

II. Ethics by Design: Potenziale, Umsetzung und Grenzen

1. Einleitung

Der Bedarf an ethischer Orientierung in der Forschung und das Interesse an *Ethics by Design* als neuem Schwerpunkt einer ange-wandten Ethik haben sich infolge einer intensiveren Wahrnehmung der gesellschaftlichen und sozialen Auswirkungen von Bio- und Digitaltechnologien, wie beispielsweise Gentechnologie, Machine learning oder Robotik, in den letzten Jahrzehnten zunehmend erhöht. Erste ethische Begleitforschungen lassen sich bereits bis in die 1960er- und 1970er-Jahre in der (US-amerikanischen) bioethischen Forschung zurückverfolgen (Gransche & Manzeschke, 2020, S. 11f.). Zentrale Auslöser waren kontroverse medizinische Experimente, insbesondere die Tuskegee-Syphilis-Studie (1932–1972), bei der afro-amerikanischen Männern bewusst eine medizinische Behandlung vorenthalten wurde (Jones, 1993; Brandt, 1978). Dieser Skandal führte zur Gründung der *National Commission for the Protection of Human Subjects* und zur Veröffentlichung des wegweisenden Belmont Reports (1979), der die Prinzipien Autonomie, Wohltun und Gerechtigkeit definierte (Emanuel et al., 2023; National Commission, 1979). Parallel sorgte die Einführung der Dialysetechnik in den 1960er-Jahren für ethische Kontroversen. Sogenannte »God Committees« mussten über begrenzte Behandlungskapazitäten entscheiden, was erstmals deutlich machte, dass technologische Innovationen ethische Begleitung erfordern (Alexander, 1962; Jonsen, 2023). Ähnliche Debatten lösten frühe Forschungen zur In-vitro-Fertilisation aus, die ethische Fragen zum Status menschlicher Embryonen aufwarfen (Kass, 1971).

Die Berücksichtigung von ethischen, sozialen und rechtlichen Aspekten in naturwissenschaftlichen Forschungsvorhaben wurde

dann Ende der 1980er-Jahre im Zuge der Genomforschung unter dem Begriff ELSI (als Akronym für »Ethical, Legal and Social Implications«) als Forschungsdesiderat in den internationalen Wissenschaftsdiskurs eingeführt (vgl. Bellon & Nähr-Wagener, 2020, S. 46). Gleichzeitig entstanden auch Einrichtungen der Technikfolgenabschätzung, die wohl Folge einer kritischeren Sichtweise auf den technischen Fortschritt im Allgemeinen waren und mithalfen, die Verantwortungsfrage des Menschen gegenüber Natur und Umwelt im Speziellen zu stellen und die gesellschaftlichen und ethischen Dimensionen der Technik stärker zu gewichten (vgl. Grunwald, 2022; Böschens et al., 2021).

Von einem »ethical turn« in den Technik- und Lebenswissenschaften ist allerdings erst seit der Hochkonjunktur der Digitaltechnik und künstlichen Intelligenz (KI) zu sprechen. So ist im Zuge der allseitigen Bewertungs- und Entscheidungsgewalt von Maschinen über Menschen, wie z. B. bei der algorithmenbasierten Berechnung von Krankheitsverläufen oder der beruflichen bzw. schulischen Leistung (vgl. Zweig, 2023, S. 171–191), und damit der maschinellen Vorhersagbarkeit und Berechenbarkeit von Menschen, der Bedarf einer ethisch informierten Begleitung der Projekte in den Technik- und Lebenswissenschaften stark gestiegen.

2. Potenziale von Ethics by Design (EbD)

2.1 Entwicklung und Ziele eines werteorientierten Design-Ansatzes

Mehr oder weniger synchron zu dem *ethical turn* zeichnete sich auch ein »design turn« in der angewandten Ethik ab. Einer der prominenten Vertreter dieses Ansatzes ist Jeroen van den Hoven, der pointiert feststellt: »*This design turn in applied ethics can be seen as the third and the most recent phase in the development of contemporary ethics.*« (van den Hoven, 2017, S. 5) Für ihn ist die Notwendigkeit einer neuen Design-Perspektive dadurch begründet, dass Handlungsfähigkeit und Zuschreibung von Verantwortlichkeit dem Handelnden gegenüber in komplexen Systemen und Prozessen kaum mehr möglich und damit auch alte moralische Konzepte überholt seien. Überwiegend finde zudem Kommunikation und Handeln

nicht mehr direkt, sondern indirekt via Technologie statt: »Much of what we do in the twenty-first century we do not *directly* to each other, but almost always *by way of technology* and *via a technological milieu* that was designed, and that came about over the course of centuries as a result of myriad contributions by others separated from us by time and place.« (van den Hoven, 2017, S. 19) Im Mittelpunkt eines Ethics by Design-Ansatzes (EbD-A) müsste deshalb das Design und die Designgeschichte eines Systems und technologischen Umfelds sowie die Rolle der Design-Agenten und der von ihnen entwickelten Wahloptionen stehen.

Während Brey und Dainow (2024) die Anfänge von Ethics by Design in die 2010er-Jahre datieren, lassen sich die Wurzeln weit früher, nämlich bis in die späten 1990er-Jahre zurückverfolgen. So lässt sich m. E. Ethics by Design im Wesentlichen als Teil des Forschungsfelds von *Values in Design* (vgl. Simon, 2016) verstehen. Im Forschungsfeld von Ethics by Design haben sich verschiedene Ansätze gebildet, die unter den Begriffen »Value Sensitive Design« (vgl. Friedman et al., 2002), »Values at Play« (vgl. Flanagan & Nissenbaum, 2014), »A Value-Based System Design Approach« (vgl. Spiekermann, 2015) bzw. »Value-Based Engineering« (vgl. Spiekermann, 2021, 2023) oder »Privacy by Design« (vgl. Cavoukian, 2011) bekannt sind

Insbesondere der Ansatz von Value Sensitive Design (VSD) ist methodisch wie auch perspektivisch als Vorreiter des Ethics by Design-Ansatzes zu deuten. Bekannt wurde der VSD mit der Veröffentlichung »Human Values and the Design of Computer Technology« von Batya Friedman (1997). Demnach ist VSD der Versuch, Werte in den Entwurfs- und Entwicklungsprozess von Beginn an unter Berücksichtigung aller indirekt und direkt betroffenen Stakeholder mit einzubeziehen.

Der VSD-Ansatz will von Beginn an und während des Entwicklungsprozesses die Gestaltung der jeweiligen Technologie bzw. deren Anwendung beeinflussen, zudem will er Design- und Technikmethoden für eine kritische Analyse der Werte anwenden und dabei ein breites Spektrum menschlicher Werte, die im menschlichen Kontext von Bedeutung sind, berücksichtigen.

Bezeichnend für die Perspektive des VSD sind dementsprechend die folgenden vier Kriterien:

- »(1) Proactive orientation toward influencing design. (...)
- (2) Carrying critical analyses of human values into design and engineering. (...)
- (3) Enlarging the scope of human values. (...)
- (4) Broadening and deepening methodological approaches.« (Friedman & Hendry, 2019, S. 4)

Hinsichtlich seiner Methoden ist der VSD-Ansatz für verschiedene Perspektiven, wie z. B. die aus der Anthropologie, Soziologie, Software Engineering, Psychologie und Philosophie offen und möchte diese noch erweitern. Im Prinzip steht VSD damit für einen interdisziplinären Forschungsansatz, wie es auch bei EbD der Fall ist. Im Wesentlichen lassen sich damit für EbD folgende Potenziale erkennen: Mit diesem Ansatz

- können die bewussten und unbewussten Wertehaltungen der an der Gestaltung und Entwicklung von Bio- und Digitaltechnologien Beteiligten und Betroffenen transparent gemacht,
- Wertekonflikte erkennbar,
- multidisziplinäre Perspektiven bezüglich des Designs reflektiert und
- Technologien werteorientiert gestaltet werden.

2.2 Ethics by Design und ELSI

Auch wenn die Schnittmenge von VSD und EbD aufgrund der Wertorientierung und empirischen Ausrichtung sehr groß ist, stellt der EbD-Ansatz, wie er im *Institut für Digitale Ethik der Hochschule der Medien Stuttgart* entwickelt wurde, einen modifizierten Ansatz des VSD dar, da er als Teil der Digitalen Ethik a) methodisch durch die Narratologie theoretisch und empirisch erweitert wurde, b) nicht allein auf den Wertefokus begrenzt ist, sondern auch deontologische, konsequentialistische und tugendethische Ansätze miteinbezieht, c) für die Reflexion und Umsetzung ethischer Ansprüche im Designprozess digitalethische Instrumente und/oder narrative Analysetools entwickelt und d) einen prinzipiell integrativen respektive interdisziplinären Ansatz verfolgt.

Dementsprechend wurden in den Forschungsprojekten, die das Institut für Digitale Ethik der Hochschule der Medien in den letzten Jahren überwiegend im Verbund mit technikwissenschaftlichen Partner*innen, Unternehmen und öffentlichen bzw. staatlichen Organisationen durchgeführt hat, ein interdisziplinäres Forschungsdesign, in das Ethics by Design mit eingebunden war, erprobt und auf unterschiedliche Forschungsfelder angewandt, wie z. B. auf Forschungen im Bereich des autonomen Fahrens, der Sicherheitstechnologien, der digitalen Erlebniswelten, der Generativen KI in den Medien, der Robotik im öffentlichen Raum, der Barrierefreiheit in der (digitalen) Lehre, der digitalen Bürgerbeteiligung usw. Zudem wurden zwei Tools entwickelt, die ELS-Aspekte (ethische, rechtliche und soziale) schon bei der Antragstellung von Forschungsprojekten in den Bereichen der Mensch-Technik-Interaktionen berücksichtigen. Hierbei handelt es sich um die ELSI-Screening- und Awareness-tools *ELSI-SAT* und *ELSI-SAT Health & Care*, die in Kapitel 5 vorgestellt werden.

Quer zu dem VSD-A und damit auch dem EbD-A liegt die ELSI-Forschung, die sich ebenfalls damit befasst, wie ethische Perspektiven in technologische Systeme und Geräte implementiert werden können. Die in den 1960er- und 70er-Jahren im Kontext der Lebenswissenschaften entstandene interdisziplinäre ELSI-/ELSA-Forschung stellt eine spezifische Form der Wissensintegration dar, eben die der ethischen, rechtlichen und sozialen Disziplinen in die Technikforschung. Allerdings scheint der Begriff der ELSI-Forschung etwas aus der Mode gekommen zu sein (vgl. Mikami, 2021). Alternativ wird von »verantwortungsvoller Forschung« (Responsible Research and Innovation (RRI), vgl. Stilgoe & Guston, 2017) oder »Integrierter Forschung« (vgl. Gransche & Manzeschke, 2020) gesprochen. Teilweise wird die Integrierte Forschung dabei mit der ELSI-Forschung mehr oder weniger gleichgesetzt (vgl. Schikowitz & Maasen, 2021; Kemmer, 2020). Teilweise wird sie aber auch als eine Weiterführung der ELSI-Forschung und umfassendere Wissensintegration in die Mensch-Technik-Forschung verstanden. Bei diesem Ansatz wird die Fokussierung auf ethische, rechtliche und soziale Aspekte aufgebrochen und für weitere Perspektiven geöffnet, wie etwa wirtschaftliche, kulturelle oder ästhetische: »Die Durchdringung unsere Lebenswelt mit vernetzter Technologie birgt sicherlich Implikationen auch für all jene Bereiche, die mit ELSI nicht aufgerufen sind.«

(Gransche & Manzeschke, 2020, S. 16) Allerdings verliert damit der Ethik- und Wertefokus, wie er bei EbD und VSD repräsentiert wird, seine zentrale Bedeutung.

3. Ethics by Design als Teil der Digitalen Ethik und der Wertebegriff

3.1 Verständnis von »Digitaler Ethik«

Ethics by Design ist ein wichtiger Baustein bei der Umsetzung der Digitalen Ethik. Letztere lässt sich als Synthese und Fortschreibung einer Medien- und Informationsethik verstehen, die den gesamten Bereich des Digitalen aus ethischer Sicht betrachtet (Grimm et al., 2019). Dazu gehören digitale Medien und digitale Artefakte wie bspw. Roboter, autonome Fahrzeuge, smarte Häuser und Städte sowie digitale Angebote, Services, künstliche Systeme und Infrastrukturen. Digitale Ethik lässt sich zudem als Metabegriff der verschiedenen, auf die Digitalisierung bezogenen Teilethiken – Computerethic, Algorithmenethik, Maschinenethic, Roboterethik, KI-Ethic, Hackerethik, Sicherheitsethik, Datenethik – begreifen. Der Begriff »Digitale Ethik« geht ursprünglich auf Charles Ess zurück, der 2009 erstmals von »Digital Media Ethics« sprach und damit insbesondere den Bereich der digitalen Medien meinte. Rafael Capurro (2009, 2017) führte den Begriff »Digitale Ethik« gleichbedeutend mit Informationsethik in den deutschsprachigen Raum ein, wobei er selbst mit seiner Publikation »Informationsethos und Informationsethik« (1988) bereits als Gründervater für die Informationsethik im deutschsprachigen Raum gilt (vgl. Lenzen, 2011, S. 211). Mit der Einrichtung des Instituts für Digitale Ethik an der Hochschule der Medien Stuttgart (2013) wurde programmatisch eine Synthese der Tech-Ethiken (IKT-orientierte Ethiken) mit der Medien- und Kommunikationsethik vorgenommen. Diese Koinzidenz der ethischen Anwendungsfelder ist nicht willkürlich. Vielmehr beruht sie auf dem digitalen Strukturwandel in der Gesellschaft und in den Medien, der zu einer medialen bzw. digitalen Durchdringung aller sozialen Systeme führte und weiter führt.

3.2 Werte für Ethics by Design

Werte sind zentrale Bezugspunkte für EbD. Was ist aber eigentlich unter Werten zu verstehen? Der Wertbegriff als solcher wurde erst mit Beginn des 19. Jahrhunderts in der neuzeitlichen Philosophie Gegenstand der Reflexion. Auch die Frage nach dem Guten bzw. guten Leben und was als Wert zu gelten hat, waren immer schon im philosophischen bzw. ethischen Kontext von Bedeutung. Als philosophisches Phänomen lassen sich für die Bedeutung von »Wert« zwei Positionen unterscheiden: So werden Werte zum einen als »Gut« und zum anderen »als Maßstab, als Kriterium oder Standards« (Krobath, 2009, S. 32) beschrieben. Entsprechend dieser Sichtweise hat Lautmann (1971) anhand einer sprachanalytischen Begriffsanalyse der Fachliteratur, in der er 180 verschiedene Wertdefinitionen fand, folgenden Wertbegriff herausgearbeitet:

»Wert ist

- ein Maßstab der guten Gegenstände,
- Kriterium zur Auswahl der Objekte, die wir anstreben sollen,
- normativer Standard zur Beurteilung von Objekten,
- Kriterium für normativ gebilligte Gegenstände.« (S. 105)

Es lässt sich daraus weiter ableiten, dass Werte als Vorstellungen, Ideen oder Ideale zu verstehen sind. Während Werte als abstrakte Entitäten und losgelöst von Personen verstanden werden können, lassen sich mit Aristoteles die ethischen Tugenden und teilweise auch die Tugenden des Denkens (dianoethischen) als Charaktermerkmale und Dispositionen beschreiben (vgl. Krobath, 2009, S. 22). Zu den Tugenden des Denkens gehören Weisheit (sophia), Kunst/Können (téchne) und Klugheit (phronesis). Im Rahmen eines EbD-Prozesses können diese Tugenden Orientierungswerte für verantwortliche Wissenschaftler*innen des jeweiligen Projekts sein. Ebenso können die ethischen Tugenden, die für Aristoteles (2023) besonders relevant sind, für EbD-Teams Maßstäbe sein: Tapferkeit, Mäßigkeit, Freigiebigkeit, Hochherzigkeit, Hochsinn, gesunder Ehrgeiz, Sanftmut, Wahrhaftigkeit, Humor, Freundlichkeit und Gerechtigkeit. Sie sind als intrinsische Werte zu verstehen, also als Werte an sich, die zudem erst in ihrer Ausübung zur Geltung kommen (vgl. Halbig, 2013, S. 63).

Einen modernen Ansatz einer explizit technikbezogenen Tugendethik (Technomoral Virtue Ethics) hat Shannon Vallor (2016) vor-

gelegt. Mit Rückbezug auf eine aristotelische, buddhistische und konfuzianische Ethik schlägt sie eine erweiterbare Taxonomie technomoralischer Tugenden vor, die für ein gutes Leben in der digitalen Welt erforderlich seien: Ehrlichkeit, Selbstkontrolle, Demut, Gerechtigkeit, Mut, Empathie, Fürsorge, Zivilität (wie Respekt, Toleranz), Flexibilität, Einsichtsvermögen, Großmut (*magnanimity*) und technomoralische Weisheit.

Welche Werte für EbD letztendlich von Bedeutung sind, kann nicht einheitlich definiert werden. Der VSD-Ansatz geht von dreizehn ethischen Werten aus, die häufig in technischen Designs eine Rolle spielen und in der Fachliteratur vielfach genannt werden: Menschliches Wohlergehen, Eigentum und Besitz, Privatsphäre, Freiheit von Voreingenommenheit, universelle Nutzbarkeit, Vertrauen, Autonomie, informierte Einwilligung, Übernahme von Verantwortung, Höflichkeit, Identität, Gelassenheit und ökologische Nachhaltigkeit (vgl. Friedman et al., 2008, S. 90–91; Friedman & Hendry, 2019, S. 28). Demgegenüber begrenzen Philipp Brey und Brandt Dainow (2024) die Werteliste von Ethics by Design für KI auf sechs Werte: Freiheit, Privatsphäre, Fairness, Transparenz, Verantwortlichkeit und Wohlbefinden (des Einzelnen, der Gesellschaft und der Umwelt), wobei sie zudem die Freiheit, Autonomie und Würde des Menschen unter dem Aspekt der Handlungsfähigkeit zusätzlich als relevant erachten (S. 1267f.). Begründet wird diese Auswahl mit Verweis auf einschlägige Dokumente und ethische Leitlinien zu KI von bekannten Institutionen und mit einer Meta-Analyse von ethischen Leitlinien (vgl. Hagendorff, 2020).

Ergänzend wird hier als Grundlage für die Werteliste eines EbD-Ansatzes auf die Wertetopografie einer Digitalen Ethik (vgl. Grimm, 2021) verwiesen. Sie kann als Ausgangspunkt für die konzeptionelle Phase in einem EbD-Prozess dienen. Teilweise überschneiden sich die Werte mit den oben genannten, teilweise berücksichtigen sie noch weitere Werte. Als Grundlage für die Werte-Topografie dienten ethische Leitlinien im Bereich der Mensch-Maschine-Interaktion (vgl. Grimm & Mönig, 2000) und zur künstlichen Intelligenz (AlgorithmWatch, 2024), das Gutachten der Datenethikkommission der Bundesregierung (2019) sowie eigene Arbeiten zur Verankerung von Ethik in der integrierten Forschung. Die Wertetopografie der Digitalen Ethik umfasst folgende zehn Werte:

Tabelle 1: Werteliste der Digitalen Ethik

Werte	Beispiele
<i>Menschenwürde</i>	<ul style="list-style-type: none"> - Achten und schützen der Würde in der digitalen Welt - Körperliche und geistige Unversehrtheit des Menschen schützen
<i>Autonomie</i>	<ul style="list-style-type: none"> - Nicht ausgeliefert sein an und abhängig sein von digitalen Systemen bzw. KI - Nicht von KI moralisch bewertet oder manipuliert werden - Wechseln können zwischen digitalen Produkten und Diensten
<i>Privatheit</i>	<ul style="list-style-type: none"> - Seine Daten schützen können und nicht überwacht werden - Personenbezogene Daten nur für festgelegte Zwecke und auf gesetzlicher Grundlage erheben und verarbeiten - Privatsphäre von Kindern, Heranwachsenden und schutzbedürftigen Menschen besonders achten - Gewährleisten von Löschung und Berichtigung von persönlichen Daten, Widerspruch, Information und Auskunft - Eine nicht-personalisierte Nutzung digitaler Angebote ermöglichen
<i>Freiheit</i>	<ul style="list-style-type: none"> - Freie Information und Kommunikation sowie freie Meinungsäußerung in digitalen Medien gewährleisten - Schranken der Meinungsfreiheit in den Vorschriften der allgemeinen Gesetze akzeptieren - Freier und gleicher Zugang zu Kommunikations- und Informationsdiensten ermöglichen - Medien- und Pressefreiheit achten
<i>Transparenz und Erklärbarkeit</i>	<ul style="list-style-type: none"> - Offenlegen der Kriterien automatisierter Entscheidungen - Transparenz hinsichtlich des Einsatzes von KI - Transparenz der Datenverarbeitung entsprechend dem Stand der Technik - Funktionsweise von KI verstehen, erklären und kontrollieren können - Betroffene entsprechend ihrem Kompetenzniveau genügend Informationen vermitteln, um ihre Rechte angemessen wahrnehmen und Entscheidungen infrage stellen zu können - Methoden zur Nachvollziehbarkeit der Ergebnisse von KI fördern - Der Vermenschlichung von KI entgegenwirken
<i>Gerechtigkeit</i>	<ul style="list-style-type: none"> - Faire Algorithmen durch Bewusstsein über mögliche Voreingenommenheit, Vorurteile und kulturelle Prägung ermöglichen - Datenbasis auf Diskriminierungspotenziale prüfen und systematische Fehler und Verzerrung – auch auf technischem Wege – vermeiden - Prinzipien der Vielfalt und Offenheit in Bezug auf digitale Angebote, Infrastrukturen und Zugänge zu Netzen umsetzen - Nicht durch digitale Systeme Personen vom Zugang zu Gütern und Dienstleistungen und der Teilhabe am gesellschaftlichen Leben ausschließen

Werte	Beispiele
	<ul style="list-style-type: none"> - Automatisierte Entscheidungen, die sich erheblich auf die Lebensführung von Personen auswirken, durch Menschen überprüfen und letztendlich entscheiden lassen
<i>Gesundheit</i>	<ul style="list-style-type: none"> - Schutz vor Sucht und Abhängigkeit (Spiele, Abhängigkeitsgefühl vom Handy) - Verhinderung von physischen Schäden (z. B. Essstörungen, Selbstverletzung, Übergewicht), - Verhinderung von psychischer Schädigung (z. B. Ängste, Krankheit und Isolation durch Cybermobbing, Einsamkeit durch fehlende soziale Einbettung, Minderwertigkeitsgefühle durch ständigen sozialen Vergleich, Cyberchondrie, Gesundheitsängste) - Schutz vor sexuellem Missbrauch und Übergriffen (Grooming, Sextortion etc.)
<i>Wahrheit und Wahrhaftigkeit</i>	<ul style="list-style-type: none"> - Verhindern von Desinformationen und Verschwörungserzählungen durch wirksame Maßnahmen (Moderation, Fact Checking etc.), die von den Anbietern digitaler Dienste und Plattformen gewährleistet werden sollen - Entgegenwirken der Gefahr, durch <i>Social Bots</i> bestimmte Meinungen massenhaft zu verbreiten und die Stimmung in sozialen Medien zu lenken
<i>Sicherheit</i>	<ul style="list-style-type: none"> - Unversehrtheit, Vertraulichkeit und Integrität von KI und Infrastrukturen sicherstellen - In der Entwicklungsphase Systemkompatibilität und die Handhabung der Technik unter realen Bedingungen beachten, um keine unnötigen Fehlerquellen entstehen zu lassen - Kontinuierlich überprüfen, wie sich die Zwecke verändern, für welche die Technik entwickelt wurde - Abschätzen, wie die Technik zweckentfremdet angewandt werden kann und ob dies bestimmte vorsorgliche Gegenmaßnahmen notwendig macht
<i>Nachhaltigkeit</i>	<ul style="list-style-type: none"> - KI und digitale Produkte, Dienste und Infrastrukturen nachhaltig designen - Emissionsreduktion, Ressourceneinsparung und Verzicht auf Anheizen des Konsums gewährleisten - Digitalisierung von Lebensräumen (z. B. smarte Städte, Häuser, autonome Fahrzeuge) nachhaltig gestalten - Lebenszyklus digitaler Geräte und Software verlängern und vorausschauend planen und deren Produktion und Entsorgung mitdenken

Werte sind grundsätzlich abstrakt und konkretisieren sich erst im Erzählen, in der Erfahrung und Umsetzung. Ein EbD-Prozess kann somit nicht nur jeweils sehr unterschiedliche Konkretisierungen der Werte, sondern auch weitere Werte durch die Kontextualisierung

und Perspektive der Betroffenen umfassen. Wie EbD umgesetzt werden kann und welche Chancen damit verbunden sind, wird im folgenden Kapitel erläutert.

4. Umsetzungsdimension I: Operationalisierung von Ethics by Design

4.1 Praktische Methode des Value Sensitive Design

Wie man Ethics by Design operationalisieren kann, zeigt die Methode des Value Sensitive Designs. Sie zeichnet sich dadurch aus, dass sie konzeptionelle, empirische und technische Forschung bzw. Untersuchungen zu vereinen sucht. Dieser triadische Ansatz ist das Bestreben, unterschiedliche Techniken proaktiv einzusetzen und im Konkreten wie im Abstrakten zu forschen, um verschiedene Ergebnisse optimal zusammenfügen zu können (vgl. Friedman et al., 2002). VSD basiert auf einer iterativen Methodik, die konzeptionelle, empirische und technische Untersuchungsphasen integriert (vgl. Friedman et al., 2008, S. 72–73; Friedman & Hendry, 2019, S. 31–35): In der konzeptionellen Phase werden die Stakeholder eruiert, die von einem technologischen Entwurf indirekt oder direkt betroffen sind. Ebenso werden die Werte des Designs konzeptionell begründet sowie etwaige ethische Wertekonflikte (z. B. Autonomie vs. Sicherheit) und Wertekonkurrenzen moralischer und nicht-moralischer Werte (z. B. ästhetische Vorlieben) identifiziert. In der empirischen Phase sollen je nach Erkenntnisziel quantitative und qualitative Methoden der Sozialforschung zum Einsatz kommen, um die Wahrnehmung der vom Design betroffenen Stakeholder und die Wirkung des Designs zu ermitteln. Ebenso sind deren Umgang mit Werten bzw. Wertekonflikten sowie deren etwaige Abwägungen und Priorisierungen von Werten zu untersuchen. Hierbei sind nicht nur Einzelpersonen zu berücksichtigen, sondern auch Organisationen, die vom Designprozess betroffen sind. In der technischen Untersuchungsphase sollen die in der empirischen und konzeptionellen Erhebung ermittelten Erkenntnisse in das technische Design implementiert und die Werte damit operationalisiert werden. Darüber hinaus beschreibt der VSD-Ansatz (vgl. Friedman et al., 2017, S. 12–15; Friedman & Hendry, 2019, S. 60–64) ein methodisches Instrumentarium von

vierzehn bzw. siebzehn VSD-Methoden, mit deren Hilfe Werte und Werterepräsentationen analysiert werden können. Dabei handelt es sich aber eher um methodische Werkzeuge (*value scenario*, *value sketch*, *value-oriented semistructured interview etc.*) als um theoriegestützte Methoden, wie z. B. narrative Theorieansätze.

In der Forschungsliteratur des Ethics by Design werden die drei Phasen des VSD häufig in folgende Einheiten weiter unterschieden: konzeptionelle Phase (Analyse, Synthese), empirische Phase (Simulation, Evaluation, Selektion) und technische Phase (Umsetzung und Testen des Prototyps) (vgl. van de Poel, 2017, S. 67). Die folgende Grafik (vgl. Abb. 1) nach Eeckels und Roozenburg (1991) veranschaulicht diese sukzessiven Phasen:

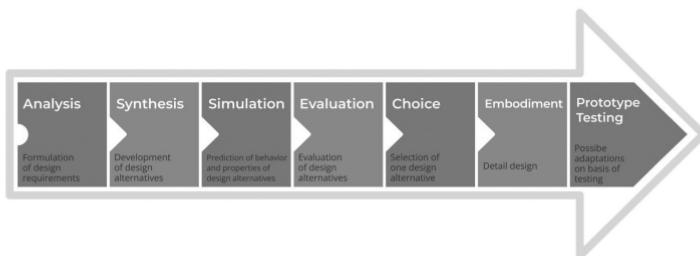


Abb. 1: Phasen des Designprozesses

Es gibt zwar keine einfache Formel für EbD, die bei jeder technischen Entwicklung gleich funktionieren würde und die diese durch eine Anwendung in eine ethisch unbedenkliche oder sogar ethisch begrüßenswerte Innovation verwandeln könnte. Allerdings lässt sich ein Prozess, der ja auch immer ein Prozess der Reflexion ist, modellhaft entwickeln. Dabei sollten allerdings drei Aspekte, die in den bisherigen Ethics by Design-Ansätzen vernachlässigt werden, methodisch und theoretisch fundiert werden: erstens wie Werte im Designprozess analysiert und empirisch ermittelt werden, zweitens wie der Dialog zwischen Ethiker*innen und Entwickler*innen systematisiert werden kann und drittens wie die Synthese (Entwicklung alternativer Designs) und die Simulation von Anwendungsfällen und Szenarien methodisch erfolgen soll. Diese Defizite können mit einem »Narrative Ethics by Design«-Ansatz ausgeglichen werden, wie er im Rahmen einer Expertise für ein Qualifizierungskonzept

für Start-ups als SEC-Methode von Grimm, Keber und Müller (2021) entwickelt wurde. Was unter Narrative Ethics by Design zu verstehen ist, soll im Folgenden erläutert werden.

4.2 Narrative Ethics by Design

Werte allein auf sich gestellt bleiben abstrakt, erst in einer Narration werden sie erfahrbar, weil ihre Bedeutung konkretisiert wird und sie kontextuell eingebunden sind. Geschichten, narrative Szenarien und Use Cases sollen deshalb in einem Narrative Ethics by Design-Ansatz genutzt werden, um in der konzeptionellen Phase die für eine konkrete Technikanwendung relevanten Werte zu ermitteln (vgl. Grimm & Kuhnert, 2018). Damit soll es den Beteiligten ermöglicht werden, abstrakte Werte wie Transparenz, Vertrauen, Privatheit usw. aus ihren eigenen Erfahrungen konkret abzuleiten und sich in praktischen Anwendungsszenarien vorstellen zu können, was mit diesen Werten gemeint ist. Werte, die aus Sicht einer Digitalen Ethik und auf der Grundlage von eigenen Forschungsarbeiten bei Informations- und Kommunikationstechnologien insbesondere zur Disposition stehen, können heuristisch in der ersten Phase der Wertanalyse fokussiert werden (vgl. hierzu die Werteliste der Digitalen Ethik in Tabelle 1).

Unter Narrative Ethics by Design (NED) ist ein praxisorientiertes Modell zu verstehen, mithilfe dessen der Designprozess auf der Grundlage einer Narrativen Ethik und mit Hilfe der Methoden der empirischen Narrationsforschung (vgl. Müller & Grimm, 2016) entwickelt und analysiert wird.

Zum Verständnis von NED ist eine kurze Erläuterung der Narrativen Ethik hilfreich. Die ersten Ansätze zu einer narrativen Ethik bildeten sich in den USA mit den Theorien von Martha Nussbaum (1985, 2001, 1992), Richard Rorty (1992, 2001) und Alasdair MacIntyre (1995 [1981]) heraus, während in Deutschland Dietmar Mieth (1976, 1998, 2000, 2002, 2007) den Begriff entscheidend prägte. Die Fortschreibung einer Narrativen Ethik im digitalen Zeitalter beschreibt Grimm (2023, 2024) in Bezug auf die drei Ausprägungen, in denen eine Narrative Ethik hilfreich ist: 1. bei der *Analyse von Narrationen* (= *Kommunikat-Analyse*), 2. bei der *Vermittlung von Kompetenzen* und 3. bei einem *methodischen Zugang zu Werten und*

Erlebniswelten der Menschen. Die erste Ausprägung analysiert Medieninhalte sowie politische, wirtschaftliche oder soziale Kommunikate (= Inhalte von Kommunikation) mittels des narratologischen Instrumentariums und reflektiert die daraus abgeleiteten Aussagen, Werte und Weltmodelle. Die zweite Ausprägung lässt sich im Sinn einer didaktisch ausgerichteten Ethik beschreiben: Hier dient die narrative Ethik zum Zwecke der Bildung von (Digital-)Kompetenzen in moralischen Lebenszusammenhängen. So kann die Interpretation literarischer Texte, Filme oder anderer Kommunikate in pädagogischen Zusammenhängen dazu dienen, ethische Konflikte, Werte und Haltungen zu reflektieren und moralische Urteilsfähigkeit zu fördern. In diesem Zusammenhang hat das Institut für Digitale Ethik (IDE) eine Vielzahl an medienethischen Materialien erstellt. Die dritte Ausprägung kann als Ethik des *Homo narrans* beschrieben werden. Sie fragt nach den Erfahrungen und subjektiven Erlebniswelten der Menschen, indem sie narrative Interviews führt und diese wiederum mittels der Narratologie u. a. hinsichtlich der Werte und Wertekonflikte auswertet. Darauf aufbauend kann dann eine ethische Reflexion erfolgen. Letztere Ausprägung wird auch im EbD-Prozess relevant, nämlich dann, wenn die Werte der Betroffenen empirisch ermittelt und reflektiert werden.

Ausgehend von einem Regelkreismodell wird hier idealtypisch von folgenden Prozessphasen des NED ausgegangen:

Phase 1: Ethical & Legal Screening: Analyse ethischer Werte und rechtlicher Vorgaben

Hier gilt es zu ermitteln, welche Werte für die Entwicklung und Anwendung einer spezifischen Technik (z. B. einer Gesundheits-App) von Bedeutung und wünschenswert sind. Dabei können mittels narrativer Anwendungsfälle mögliche Szenarien mit den Betroffenen entwickelt werden. Ebenso ist zu prüfen, welche rechtlichen Vorgaben und Rahmenbedingungen bestehen und wie diese in die Praxis umgesetzt werden können. Darüber hinaus ist zu analysieren, welche (ggf. verdeckten) subjektiven Standpunkte und tatsächlichen Wertvorstellungen der Beteiligten existieren, welche potenziellen Wertekonflikte (z. B. Sicherheit versus Selbstbestimmung) mit der spezifischen Technik verbunden sind und welche Anforderungen an Usability und Funktionalisierung bestehen.

Phase 2: Bewusstseinsbildung und Haltung

Im zweiten Schritt sollte ein Transfer der Ergebnisse aus der Werteanalyse stattfinden mit dem Ziel, Bewusstsein für Wertefragen zu schaffen und die Entwicklung einer Haltung der Beteiligten zu motivieren. Narrationen sind hierzu fundamental, weil sie Zugänge zu ethischen Fragen und Konflikten darstellen, die Situationen und Perspektiven der Anderen »erklären« und für ethische Aspekte des Lebens sensibilisieren, (emotionale) Erfahrungen mit rationalem Handeln verbinden und zur Bildung einer Haltung beitragen, einen Perspektivwechsel herbeiführen und eine ethische Haltung emotional verankern sowie Handlungsspielräume und Zukunftsszenarien ausloten helfen (vgl. Grimm, 2024).

Phase 3: Entwicklung des Designs

Im dritten Schritt sollen Use Cases und Anwendungsszenarien u. a. mittels narrativer Methoden wie dem Zukunftsrad (*Futures Wheel*) (vgl. Glenn, 2009) und dem narrativen Aktantenmodell (vgl. Müller & Grimm, 2016, S. 86–91) reflektiert werden, um die im Ethical & Legal Screening in Phase 1 ermittelten Befunde und die in Phase 2 erworbenen Perspektiven in einer Simulation des Designs umzusetzen und zu definieren, welche Kriterien für die Bewertung eines werteorientierten Designs gelten sollen.

Phase 4: Evaluation und Steuerung

In der vierten Phase sollen ein Ethical Assessment, also eine Bewertung der Risiken und deren Adressierung sowie eine Reflexion, z. B. mittels der beiden ELSI-SAT-Tools (vgl. Kap. 5), erfolgen und eventuell alternative Designs entwickelt werden.

Phase 5: Innovationsoptimierung

In der fünften Phase sind die Befunde der Evaluation und Risiko-adressierung in ein detailliertes Design zu implementieren und die Prototypen zu testen.

Der Gesamtprozess ist als Rückkoppelungsprozess zu verstehen, indem die Befunde und Erkenntnisse der jeweiligen Phasen an vorherige rückgebunden werden. Das NED-Modell wendet für die Befragungen die narrative Interviewtechnik an, für die Werte- und Funktionsanalysen die narrative-semiotische Methode und für die

Entwicklung von Szenarien und Use Cases die narrative Szenario-methode. Damit soll sichergestellt werden, dass nicht etwaige erwünschte Meinungen, sondern de facto vorliegende Einstellungen und Verhaltensweisen erfasst und tieferliegende Bedeutungen und Funktionszusammenhänge in der Analyse ermittelt werden können (vgl. Müller & Grimm, 2016).

Die folgende Grafik (vgl. Abb. 2) veranschaulicht das oben beschriebene NED-Modell.



Abb. 2: Rückkopplungsprozess im Narrative Ethics by Design

Es lassen sich nun zwei unterschiedliche Anwendungsszenarien von EbD bzw. NED aufzeigen: Im ersten Fall werden zwei Ethics by Design-Instrumente erläutert, die antragstellenden Forschenden bei der Implementierung von Ethics by Design helfen sollen. Die Zielgruppe

sind alle Mitwirkenden in der Forschung, die sich mit dem Design von Forschungsprojekten befassen, aber auch Interessierte an der ELSI-Forschung. Im zweiten Fall wird die Anwendung von NED anhand des unternehmensorientierten Projekts »Ethics by Design für Start-ups« erläutert. Hierbei wird sowohl die Entwicklung der SEC-Methode (*Start-up with Ethics Compass*) sowie deren Umsetzung in Form eines Workbooks, das den NED-Prozess mit Start-Ups Schritt für Schritt durchspielt, demonstriert. Die Zielgruppe sind Unternehmen, resp. Start-ups, die sich praxisbezogen Orientierung und Anleitung für die Implementierung von Werten in ihr Geschäftsmodell und ihre Unternehmenskultur wünschen.

5. Umsetzungsdimension II: Ethics by Design-Tools

5.1 Das Risikobewertungstool ELSI-SAT

Welche ethischen, rechtlichen und sozialen Aspekte sind in Forschungsprojekten zur Mensch-Technik-Interaktion (MTI) zu berücksichtigen? Ausgehend von dieser Frage wurde im Forschungsprojekt »Automatisiertes ELSI-Screening & Assessment-Tool für MTI-Forschungsvorhaben (ELSI-SAT)« vom Institut für Digitale Ethik der Hochschule der Medien Stuttgart eine Software entwickelt, die Forschungsteams für ethische, rechtliche und soziale Implikationen bei der Antragstellung und Gestaltung ihres Forschungsdesigns sensibilisiert. Zugleich dient es als Orientierungshilfe für bislang unberücksichtigte Forschungssaspekte. Das ELSI-Tool kann zudem von Forschenden für die Gestaltung eines EbD-Prozesses als hilfreiches Werkzeug eingesetzt werden. Das mit Mitteln vom Bundesministerium für Bildung und Forschung (BMBF) geförderte interdisziplinäre Projekt (Förderkennzeichen: 16SV8204) wurde vom Institut für Digitale Ethik an der Hochschule der Medien Stuttgart (HdM) über den Zeitraum April 2019 bis Oktober 2020 in einem interdisziplinären Forschungsteam (Petra Grimm, Tobias Keber und Simon Wiest) umgesetzt. Das Tool ist im Internet zusammen mit einem Video zur kurzen Einführung in die Handhabung abrufbar unter <https://app.el-si-sat.de/#/>.

Ziel des Projekts war es nicht nur, eine ethische Optimierung von Technologien und deren Entwicklungsstrategien zu ermöglichen,

sondern auch ein EbD-Konzept zu entwickeln, das im Rahmen von *Responsible Research and Innovation* (vgl. Yaghmaei & van de Poel, 2021) oder konstruktiver Technikfolgenabschätzung (vgl. Grunwald, 2022) nutzbar wird.

ELSI-SAT ist eine Software-Anwendung, bei der Antragstellende in Form eines geführten Dialogs Fragen zum geplanten Vorhaben beantworten und vorgesehene Maßnahmen mit ELSI-Relevanz erfassen. Inhaltlich werden zunächst vor allem Aspekte der Mensch-Technik-Interaktion (MTI) betrachtet. Die gestellten Fragen passen sich dabei dynamisch an bisher gegebene Antworten an. Dadurch kann auf effiziente Weise eine Beschreibung ELSI-relevanter Aspekte eines Forschungsvorhabens erstellt und dem Antrag beigelegt werden, idealerweise so knapp wie möglich, aber so umfangreich wie nötig. Als Projektziele standen die Verbesserung der Antragsqualität (Vollständigkeit, Reflexionstiefe, Risikobetrachtung) durch systematische Betrachtung ethischer und rechtlicher Implikationen sowie die Steigerung der Effizienz der Antragerstellung durch Vorstrukturierung eines sequenzierten Fragenkatalogs im Vordergrund.

Im Detail funktioniert ELSI-SAT wie folgt: Wissenschaftler*innen starten die ELSI-SAT-Anwendung auf ihrem Rechner und beantworten 57 Fragen, die in einer vordefinierten Sequenz präsentiert werden (Assessment). Die Fragen sind in drei Hauptkategorien unterteilt: »Allgemein« (Hintergrundinformationen zum Forschungsvorhaben), »Datenschutz« (rechtliche Aspekte) und »Werte« (ethische Aspekte in Bezug auf fünf Wertekategorien). Die Auswahl genau dieser Wertekategorien erfolgte induktiv, indem auf der Analyse von 53 Richtlinien im Bereich der Mensch-Technik-Interaktion und Künstlichen Intelligenz eine Extraktion von Wertbedeutungen und anschließende Subsummierung in fünf übergreifende »Big 5«-Wertekategorien vorgenommen wurde: Fürsorge, Schadensvermeidung, Autonomie, Gerechtigkeit und Transparenz.

In Abhängigkeit des geplanten Forschungsvorhabens (betrachtete Forschungsgebiete, Methodik etc.) wird der Fragenkatalog dabei dynamisch angepasst, d. h. bestimmte Antworten lösen weitere Folgefragen aus. Das Assessment kann in mehreren Sitzungen durchlaufen werden, d. h. Bearbeitungsstände können jederzeit gespeichert und später wieder geladen werden. So kann das Assessment sukzessive aktualisiert und vervollständigt werden. Am Ende erhalten die Nutzer*innen eine Übersicht der Ergebnisse in Form von Grafiken

II. Ethics by Design: Potenziale, Umsetzung und Grenzen

und Risikoampeln für die einzelnen Kategorien (vgl. Abb. 3). Abschließend wird ein PDF-Dokument erzeugt, das sowohl präsentierte Fragen als auch die Antworten protokolliert und die Bewertung und Risikoadressierung zeigt (*Scoring*). Bewertet werden Vollständigkeit der Antworten, die ELSI-Reflexionstiefe, das ELSI-Risikopotenzial sowie die Adressierung dieser Risiken. Das PDF-Dokument kann dann den Antragsunterlagen beigelegt werden. Während bei diesem ersten ELSI-Tool die Risikobewertung und deren Adressierung im Mittelpunkt des Projektes standen, sollte im zweiten Ethics by Design-Tool – ELSI-SAT Health & Care – die Befähigung zur Reflexion ethischer Aspekte das zentrale Anliegen sein.

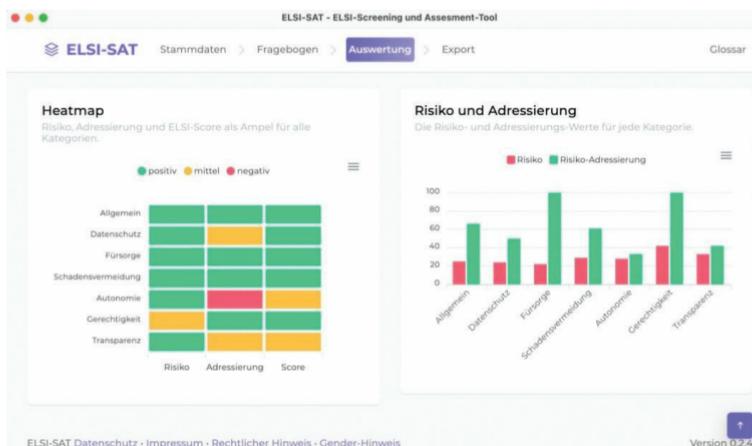


Abb. 3: Screenshot ELSI-SAT Risikoadressierung

5.2 Das Reflexionstool ELSI-SAT Health & Care

Die Evaluation von ELSI-SAT hatte ergeben, dass Nutzende offen für ein Instrumentarium sind, das ihnen sinnvolle Anregungen für einen ethischen und rechtlichen Reflexionsprozess bieten kann. ELSI-SAT Health & Care stellt nicht mehr die Screening- und Assessment-Funktion in den Vordergrund, sondern die Reflexions- und Inspirationsfunktion, womit gleichzeitig ethisches und rechtliches Wissen vermittelt werden soll. Die neue Anwendung legt zum einen den Fokus auf Nutzerfreundlichkeit und Anwendbarkeit und ist inhaltlich explizit

zit auf den Bereich Forschung in Gesundheitswesen, Medizin und Pflege ausgerichtet. So wurde das Forschungsteam nicht nur um Medizinethiker*innen, sondern auch mit UX- und Usability-Expert*innen erweitert, um einer menschzentrierten Gestaltung der Anwendung Rechnung tragen zu können. Auf Basis der Untersuchungsergebnisse wurde das neue Tool mit erweiterten Funktionen und Bereichen entwickelt. Zudem stellte sich im weiteren Projektverlauf heraus, dass ein adäquates Erwartungsmanagement für die Anwendenden wichtig ist, bevor sie mit der Nutzung beginnen. Daher wurde auf der Startseite gezielt darauf eingegangen. Das Projekt entstand im Zeitraum 2021 bis 2023 und wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 16SV8793 gefördert. Die ausführliche Beschreibung des gesamten Entwicklungsprozesses sowie detaillierte Angaben zu dem Konzept und der Ausgestaltung des Tools finden sich in der Veröffentlichung des Konsortialverbundes (vgl. Grimm & Zöllner, 2025), der von der Hochschule der Medien und der Universität Bonn gebildet wurde. Eine ausführliche Darstellung hinsichtlich der ethischen Perspektive beschreiben Susanne Kuhnert (2025) und Jan Mehlich (2023, 2025), hinsichtlich der Usability und User Experience Sarah Bacher, Patrizia Schiffner und Michael Burmester (2025) sowie hinsichtlich der rechtlichen Perspektive Tobias List und Jörn Hoffmann (2025).

Im Überblick ist hier zu nennen, dass ELSI-SAT Health & Care vier verschiedene Toolbereiche anbietet: 1. ELSI-Wertereflection, 2. Erstellung eines ELSI-Profils, 3. ELSI-Karten, in denen Informationen zu den ELSI-Kategorien und narrative Fallbeispiele enthalten sind, und 4. Projektreflexion, bei dem das eigene Projekt anhand einer Checkliste reflektiert wird und ein Transfer von den ELSI-Inhalten zum eigenen Forschungsvorhaben ermöglicht wird.

In der Softwarekonzeption wurden zunächst die relevanten »ELSI-Felder« identifiziert, die im Anwendungskontext hohe Relevanz besitzen. Hierzu wurde eine umfassende Literaturanalyse vorgenommen und eine Unterscheidung nach systemischen und individuellen Implikationen vorgenommen, die in folgender Tabelle dargestellt sind:

Tabelle 2: ELSI-Felder

Systemische Implikationen	Persönliche Implikationen
Gesundheitspolitik und Regulierung	Daten
Ökonomische Faktoren	Persönliche Integrität
Medizinische Berufe und Arbeitsmarkt	Patient*in-Fachkraft-Verhältnis
Umwelt	Professionelles Handeln
Gesellschaft	Individuelle Bedarfe

Mithilfe dieser ELSI-Felder wurde zunächst ein Fragenkatalog mit insgesamt 37 Fragen erarbeitet, dessen Ergebnis die Ausgabe eines sogenannten »ELSI-Profil« ist. Das ELSI-Profil ist die Auswertung der Implikationsrelevanzen der einzelnen Felder anhand einer Punkteskala pro gegebener Antwort auf eine der Fragen. Ein Beispiel hierfür findet sich in Abbildung 4. Dabei werden im Fragenkatalog keine ELS-Implikationen, sondern Innovations- bzw. Technologiecharakteristika abgefragt, um detaillierte Informationen zum geplanten Innovationsprojekt zu erhalten. Die Antworten können dabei je nach Frage sowohl als Multiple-Choice-Antworten als auch als einzelne Auswahlmöglichkeiten vorliegen. Je nach gegebener Antwort werden im Software-Backend Punkte (2 Punkte für eine wahrscheinlich vorliegende Implikation, 1 Punkt für eine mögliche Implikation und 0 Punkte für ein sehr geringes Implikationsrisiko) vergeben und diese Punkte pro Implikationsfeld addiert.

So ergibt sich nach Ausfüllen des Fragebogens ein ELSI-Profil, das die möglichen Implikationswahrscheinlichkeiten der einzelnen Felder darstellt. Mithilfe von Design-Elementen wird das Profil übersichtlich dargestellt und leitet direkt zu den passenden »ELSI-Karten«.

Die ELSI-Karten bilden die vertiefende Lektüre und inhaltliche Sensibilisierung der Profil-Ergebnisse in der Software. Dabei werden die relevanten Felder, die im ELSI-Profil als einschlägig extrahiert wurden, nochmals gezielt hervorgehoben und ermöglichen so eine intensive Auseinandersetzung mit den Themen. Die ELSI-Karten befassen sich neben inhaltlichen Beschreibungen zum Implikationsfeld auch mit Use-Cases und Problemszenarien aus der Praxis. Damit leiten die ELSI-Karten zur Selbstreflexion an, damit die Transferleistung zum eigenen Projekt in der Projektreflexion erfolgen kann. Ein Beispiel für die ELSI-Karten findet sich in Abbildung 5.

Petra Grimm



Abb. 4: Screenshot eines ELSI-Profil aus ELSI-SAT Health & Care

Da die Nutzenden unterschiedliche Berührungspunkte mit ELSI-Themen haben und über unterschiedliche Wissensstände verfügen, beinhaltet ELSI-SAT Health & Care zudem eine Wertereflexion. Diese ist dem Fragenkatalog und den ELSI-Karten vorangestellt, wenngleich die Software sich in beliebiger Reihenfolge nutzen lässt.

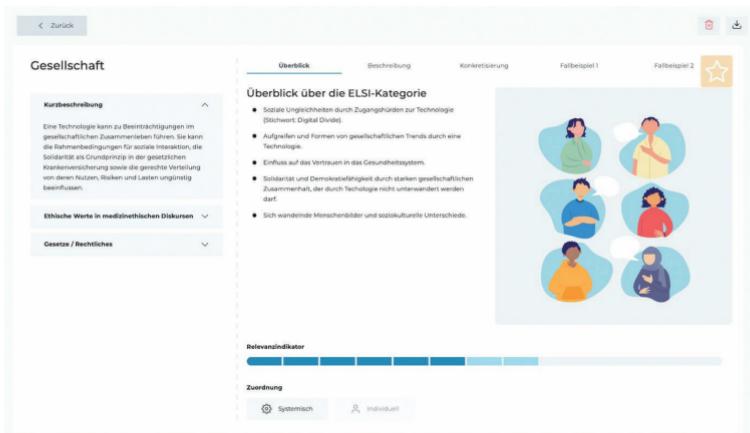


Abb. 5: Screenshot einer ELSI-Karte (Ausschnitt) aus ELSI-SAT Health & Care

II. Ethics by Design: Potenziale, Umsetzung und Grenzen

Die Werterefexion ist als eigenständige Übung zur ersten Auseinandersetzung mit zentralen Werten im eigenen Forschungsprojekt gedacht. Die Nutzenden können hier die ethischen Kernthemen ihrer Innovation auf einer Werte-Zielscheibe selbst einordnen und die Werte dabei aus unterschiedlichen Bedeutungsebenen betrachten. Hierzu sind Kontextbeschreibungen zu den vorgegebenen Werten in die Software integriert, zudem gibt es die Möglichkeit, eigene Notizen hinzuzufügen. Auch kann die Werterefexion um eigene Werte ergänzt werden, die für spezifische Projekte relevant sein können. Ein Beispiel für die Werterefexion findet sich in Abbildung 6.

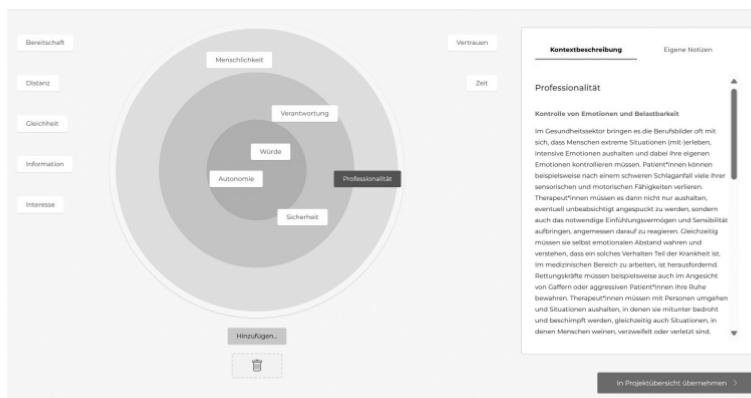


Abb. 6: Screenshot der Werterefexion (Ausschnitt) aus ELSI-SAT Health & Care

Das finale Modul ist die Projektreflexion. Hier leitet ELSI-SAT Health & Care zum Wissenstransfer auf das eigene Projekt an. Die Ergebnisse der Werterefexion, des ELSI-Profil und das durch die ELSI-Karten hinzugewonnene Wissen soll nun auf das eigene Innovationsprojekt übertragen werden und im Forschungsteam so zur Integration von passenden Methoden zur Sicherstellung der ethischen Projektentwicklung beitragen. Dabei werden anhand von Leitfragen in vier Schritten konkrete Schlussfolgerungen angeregt, die im weiteren Projektverlauf integriert werden können, um das Risiko von ELS-Implikationen im Entwicklungsprozess zu minimieren. Ein Beispiel für die Darstellung der Projektreflexion findet sich in Abbildung 7.

Favorisierte ELSI-Kategorien für zusätzliche Reflexionsfragen

Daten ★ Gesellschaft ★ Gesundheitspolitik und -regulierung ★ Individuelle Bedarfe ★ Medizinische Berufe & Arbeitsmarkt ★ Patient-Fachkraft-Verhältnis ★

Persönliche Integrität ★ Professionelles Handeln ★ Umwelt ★ Ökonomische Faktoren ★

Reflexion und Transfer in 4 Schritten

01 Zentrale Werte, Vorannahmen und Auswirkungen relevanter ELSI-Kategorien überprüfen

Sehen Sie sich das ELSI-Wertosystem erneut an und überprüfen Sie Ihre gewählten Werte. Möchten Sie Änderungen vornehmen, basierend auf neu gewonnenen Erkenntnissen?

Betrachten Sie Ihre Werte nochmals aus der Sicht unterschiedlicher Personengruppen. Lassen sich die verschiedenen Interessen ohne Probleme in Einklang bringen? Oder gibt es kritische Punkte, die im Blick behalten müssen?

Vertreten Sie im Forschungsteam gemeinsame Werte und haben Sie ein geteiltes Verständnis von diesen Werten? Oder gibt es Abweichungen, die vielleicht entscheidend sind?

Überprüfen Sie auch Annahmen, die dabei bewusst oder unbewusst gemacht werden, beispielsweise über den Status des Gesundheitssystems oder das Einsatzgebiet Ihrer Technologie.

Notieren und markieren Sie sich kritische Punkte, die Sie mit dem Team besprechen können.

Inscreieren breitete Ihre Technologie die gesellschaftlich wichtigen Themen der Diversität, Inklusion und Integration? Berücksichtigen Sie diese Vorstellungen auch ausdrücklich in Ihrer Projektidee! | Individuelle Bedarfe

Zusätzliche Informationen

Eigene Notizen
z.B. Themen die noch gelärt werden müssen.

Abb. 7: Screenshot einer Projektreflexion (Ausschnitt) aus ELSI-SAT Health & Care

Beide EbD-Tools haben den Vorteil, dass sie offen zugänglich, automatisiert und ohne großen Aufwand durchführbar und damit die Hürden für ihre Umsetzung in Forschungsteams möglichst gering sind.

6. Umsetzungsdimension III: Unternehmensbezogenes Anwendungsszenario für Narrative Ethics by Design

6.1 Start-up with Ethics Compass

Ein weiteres Anwendungsszenario für NED legt den Schwerpunkt auf die Unternehmenspraxis. Im Rahmen des Forschungsprojekts »Ethics by Design«, das die *Bayerische Landeszentrale für neue Medien* gefördert hat, wurde ein NED-Prozess für Start-ups entwickelt. Im ersten Teil des Projekts erfolgte eine empirische Analyse zur Ermittlung der Bedarfe, der Werterelevanz und der Anforderungen an ein praktisches EbD-Tool für Start-ups. Dafür wurden achtzehn Startups aus den Branchen *Medien im Gesundheitsbereich, Journalismus & Unterhaltung, IT & Services, Games, Socializing & Kultur* sowie *Pädagogik* mittels narratologischer Methoden befragt. *In concreto* wurden narrative Interviews sowohl mit Fokusgruppen als auch mit einzelnen Start-ups zu ihrer Gründungsgeschichte durchgeführt

und deren Erlebnisse und Erfahrungen visualisiert. Der Vorteil dieser Methode ist, dass sie die tatsächlichen Ereignisse und Konflikte ersichtlich macht und nicht nur Meinungen und Einstellungen erfasst. Zusammen mit den Start-ups wurde reflektiert, an welchen Stellen explizite oder implizite ethische Entscheidungen eine Rolle spielen. Des Weiteren wurde die Bedeutung von EbD in der Praxis mit ihnen diskutiert. Dabei wurden auch Fortbildungs- und Beratungsbedarfe identifiziert. Des Weiteren wurde mit den Start-Ups ein narratives Planspiel in mehreren Workshops durchgeführt. So entwickelten die Teilnehmer*innen in Kleingruppen Ideen zu einer fiktiven Nanny-App, also einem Online-Service, bei dem die unterschiedlichen Perspektiven der indirekt und direkt Betroffenen sowie Risiken und Chancen »durchgespielt« und dann im Plenum erzählt wurden. Dabei sollten insbesondere auch Wertefragen, wie z. B. zur Privatsphäre, Sicherheit und Kontrollierbarkeit identifiziert werden sowie ethische Fragestellungen erkannt und ethische Orientierungskriterien entwickelt werden. Das Ergebnis der Befragung war, dass aus Sicht der interviewten Start-ups ethische Fragestellungen schon beim oder sogar noch vor dem Gründungsprozess mit einbezogen werden sollten. Als ein weiteres wichtiges Desiderat kristallisierte sich die Forderung nach einer Anschlussfähigkeit von EbD an bekannte Entwicklungsprozesse von Business-Modellen heraus, etwa an das »Business Model Generation Canvas« (vgl. Osterwalder & Pigneur, 2011). Als Ansprechpartner für ein ethisches Coaching, die den Start-ups bei der Entwicklung eines werteorientierten Designs helfen sollten, wurden die Berater- und Gründungsinstitutionen sowie die Inkubatoren genannt. Der Bedarf an Qualifizierung betrifft also einerseits die Start-ups selbst und andererseits die Berater*innen.

Im zweiten Teil des Forschungsprojektes wurde auf der Basis der erzählten Erfahrungen, Herausforderungen, Konflikte und Handlungsfelder die sogenannte SEC-Methode (Start-up-with-Ethics-Compass-Methode) entwickelt. Die Kompass-Metapher wurde deshalb gewählt, weil damit eine individuelle Anpassungsmöglichkeit der Methode an die Bedingungen der jeweiligen Unternehmen zum Ausdruck kommen soll. Denn ein Kompass kann die Richtung anzeigen, er schreibt sie jedoch nicht vor. Die Methode kann also an die individuellen Bedarfe des Start-ups angepasst werden und soll

sie Schritt für Schritt durch einen ethischen Reflexionsprozess des Geschäftsmodells und dessen Operationalisierung führen.

6.2 Das Arbeitsbuch zur SEC-Methode

Das Arbeitsbuch enthält genaue Handlungsanweisungen, wie Start-ups selbstständig die SEC-Methode anwenden können. Begleitend dazu wird eine fiktive Geschichte eines Start-ups mit verschiedenen Persona erzählt, die parallel die einzelnen Schritte des Prozesses durchspielen und diskutieren. Die Vorgehensweise bei der Anwendung der Methode erfolgt in sechs Schritten (vgl. Abb. 8):

Schritt 1 *Core-Story*: Es wird im Team gemeinsam die jeweilige zentrale Identitätsgeschichte, die sogenannte Core-Story, entwickelt, die die Ziele, Motivationen und Geschäftsidee kurz und in narrativer Form schlüssig beschreibt und erste Hinweise darauf gibt, welche ethischen Fragen relevant sein könnten.

Schritt 2 *Werteprofil*: In dieser Phase wird gemeinsam herausgefunden, welche Werte und Wertekonflikte im Team virulent sind und welche Werte mit der Geschäftsidee verbunden sind, um darauf aufbauend den Wertehorizont des Start-ups festzulegen. Hierzu wird das eigens entwickelte Tool »Value Navigation« genutzt, bei dem ein Werte-Logbuch erstellt und der Werte-Horizont des Teams definiert werden. Im Anschluss daran wird der Wertekurs des Start-ups mittels des narrativen Aktenmodells und einer Wertekarte festgelegt.

Schritt 3 *Sinnerzählung*: In diesem Schritt wird die Sinnerzählung des Start-ups mit dem Team entwickelt. Hierzu wird die Passung von Core-Story und Sinnerzählung, die das Ausmalen einer Vision impliziert, überprüft und ggf. angepasst.

Schritt 4 *Netzwerke*: In dieser Phase findet ein Netzwerk-Scan statt, der sowohl die unmittelbaren Stakeholder als auch mittelbaren Akteure definiert. Letztere lassen sich nach dem PESTEL-Prinzip identifizieren, das externe Einflussfaktoren (*political, economic, social, technological, ecological, legal*) auf ein Unternehmen systematisiert. Eine Netzwerk-Matrix, die die (positive/ negative) Einstellung der Akteure zum Projekt und deren Einflussstärke erfasst, soll für eine Reflexion der betroffenen Akteure dienen.

Schritt 5 *Operationalisierung*: Die operationale Ebene überführt die Erkenntnisse der vorherigen Schritte *in concreto*, indem Hand-

lungen (was gemacht wird) hinsichtlich ihrer normativen Optionalität (was getan werden soll) mittels einer Bewertungsmatrix überprüft, bewertet und eventuell korrigiert werden.

Schritt 6 *Kontinuität*: Ethische Reflexion ist keine Momentaufnahme, sondern eine permanente organisatorische Gestaltungsaufgabe. Deshalb beschränkt sich auch das Konzept von EbD nicht nur auf initiale Maßnahmen in der Planungsphase eines Geschäftsmodells. So sollte nicht nur eine stetige Überwachung laufender Prozesse (*Monitoring*) stattfinden, sondern auch der gesamte Lebenszyklus eines Produkts oder einer Dienstleistung (*End-of-Life-Management*) bewertet werden.

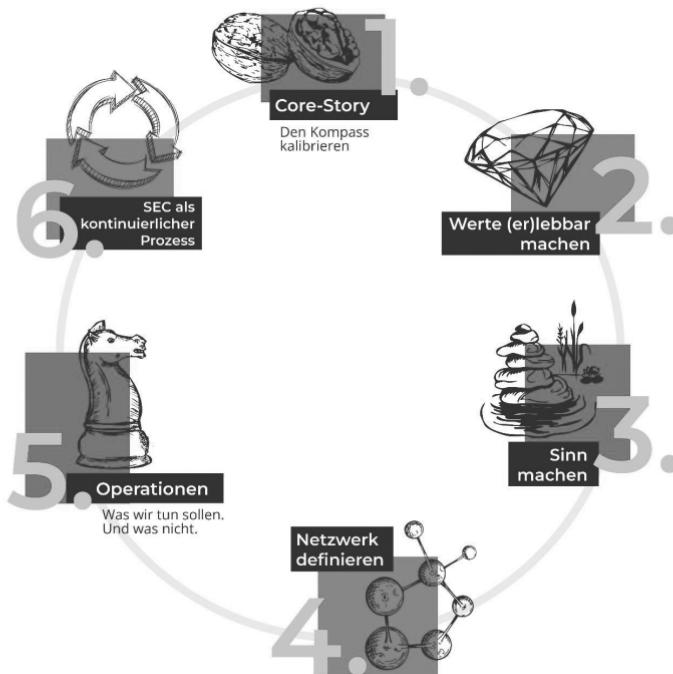


Abb. 8: SEC-Methode

Diese Vorgehensweise beschreibt die idealtypische; in der Praxis können sich jedoch auch andere Reihenfolgen ergeben. Beispielsweise könnten bestimmte Operationen von vorneherein festgelegt sein und die Sinnerzählung muss dann davon abgeleitet werden. In dem hier vorgestellten SEC-Prozess kommt ein wertebasiertes Design nicht nur auf der inhaltlichen, sondern auch auf der kommunikativen Ebene zur Anwendung. So wird nicht nur ausreichend Raum für das Erzählen (*Storytelling*) der Teams gegeben, sondern auch ein Schwerpunkt auf das Zuhören (*Storylistening*) gelegt. Damit soll es ermöglicht werden, eigene, eventuell verengte Standpunkte zu überwinden und gemeinsam Ideen und neue Perspektiven zu entwickeln.

7. Grenzen bzw. Herausforderungen für Ethics by Design

Es lassen sich im Wesentlichen vier Problembereiche von EbD identifizieren: a) die Universalisierung von Werten, b) die Identifikation der relevanten Werte, c) die Favorisierung eines ethischen Ansatzes und d) der menschliche Faktor für ein interdisziplinäres resp. integriertes EbD-Team.

Ein Grundproblem bei dem Versuch, sich auf einen Wertekanon für EbD zu einigen, ist der Anspruch, diese kulturübergreifend und universell zu vereinbaren. Hier wird man sehr schnell mit dem Einwand des Werterelativismus und Wertpluralismus konfrontiert. Das Argument, es gebe keine universalen Werte und deshalb könne man auch keine Werte für global wirksame Technologien ausmachen, scheint ein solches Vorhaben ins Leere laufen zu lassen. Der Einwand ist nicht ohne weiteres von der Hand zu weisen, weil Technologien ja in unterschiedlichen Gemeinschaften, Kulturen etc. zum Einsatz kommen. Dem ist allerdings zu entgegnen, dass Ethics by Design die Werte umfasst, die als normative Standards kulturübergreifend für verbindlich erachtet werden, auch wenn diese kulturspezifisch konkretisiert werden, wie z. B. der Wert der Privatheit. Privatheit ist ein universal gültiger und gleichzeitig in der Anwendung relativierer Begriff, abhängig von dem jeweiligen Kontext und kulturspezifischen Handlungspraxen (vgl. Grimm & Krah, 2016). Trotz kultureller Unterschiede in der Gewichtung von Werten gibt es also Werte, die alle Menschen betreffen. Beispielhaft hierfür sind diejenigen Werte, die durch die in der Allgemeinen Erklärung der

Menschenrechte festgeschriebenen Grundrechte deutlich werden. Diese Werte gilt es zu berücksichtigen, ohne sie gegeneinander aufzuwiegen oder aufgrund von vermeintlichen Sachzwängen oder kulturellen Empfindlichkeiten außer Kraft zu setzen. Gegen den Werterelativismus, der eine Verbindlichkeit von Werten nur systembezogen für möglich hält, oder den Werteneihilismus, der die Existenz moralischer Werte an sich negiert, wehrt sich der Werteuniversalismus. Demgemäß gelten moralische Werte – wenn auch in einer minimalistischen Form – universal. Gleichwohl weist Markus Gabriel (2020), der einen Werteuniversalismus vertritt, darauf hin, dass in den lebenspraktischen Gegebenheiten »eine prinzipiell niemals vollständig zu überwindende Kluft zwischen den universalen Werten und ihren Anwendungsbedingungen in komplexen Handlungsfeldern und einzelnen Situationen« (S. 141) bestehe.

Diese Sichtweise ist auch für die Praxis des EbD wichtig. Um diesem Aspekt gerecht zu werden, sind die Befunde des empirischen Untersuchungsteils eines EDB-Prozesses zu berücksichtigen, da durch die Stakeholder-Befragungen ja die spezifischen Interpretationen von Werten, Werturteile und Wertekonflikte ermittelt werden können. Dennoch ist die Frage nach der Universalisierbarkeit von Werten in EbD-Prozessen ein Aspekt, der eine immer wiederkehrende Beachtung erfordert und den Versuch einer Konsensfindung bedarf.

Ein weiterer kritischer Aspekt von EbD ist die empirisch begründete Definition relevanter Werte: Welchen Stellenwert sollen die durch die Befragung der Stakeholder gewonnenen Werte haben? Glenn Parsons (2016) problematisiert, dass die Befunde von Befragungen nicht eindeutig und deshalb auch keine eindeutigen Handlungsfolgen ableitbar sind. Susanne Kuhnert (2025) folgert deshalb: »Die zentrale Frage, die die größten Probleme aufwirft, lautet dabei schlicht: Wie soll mit den Ergebnissen in Bezug auf Wertvorstellungen umgegangen werden?« (S. 154) Ebenso problematisiert sie mit Verweis auf Parsons, dass die Ergebnisse der Stakeholder-Befragung deskriptiv und nicht normativ seien und dies zu einem naturalistischen Fehlschluss verleiten könne, womit ein Sein mit dem Sollen verwechselt werde. Denn was empirisch von Befragten wünschenswert erscheine, »muss nicht ethisch wünschenswert sein, und deshalb können empirische Ergebnisse keine ethisch-normativen Vorgaben liefern« (Kuhnert, 2025, S. 154). Diese Kritik ist gerechtfertigt.

Denn die Befragung von Betroffenen hinsichtlich ihrer Werte kann nicht automatisch die zu erstrebenden Werte eines EbD-Prozesses festlegen. Wie die Verknüpfung bzw. Abgleichung von konzeptionell definierten und empirisch ermittelten Werten erfolgen soll, ist weder theoretisch noch methodisch in der EbD-Forschung ausreichend begründet.

Eine weitere methodisch immanente Kritik zu EbD bezieht sich auf die Frage, welche Ansätze der Ethik dem Konzept eines EbD zugrunde gelegt werden sollen. Schließlich gibt es unterschiedliche ethische Theorien und Begründungsmuster. Mit einem tugendethischen Ansatz kann man zu völlig anderen Schlüssen gelangen als mit einem konsequentialistischen oder deontologischen Ansatz (vgl. Grimm, 2021).¹ Eine mögliche Lösung wäre, die jeweiligen Perspektiven der Ansätze transparent im EbD-Prozess aufzuzeigen und den Akteuren die Entscheidung darüber, welche Perspektive sie favorisieren, zu überantworten.

Alle bislang aufgezeigten Herausforderungen für die Verwirklichung eines EbD-Ansatzes sind Makulatur gegenüber der Anforderung, EbD interdisziplinär bzw. integrativ umzusetzen. Denn hier wird eine oftmals unüberwindbare Hürde offensichtlich: der menschliche Faktor bzw. die persönliche Divergenz der verantwortlichen Personen, die das EbD-Projekt zu verantworten und zu gestalten haben. So sollte mehr als bisher der Fokus auf die menschlichen Aspekte gelegt werden, da sie für das Gelingen von Ethics by Design in Unternehmen oder Forschung maßgeblich von Bedeutung sind. Bei dieser Dimension handelt es sich um einen weitgehend blinden Fleck in der Literatur, wenn man sich die vielfachen Publikationen zur Inter- und Transdisziplinarität ansieht, die überwiegend die fachlichen, methodischen und formalen Aspekte mit wenigen Ausnahmen (vgl. Gethmann et al., 2015; Hanschitz et al., 2009) in den Vordergrund rücken. »Menschliche Aspekte« meint hier zum einen Haltung, Werte, Tugenden, persönliche Dispositionen, Charaktereigenschaften, Identitäten und Fähigkeiten, die sich in der Persönlichkeit der Akteur*innen manifestieren, letztere aber auch in ihren sozialen und professionellen Rollen zum Ausdruck bringen. Dazu gehören weiterhin auch kommunikative und soziale

¹ Vgl. auch Abschnitt 1. (»Einleitung«) des dritten Teils (Ausbildung sozial-verantwortlicher Ingenieur*innen) des vorliegenden Sachstandsberichts.

Kompetenzen. Hürden für das interdisziplinäre Unterfangen lassen sich auf (mindestens) zwei Dimensionen beschreiben: zum einen auf der wissensbezogenen, zum anderen auf der personenbezogenen Dimension. Während aber die wissensbezogenen Herausforderungen in der Fachliteratur adressiert werden, bleiben die Hürden auf der personenbezogenen Dimensionsebene weitgehend unberücksichtigt. Diese betreffen die menschliche und kommunikative Seite sowie die subjektiv wahrgenommenen Erfahrungswelten der forschenden oder lehrenden Akteure. Sie sind aber maßgebliche Faktoren, die die Motivation und Befähigung der an EbD Beteiligten beeinflussen und deren Erfolg ermöglichen – oder aber auch verhindern können. Um EbD erfolgreich umsetzen zu können, bedarf es einer Ethik der Kommunikation und einer interdisziplinären Kompetenz. Nur unter Berücksichtigung dieser Voraussetzungen und der reflektierten Adressierung der Grenzen von Ethics by Design können die Anwendungsprojekte ihr Potenzial in der Praxis erfolgreich ausschöpfen.

Literaturverzeichnis

- Alexander, S. (1962). They Decide Who Lives, Who Dies. *LIFE Magazine*, 53(19), 102–125.
- AlgorithmWatch. (2024). *AI Ethics Guidelines Global Inventory*. <https://inventory.algorithmwatch.org/>
- Bacher, S., Schiffrer, P., & Burmester, M. (2025). Das neue Tool ELSI-SAT Health & Care: Projektteams mit ELSI inspirieren. In P. Grimm & O. Zöllner (Hrsg.), *Ethik der Digitalisierung in Gesundheitswesen und Pflege. Analysen und ein Tool zur Integrierten Forschung* (S. 157–174). Franz Steiner Verlag. <https://doi.org/10.25162/9783515135528>
- Bellon, J., & Nähr-Wagener, S. (2020). Interdisziplinarität, ELSI und Integrierte Forschung – aus Einem Vieles und aus Vielem Eines? In B. Gransche & A. Manzeschke (Hrsg.), *Das geteilte Ganze. Horizonte integrierter Forschung für künftige Mensch-Technik-Verhältnisse* (S. 37–52). Springer VS. https://doi.org/10.1007/978-3-658-26342-3_2
- Böschen, S., Grunwald, A., Krings, B.-J., & Rösch, C. (2021). Technikfolgenabschätzung – neue Zeiten, neue Aufgaben. In S. Böschen, A. Grunwald, B.-J. Krings, & C. Rösch (Hrsg.), *Technikfolgenabschätzung. Handbuch für Wissenschaft und Praxis* (S. 15–40). Nomos. <https://doi.org/10.5771/978378901990-13>

- Brandt, A. M. (1978). Racism and Research: The Case of the Tuskegee Syphilis Study. *Hastings Center Report*, 8(6), 21–29. <https://www.jstor.org/stable/3561468?origin=crossref>
- Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. *AI and Ethics*, 4(4). <https://doi.org/10.1007/s43681-023-00330-4>
- Capurro, R. (2017). *Homo Digitalis. Beiträge zur Ontologie, Anthropologie und Ethik der digitalen Technik*. Wiesbaden.
- Capurro, R. (2009). Digital Ethics. In The Academy of Korean Studies and Korean National Commission for UNESCO (Hrsg.), *Civilization and Peace* (S. 207–216). The Academy of Korean Studies.
- Cavoukian, A. (2011). *Privacy by Design*. <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>
- Datenethikkommission der Bundesregierung. (2019). *Gutachten der Datenethikkommission der Bundesregierung*. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6
- Eeckels, J., & Roozenburg, N. F. M. (1991). A Methodological Comparison of the Structures of Scientific Research and Engineering Design: Their Similarities and Differences. *Design Studies*, 12(4), 197–203. [https://doi.org/10.1016/0142-694X\(91\)90031-Q](https://doi.org/10.1016/0142-694X(91)90031-Q)
- Emanuel, E. J., Grady, C., Crouch, R. A., Lie, R., Miller, F. G. & Wendler, D. (Hrsg.). (2023). *The Oxford Textbook of Clinical Research Ethics*. Oxford University Press. <https://doi.org/10.1093/med/9780195168655.001.0001>
- European Commission. (2021). *Ethics By Design and Ethics of Use Approaches for Artificial Intelligence*. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf.
- Flanagan, M., & Nissenbaum, H. (Hrsg.). (2014). *Values at Play in Digital Games*. <https://doi.org/10.7551/mitpress/9016.001.0001>
- Friedman, B. (Hrsg.). (1997). *Human values and the design of computer technology*. CSLI Publications/Cambridge University Press.
- Friedman, B. (Hrsg.). (1997). *Human Values and the Design of Computer Technology*. Cambridge University Press.
- Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design. Shaping Technology with Moral Imagination*. The MIT Press.
- Friedman, B., Kahn, P. H., & Borning, A. (2002). Value Sensitive Design: Theory and Methods. *UW CSE Technical Report*, 02-12-01, 1–8.
- Friedman, B., Kahn P., & Borning A. (2008). Value Sensitive Design and Information Systems. In K. E. Himma & H.T. Tavani (Hrsg.), *The Handbook of Information and Computer Ethics* (S. 69–101). John Wiley & Sons. <https://doi.org/10.1002/9780470281819.ch4>
- Gabriel, M. (2020). *Moralischer Fortschritt in dunklen Zeiten*. Ullstein.

II. Ethics by Design: Potenziale, Umsetzung und Grenzen

- Gethmann, C. F., Carrier, M., Hanekamp, G., Kaiser, M., Kamp, G., Lingner, S., Quante, M., & Thiele, F. (2015). *Interdisciplinary Research and Transdisciplinary Validity Claims*. Springer. <https://doi.org/10.1007/978-3-319-11400-2>
- Glenn, J. C. (2009). Introduction to the futures research. In J. C. Glenn, & T. J. Gordon (Hrsg.), *Futures research methodology. Version 3.0. The Millennium Project*. <https://millennium-project.org/publications-2/futures-research-methodology-version-3-0-2/>
- Gransche, B., & Manzeschke, A. (2020). Das geteilte Ganze. Einleitende Überlegungen zu einem Forschungsprogramm. In B. Gransche & A. Manzeschke (Hrsg.), *Das geteilte Ganze. Horizonte integrierter Forschung für künftige Mensch-Technik-Verhältnisse* (S. 1–33). Springer VS. https://doi.org/10.1007/978-3-658-26342-3_1
- Gransche, B., & Manzeschke, A. (Hrsg.). (2020). *Das geteilte Ganze. Horizonte integrierter Forschung für künftige Mensch-Technik-Verhältnisse* (S. 53–70). Springer VS. <https://doi.org/10.1007/978-3-658-26342-3>
- Grimm, P. (2021). Was können ethische Ansätze für eine werteorientierte Digitalisierung leisten? In C. Piallat (Hrsg.), *Der Wert der Digitalisierung. Gemeinwohl in der digitalen Welt* (S. 55–96). Transcript. <https://doi.org/10.14361/9783839456590-003>
- Grimm, P. (2024). Ansätze der Narrativen Ethik. In P. Grimm, K. E. Trost & O. Zöllner (Hrsg.), *Digitale Ethik* (S. 63–76). Nomos.
- Grimm, P. (2025). Künstliche Intelligenz: Prometheus oder Büchse der Pandora? Herausforderungen für eine narrative Ethik der KI. In Z. Mahayni (Hrsg.), *Ethische Fragen im Digitalzeitalter* (S. 135–154). Aisthesis Verlag.
- Grimm, P., & Krah, H. (2016). Privatsphäre. In J. Heesen (Hrsg.), *Handbuch Medien- und Informationsethik* (S. 178–185). J. B. Metzer. https://doi.org/10.1007/978-3-476-05394-7_24
- Grimm, P., Keber, T., & Müller, M. (2021). *Start-up with Ethics. Mit der SEC-Methode ein werteorientiertes Unternehmen entwickeln*. Koepäd.
- Grimm, P., & Kuhnert, S. (2018). Narrative Ethik in der Forschung zum automatisierten und vernetzten Fahren. In P. Grimm & O. Zöllner (Hrsg.), *Mensch – Maschine. Ethische Sichtweisen auf ein Spannungsverhältnis* (S. 93–110). Franz Steiner Verlag. <https://doi.org/10.25162/9783515122726>
- Grimm, P., Keber, T., & Zöllner, O. (2019). *Digitale Ethik. Leben in vernetzten Welten*. Reclam.
- Grimm, P., & Mönig, J. (2020). Ethical recommendations for cooperative driver-vehicle interaction – Guidelines for highly automated driving. In G. Meixner, (Hrsg.), *Smart Automotive Mobility: Reliable Technology for the mobile Human.* (S. 213–229). Springer. https://doi.org/10.1007/978-3-030-45131-8_4
- Grunwald, A. (2022). *Technikfolgenabschätzung. Einführung* (3. Aufl.). Nomos. <https://doi.org/10.5771/9783748928775>

- Halbig, C. (2013). *Der Begriff der Tugend und die Grenzen der Tugendethik*. Suhrkamp.
- Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds & Machines*, 30, 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Hanschitz, R.-C., Schmidt, E., & Schwarz, G. (2009). *Transdisziplinarität in Forschung und Praxis. Chancen und Risiken partizipativer Prozesse*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-91451-0>
- Hoffmann, J., & List, T. (2025). Konzeption und Umsetzung eines Digital-Tools aus rechtlicher Sicht. In P. Grimm & O. Zöllner (Hrsg.), *Ethik der Digitalisierung in Gesundheitswesen und Pflege. Analysen und ein Tool zur Integrierten Forschung* (S. 175–196). Franz Steiner Verlag. <https://doi.org/10.25162/9783515135528>
- Jones, J. H. (1993). *Bad Blood: The Tuskegee Syphilis Experiment*. Free Press.
- Jonsen, A. R. (2023). *The Birth of Bioethics* (online edition). Oxford Academic. <https://doi.org/10.1093/oso/9780195103250.001.0001>
- Kass, L. R. (1971). Babies by Means of In Vitro Fertilization: Unethical Experiments on the Unborn? *New England Journal of Medicine*, 285(21), 1174–1179. <https://doi.org/10.1056/NEJM197111182852105>
- Kemmer, D. (2020). A → B („Wenn A, dann B“) Zur zentralen Rolle von Implikationen für die Konzeption und Praxis einer Integrierten Forschung. In B. Granske & A. Manzeschke (Hrsg.), *Das geteilte Ganze. Horizonte integrierter Forschung für künftige Mensch-Technik-Verhältnisse* (S. 53–70). Springer VS. https://doi.org/10.1007/978-3-658-26342-3_3
- Krobath, H. T. (2009). *Werte. Ein Streifzug durch Philosophie und Wissenschaft. Mit einem Vorwort von Hans Albert*. Königshausen & Neumann.
- Kuhnert, S. (2025). BMBF-Projekt ELSI-SAT Health & Care: Ethische Fundierung und Hintergründe zur Konzeption des Moduls Wertereflexion. In P. Grimm & O. Zöllner (Hrsg.), *Ethik der Digitalisierung in Gesundheitswesen und Pflege. Analysen und ein Tool zur Integrierten Forschung* (S. 141–156). Franz Steiner Verlag. <https://doi.org/10.25162/9783515135528>
- Lautmann, R. (1971). *Wert und Norm. Begriffsanalysen für die Soziologie* (2. Aufl.). Westdeutscher Verlag.
- Lenzen, M. (2011). Informationsethik. In R. Stoecker, C. Neuhäuser, & M.-L. Raters (Hrsg.), *Handbuch Angewandte Ethik* (S. 210–215). J. B. Metzler.
- MacIntyre, A. (1995). *Der Verlust der Tugend. Zur moralischen Krise der Gegenwart*. Suhrkamp.
- Mehlich, J. (2023). ELSI-SAT Health & Care: Ein Ethics-By-Design-Tool für Integrierte Forschung und Technikentwicklung im Frühstadium. In P. Grimm, H. Pechlaner & O. Zöllner (Hrsg.), *Medien – Ethik – Digitalisierung. Aktuelle Herausforderungen* (S. 59–75). Franz Steiner Verlag. <https://doi.org/10.25162/9783515136013>

II. Ethics by Design: Potenziale, Umsetzung und Grenzen

- Mehlich, J. (2025). Jenseits von Bereichsethiken: Die Navigation epistemischer Regime der Angewandten Ethik im ELSI-SAT-H&C-Projekt. In P. Grimm & O. Zöllner (Hrsg.), *Ethik der Digitalisierung in Gesundheitswesen und Pflege. Analysen und ein Tool zur Integrierten Forschung* (S. 127–140). Franz Steiner Verlag. <https://doi.org/10.25162/97835135528>
- Mieth, D. (1976). *Dichtung, Glaube und Moral. Studien zur Begründung einer narrativen Ethik. Mit einer Interpretation zum Tristanroman Gottfrieds von Strassburg*. Matthias-Grünewald-Verlag.
- Mieth, D. (1998). *Moral und Erfahrung II. Entfaltung einer theologisch-ethischen Hermeneutik*. Verlag Herder.
- Mieth, D. (1999). *Moral und Erfahrung I. Grundlagen einer theologisch-ethischen Hermeneutik*. Verlag Herder.
- Mieth, D. (2000). Identität – wie wird sie erzählt? In D. Mieth (Hrsg.), *Erzählen und Moral. Narrativität im Spannungsfeld von Ethik und Ästhetik* (S. 67–82). Attempto Verlag.
- Mieth, D. (2002). Rationalität und Narrative Ethik. Eine Erweiterung der rationalen Zugänge in der Ethik. In N. C. Karafyllis, & J. C. Schmidt (Hrsg.), *Zugänge zur Rationalität der Zukunft* (S. 277–303). Metzler. https://doi.org/10.1007/978-3-476-02903-4_12
- Mieth, D. (2007). Literaturethik als narrative Ethik. In K. Joisten (Hrsg.), *Narrative Ethik. Das Gute und das Böse erzählen* (S. 214–234). Akademie Verlag. <https://doi.org/10.1524/9783050047287.215>
- Mikami, K., Ema, A., Minari, J., & Yoshizawa, G. (2021). ELSI is Our Next Battlefield. *East Asian Science, Technology and Society: An International Journal*, 15, 86–96. <https://doi.org/10.1080/18752160.2021.1881279>
- Müller, M., Grimm, P. (2016). *Narrative Medienforschung. Einführung in Methodik und Anwendung*. UVK Verlagsgesellschaft.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. U.S. Government Printing Office. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>
- Nussbaum, M. C. (1985). Finely Aware and Richly Responsible: Moral Attention and the Moral Task of Literature. *The Journal of Philosophy*, 82(10), 516–529. <https://doi.org/10.2307/2026358>
- Nussbaum, M. C. (2001). *The Fragility of Goodness. Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511817915>
- Nussbaum, M. C. (1992). *Love's Knowledge. Essays on Philosophy and Literature*. Oxford University Press. <https://doi.org/10.1093/oso/9780195074857.01.0001>

- Osterwalder, A., & Pigneur, Y. (2011). *Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer*. Campus Verlag.
- Parsons, G. (2016). *The Philosophy of Design*. Polity Press.
- Rorty, R. (1992). *Kontingenz, Ironie und Solidarität* (1. Aufl.). Suhrkamp.
- Rorty, R. (2001). *Philosophie & die Zukunft. Essays*. Fischer Taschenbuch.
- Schikowitz, A., & Maasen, S. (2021). Integrative Forschung. In T. Schmohl & T. Philipp (Hrsg.), *Handbuch Transdisziplinäre Didaktik* (S. 151–165). Transcript. <https://doi.org/10.14361/9783839455654-015>
- Simon, J. (2016). Values in Design. In J. Heesen (Hrsg.), *Handbuch Medien und Informationsethik* (S. 357–364). J. B. Metzler. https://DOI 10.1007/978-3-476-05394-7_1
- Spiekermann, S. (2015). *Ethical IT Innovation: A Value-Based System Design Approach*. Taylor & Francis Group. <https://doi.org/10.1201/b19060>
- Spiekermann, S. (2021). Value-based Engineering: Prinzipien und Motivation für bessere IT-Systeme. *Informatik Spektrum*, 44, 247–256. <https://doi.org/10.1007/s00287-021-01378-4>
- Spiekermann, S. (2023). *Value-Based Engineering: A Guide to Building Ethical Technology for Humanity*. De Gruyter. <https://doi.org/10.1515/9783110793383>
- Stilgoe, J., & Guston, D. H. (2017). Responsible Research and Innovation. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Hrsg.), *The Handbook of Science and Technology Studies* (4. Aufl., S. 853–880). MIT Press.
- Vallor, S. (2016). *Technology and the Virtues. A Philosophical Guide to a Future Worth Wanting*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190498511.001.0001>
- van den Hoven, J. (2017). The Design Turn in Applied Ethics. In J. Van den Hoven, S. Miller, & T. Pogge (Hrsg.), *Designing in Ethics* (S. 11–31). Cambridge University Press. <https://doi.org/10.1017/9780511844317.004>
- van de Poel, I. (2017). Dealing with Moral Dilemmas through Design. In J. Van den Hoven, S. Miller & T. Pogge (Hrsg.), *Designing in Ethics* (S. 57–77). Cambridge University Press. <https://doi.org/10.1017/9780511844317.004>
- Yaghmaei, E., & Van de Poel, I. (Hrsg.). (2021). *Assessment of Responsible Innovation. Methods and Practices*. Routledge. <https://doi.org/10.4324/9780429298998>
- Zweig, K. (2023). *Die KI war's. Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz*. Wilhelm Heyne.