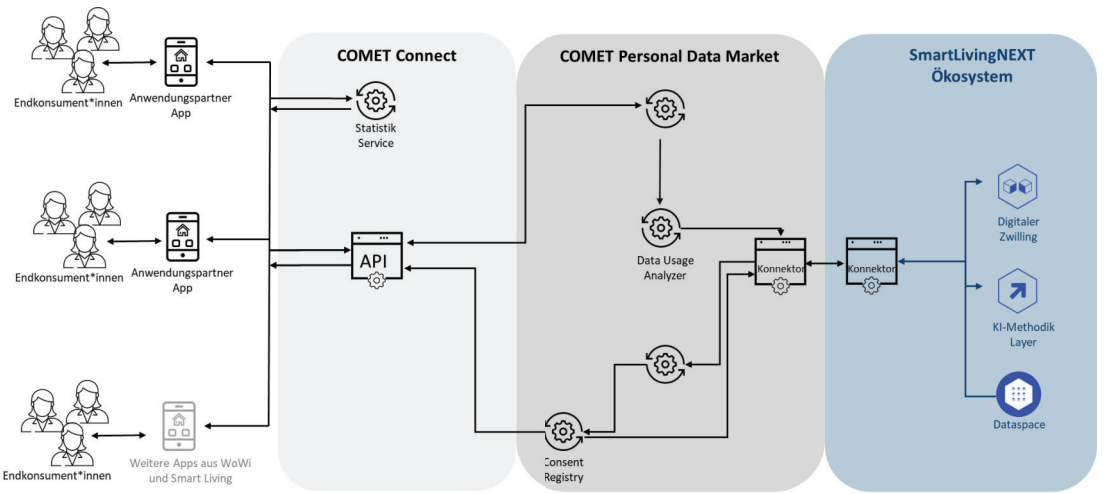


ABBILDUNG 3: PILOTIERTE MECHANISMEN IM PROJEKT COMET (QUELLE: EIGENE DARSTELLUNG IN ANLEHUNG AN COMET 2025)

Die genannten sechs Schlüsselfragen (siehe Kapitel 2) waren zentral bei der Entwicklung entsprechender Mechanismen im Projekt COMET. Die Sicherstellung der fairen Datensammlung über verschiedene Akteure hinweg (SF6) wird primär durch die aktive Einbindung von Endkonsument*innen als souveräne Entitäten adressiert. Endkonsument*innen sollen im Rahmen dieser Mechanismen selbst entscheiden können, wer ihre Daten erhält und was mit ihnen passiert und treten damit an die Stelle der bisher häufigen Entscheidungsträger: organisationale Datenbereitsteller wie App-Anbieter. Diese Souveränität soll möglichst niederschwellig und einfach ausübbar sein. Endkonsument*innen erhalten eine aktive und selbstbestimmte Rolle im digitalen Datenökosystem, sodass akquirierende Unternehmen direkt mit Endkonsument*innen in Kontakt kommen und nicht wie bisher zumeist mit einem

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

zwischengeschalteten Unternehmen, mit dem die Datenerzeuger*innen in direkter Interaktion stehen. Durch diese direkte Interaktion mit Endkonsument*innen kann von einer faireren Datenakquise ausgegangen werden, da Endkonsument*innen ihre Daten bewusst, informiert, fair und freiwillig teilen. Entsprechend übernimmt das akquirierende Unternehmen auch über Unternehmensgrenzen hinweg Verantwortung, indem eine faire Datensammlung Voraussetzung für die Akquise der Daten ist. Es werden also nur Daten akquiriert, die fair, informiert und freiwillig gewonnen wurden durch die direkte Einbindung von Endkonsument*innen. Im Gegensatz dazu, sind beispielsweise unbewusst erhobene Scrolling-Daten weniger verantwortungsvoll erhobene Daten, worüber sich auch das akquirierende Unternehmen bewusst sein und dafür Verantwortung übernehmen sollte.

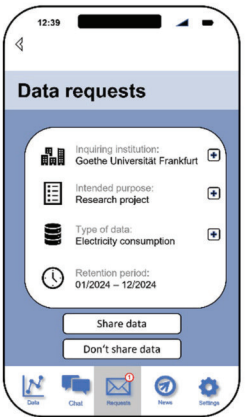
Darüber hinaus sollen in diesem exemplarischen Ökosystem Endkonsument*innen für die freiwillige und faire Datenfreigabe einen persönlichen Vorteil erhalten, unabhängig davon, wie oft oder regelmäßig sie ihre Daten freigeben, um die tatsächliche Freiwilligkeit der Freigabe sicherzustellen. Derart freigegebene Daten sind Innovations-Enabler, sodass Endkonsument*innen von innovativeren Produkten und Services profitieren. Darüber hinaus wird ihnen von einer zentralen Entität des digitalen Ökosystems SmartLivingNEXT ein weiterer Service zur Verfügung gestellt, der illustriert, was mit freigegeben Daten erreicht werden kann. Im Falle des SmartLivingNEXT Ökosystems und des Projektes COMET ist dies ein sogenannter „Statistik-Service", der Energie- und Assisted Living-Daten auswertet und mit Hilfe von KI personalisierte Empfehlungen, Vorhersagen und Warnungen (in den Bereichen Energie und Assisted Living) ausgibt. So sollen Endkonsument*innen durch das Ökosystem einen persönlichen Vorteil erhalten als Motivation für mögliche freiwillige Datenfreigaben. Gleichzeitig wird mit diesem Statistik-Service gezeigt, welche Systeme mit Hilfe derartiger Datenfreigaben zukünftig durch die akquirierenden Unternehmen entwickelt werden könnten. Damit geht das digitale Ökosystem in Vorleistung gegenüber den Endkonsument*innen ohne erwartete Gegenleistung. Wichtig dafür ist, dass Datenfreigaben keine Voraussetzung für die Nutzung dieses Service sind. Vielmehr soll der Service illustrieren, was mit Hilfe von Datenfreigaben möglich ist und einen Mehrwert für das digitalen Datenökosystem als Ganzes darstellen. In verschiedenen Use Cases in der Domäne Smart Living konnte bereits gezeigt werden, dass durch derartige Statistik-Services ein Mehrwert für Endkonsument*innen entstehen kann (z. B. Brîncoveanu et al. 2024).

Ein weiterer zentraler pilotierter Mechanismus ist die sogenannte „Consent Registry". Dieser Mechanismen hat Einzelanfragen für die Datenfreigabe zum Zweck, die von Endkonsument*innen im Sinne einer gestärkten Souveränität geschätzt werden (vgl. Hanneke et al. 2023). Das heißt

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Endkonsument*innen sollen nicht durch (einmalige) Zustimmung zu einer Datenschutzerklärung ihre Daten freigeben. Vielmehr sollen sie je Verwendungszweck und akquirierendem Unternehmen entscheiden können, ob sie ihre Daten für diesen Zweck freigeben wollen (vgl. Witzki et al. 2025a) (siehe Abbildung 4).

(a) Initiale Ansicht einer Datenanfrage für Endkonsument*innen



(b) Detaillierte Ansicht einer Datenanfrage für Endkonsument*innen (ausklappbare Felder)

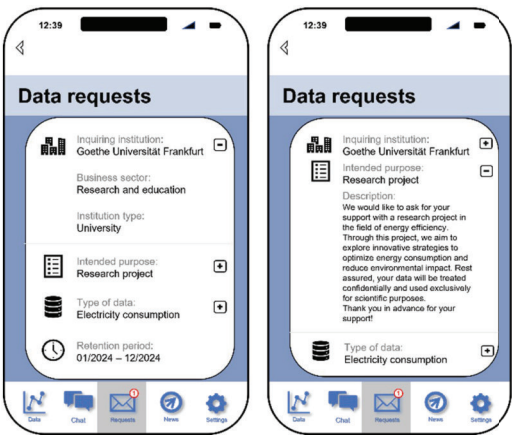


ABBILDUNG 4: ANSICHT VON DATENANFRAGEN FÜR ENDKONSUMENT*INNEN IM RAHMEN DER CONSENT REGISTRY (QUELLE: WITZKI ET AL. 2025A)

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Die Consent Registry soll die Souveränität von Endkonsument*innen stärken und eine Personalisierung der Datenfreigaben soll einfach und intuitiv möglich sein (SF4). Durch diese Einzelanfragen soll sichergestellt werden, dass jeder Freigabe bewusst und informiert zugestimmt wird (SF1) und somit der Grundstein für faire Datenweitergabe und damit faire KI gelegt wird. Durch die Verwendung der Consent Registry für Datenweitergabe und die explizit benötigte Zustimmung zu jeglicher Verwendung, entsteht bei Endkonsument*innen ein größeres Bewusstsein und Verständnis von wem und wofür ihre Daten verwendet werden (SF1). So werden Endkonsument*innen in ihrer Datenhoheit und Souveränität gestärkt.

Durch die proaktive Bereitstellung eines Mechanismus zur Datenfreigabe auf Basis einer Opt-in-Lösung, kann eine vollständige Personalisierung der Datenfreigabe erfolgen. Die Datenfreigabe ist keine Voraussetzung für die Nutzung von Diensten oder Produkten, sodass kein Zwang besteht, der Verarbeitung zuzustimmen. Durch eine möglichst einfache und intuitive Einbindung in neue und bestehende Applikationen soll zudem die Nutzungsschwelle so gering wie möglich gehalten werden, um eine Personalisierung möglichst einfach zu machen. So bleibt auch die Kontrolle der Datenfreigabe komplett bei Endkonsument*innen. Die Datenfreigabe erfolgt bewusst und informiert ohne Verpflichtung oder als Trade-off gegenüber der Nutzung von Produkten und Services und somit freiwillig. Dadurch erhalten Endkonsument*innen volle Souveränität und behalten die Datenhoheit (SF5).

Insbesondere die Datenweitergabe über Unternehmensgrenzen hinweg (SF2) wird für Endkonsument*innen dadurch transparenter und eine souveräne Entscheidung kann je Einzelfall getroffen werden. Das heißt, bei gleichem Zweck, Dauer, etc. kann trotzdem zwischen Unternehmen unterschiedlich entschieden werden, weil beispielsweise das Vertrauen zu Unternehmen A größer ist als zu Unternehmen B oder die erwartete Technologie interessanter. Auch im Zeitverlauf kann zwischen gleichen Zweck, Datentyp und Unternehmen abweichend entschieden werden. Es erfolgt eine Zustimmung auf Einzel-Case-Basis, die nicht auf zukünftige Anfragen angewendet wird. Insbesondere wird nicht auf Unternehmensbasis, sondern auf Use-Case-Basis eine Entscheidung zur Datenfreigabe getroffen, sodass kein Unternehmen Blanko-Daten bekommt, mit denen beliebige Zwecke verfolgt werden können (z. B. das Training mehrere verschiedener KI-basierter Services). Unternehmen müssen für jeden geplanten Dateneinsatz eine eigene Anfrage an die Endkonsument*innen stellen. Insgesamt wird so mehr Bewusstsein und Transparenz sichergestellt, was schlussendlich zu einer faireren und bewussteren Datenbereitstellung beiträgt.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Voraussetzung für Anfragen von akquirierenden Unternehmen über die Consent Registry ist zudem, dass Anfragen leicht verständlich formuliert werden. Zudem wird den akquirierenden Unternehmen ein Formular bereitgestellt, das genutzt werden muss, sodass Endkonsument*innen Anfragen immer in der gleichen Struktur erhalten und alle benötigten Informationen für eine informierte Entscheidung bereitgestellt werden (SF3). So ist eine einfachere Vergleichbarkeit zwischen Anfragen gegeben. Eine gleichbleibende Struktur reduziert zudem für Endkonsument*innen den nötigen Aufwand, sich in neue Anfragen einzuarbeiten und entsprechende Entscheidungen zu treffen.

Des Weiteren sollen Endkonsument*innen nicht nur souverän ihre Daten freigeben können, sondern ihre Souveränität und Datenhoheit soll auch langfristig geschützt werden, mit Hilfe des sogenannten „Data Usage Analyzer". Dieser Mechanismus soll Endkonsument*innen die langfristige Nachverfolgbarkeit, was mir ihren Daten geschieht, ermöglichen und ihnen somit die Kontrolle über ihre Daten vereinfachen (SF5). So soll darüber jede getätigte Freigabe ersichtlich sein, wofür diese Daten im akquirierenden Unternehmen genutzt wurden, ob diese bereits gelöscht wurden oder wann diese gelöscht werden und weitere Informationen zur Freigabe langfristig und zugänglich gespeichert werden. So müssen nicht wie nach DSGVO möglich Anfragen zur Datennutzung gestellt werden. Vielmehr sind diese Informationen jederzeit über die angebundenen Apps abrufbar. So wird für Endkonsument*innen die Ausübung ihres Rechts auf Information, Anpassung und Löschung vereinfacht durch einen konsolidierten, einfach nachvollziehbaren Überblick über die Datenverwendung und so ihre Souveränität gestärkt. Damit wird auch nach Datenfreigabe zur Personalisierbarkeit beigetragen (SF4), indem eine persönliche Übersicht der genutzten Daten bereitgestellt wird. Entsprechend hört Verantwortung für implementierende Unternehmen nicht bei der Datenfreigabe selbst auf, sondern auch danach wird weiter Verantwortung im Sinne fairer KI übernommen. Derartige nachträgliche Transparenz über die tatsächliche Verwendung der Daten kann das Vertrauen von Endkonsument*innen stärken und so schlussendlich gegebenenfalls mehr Datenfreigaben begünstigen, da die Transparenz hinsichtlich der Verwendung ermöglicht wird und so die Kontrolle über die Daten bei den Endkonsument*innen unterstützt und leichter ausübbar ist.

Die pilotierten Mechanismen können sowohl innerhalb eines Datenökosystems als auch innerhalb eines einzelnen Unternehmens eingesetzt werden. Unternehmen können auch ihren eigenen Kund*innen Souveränität geben, einfach, verständlich und fair ihre Daten zu verwalten und

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Datenhoheit über eigens erzeugte Daten zu behalten. So kann beispielsweise auch an eigene Kund*innen (nicht nur an unternehmensexterne Endkonsument*innen) eine Anfrage gestellt werden, ob Daten für das Training eines bestimmten KI-basierten Systems genutzt werden können. So können Unternehmen die in COMET pilotierten Mechanismen auch für den internen Einsatz weiterdenken und so eine fairere Datenakquise intern verfolgen.

4. Ausblick

Zusammengefasst sollte ein möglicher Ansatz für faire Datenakquise gezeigt und damit ein Beitrag zu fairer KI und der Übernahme von Verantwortung im Sinne von Corporate Digital Responsibility (CDR) geleistet werden. Faire KI umfasst eine Vielzahl verschiedener Bausteine wie diverse und diskriminierungsfreie Trainingsdaten, faire Systeme, bestenfalls erklärbare Entscheidungen und ein ethischer und fairer Einsatz solcher Systeme.

Die hier vorgestellten pilotierten Mechanismen und Anwendungen sollen einen Beitrag dazu leisten, faire KI auch aus einer anderen, weniger diskutierten Brille zu beleuchten: der Fairness in der Datenakquise und -nutzung für das Training von KI. Insbesondere aufgrund des Bedarfs, Daten über Unternehmensgrenzen hinweg verfügbar zu machen, können Endkonsument*innen als souveräne Akteure in digitalen Datenökosystemen eine wichtige Rolle spielen. Einerseits können Endkonsument*innen so souveräner ihre Daten kontrollieren sowie deren Verwendung. Andererseits werden Daten so über Unternehmensgrenzen hinweg eher verfügbar gemacht, sodass insbesondere KMUs und Startups Zugang zu ihnen erhalten bei gleichzeitiger Übernahme von Verantwortung und Verfolgung des Ziels fairer KI. Faire KI bedeutet auch, dass Endkonsument*innen souverän mit ihr interagieren. Und das beginnt bereits bei der Art und Weise, wie Daten für das Training akquiriert werden. Die vorgestellten Mechanismen und Governance-Aspekte sollen einen möglichen Weg aufzeigen, wie die Souveränität von Endkonsument*innen gestärkt werden kann bei gleichzeitiger Stärkung von KMUs und Startups in ihrer Möglichkeit, Daten zu akquirieren und KI-basierte Systeme zu trainieren.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Acquisti, A. / Adjerid, I. / Balebako, R. / Brandimarte, L. / Cranor, L. F. / Komanduri, S., Leon, P. G. / Sadeh, N. / Schaub, F. / Sleeper, M. / Wang, Y. / Wilson, S. (2018): Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online, in: ACM Computing Surveys, Jg. 50 / Nr. 3, 1–41, DOI: 10.1145/3054926.

Aitken, M. / de St. Jorre, J. / Pagliari, C. / Jepson, R. / Cunningham-Burley, S. (2016): Public Responses to the Sharing and Linkage of Health Data for Research Purposes: A Systematic Review and Thematic Synthesis of Qualitative Studies, in: BMC Medical Ethics, Jg. 17 / Nr. 1, 73, DOI: 10.1186/s12910-016-0153-x.

Brîncoveanu, C. / Carl, K. V. / Binz, S. / Weiher, M.-A. / Thomas, O. / Hinz, O. (2024): Artificial Intelligence-Based Assistance Systems for Environmental Sustainability in Smart Homes: A Systematic Literature Review on Requirements and Future Directions, in: Klein, M. / Krupka, D. / Winter, C. / Gergeleit, M. / Martin, L. (Hrsg.): INFORMATIK 2024. Lecture Notes in Informatics (LNI), Bonn: Gesellschaft für Informatik, 1165–1182.

Carl, K. V. / Hinz, O. (2024): What We Already Know About Corporate Digital Responsibility in IS Research: A Review and Conceptualization of Potential CDR Activities, in: Electronic Markets, Jg. 34 / Nr. 27, 1–30 DOI: 10.1007/s12525-024-00708-0.

COMET (2025): COMET – SmartLivingNEXT, URL: https://smartlivingnext.de/smartlivingnext-comet/ (aufgerufen am: 16/06/2025).

Dreesbach, T. / Dankert, L. / Ruhe, J. / Carl, K. V. / Hinz, O. / Thomas, O. (2025): Data Unlocked — Ein Modell zur anwendungsübergreifenden Nutzung persönlicher Daten im Smart-Living-Kontext, in: HMD Praxis der Wirtschaftsinformatik [im Erscheinen].

Feuerriegel, S. / Dolata, M. / Schwabe, G. (2020): Fair AI, in: Business & Information Systems Engineering, Jg. 62 / Nr. 4, 379–384, DOI: 10.1007/s12599-020-00650-3.

Hanneke, B. / Baum, L. / Schlereth, C. / Hinz, O. (2023): Consumer Preferences for Privacy Management Systems, in: De', R. / Paul, S. / Sarker, S. / Tuunainen, V. K. / Fernández W. D. / Nandhakumar, J. / Santhanam, R. (Hrsg.): Proceedings of the 44th International Conference on Information Systems, ICIS 2023, Rising Like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies, Atlanta: Association for Information Systems.

Mihale-Wilson, C. / Zibuschka, J. / Carl, K. V. / Hinz, O. (2021): Corporate Digital Responsibility – Extended Conceptualization and a Guide to Implementation, in: ECIS 2021 Research Papers, Nr. 80.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Oliveira, M. I. S. / Barros Lima, G. de F. / Farias Lóscio, B. (2019): Investigations into Data Ecosystems: A Systematic Mapping Study, in: Knowledge and Information Systems, Jg. 61 / Nr. 2, 589–630, DOI: 10.1007/s10115-018-1323-6.

Scheider, S. / Lauf, F. / Möller, F. / Otto, B. (2023): A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems, in: Business & Information Systems Engineering, Jg. 65 / Nr. 5, 577–595, DOI: 10.1007/s12599-023-00816-9.

SmartLivingNext (2025): COMET – Consumer Empowerment und Teilhabe in der Domäne Smart Living, URL: https://smartlivingnext.de/smartlivingnext-comet/ (aufgerufen am: 10/07/2025).

Tallon, P. P. (2013).: Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost, in: Computer, Jg. 46 / Nr. 6, 32–38, DOI: 10.1109/MC.2013.155.

van Houten, H. (2020): For Fair and Equal Healthcare, We Need Fair and Bias-Free AI, URL: https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/2020/20201116-for-fair-and-equal-healthcare-we-need-fair-and-bias-free-ai.html (aufgerufen am: 17/06/2025).

von Zahn, M. / Feuerriegel, S. / Kuehl, N. (2022): The Cost of Fairness in AI: Evidence from E-Commerce, in: Business & Information Systems Engineering, Jg. 64 / Nr. 3, 335–348, DOI: 10.1007/s12599-021-00716-w.

Witzki, A. / Carl, K. V. / Dankert, L. / Thomas, O. / Hinz, O. (2025a): What Are Your Intentions With My Data? A Tool to Enhance Consumer Data Sovereignty in Digital Ecosystems, in: Roßnagel, H. / Schunck, C. H. / Pöhn, D. (Hrsg.): OID 2025, Bonn: Gesellschaft für Informatik e.V., DOI: 10.18420/OID2025_01.

Witzki, A. / Carl, K. V. / Hinz, O. (2025b): Strategische Entscheidungsprozesse in digitalen Startups: Strukturierung von Faktoren für den Beitritt zu einem digitalen Ökosystem, in: HMD Praxis der Wirtschaftsinformatik; Jg. 62 / Nr. 1, 531–549, DOI: 10.1365/s40702-025-01155-x.

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

https://doi.org/10.5771/9783748965039-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY