

---

Ciplus  
Band 2/2023

# Konviviale Künstliche Intelligenz: Definition und Entwicklung eines Vorgehensmodells

Markus Dusdal

Richard Schulz

Christoph Haag

Thomas Bartz-Beielstein



# Inhalt

|  |           |
|--|-----------|
| <b>1. EINFÜHRUNG</b>   | <b>1</b>  |
| <b>2. THEORETISCHE GRUNDLAGEN</b>  | <b>3</b>  |
| 2.1. KONVIVAL KONZIPIERTE WERKZEUGE  | 3         |
| 2.2. VORGEHENSMODELLE FÜR DIE EINFÜHRUNG VON KI-SYSTEMEN                           | 4         |
| 2.2.1. <i>CRISP-DM</i>   | 4         |
| 2.2.2. <i>Erweiterungen des CRISP-DM</i>   | 5         |
| 2.2.3. <i>Alternative Vorgehensmodelle</i>   | 5         |
| <b>3. ABLEITUNG VON PRINZIPIEN FÜR EINE „KONVIVIALE KI“</b>                        | <b>7</b>  |
| 3.1. ANWENDUNGS- / NUTZENZENTRIERTE VORGEHENSWEISE                                 | 7         |
| 3.2. BETONUNG DER MENSCH-TECHNIK-INTERAKTION UND SELEKTIVE (VOLL-) AUTOMATISIERUNG | 7         |
| 3.3. REDUZIERUNG DER TECHNOLOGISCHEN KOMPLEXITÄT VON KI (ERKLÄRBARKEIT)            | 8         |
| <b>4. ERSTELLUNG EINES KONVIVIALEN KI-MODELLS</b>                                  | <b>9</b>  |
| 4.1. PHASE: INITIAL OPPORTUNITY RECOGNITION  | 9         |
| 4.2. PHASE: REQUIREMENTS ENGINEERING   | 10        |
| 4.3. PHASE: MODEL DEVELOPMENT  | 11        |
| 4.4. PHASE: MODEL EVALUATION   | 11        |
| 4.5. OUTCOMES / BOUNDARY OBJECTS   | 11        |
| <b>5. ZUSAMMENFASSUNG</b>  | <b>13</b> |
| <b>LITERATURVERZEICHNIS</b>  | <b>14</b> |



# 1. Einführung

Die Entwicklung künstlicher Intelligenz (KI) und ihre wirtschaftliche Verbreitung wurden seit Beginn der 2010er Jahren bis heute vorangetrieben. Das zeigt u.a. ein Blick auf die zur Verfügung gestellten Finanzmittel für KI-Startups: Diese sind in der Zeit zwischen 2011 und 2017 um den Faktor 50 auf mehr als 15 Mrd. US\$ angewachsen (CB Insights 2018). Die Einsatzgebiete für KI reichen von intelligenten Produktlösungen über innerbetriebliche Prozessoptimierungen bis hin zu automatisierten Kundenservices und Pricing-Vorgängen (Gentsch 2018, S. 43). Heutzutage beeinflussen KI-Anwendungen und deren algorithmische Implikationen den beruflichen und privaten Alltag auf ganz selbstverständliche und mitunter subtile Art und Weise. Es ist absehbar, dass zukünftig Entwicklungssprünge in immer kürzeren Innovationszyklen und ein dynamisch zunehmender Einzug von KI in die Gesellschaft folgen werden.

Trotz des zweifelsfrei vorhandenen Nutzenpotenzials von KI wird deren zunehmende Verbreitung mittlerweile kritisch diskutiert. Zweifel bezüglich der Auswirkungen und Vorteilhaftigkeit von KI-Anwendungen wurden zuletzt öffentlichkeitswirksam durch ein allgemeines Positionspapier führender KI-Experten, wie ChatGPT-Erfinder Sam Altman, zum Ausdruck gebracht. Hierin werden insbesondere die durch KI entstehenden Gefahren für die Menschheit betont (Forschung & Lehre 2023).

Andere Studien warnen vor einem unkontrollierten Einsatz von KI, da dieser im Endeffekt kontraproduktiv sein und zu einer „Verdummung“ von Mitarbeitern führen kann (Fügener et al. 2021; Granados 2022). Um potentielle Gefahren eines unkontrollierten KI-Einsatzes abzuwenden hat das EU-Parlament zuletzt den sog. „AI-Act“ ins Leben gerufen, der u.a. eine Risikobewertung von generativer (d.h. künstliche Inhalte generierender) KI vorsieht (News Europäisches Parlament 2023). Die zunehmende Sorge um einen ausufernden KI-Einsatz, der die menschliche Kontrollfähigkeit übersteigt, hat innerhalb der Wissenschaft zu neuen Forschungsrichtungen geführt, die als „Responsible AI“, „Trustworthy AI“ oder „Explainable AI“ bezeichnet werden. Diese ambitionierten und auf großes Forschungsinteresse stoßenden Denkansätze sind bislang jedoch wenig ausgereift oder gar für den Praxiseinsatz operationalisiert (Linardatos et al. 2020, S. 36).

Die Entwicklung und Anwendung von KI erzeugt ein zwiespältiges Meinungsbild: Dem offenkundigen Nutzen für die Gesellschaft und Industrie stehen Sorgen über unerwünschte Nebenwirkungen durch fehlerhafte oder missbräuchliche Anwendungen gegenüber. Folglich bedarf es neuer Lösungsansätze zur Erschließung dieser Potentiale bei gleichzeitiger Minimierung möglicher Risiken. Zur Lösung dieses Konflikts schlagen wir einen Ansatz vor, der als „konviviale künstliche Intelligenz“ (konviviale KI) bezeichnet und im Folgenden näher erläutert werden soll.

Im allgemeinen Sprachgebrauch bezeichnet „Konvivialität“ das gemeinschaftliche Miteinander (so etwa zwischen Menschen und Technik). Konvivialität beschreibt Vorgehensweisen, bei denen die Gemeinschaft im Vordergrund steht. Im engeren Sinne lehnt sich der hier vorgestellte Ansatz an den durch Ivan Illich geprägten Begriff der „konvivial konzipierten Werkzeuge“ (engl. „Tools for Conviviality“) an. Danach sind unter konvivialen Werkzeugen solche Technologien und Hilfsmittel zu verstehen, „welche zwar die Produktivität menschlicher Arbeitskraft erhöhen, diese aber nicht ersetzen“ (Paech 2013, S. 59). Die Bezeichnung „konviviale KI“ soll folglich auf ein harmonisches Zusammenspiel zwischen KI und Mensch hindeuten. Hiermit wird nicht etwa die Entwicklung einer neuen

Technologie angestrebt. Vielmehr werden damit eine neue Anspruchshaltung und damit einhergehende, menschenzentrierende Gestaltungselemente formuliert, die es bei der Entwicklung und Implementierung von KI-Modellen zu berücksichtigen gilt.

Um diesen neuen Ansatz darzulegen, werden nachfolgend zunächst Illichs Gedanken zu „konvivial konzipierten Werkzeugen“ erläutert. Es folgt die Vorstellung bereits existierender und in der Industrie etablierter Vorgehensmodelle für die Entwicklung und Implementierung von künstlicher Intelligenz. Anschließend werden – ausgehend von Illichs Konvivialitätsverständnis – die Prinzipien eines „konvivialen KI“-Systems abgeleitet und ausführlich dargestellt. Schließlich wird ein neues Vorgehensmodell vorgestellt, welches die Entwicklung von KI-Systemen unter Berücksichtigung der konvivialen Leitprinzipien befördern und operationalisieren soll.

## 2. Theoretische Grundlagen

Im folgenden Kapitel werden die vom Ökonomen Ivan Illich konvivial konzipierten Werkzeuge erläutert. Anschließend werden mögliche Vorgehensmodelle aus der gegenwärtigen KI-Landschaft vorgestellt.

### 2.1. Konvivial konzipierte Werkzeuge

Nachdem sich bereits Autoren wie Günther Anders (1956) und Lewis Mumford (1968) aus unterschiedlichen Beweggründen gegen einen starken Technologie-Einzug in die Gesellschaft positioniert hatten, verstärkten sich in den frühen 1970er Jahren entsprechende Bewegungen, die sich insbesondere vor dem Hintergrund der Veröffentlichung „The Limits to Growth“ des Club of Rome im Jahr 1972 kritisch mit der Gestaltung von Technologien befassten (Vetter 2018).

Zu den Gründungsvätern derartiger Denkströmungen zählt Ivan Illich mit dem Ansatz des „konvivial konzipierten Werkzeugs“. In seinem Essay „Selbstbegrenzung“ übt Ivan Illich massiv Kritik an dem Wachstumsgedanken hochindustrialisierter Gesellschaften und warnt vor den Gefahren einer übertechnisierten Lebenswelt. Als Gegenmodell zum Wachstumsparadigma zeichnet Illich das Bild einer „konvivialen“ Gesellschaft. Hierbei ist der Gedanke der „Konvivialität“ durch Autonomie und schöpferischer Kreativität charakterisiert und soll als Gegensatz zu den konditionierten Reaktionen von Menschen auf Anforderungen durch eine künstliche Umwelt verstanden werden (Illich 1998, S. 28). Die Notwendigkeit eines entsprechenden Gesellschaftswandels begründet Illich mit der Überzeugung, dass die Mitglieder eines Systems bei mangelnder Konvivialität durch keine noch so hohe industrielle Produktivität befriedigt werden können (Illich 1998, S. 29).

Ob eine Gesellschaft als konvivial bezeichnet werden kann, hängt laut Illich maßgeblich von der Gestaltung und dem Umgang mit Werkzeugen ab. Menschen empfinden nicht nur Befriedigung, sondern Freude, wenn sie schöpferisch tätig sein können. Hingegen entstünden Abhängigkeiten, Ausbeutung und Ohnmacht, sobald Werkzeuge zu leistungsfähig und über das menschliche Fassungsvermögen hinauswachsen würden. Werkzeuge gelten nach Ansicht von Illich somit als konvivial, wenn sie jedem, der sie benutzt, die bestmögliche Gelegenheit bieten, die Umwelt mit den Ergebnissen seiner Visionen zu bereichern (Illich 1998, S. 41). Anhand der Beschreibung „konvivialer Werkzeuge“ fällt auf, dass der Begriff der Werkzeuge von Illich weit gefasst wird, weshalb in der Fachliteratur häufig auch von „konvivialen Technologien“ gesprochen wird, wenngleich Illich diesen Begriff nicht verwendete (Vetter 2018, S. 1781).

Das Gegenstück zu konvivialen Werkzeugen bezeichnet Illich als „manipulative Werkzeuge“. Diese verursachen mehr Kosten als Nutzen, sind exklusiv und erzeugen Abhängigkeiten. Hervorzuheben ist, dass nicht jedes in einer konvivialen Gesellschaft vorkommende Produktionsmittel die Kriterien der Konvivialität erfüllen muss, das Verhältnis zwischen den jeweiligen Werkzeugen jedoch ausgewogen sein sollte (Illich 1998, S. 44–45).

Weiterhin hebt Illich sechs Faktoren hervor, die durch industrielle Entwicklungsprozesse entstanden sind und Bedrohungen für den Menschen darstellen. Diese umfassen unter anderem eine durch fehlerhafte Technik unbewohnbar werdende Umwelt, die Dominanz

einzelner Produkte („radikales Monopol“) und die Obsoleszenz von Produkten (Illich 1973, S. 57–58). Illich spricht von sogenannten „Wasserscheiden“ als markanten Zeitpunkten („Wendejahren“), an denen technologische und wissenschaftliche Errungenschaften begannen, kontraproduktive Wirkungen auf die Gesellschaft und Umwelt zu entfalten (Leggewie et al. 2012, S. 307). Mit Blick auf die technologischen, wirtschaftspolitischen und klimatischen Entwicklungen seit Illichs Ausführungen in den 1970er Jahren darf festgestellt werden, dass seine Grundgedanken heutzutage aktueller und relevanter denn je sind (Leggewie et al. 2012, S. 307; Vetter 2018, S. 1781).

## 2.2. Vorgehensmodelle für die Einführung von KI-Systemen

In der Fachliteratur existieren verschiedene KI-Vorgehensmodelle, von denen einige nachfolgend exemplarisch dargestellt werden sollen. Neben einer ausführlichen Beschreibung des renommierten CRISP-DM und einer Auflistung von Nachfolge-Modellen, die auf den Grundgedanken des CRISP-DM basieren, sollen dazu auch die Ansätze alternativer Vorgehensmodelle aufgegriffen werden.

### 2.2.1. CRISP-DM

Ein verbreitetes Vorgehensmodell für Einführung von KI-Prozessen ist der „Cross-Industry Standard Process for Data Mining“ (Abk.: CRISP-DM; siehe Abb. 1). CRISP-DM ist ein branchenunabhängiges iteratives, sechs-phasiges Vorgehensmodell für Data Mining (Schröder et al. 2021, S. 527) und wird in der Fachliteratur als die am weitesten verbreitete Analytik-Methode bzw. als De-facto-Standard für die Durchführung von Datenanalysen in industriellen Anwendungen gehandelt (Huber et al. 2019, S. 403; Martinez-Plumed et al. 2021, S. 3048; Schröder et al. 2021, S. 526), was wesentlich durch die hohe Flexibilität bei der Verwendung des Ansatzes begründet ist.

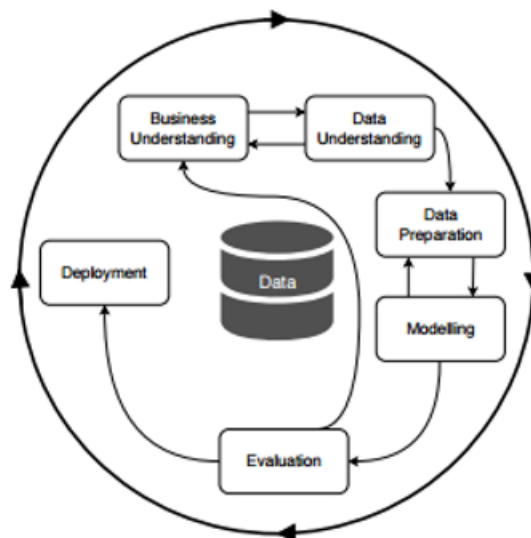


Abbildung 1: CRISP-DM Prozessmodell (Martinez-Plumed et al. 2021, S. 3049)

Die initiale Modellphase wird als „Business Understanding“ bezeichnet und zielt darauf ab, den Projektumfang und konkrete Projektziele, wie z.B. Kriterien zur Erfolgsmessung, zu definieren. In Abhängigkeit vom vorliegenden Problem erfolgt in der Phase „Data Understanding“ die Erfassung und Sammlung von Daten, wobei diese bereits durch ihre Quelle, Eigenschaften oder Qualität eine unterschiedliche Eignung aufweisen und



entsprechend vorselektiert werden können (Maulana et al. 2020, S. 2). Im Rahmen der anschließenden Phase „Data Preparation“ werden die erhobenen Daten mittels Datenauswahl, -bereinigung, -normalisierung und -transformation (Maulana et al. 2020, S. 2; Gonçalves et al. 2020, S. 558) für die weitere Verarbeitung vorbereitet. In der Phase „Modeling“ erfolgt die Auswahl einer in Abhängigkeit der genauen Projektfragestellung geeigneten Modellierungstechnik. Abschließend werden die durch das Modell erzielten Ergebnisse hinsichtlich ihrer Genauigkeit geprüft und im Hinblick auf die anfangs definierten Geschäftsziele bewertet (Maulana et al. 2020, S. 2; Gonçalves et al. 2020, S. 559). Die abschließende „Deployment“-Phase umfasst bspw. die Bereitstellung einer Softwarekomponente oder eines Abschlussberichts, sowie die Implementierung neuer Architekturen oder Modelle in bestehende Systeme (Schröder et al. 2021, S. 531).

### 2.2.2. Erweiterungen des CRISP-DM

In den vergangenen Jahren war zu beobachten, wie das CRISP-DM zahlreichen Erweiterungs- und Optimierungsversuchen unterzogen worden ist. Beispielhaft sollen einige dieser Ansätze nachfolgend erläutert werden:

- **APRED-DM:** Nagashima und Kato (2019) schlagen das APREP-DM-Framework vor, das einen besonderen Fokus auf die Vorverarbeitung von Sensordaten legt.
- **CRISP-ML (Q):** Studer et al. (2020) sehen eine zusätzliche Phase der Modell-Überwachung und -Wartung im CRISP-DM als erforderlich an und integrieren diese im CRISP-ML(Q).
- **LDTM:** Um die Stärken des CRISP-DM mit Lean-Startup- und Design-Thinking-Strategien zu verknüpfen, entwickeln Ahmed et al. (2018) die Lean Design Thinking Methodology (LDTM).
- **DMME:** Auch Huber et al. (2019, S. 406) gehen mittels des Data Mining Methodology for Engineering Applications (DMME) über die ursprüngliche Methodik des CRISP-DM hinaus. Die zu definierenden Geschäftsziele werden um technische Parameter erweitert und eine Prüfung zusätzlich erforderlicher Sensoren und Schnittstellen vorgesehen ist (siehe Abb. 2).

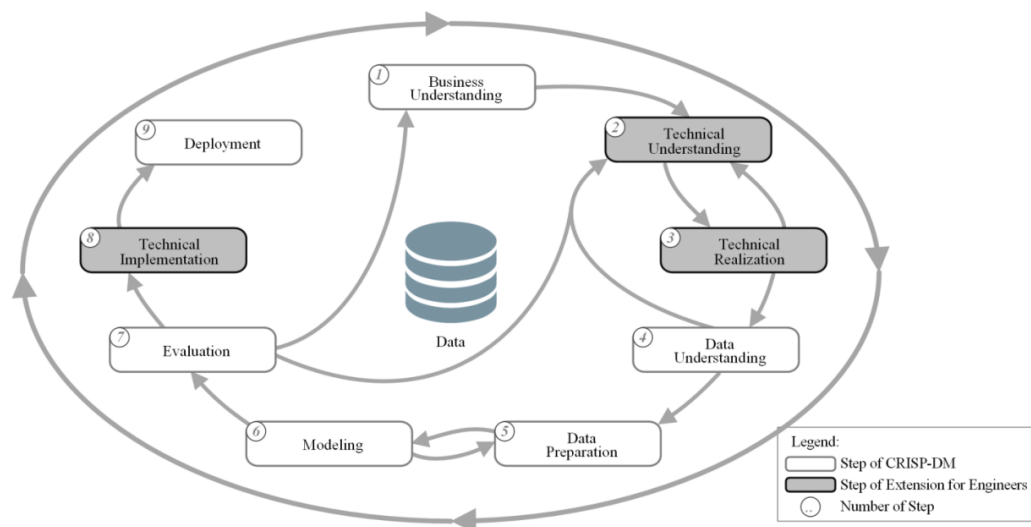


Abbildung 2: DMME-Prozess (Huber et al. 2019, S. 406)

### 2.2.3. Alternative Vorgehensmodelle

Alternativ zum CRISP-DM-Prozessmodell existieren in der Fachliteratur weitere Vorschläge, die Parallelen zu einzelnen Aspekten des CRISP-DM-Modells aufweisen oder

gänzlich neue Denkansätze integrieren. Beispielhaft sollen dazu drei alternative Vorgehensmodelle vorgestellt werden:

- **„Machine Learning Workflow“** von Amershi et al. (2019): Im Unterschied zum CRISP-DM werden, gestalten sich die Anfangs- und Endaktivitäten im 9-stufigen Prozessmodell anders. Während zu Beginn eine Phase des „Model Requirements“ ergänzt wird, endet das Modell mit einer kontinuierlichen „Model Monitoring“-Phase.
- **„Process Model for AI Systems Engineering“** vom Kompetenzzentrum für KI-Engineering Karlsruhe (CC-KING 2021): Charakteristisch für das PAISE-Modell ist die Zerlegung des Gesamtsystems in einzelne Systemkomponenten und das Anstoßen eines zyklischen Durchlaufens von Komponentenentwicklung, Bewertungen (sog. Checkpoints) und Verfeinerungen, um den Reifegrad aller Komponenten und damit des Gesamtsystems kontinuierlich zu erhöhen.
- **„Menschenzentrierte KI-Anwendungen in der Produktion“** von Pokorni et al. (2021): Der Leitfaden legt den Betrachtungsfokus im Unterschied zu anderen Vorgehensmodellen nicht nur auf technologische Herausforderungen, sondern auch auf organisatorische und menschliche Fragestellungen für die Implementierung von künstlicher Intelligenz. Aufgrund des erweiterten Betrachtungsfeldes werden neue Aspekte, wie bspw. die Integration von Veränderungsprozessen, im Leitfaden berücksichtigt.

### 3. Ableitung von Prinzipien für eine „Konviviale KI“

Aus den Ausführungen von Ivan Illich lassen sich allgemeine Prinzipien ableiten, die auf Technologien angewendet und insbesondere für die Entwicklung und Gestaltung einer künstlichen Intelligenz von Bedeutung sein können. Dieses Kapitel befasst sich mit den Implikationen dieser Prinzipien zur Entwicklung wünschenswerter KI-Systeme.

#### 3.1. Anwendungs- / nutzenzentrierte Vorgehensweise

Dem Gedanken der Konvivialität folgend, sollte KI nicht ungeprüft für jede denkbare Anwendung genutzt werden, sondern maßgeschneidert auf die jeweiligen Unternehmenssituationen angepasst eingesetzt werden. Anstatt einer „blinden“ Einführung von KI, sind daher zunächst die konkreten Bedarfe eines Unternehmens zu ermitteln. Entsprechend des festgestellten Bedarfs sollte sich die Schaffung von „Infrastrukturen“ anschließen, was in Form neuer Denkweisen und kultureller Organisationsanpassungen, optimierter Geschäftsprozesse, aber auch benötigter Kommunikations-, Datenerhebungs- und Datenerfassungstools in den Unternehmen erfolgen kann. Erst im Anschluss sollte die Implementierung des unternehmensindividuell zugeschnittenen KI-Systems erfolgen.

#### 3.2. Betonung der Mensch-Technik-Interaktion und selektive (Voll-) Automatisierung

Nach Illich ist die Forderung nach Konvivialität eng verknüpft mit der Art der Interaktion zwischen dem Menschen und seinem „Werkzeug“. Hierbei lehnt er das pauschale Streben nach möglichst hoher Automatisierung (und damit erwarteter Optimierung) von Geschäftsprozessen ab: „Nicht Werkzeuge, die ihnen die Arbeit abnehmen, brauchen die Menschen, sondern neue Werkzeuge, mit denen sie arbeiten können. Nicht weitere gut programmierte Energiesklaven brauchen sie, sondern eine Technologie, die ihnen dabei hilft, das Beste zu machen aus der Kraft und Phantasie, die jeder besitzt“ (Illich 1998, S. 27). Illich begründet diese Sichtweise damit, dass Menschen nicht nur Befriedigung, sondern Freude empfinden würden, wenn sie schöpferisch tätig sein können, was dem Gedanken einer kompletten Substitution menschlicher Arbeit entgegensteht (Illich 1998, S. 41). Die Gefahr, dass ursprüngliche Entscheidungsträger zu Konsumenten rein maschinell getroffener Absatz- und Produktionsentscheidungen werden und menschliches Prozess-Know-How zunehmend verloren geht, ist mit Blick auf den KI-Einsatz und den dadurch gegebenen Automatisierungsmöglichkeiten hochrelevant. Eine konviviale KI sollte daher vornehmlich als unterstützendes Medium entwickelt werden, das Entscheidungsgrundlagen für Informationsbedarfsträger schafft und die Kreativität und Produktivität dieser bereichert. Folglich stellt die Berücksichtigung entsprechender Kollaborations- bzw. Einwirkungsmöglichkeiten durch den Menschen einen wesentlichen Entwicklungsschwerpunkt konvivialer KI dar.

Die Forderung danach, ein KI-System als unterstützendes Medium zur Schaffung von Entscheidungsgrundlagen für Informationsbedarfsträger zu entwickeln, wirft die Frage auf, ob eine vollständige Automatisierung von Geschäftsprozessen mittels KI pauschal abzulehnen ist. Eine derartige Herangehensweise könnte im Hinblick auf Phänomene, wie einem aufkommenden Fachkräftemangel, fragwürdig erscheinen.

An dieser Stelle sei darauf hingewiesen, dass Illich eine konviviale Gesellschaft nicht durch die ausschließliche Anwendung von konvivialen Werkzeugen charakterisiert,

sondern durch einen angemessenen Mix zwischen konvivialen und nicht-konvivialen Werkzeugen (Illich 1998, S. 44–45). Es sollte somit unterschieden werden, in welchen Fällen sich ein KI-Modell als unterstützendes Entscheidungstool eignet und die Kreativität, sowie die Entscheidungsfähigkeit des Menschen durch ein Mensch-Technologie-Zusammenspiel verbessert und in welchen Fällen eine vollständige Automatisierung angebracht ist. Auch Ernst Schumacher betont in diesem Zuge die Notwendigkeit einer Fallunterscheidung und kritisiert, dass hochindustrielle Gesellschaften dazu neigen, die falschen Prozesse zu automatisieren: „Wir können daher sagen, dass moderne Technologie dem Menschen die Art von Arbeit genommen hat, die er am liebsten tut, nützliche, schöpferische Arbeit mit Händen und Kopf, und ihm viele arbeitsteilige Aufgaben gegeben hat, die ihm zum größten Teil keine Freude machen“ (Schumacher 2013, S. 159).

Als Konsequenz dieser Feststellung sollte sich eine vollständige Automatisierung mittels KI auf jene Prozesse beziehen, die kein Potential zur schöpferisch-kreativen und unabhängigen Arbeit des Menschen bieten. Dies trifft insbesondere auf repetitive, kleinteilige und routinemäßigen Aufgabengebiete zu.

### 3.3. Reduzierung der technologischen Komplexität von KI (Erklärbarkeit)

KI gilt als komplexe Technologie, wobei die Entwicklung und die Funktionsweise von KI-Modellen für Teile der Gesellschaft und der Unternehmenswelt unverständlich sind. Die Konsequenz ist, dass die mit dem Einsatz von KI einhergehenden Vorteile nicht genutzt werden, wenn das entsprechende Expertenwissen unternehmensintern nicht gegeben ist und eine KI-Nutzung daher abgelehnt wird (Hansen und Bogh 2020, S. 8–9).

Ein Vorgehensmodell zur Entwicklung konvivialer KI-Modelle sollte es sich daher zur wesentlichen Zielsetzung machen, die unbestrittene Komplexität von KI auf die individuellen Unternehmensbedürfnisse herunterzubrechen und zu einem nahbaren Themenfeld werden zu lassen. In diesem Sinne sind Schnittmengen zwischen der Entwicklung konvivialer KI und dem Grundgedanken erklärbarer KI („Explainable AI“) erkennbar, da ein als „Black Box“ agierendes KI-System abgelehnt wird. Über den Anspruch erklärbarer KI hinausgehend, sollen KI-Algorithmen nicht nur für Theoretiker und Experten, sondern auch für „Praktiker“ nachvollziehbar gestaltet werden. Ein Vorgehensmodell zur Entwicklung konvivialer KI-Modelle sollte folglich vorsehen, dass Anwender ein fundiertes Technologie- und Modellverständnis aufbauen können und Vertrauen in die erzielten Modell-Ergebnisse gewinnen.

Durch den erhöhten Anspruch an die Erklärbarkeit von KI-Modellen kann zusätzlich erreicht werden, dass Anwender bzw. Entscheidungsträger ihre Kenntnisse über prozessuale Zusammenhänge erweitern und aus vergangenen Störereignissen selbstständig Lerneffekte generieren.

## 4. Erstellung eines Konvivialen KI-Modells

Um die in Kapitel 3 dargestellten Leitprinzipien für den Einsatz in KI-Entwicklungsprojekten zu operationalisieren, wird nachfolgend ein Vorgehensmodell vorgestellt, das die aus Sicht der Autoren erforderlichen Aktivitäten zur Entwicklung eines konvivialen KI-Systems berücksichtigt (siehe Abb. 3).

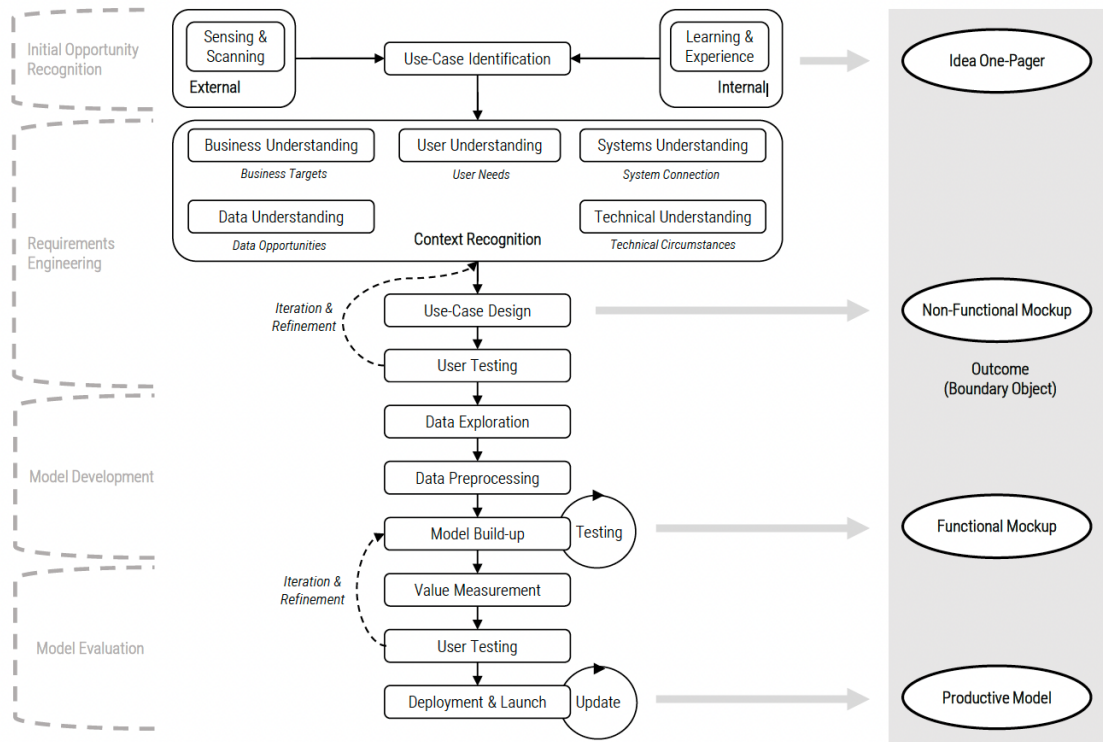


Abbildung 3: Konviviales KI-Vorgehensmodell (eigene Darstellung)

### 4.1. Phase: Initial Opportunity Recognition

Die „Initial Opportunity Recognition“-Phase hat die Identifikation bestehender Geschäftsprozesse und Anwendungsfälle zum Gegenstand, die aufgrund der Prozess- bzw. Anwendungsfallcharakteristika eine grundsätzliche Eignung für den KI-Einsatz aufweisen. Diese kann als gegeben betrachtet werden, wenn bspw. die Möglichkeit besteht, regelmäßig quantitative Prozessdaten erheben zu können und die Prozessleistung mithilfe von datenbasierten Auswertungsmethoden verbessern zu können.

Die Anwendungsfallidentifikation kann sowohl interne als auch externe Treiber zum Ursprung haben. Externe Treiber kommen durch effektive Marktbeobachtungen zur Anwendung und werden durch eine ausgeprägte „Sensing“-Fähigkeit des Unternehmens gefördert. Unter „Sensing“ lässt sich die Fähigkeit verstehen, Chancen, Risiken und Trends im Unternehmensumfeld zu erkennen (Teece 2007, S. 1322–1326; Baškarada und Koronios 2018, S. 337). Interne Treiber für eine Anwendungsfallidentifikation liegen vor, wenn Änderungsbedarfe und Potentiale durch Mitglieder des eigenen Unternehmens erkannt werden.

Die wesentliche Aufgabe der Initial Opportunity Recognition-Phase besteht somit in der Beantwortung der Frage, welche Anwendungsfälle sich grundsätzlich für einen KI-

Einsatz eignen. Die Phase wird durch die Erstellung eines „Idea One-Pager“ abgeschlossen, der einen ersten Prozessentwurf unter Einbindung von KI beinhaltet.

## 4.2.Phase: Requirements Engineering

Die „Requirements Engineering“-Phase startet mit der „**Context Recognition**“, die eine eingehende Prüfung des identifizierten Anwendungsfalls umfasst, um die wesentlichen Prozessanforderungen und Zielgrößen zu identifizieren und diese durch den geplanten KI-Einsatz mit einer höheren Wirksamkeit erreichen zu können.

Die Identifikation der Prozessanforderungen vollzieht sich dabei auf fünf Ebenen:

- Auf Ebene des **Business Understanding** werden die adressierten Geschäftsziele ermittelt. Die Projektbeteiligten werden auf diesem Weg befähigt, ein tieferes Verständnis der zugrundeliegenden Unternehmensabläufe zu entwickeln.
- Die Ebene des **Data Understanding** hat die Analyse der Datengrundlage zum Gegenstand. Die Projektbeteiligten entwickeln ein Verständnis für den vorliegenden Datenbedarf und analysieren die Möglichkeiten der internen und externen Datenerhebung, der Datenspeicherung und des Schnittstellen-Managements.
- In der Ebene des **User Understanding** wird ein Verständnis für die Bedarfe des Anwenders entwickelt. Im Vordergrund steht die Frage, wie der Endanwender durch den KI-Einsatz in seiner Kreativität und Produktivität gefördert und bei der Erfüllung der adressierten Geschäftsziele unterstützt werden kann.
- Die Ebene des **Technical Understanding** umfasst die Frage, welcher technischer Sachverhalt dem Anwendungsfall zugrunde liegt. Ziel ist die Identifikation relevanter technischer Einflussgrößen.
- Schließlich wird auf Ebene des **Systems Understanding** die bestehenden System- und Datenstrukturen analysiert. Im Vordergrund steht die Analyse der Datenherkunft, sowie das Schnittstellen-Management.

Die Untersuchung der einzelnen Ebenen dient dem Aufbau eines grundlegenden Prozessverständnisses. Insbesondere die Ebenen des „Data Understanding“ und „Systems Understanding“ werden im Rahmen der späteren „Model Development“-Phase einer detaillierteren Betrachtung unterzogen.

Nach Abschluss der Anwendungsfall-Analyse, startet die Aktivität des „**Use-Case Design**“, das zum Ziel hat, den KI-Einsatz unter allen Projektbeteiligten zu planen. Um der Forderung nach einer konvivialen Interaktion von menschlicher Arbeitskraft und KI-Modell nachzukommen, gelten die anvisierte Stufe der KI-Entscheidungsunterstützung, die Abstimmung der Mensch-Maschinen-Schnittstelle, sowie die Gestaltung der Benutzeroberfläche als Kernelemente in der Diskussion um die Neugestaltung des Anwendungsfalls. Da die detaillierte Auseinandersetzung mit dem identifizierten Anwendungsfall das Aufzeigen von bisher unerkannten Prozess-Schwachstellen zur Folge haben kann, empfiehlt es sich zudem, die „Business Process Optimization“ als integrativen Bestandteil des „Use-Case Designs“ zu betrachten, um eine „blinde“ Übernahme prozessualer Schwachstellen zu verhindern.

Das „Use-Case-Design“ mündet schließlich im „Non-Functional Mockup“. Eine strukturierte Vorgehensweise zum Aufbau des Mockups ist für die gemeinsame Abstimmung zwischen den Projektbeteiligten von hohem Nutzen, was durch einen Rückgriff auf entsprechende Vorlagen aus der Literatur gewährleistet werden kann (siehe z.B. Haag 2022). Um sicherzustellen, dass der neudefinierte Prozess den Vorstellungen des

Endanwenders entspricht, endet die „Requirements Engineering“-Phase mit dem „User Testing“.

### 4.3. Phase: Model Development

Die „Model Development“-Phase teilt sich in die Aktivitäten der „Data Exploration“, des „Data Preprocessing“ und des „Model Build-Up“. Diese weisen Ähnlichkeiten zu bestehenden KI-Vorgehensmodellen auf (siehe Kapitel 2.2.) und umfassen u.a. das Überprüfen der verfügbaren Datenlage, die Datenauswahl und -bereinigung sowie die Planung der Datenauswertung und der zu verwendenden Algorithmen. Als Hilfestellung kann der Rückgriff auf die initiale Analyse der Datengrundlage im Rahmen des „Data Understanding“ dienen.

Im Zuge des „Model Build-Up“ trägt das möglichst frühzeitige Erstellen eines Modell-Prototypen bzw. eines Minimum Viable Product (MVP) dazu bei, die spätere Leistungsfähigkeit und Funktionsweise des Modells für den Endanwender zu veranschaulichen und ggf. gewünschte Änderungen frühzeitig zu implementieren (Ahmed et al. 2018). Sobald das entwickelte Modell einen Stand erreicht, der die wesentlichen Anforderungen des Non-Functional Mockup erfüllt, gilt dieser als „Functional Mockup“ und kann in der abschließenden „Model-Evaluation“-Phase einer ausführlichen Bewertung unterzogen werden.

### 4.4. Phase: Model Evaluation

Die „Model Evaluation“-Phase beinhaltet den Abgleich der erreichten Prozessverbesserungen mit den anfangs definierten Zielgrößen. Hierbei gelten insbesondere die konvivialen Zielwerte als Ausschlusskriterien für die Validierung des entwickelten Modells. Kann das entwickelte Modell die Zielwerte und die Erwartungen des Endanwenders erfüllen, erfolgt die Einführung des Modells in den laufenden Betrieb.

### 4.5. Outcomes / Boundary Objects

Im präsentierten Vorgehensmodell kommen den sog. „Outcomes“ eine entscheidende Rolle zu. Outcomes sind Elemente, die das Ergebnis einer bestimmten Projektstufe darstellen und als „Boundary Objects“ betrachtet werden können, da sie über alle Abteilungen hinweg als Diskussionsgrundlage und als verbindendes Element für eine gemeinsam gesprochene Sprache der Projektbeteiligten dienen.

In diesem Sinne können die Outcomes zunächst zur Abstimmung bestimmter Inhalte dienen, darüber hinaus jedoch auch als „Checkpoints“ bzw. Entscheidungspunkte betrachtet werden, um die Erfüllung festgelegter Kriterien sicherzustellen.

Die jeweiligen Outcomes sind auf unterschiedlichen Projektstufen angesiedelt und können zur Klärung inhaltlicher Abstimmungsbedarfe und zur Prüfung festgelegter Checkpoint-Kriterien dienen. Die erforderlichen Inhalte der Outcomes sowie die notwendigerweise zu erfüllenden Checkpoint-Kriterien sind in Tabelle 1 aufgelistet. Diese können situativ in Abhängigkeit der jeweiligen Unternehmensbedarfe erweitert werden.

| Outcome                      | Inhaltliche Abstimmungsbedarfe  | Checkpoint-Kriterien  |
|------------------------------|---|---|
| <b>Idea One-Pager</b>        | <ul style="list-style-type: none"> <li>- Fixierung auf einen Anwendungsfall</li> <li>- Definition des Kernproblems</li> <li>- Definition der Zielsetzung</li> <li>- Bestimmung der Projektbeteiligten und Festlegung ihrer jeweiligen Einbindung</li> </ul>   | <ul style="list-style-type: none"> <li>- Relevanz des Anwendungsfalls für Unternehmen und Anwender</li> <li>- Vorhandensein von datenbasierten Optimierungspotentialen</li> </ul>   |
| <b>Non-Functional Mockup</b> | <ul style="list-style-type: none"> <li>- Festlegung von modellbezogenen Zielkriterien (wirtschaftlich, konvivial)</li> <li>- Definition des Soll-Prozesses inkl. anvisierter Unterstützungsstufe</li> <li>- Festlegung der Mensch-Modell-Schnittstelle und Benutzeroberfläche</li> <li>- Festlegung der erforderlichen Datengrundlage und Infrastruktur</li> <li>- Bewertung erforderlicher Anwender-Kompetenzen und Festlegung von Qualifizierungsmaßnahmen</li> <li>- Festlegung von Projektzeitraum und -budget</li> </ul> | <ul style="list-style-type: none"> <li>- Wirtschaftliche Machbarkeit des Vorhabens</li> <li>- Technologische Machbarkeit des Vorhabens (aus Datensicht)</li> <li>- Kompatibilität zu bereits vorhandenen Systemen</li> <li>- Möglichkeit der Verbesserung und Berücksichtigung konvivialer Aspekte</li> </ul> |
| <b>Functional Mockup</b>     | <ul style="list-style-type: none"> <li>- Festlegung der akzeptablen Fehlertoleranz bei erzielten Modell-Ergebnissen</li> </ul>  | <ul style="list-style-type: none"> <li>- Anwendbarkeit bzw. nutzerzentrierte Modell-Gestaltung</li> <li>- Einhaltung der Modell-Zielkriterien</li> <li>- Erzielung der anvisierten Unterstützungsstufe</li> </ul>   |
| <b>Productive Model</b>      | <ul style="list-style-type: none"> <li>- Erfahrungsaustausch für zukünftige Projekte</li> <li>- Planung von modellbezogenen Follow-Up-Aktivitäten (u.a. kontinuierliches Monitoring)</li> <li>- Planung einer Mitarbeiter-Integration in die Identifikation weiterer Anwendungsbereiche mit datenbasiertem Optimierungspotential</li> </ul>   | <ul style="list-style-type: none"> <li>- Übereinstimmung von Modell-Komplexität und Anwender-Kompetenzen</li> <li>- Genauigkeit der erzielten Modell-Ergebnisse</li> </ul>  |

Tabelle 1: Inhalte und Checkpoint-Kriterien der einzelnen Outcomes



## 5. Zusammenfassung

Das Ziel der vorliegenden Ausarbeitung war es, ein Vorgehensmodell zur Entwicklung von KI-Modellen unter Berücksichtigung konvivialer Elemente zu erstellen. Um dies zu erreichen, konnten die Ausarbeitungen von Ivan Illich genutzt werden, um grundsätzliche Prinzipien konvivial konzipierter Werkzeuge zu definieren und diese im Anschluss auf drei Anforderungen einer konvivialen KI zu konkretisieren: Die Betonung der Anwendungszentrierung und der Mensch-Technik-Interaktion sowie die Reduzierung der technologischen Komplexität von KI im Sinne der Erklärbarkeit und Nachvollziehbarkeit von erzielten Modell-Ergebnissen.

Das präsentierte Vorgehensmodell legt einen starken Fokus auf die Berücksichtigung der genannten Prinzipien: Während der Aspekt der Anwendungszentrierung durch die initiale Anwendungsfallidentifikation und -analyse betont wird, erfüllt das gemeinsam erstellte Use-Case Design den Anspruch an eine intensive Planung der Mensch-Technik-Interaktion. Die in gemeinsamer Absprache aller Projektbeteiligten erstellten Outcomes sorgen für ein einheitliches und frühzeitiges Verständnis der grundsätzlichen Projektausrichtung und der späteren Modell-Leistungsfähigkeit, wobei dieses Vorhaben durch das Erstellen von Modell-Prototypen in der frühen Model Development-Phase zusätzlich unterstützt wird. Bei alledem stehen die Anforderungen des Endanwenders durch die mehrfach integrierte Aktivität des „User Testings“ im Mittelpunkt der Modellentwicklung.

Im Vorgehensmodell der vorliegenden Ausarbeitung werden somit Aspekte in den Vordergrund gerückt, die in den bisher verfügbaren Vorgehensmodellen der KI-Landschaft kaum Berücksichtigung finden. Seine Anwendung kann die gezielte Identifikation geeigneter Anwendungsfälle fördern und die Planung eines KI-Einsatzes aus Unternehmens- und Anwendersicht optimieren.

## Literaturverzeichnis

- Ahmed, Bakhtiyar; Dannhauser, Thomas; Philip, Nada (2018): A Lean Design Thinking Methodology (LDTM) for Machine Learning and Modern Data Projects. In: 2018 10th Computer Science and Electronic Engineering (CEECE). 2018 10th Computer Science and Electronic Engineering (CEECE). Colchester, United Kingdom, 19.09.2018 - 21.09.2018: IEEE, S. 11–14.
- Amershi, Saleema; Begel, Andrew; Bird, Christian; DeLine Robert; Gall, Harald; Kamar, Ece et al. (Hg.) (2019): Software Engineering for Machine Learning: A Case Study. 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). Montreal, QC, Kanada, 25.-31. Mai 2019.
- Anders, Günther (1956): [The Outdateness of Human Beings] Die Antiquiertheit des Menschen. CH Beck München.
- Başkarada, Saša; Koronios, Andy (2018): The 5S organizational agility framework: a dynamic capabilities perspective. In: *IJOA* 26 (2), S. 331–342. DOI: 10.1108/IJOA-05-2017-1163.
- CB Insights (2018): Top artificial intelligence trends to watch in 2018. Online verfügbar unter <https://www.cbinsights.com/research/report/artificial-intelligence-trends-2018/>.
- CC-KING (2021): PAISE (Process Model for AI Systems Engineering). Das Vorgehensmodell für KI-Engineering.
- Ellul, Jacques (1964): The Technological Society. A penetrating analysis of our technical civilization and of the effect of an increasingly standardized culture on the future of man: Alfred A. Knopf, Inc. and RandomHouse, Inc.
- Forschung & Lehre (2023): Experten warnen vor Risiken durch Künstliche Intelligenz. Online verfügbar unter <https://www.forschung-und-lehre.de/zeitfragen/experten-warnen-vor-risiken-von-kuenstlicher-intelligenz-5658>.
- Fügener, Andreas; Grahl, Jörn; Gupta, Alok; Ketter, Wolfgang (2021): Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI. In: *Management Information Systems Quarterly (MISQ)* 45 (3), S. 1527–1556.
- Gentsch, Peter (2018): Künstliche Intelligenz für Sales, Marketing und Service: Springer Fachmedien Wiesbaden.
- Gonçalves, Carina; Ferreira, Diana; Neto, Cristiana; Abelha, António; Machado, José (2020): Prediction of Mental Illness Associated with Unemployment Using Data Mining. In: *Procedia Computer Science* 177, S. 556–561. DOI: 10.1016/j.procs.2020.10.078.
- Granados, Nelson (2022): Human Borgs: How Artificial Intelligence Can Kill Creativity And Make Us Dumber. Forbes.
- Haag, Christoph (Hg.) (2022): Value Proposition of AI-based Innovation – A Utility-centred Framework. The XXXIII ISPIM Innovation Conference "Innovating in a Digital World". Copenhagen, 05.-08.06.2022.
- Hansen, Emil Blixt; Bogh, Simon (2020): Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. In: *Journal of Manufacturing Systems*, S. 1–11.
- Hine, H. J.; Mumford, Lewis (1968): The Myth of the Machine: Technics and Human Development. In: *Man* 3 (3), S. 519. DOI: 10.2307/2798943.

- Huber, Steffen; Wiemer, Hajo; Schneider, Dorothea; Ihlenfeldt, Steffen (2019): DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model. In: *Procedia CIRP* 79, S. 403–408. DOI: 10.1016/j.procir.2019.02.106.
- Illich, Ivan (1973): *Tools for conviviality*. London: Calder and Boyars (Open forum).
- Illich, Ivan (1998): *Selbsbegrenzung. Eine politische Kritik der Technik*. 1. Auflage. München: Beck (Beck'sche Reihe [Formerly: Beck'sche Schwarze Reihe], 1167).
- Leggewie, Claus; Zifonun, Darius; Lang, Anne; Siepmann, Marcel; Hoppen, Johanna (2012): *Schlüsselwerke der Kulturwissenschaften*. Bielefeld: Transcript (Edition Kulturwissenschaft, Bd. 7).
- Linardatos, Pantelis; Papastefanopoulos, Vasilis; Kotsiantis, Sotiris (2020): Explainable AI: A Review of Machine Learning Interpretability Methods. In: *Entropy (Basel, Switzerland)* 23 (1). DOI: 10.3390/e23010018.
- Martinez-Plumed, Fernando; Contreras-Ochando, Lidia; Ferri, Cesar; Hernandez-Orallo, Jose; Kull, Meelis; Lachiche, Nicolas et al. (2021): CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. In: *IEEE Trans. Knowl. Data Eng.* 33 (8), S. 3048–3061. DOI: 10.1109/TKDE.2019.2962680.
- Maulana, Yudhi Dwi Fajar; Ruldeviyani, Yova; Sensuse, Dana Indra (2020): Data Mining Classification Approach to Predict The Duration of Contraceptive Use. In: *IEEE Xplore*, 2020.
- Nagashima, Hiroko; Kato, Yuka (2019): APREP-DM: a Framework for Automating the Pre-Processing of a Sensor Data Analysis based on CRISP-DM. In: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). Kyoto, Japan, 11.03.2019 - 15.03.2019: IEEE, S. 555–560.
- News Europäisches Parlament (2023): MEPs ready to negotiate first-ever rules for safe and transparent AI. Online verfügbar unter <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai>.
- Paech, Niko (2013): *Befreiung vom Überfluss. Auf dem Weg in die Postwachstumsökonomie*. 4e dr. München: Oekom.
- Pokorni, Bastian; Braun, Martin; Knecht, Christian (2021): *Menschzentrierte KI-Anwendungen in der Produktion. Praxiserfahrungen und Leitfaden zu betrieblichen Einführungsstrategien*. Fraunhofer IAO.
- Schröer, Christoph; Kruse, Felix; Gómez, Jorge Marx (2021): A Systematic Literature Review on Applying CRISP-DM Process Model. In: *Procedia Computer Science* 181, S. 526–534. DOI: 10.1016/j.procs.2021.01.199.
- Schumacher, Ernst Friedrich (2013): *Small is beautiful. Die Rückkehr zum menschlichen Maß*. Neuaufl. München: Oekom.
- Studer, Stefan; Bui, Thanh Binh; Drescher, Christian; Hanuschkin, Alexander; Winkler, Ludwig; Peters, Steven; Mueller, Klaus-Robert (2020): Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology. Online verfügbar unter <http://arxiv.org/pdf/2003.05155v2>.

Teece, David J. (2007): Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. In: *Strat. Mgmt. J.* 28 (13), S. 1319–1350. DOI: 10.1002/smj.640.

Vetter, Andrea (2018): The Matrix of Convivial Technology – Assessing technologies for degrowth. In: *Journal of Cleaner Production* 197, S. 1778–1786. DOI: 10.1016/j.jclepro.2017.02.195.

---

## Hinweise / Notes

Diese Veröffentlichungen erscheinen im Rahmen der Schriftenreihe "C|plus". Alle Veröffentlichungen dieser Reihe können unter <https://cos.bibl.th-koeln.de/home> abgerufen werden.

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Datum der Veröffentlichung: 12.9.2023

## Herausgeber / Editorship

Prof. Dr. Thomas Bartz-Beielstein,

Prof. Dr. Wolfgang Konen,

Prof. Dr. Boris Naujoks,

Faculty of Computer Science and Engineering Science,

TH Köln,

Steinmüllerallee 1,

51643 Gummersbach

## Schriftleitung und Ansprechpartner / Contact editor's office

Prof. Dr. Thomas Bartz-Beielstein,

Faculty of Computer Science and Engineering Science,

TH Köln,

Steinmüllerallee 1,

51643 Gummersbach

phone: +49 2261 8196 6391

url: <http://www.th-koeln.de/ki>

eMail: [thomas.bartz-beielstein@th-koeln.de](mailto:thomas.bartz-beielstein@th-koeln.de)

ISSN (online) 2194-2870

<https://doi.org/10.57684/COS-1158>

<https://nbn-resolving.org/html/urn:nbn:de:hbz:832-cos4-11587>

---

Technology  
Arts Sciences  
**TH Köln**