

Künstliche Intelligenz als System

Welche Perspektive zeigt die Systemtheorie auf?

erste Ausgabe, ver. 1.01 (de)

30.12.2022

Igor Furgel
(office@furgel.com)

In der vorliegenden Abhandlung möchten wir den systemtheoretischen Ansatz auf Systeme mit ‚künstlicher Intelligenz‘ (KI) anwenden. Ein Ziel, das wir dabei verfolgen, ist, die grundsätzlichen Unterscheidungskriterien zwischen verschiedenen Arten von KI-Systemen festzustellen.

Wir möchten auch der Frage nachgehen, welche grundlegenden Voraussetzungen aus systemtheoretischer Sicht ein autonomes technisches System erfüllen muss, damit es mit den gleichen (oder stärkeren und umfassenderen) intellektuellen und kreativen Fähigkeiten wie ein Mensch universell einsetzbar wäre.

Diese Betrachtung wird uns helfen, den Platz der verschiedenen Arten von Systemen der Künstlichen Intelligenz im Kontext von ‚das Unbelebte – das Belebte – der Mensch‘ zu verstehen.

Die vorliegende Abhandlung könnte die Aufmerksamkeit eines Zielpublikums auf sich ziehen, das sich sowohl für Fragen der künstlichen Intelligenz und ihrer Einordnung in den Kontext ‚das Unbelebte – das Belebte – der Mensch‘ als auch für den Systemansatz im Allgemeinen interessiert.

Das Original der vorliegenden ersten Ausgabe dieses Aufsatzes (V. 1.00 (de) wurde am 30.12.2022 veröffentlicht,
Deutsche Nationalbibliothek,
<https://d-nb.info/1276877072/>.

The first English edition (V. 1.00 (en) – „Artificial Intelligence as a System: What Perspective does Systems Theory Reveal?“) was published on 30.12.2022, Deutsche Nationalbibliothek,
<https://d-nb.info/1276876920/>.

Первое издание этой работы на русском (V. 1.00 (ru) – „Искусственный интеллект как система: Какую перспективу открывает теория систем?“) опубликовано 30.12.2022, Deutsche Nationalbibliothek,
<https://d-nb.info/1276877293/>.

Die wichtigsten Ideen dieser Arbeit, wie sie auf KI angewendet werden, entstanden zwischen Juni und Dezember 2022.

Inhaltsverzeichnis

1	Einleitung	5
2	Regelbasierte KI (Expertensysteme)	5
3	Schwache KI mit maschinellem Lernen	9
3.1	Lernsubsystem der technischen KI (TSKI-ML)	11
3.2	Lernsubsystem des ‚automatisierten Entscheidungssystems‘ (ADMS-ML)	15
3.3	Entscheidungssystem der technischen KI (TSKI-DM)	17
3.4	Entscheidungssystem des ‚automatisierten Entscheidungssystems‘ (ADMS-DM)	19
4	Starke KI	22
5	Enmorphie des ‚Selbstbewusstseins‘ von KI: Unterscheidungskriterien zwischen den KI-Arten und Ausblick	26
6	Glossar	29
7	Referenzen	36
8	Danksagung	36

1 Einleitung

In der vorliegenden Abhandlung möchten wir den in [5] entwickelten systemtheoretischen Ansatz auf Systeme mit ‚künstlicher Intelligenz‘ (KI) anwenden. Ein Ziel, das wir dabei verfolgen, ist, die grundsätzlichen Unterscheidungskriterien zwischen drei Arten von KI-Systemen festzustellen, nämlich zwischen

- *regelbasierten* Expertensystemen, die auf regelbasierter Programmierung beruhen und in den 1970er und 1980er Jahren entwickelt wurden,
- einer ‚Schwachen KI‘¹, die auf *maschinellern Lernen* basiert und lediglich für die Erledigung von spezifischen, eingeschränkten Aufgaben einsetzbar ist, und
- einer ‚Starken KI‘², die mit gleichen (oder stärkeren und umfangreicheren) intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre; die ‚Starke KI‘ ist aktuell lediglich eine Vision.

Wir möchten auch der Frage nachgehen, welche grundlegenden Voraussetzungen aus systemtheoretischer Sicht ein autonomes technisches System erfüllen muss, damit es mit den gleichen (oder stärkeren und umfassenderen) intellektuellen und kreativen Fähigkeiten wie ein Mensch universell einsetzbar wäre.

Diese Betrachtung wird uns helfen, den Platz der verschiedenen Arten von Systemen der Künstlichen Intelligenz im Kontext von ‚das Unbelebte – das Belebte – der Mensch‘ zu verstehen.

Diese Arbeit kann für sich allein gelesen werden. Die grundlegenden Begriffe sind in Kap. 6 „Glossar“ ausgeführt. Da jedoch die Ansätze, die wir in [5], Part A, CHAPTER I (insbesondere in Kap. 3 “Being, Existential Triads and Enmorphya”) und CHAPTER II entwickelt haben, für diese Abhandlung von grundlegender Bedeutung sind, empfehlen wir Lesern, die an den zugrunde liegenden Entwicklungen interessiert sind, auch [5] zu lesen.

2 Regelbasierte KI (Expertensysteme)

Für die Erstellung von *regelbasierten* Expertensystemen wird traditionelle, auch als regelbasierte bezeichnete Programmierung verwendet. Solche Expertensysteme verwenden festcodierten Wissensdatenbanken zu einem bestimmten Themengebiet sowie festcodierten Entscheidungsregeln. Wenn ein regelbasiertes Expertensystem eine Anfrage samt zu analysierender Daten bekommt, verwendet sie die festcodierten Entscheidungsregeln, um diese zu analysierenden Daten mit der Wissensdatenbank abzugleichen, und generiert eine Antwort (Entscheidung/Empfehlung), s. [4], Kap.1 und [3], Kap. 2.

Betrachten wir nun, wie die abstrakten Elemente der *existenziellen Triade* (siehe Glossar) in dieser Art von KI zum Ausdruck kommen.

Zu analysierende Daten/Anfragen sind das ‚Substrat‘ jedes *regelbasierten* Expertensystems. Die ‚Eigenschaft‘ des Systems sind die Beschaffenheit des regelbasierten Programms (inkl. der Operationalisierung/Messbarmachung des Zwecks des Programms, des Modells des

¹ EN: Narrow (weak) artificial intellect (AI), abgekürzt mit ANI; RU: слабый (узкий) искусственный интеллект (ИИ)

² EN: General (strong) AI, abgekürzt mit AGI; RU: сильный (общий) ИИ

Problems und des Algorithmus)³, welche in Programmieranweisungen (Instruktionen) inkl. der Wissensdatenbank festcodiert ist, sowie die Eigenschaften/Charakteristika der zu analysierenden Daten. Die ‚Relation‘ ist der Prozess der Anwendung der festcodierten Anweisungen auf die zu analysierenden Daten durch die Programmausführung im Rahmen der Verwendung des regelbasierten Expertensystems. Die zu analysierenden Daten ändern sich durch diese Anwendung der Anweisungen nicht.

Der Prozess der Anwendung der festcodierten Anweisungen auf die zu analysierenden Daten ist *deterministisch*, nicht *stochastisch*. Deswegen ist das Prinzip der Suffizienz der Existenziellen Triade⁴ hier nicht anwendbar: die *regelbasierten* Expertensysteme sind *deterministisch*.

Für *regelbasierte* Expertensysteme, repräsentieren umgesetzte *Prinzipien der Softwareentwicklung* inkl. der angewendeten Verfahrensregeln/-normen (also das ‚*Programmierhandbuch*‘) die ‚Information über die Steuerung der Relation‘ (die *Enmorphie der Relation*). Als Enmorphie der Relation zwischen dem Substrat (den zu analysierende Daten) und der Eigenschaft (der Beschaffenheit des regelbasierten Programms) definieren die umgesetzten Prinzipien der Softwareentwicklung (also das ‚*Programmierhandbuch*‘) den Charakter dieser Relation (Interaktion).

Wenn wir nun eine Parallele zu *stochastischen* Systemen ziehen, für welche im Allgemeinen das Prinzip der Suffizienz der Existenziellen Triade⁴ gilt, stellen wir fest, dass dieses Prinzip durch das ‚*Programmierhandbuch*‘ ersetzt wird, wenn es sich um ein *deterministisches* Expertensystem handelt.

Das ‚*Programmierhandbuch*‘ bestimmt den Charakter der Anwendung der festcodierten Anweisungen auf die zu analysierenden Daten. Das ‚*Programmierhandbuch*‘ prägt dabei verschiedene *Optimierungsfunktionen* wie z.B. das Minimieren des Stromverbrauchs oder der Ausführungszeit des Programms. Im Unterschied zu *stochastischen* Systemen im Allgemeinen, kann das ‚*Programmierhandbuch*‘ die zu analysierenden Daten nicht prägen, sondern lediglich die Beschaffenheit des regelbasierten Programms (inkl. der Operationalisierung/Messbarmachung des Zwecks des Programms, des Modells des Problems und des Algorithmus).

Veranschaulichen wir die Relation zwischen dem primären System und dem Metasystem am Beispiel von regelbasierten Expertensystemen:

³ Diese Begriffe sind in [2] ausführlich erklärt, siehe z.B. ‚Glossar‘ dort. Einer besseren Lesbarkeit halber geben wir hier lediglich den jeweiligen Kern der Definitionen wieder:

- „Unter einer Modellierung verstehe ich jegliche Form der Vereinfachung und Abstraktion einer Situation, die trotzdem noch so genau ist, dass sie Vorhersagen oder analytische Schlüsse über diese Situation zulässt.“
- „Eine Operationalisierung stellt die Messbarmachung eines (sozialen) Konzeptes dar – sie basiert immer auf einem Modell des Konzeptes.“
- „Ein Algorithmus ist eine für jede erfahrene Programmiererin und jeden erfahrenen Programmierer ausreichend detaillierte und systematische Handlungsanweisung, um ein mathematisches Problem zu lösen, sodass bei korrekter Implementierung (Übersetzung in Code) der Computer für jede korrekte Inputmenge den korrekten Output berechnet.“

⁴ [5], CHAPTER I, STM. 9 ‘The principle of sufficiency of the existential triad’ (Übersetzung des Autors): „Wenn ‚Relation‘ in einer existenziellen Triade grundsätzlich einen *stochastischen* Charakter hat und statistisch einem bestimmten Gesetz gehorcht, dann ist diese existenzielle Triade nicht nur notwendig, sondern auch hinreichend, um die Beobachtbarkeit und damit die Schaffung des Zustands des ‘Seins’ des Systems zu schaffen, das auf dieser existenziellen Triade beruht. Die Evolution dieses Systems wird dem Charakter der ‘Relation’ in der existenziellen Triade folgen.“

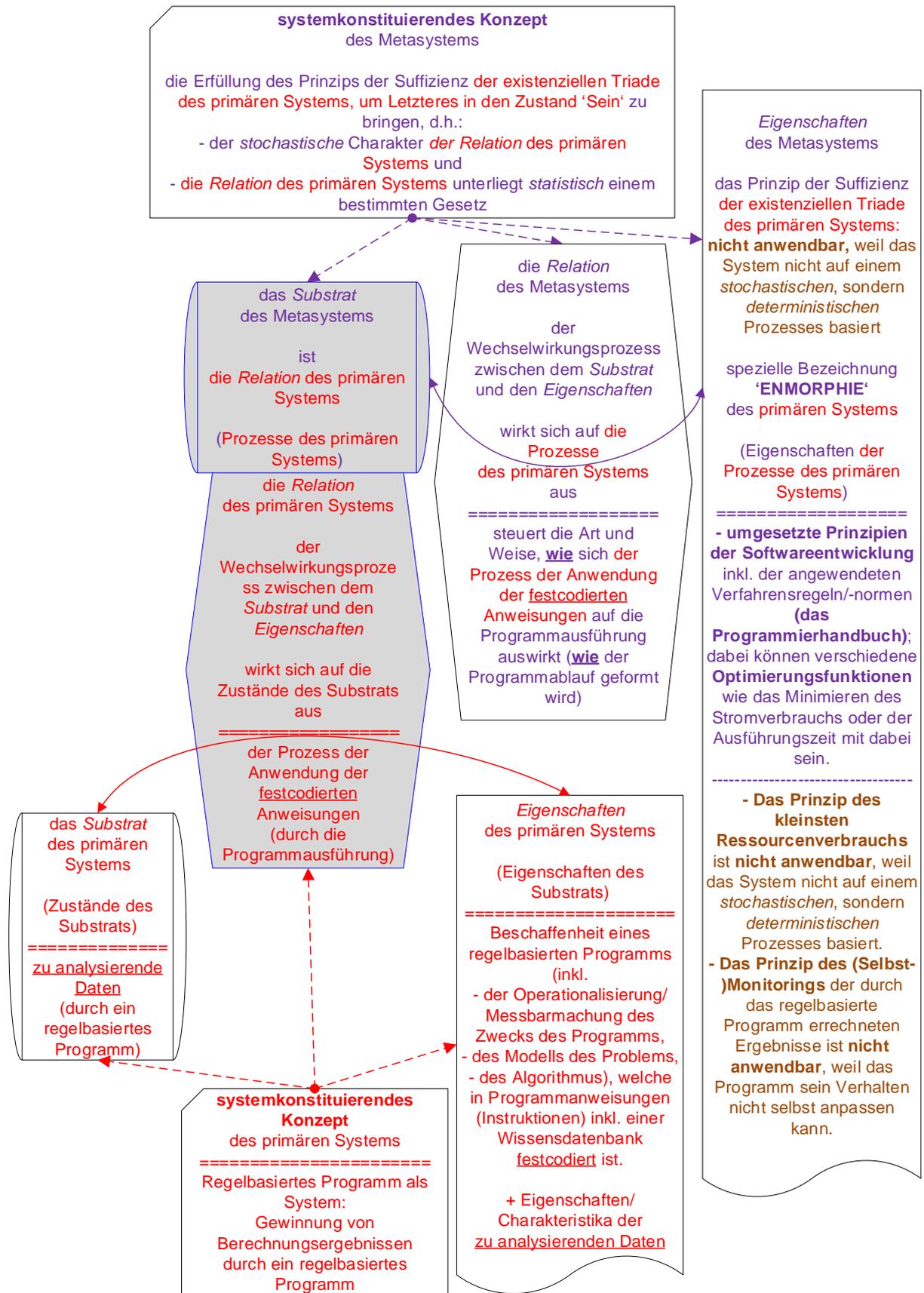


Abbildung 1: Relation zwischen dem primären System 'regelbasiertes Expertensystem' und dem entsprechenden Metasystem

Die systemtheoretische Betrachtung von *regelbasierten* Expertensystemen ermöglicht uns eine nützliche Verallgemeinerung des systemtheoretischen Unterschieds zwischen *deterministischen* und *stochastischen* Systemen:

Typ des Systems →	<i>stochastische</i> Systeme	<i>deterministische</i> Systeme (können ausschließlich als <u>Artefakte</u> existieren ⁵)
Parameter ↓		
das Prinzip der Suffizienz der Existenziellen Triade ⁴	gilt immer	nicht anwendbar
Enmorphie der Relation ('Information über die Steuerung der Relation')	ist eine systemspezifische Umsetzung des <i>Prinzips der Suffizienz der existenziellen Triade</i> , wobei das <i>Prinzip des kleinsten Ressourcenverbrauchs</i> <u>immer</u> ein Bestandteil der Enmorphie ist ⁶	ist die praktisch umgesetzten <i>Prinzipien der Herstellung</i> des jeweiligen Systems inkl. der angewendeten Verfahrensregeln/-normen; ist das <i>Pflichtenheft</i> (oder, bildlich gesprochen, das <i>„Programmierhandbuch“</i>) für die Herstellung des jeweiligen Systems.
Auswirkung der Enmorphie der Relation auf das primäre System	Die jeweilige systemspezifische Umsetzung des <i>Prinzips der Suffizienz der existenziellen Triade</i> bestimmt den Charakter der 'Relation' des primären Systems. Dadurch prägt die jeweilige systemspezifische Umsetzung des <i>Prinzips der Suffizienz der existenziellen Triade</i> <u>sowohl</u> die Zustände des ‚Substrats‘ <u>als auch</u> die ‚Eigenschaften‘ des primären Systems.	Die praktisch umgesetzten <i>Prinzipien der Herstellung</i> des jeweiligen Systems können Zustände des ‚Substrats‘ <u>nicht</u> prägen, sondern lediglich die ‚Eigenschaften‘ des primären Systems.
Auswirkung der 'Relation' des primären Systems auf	Der Interaktionsprozess der ‚Eigenschaften‘ des	Der Interaktionsprozess der ‚Eigenschaften‘ des

⁵ siehe [5], CHAPTER I, ch. 3.1 "Being and Existential Triads" or CHAPTER VII, ch. 2.1.3 "Indeterminacy and Action Quanta: Complementary Characters of the Past and the Future"

⁶ siehe [5], CHAPTER I, ch. 3.2 "Enmorphia", STM. 11 'the Principle of Least Resources Consumption' (Übersetzung des Autors): "**Das Prinzip des Kleinsten Ressourcenverbrauchs** ist eine 'Information über die Steuerung der Relation' (d. h. eine Enmorphie der Relation) und steuert nicht nur den Interaktionsprozess zwischen Materie und Information in der Natur, sondern auch zwischen dem *Substrat* und dem *Strukturfaktor* eines jeden Systems – physikalischen, sozialen, kommunikativen, etc. –, das auf einem *stochastischen* Prozesses basiert. "

Typ des Systems →	<i>stochastische</i> Systeme	<i>deterministische</i> Systeme (können ausschließlich als <u>Artefakte</u> existieren ⁵)
Parameter ↓		
Zustände des ‚Substrats‘	primären Systems mit dem ‚Substrat‘ des primären Systems <u>ändert</u> Zustände des ‚Substrats‘.	primären Systems mit dem ‚Substrat‘ des primären Systems kann Zustände des ‚Substrats‘ <u>nicht</u> ändern.

Das **Prinzip des kleinsten Ressourcenverbrauchs**, welches für alle *stochastischen* Systeme gilt, ist für ein *regelbasiertes* Expertensystem nicht anwendbar, weil das System nicht auf einem *stochastischen*, sondern *deterministischen* Prozesses basiert⁶.

Das **Prinzip der Selbsterhaltung des Systems**, welches für alle *quasi-stochastischen* Systeme gilt⁷, und welches sich für viele solche Systeme als das Prinzip eines (Selbst-)Monitorings ausprägt, ist ebenso nicht anwendbar, weil das System nicht auf einem *stochastischen*, sondern *deterministischen* Prozesses basiert, und das Programm deswegen sein Verhalten nicht selbst anpassen kann. Anhand eines durch ein *regelbasiertes* Expertensystem errechneten Ergebnisses kann das Expertensystem sein Verhalten nicht autonom anpassen.

3 Schwache KI mit maschinellem Lernen

Eine ‚Schwache KI‘ basiert – im Unterschied zu regelbasierten Expertensystemen – auf *maschinellm Lernen* (machine learning: ML). Eine ‚Schwache KI‘ ist lediglich für die Erledigung von spezifischen, eingeschränkten Aufgaben einsetzbar, s. [2], Glossar, [3], Kap. 1, [4], Kap. 2 („ANI“) oder im Internet.

Der Ansatz des *maschinellen Lernens* unterscheidet sich vom Ansatz der *regelbasierten* Programmierung technologisch darin, dass Entscheidungsregeln und Wissensdatenbanken nicht mehr im Programm *festcodiert* sind, wie im Fall der *regelbasierten* KI, sondern das Programm, welches das *maschinelle Lernen* implementiert, ein Set von Entscheidungsregeln in Form eines *statistischen Modells* selbst festlegt, und zwar auf Basis von Daten, die dem Programm mit ML zugeführt werden⁸.

Es gibt etliche technologische Arten des maschinellen Lernens, wie zum Beispiel ‚Deep Learning‘ basierend auf der Technologie ‚Künstliche Neuronale Netzwerke (KNN)‘, ‚Random Forest‘ basierend auf der Architektur von Entscheidungsbäumen, ‚Support Vector Machine (SVM)‘⁹ und einige andere, s. [4], Kap. 3, [2], Kap. 5.

⁷ siehe [5], CHAPTER I, ch. 3.5 “Enmorphya for Quasi-Stochastic Systems”, STM. 13 ‘the Principle of Self-Preservation of System‘ (Übersetzung des Autors): “Um die Stabilität der *quasi-stochastischen* Systeme zu gewährleisten, muss ihre Enmorphie mindestens ein weiteres Prinzip enthalten, das wir das **Prinzip der Selbsterhaltung des Systems** nennen.”

⁸ Maschinelles Lernen (ML) wird in [2], Glossar wie folgt definiert: “Eine Sammlung von Methoden, die in Daten der Vergangenheit nach Mustern suchen, die für die Zukunft Vorhersagen erlauben.”

⁹ RU: машина опорных векторов

Es gibt auch einige organisatorische Arten des maschinellen Lernens, wobei organisatorische und technologische Arten des maschinellen Lernens zueinander passen sollen, damit ‚Schwache KI‘ das vordefinierte Ziel effizient erreichen kann. Als Beispiele der organisatorischen Arten des maschinellen Lernens nennen wir hier (s. [3], Kap. 4):

- Überwachtes Lernen¹⁰: Dem Programm mit ML werden zuerst vom Menschen (in der Rolle des ‚Data Scientist‘, s. [2]) ausgewählte *Trainingsdaten* zugeführt, anhand derer das Programm mit ML nach Korrelationen sucht (und sie auch meistens findet) und auf dieser Basis ein Set von Entscheidungsregeln in Form eines statistischen Modells festlegt. Anschließend werden dem Programm mit ML *Testdaten* (also Daten mit bereits bekannten Ergebnissen¹¹) zugeführt. Die ‚Schwache KI‘ analysiert diese Testdaten anhand der von ihr selbst während des Trainings festgelegten Entscheidungsregeln und liefert ein Ergebnis dieser Analyse zurück. Der Mensch (in der Rolle des ‚Data Scientist‘) vergleicht das bereits zuvor bekannte Ergebnis der Testdaten mit dem zurückgelieferten Ergebnis der ‚Schwache KI‘ und wendet auf diesen Vergleich ein vom Menschen definiertes *Qualitätsmaß*¹² an. Genügt die Abweichung zwischen den verglichenen Ergebnissen diesem *Qualitätsmaß*, wird das Training als erfolgreich angesehen, vgl. [2], Kap. 5, [4], Kap. 3 („Neuronale Netzwerke“).
- Unüberwachtes Lernen¹³: Dem Programm mit ML werden unmarkierte („unlabelled“) Daten, d.h. ohne jegliche Vorgabe des Lernziels, zugeführt. Das Programm mit ML sucht in dieser Datenmenge nach Korrelationen, d.h. nach Abhängigkeiten und Mustern in diesen Daten und leitet aus den gefundenen Korrelationen¹⁴ ein Set der Entscheidungsregeln ab. Wenn diesem Programm mit ML weitere Daten zur Analyse zugeführt werden, setzt das Programm mit ML die Suche nach weiteren Korrelationen fort, justiert die Entscheidungsregeln nach und wendet sie auf die Inputdaten an. Als Ergebnis dieser Anwendung werden die Inputdaten den so selbstgelernten Clustern (Kategorien) zugeordnet. Eine der Varianten des unüberwachten Lernens sind ‚Generative Gegnerische Netze (GAN: Generative Adversarial Networks)‘. Da beim unüberwachten Lernen keine ex-ante ‚Grundwahrheit‘¹¹ vorliegt, muss der Mensch ex-post nachprüfen, ob die Entscheidungen/Empfehlungen, die durch diese Art von ‚Schwacher KI‘ als Ergebnis generiert werden, für das Ziel des Betreibers des KI-Systems adäquat sind, s. [3], Kap. 4.
- Bestärkendes Lernen¹⁵: Im ersten Schritt definiert der Mensch (in der Rolle des ‚Data Scientist‘) eine *Belohnungsfunktion* für das Programm mit ML und gibt dem Programm vor, es soll bei der Suche nach den möglichen Ergebnissen der Analyse der Inputdaten den Wert der ‚Belohnung‘ maximieren. Die ‚Schwache KI‘ probiert dann mögliche Lösungsoptionen durch (Versuch und Irrtum) und sucht sich als passende Entscheidung die Option aus, die den Wert der ‚Belohnung‘ im Vergleich zur vorangegangenen Lösung steigert. So optimiert sich die ‚Schwache KI‘ bei jeder

¹⁰ EN: Supervised Learning; RU: обучение под наблюдением (с учителем; контролируемое обучение)

¹¹ Die bereits im Voraus bekannten Eigenschaften der *Trainings-* und *Testdaten* nennt man ‚Grundwahrheit (Ground Truth)‘.

¹² „Qualitätsmaß: Eine Funktion, die bewertet, wie gut eine (algorithmische) Lösung eines Problems ist.“, s. [2], Glossar.

¹³ EN: Unsupervised Learning; RU: обучение без наблюдения (без учителя; неконтролируемое обучение)

¹⁴ Es ist wichtig anzumerken, dass gefundene Korrelationen nicht notwendigerweise kausal sind, was eine ‚Schwache KI‘ grundsätzlich nicht feststellen kann.

¹⁵ EN: Reinforcement Learning; RU: обучение с подкреплением

nächsten Lösung immer weiter, so dass der Wert der ‚Belohnung‘ im statistischen Schnitt ansteigt und irgendwann den maximalen Wert erreicht, s. [3], Kap. 4, [4], Kap. 3 („Evolutionäre Algorithmen“).

Man sieht an den Beispielen oben, dass eine ‚Schwache KI‘ auf ein ex-ante oder ex-post Eingreifen des Menschen angewiesen ist. Diese Tatsache gilt auch für alle anderen organisatorischen Arten des maschinellen Lernens.

Deswegen sind die Systeme, die auf maschinellem Lernen basieren, immer *soziotechnische* Systeme. Das bedeutet, dass sie notwendigerweise über einen rein technischen Kern, d.h. ein rein technisches Subsystem verfügen, welches in eine wohl organisierte menschliche Umgebung eingebunden ist, s. [2], Kap. 1, Abbildung 3.

Das entsprechende gesamte soziotechnische System – also das rein technische Subsystem samt der organisierten menschlichen Umgebung – nennt sich ‘Automatisiertes (Algorithmisches) Entscheidungssystem‘ (ADMS¹⁶), s. [2], Kap. 1, [3], Kap. 1.

Das rein technische Subsystem (wir werden es als TSKI abkürzen: *technisches System der künstlichen Intelligenz*) und das gesamte ADMS (*Automated Decision-Making System*) gehören – systemtheoretisch betrachtet – unterschiedlichen Systemkategorien an (mehr dazu – weiter unten in diesem Abschnitt). Deshalb ist es notwendig, zwischen TSKI und ADMS systemtheoretisch zu unterscheiden, s. [2], Kap. 1, Abbildung 2.

Einen weiteren systemtheoretisch relevanten Unterschied gibt es innerhalb eines jeden TSKI. Die Architektur eines jeden TSKI umfasst zwei technische Subsysteme: Das Subsystem des maschinellen Lernens (das ML-Subsystem), welches *stochastisch* funktioniert, und das Subsystem der Entscheidungsfindung, welches auf Basis des Sets der Entscheidungsregeln, das durch das ML-Subsystem festgelegt wurde, *deterministisch* funktioniert, s. [2], Kap. 5, Abbildung 22.

Dementsprechend müssen wir – systemtheoretisch betrachtet – zwischen den folgenden Subsystemen einer ‚Schwachen KI‘ unterscheiden:

- TSKI-ML: Lernsystem der technischen KI,
- ADMS-ML: Lernsystem des automatisierten Entscheidungssystems,
- TSKI-DM: Entscheidungssystem der technischen KI, und
- ADMS-DM: Entscheidungssystem des automatisierten Entscheidungssystems.

Im Folgenden betrachten wir jedes dieser Subsysteme separat.

3.1 Lernsystem der technischen KI (TSKI-ML)

Zu lernende Entscheidungsregeln (das statistische Modell) des TSKI sind das ‚Substrat‘ eines TSKI-ML-Subsystems einer ‚Schwachen KI‘. Die ‚Eigenschaft‘ des Subsystems sind die Beschaffenheit des TSKI (inkl. der Operationalisierung/Messbarmachung des Zwecks des TSKI, des Modells des Problems und des Algorithmus)³ sowie die Eigenschaften/Charakteristika der *Trainingsdaten*. Die ‚Relation‘ ist der Interaktionsprozess der Beschaffenheit des TSKI und der Charakteristika der *Trainingsdaten* einerseits mit den zu lernenden Entscheidungsregeln andererseits, d. h. der eigentliche Lernprozess.

¹⁶ Automated (or Algorithmic) Decision-Making System

Die Operationalisierung des Zwecks des TSKI inkludiert u.a. das *Fairnessmaß*¹⁷, welches, wenn relevant, vom Menschen (in der Rolle des Data Scientist) definiert wird, z.B. Equality vs. Equity, s. [2], Kap. 8.

Wir betrachten nun die ‚Deep Learning‘-Technologie als Beispiel. In diesem Fall beginnt das KNN (künstliches neuronales Netzwerk) den Lernprozess damit, dass die Gewichtungen/Wahrscheinlichkeiten der Übergänge zwischen den Einzelzuständen der benachbarten Schichten der künstlichen ‚Neuronen‘ vom KNN auf zufällige Werte gesetzt werden. Nach jedem Lerndurchgang unter der Nutzung von *Trainingsdaten* (als Teil der *Grundwahrheit*¹¹) justiert das KNN diese Gewichtungen *probabilistisch* nach, mit dem Ziel, sein Entscheidungsergebnis der Grundwahrheit näher zu bringen¹⁸. Dementsprechend ist auch die Abfolge von Sets der zu lernenden Entscheidungsregeln, die das statistische Modell ausmachen, grundsätzlich *probabilistisch*.

Da jeder nächste Zustand des Sets der zu lernenden Entscheidungsregeln probabilistisch allein von seinem aktuellen Zustand (und nicht von vorangegangenen Zuständen) abhängt, besitzt der Lernprozess die *Markow-Eigenschaft*. Deswegen ist der Lernprozess ein *echt-stochastischer* Prozess, s. Glossar.

Wie wir in [5], CHAPTER I, ch. 3.4 „Enmorphya for Truly-Stochastic Systems“ festgestellt haben, ist die ‚Information über die Steuerung der Relation‘ (Enmorphie der Relation) aller *echt-stochastischen* Systeme immer durch das Prinzip des kleinsten Ressourcenverbrauchs (der maximalen Entropie) repräsentiert. Das bedeutet, dass auch die Enmorphie der Relation des TSKI-ML-Subsystems ebenfalls durch das Prinzip des kleinsten Ressourcenverbrauchs (PKR) repräsentiert sein muss, konkret – durch *das Prinzip der Lernökonomie*.

Since each next state of the set of decision rules to be learned depends probabilistically solely on its current state (and not on previous states), the learning process has the *Markov property*. Therefore, the learning process is a *true-stochastic* process, see Glossar.

Manuela Lenzen merkt in [3], Kap. 4 an:

„Ein gutes KNN ist sicher darin, Daten den gewünschten Kategorien zuzuordnen. Es ist weder zu empfindlich noch zu unempfindlich gegenüber Variationen in den Daten. Und es ist sparsam, was die Zeit, die Menge an Daten und die Hardware angeht, die man zum Training benötigt.“

Janelle Shane zitiert in [4], Kap. 5 Alex Irpan, KI-Forscher bei Google:

„Ich habe mir angewöhnt, [die KI] als einen Dämon zu betrachten, der seine Belohnung absichtlich falsch interpretiert und aktiv nach dem Optimum sucht, bei dem sie möglichst faul sein kann. Das klingt lächerlich, aber die Einstellung kann tatsächlich recht produktiv sein.“¹⁹

¹⁷ „Fairnessmaß: Eine mathematische Funktion, die bewertet, inwieweit unterschiedliche Bevölkerungsgruppen gleichermaßen von Entscheidungen betroffen sind“, s. [2], Glossar.

¹⁸ Die Qualität des Entscheidungsergebnisses wird dabei durch das vom Menschen definierten *Qualitätsmaß* bestimmt.

¹⁹ Im Original: “I’ve taken to imagining deep RL as a demon that’s deliberately misinterpreting your reward and actively searching for the laziest possible local optima. It’s a bit ridiculous, but I’ve found it’s actually a productive mindset to have.”, Alex Irpan *Deep Reinforcement Learning Doesn’t Work Yet*, <https://www.alexirpan.com/2018/02/14/rl-hard.html>.

Angesichts der Erkenntnis, dass das TSKI-ML-Subsystem einer ‚Schwachen KI‘ immer dem *Prinzip der Lernökonomie* (des kleinsten Ressourcenverbrauchs) folgt, klingt diese Feststellung von Alex Irpan nicht nur nicht mehr „lächerlich“, sondern geradezu folgerichtig. In [4], Kap. 6 stellt Janelle Shane immer wieder fest, dass eine KI während des Trainings immer wieder versucht, die vorgegebene ‚Matrix‘ zu hacken, um an ‚kostenlose Energie/Nahrung‘ zu kommen.

Auch das ist aus dem *Prinzip der Lernökonomie* sofort erklärbar: Das Hacken der ‚Matrix‘ minimiert immer den Ressourcenverbrauch.

Das zweite Prinzip der Enmorphie der Relation des TSKI-ML-Subsystems ist ein Set von *Optimierungsfunktionen* nach dem Zweck des TSKI inkl. ihrer Priorisierung, wobei *Qualitätsmaß* und *Fairnessmaß* außerhalb des TSKI, d.h. TSKI-extern durch den Data Scientist vorgegeben werden.

Innerhalb des TSKI-ML-Subsystems stellen das *Prinzip der Lernökonomie* zusammen mit *Optimierungsfunktionen* die ‚Information über die Steuerung der Relation‘ (Enmorphie der Relation) dieses Systems dar. Als Enmorphie der Relation zwischen dem Substrat (den zu lernenden Entscheidungsregeln) und der Eigenschaft (der Beschaffenheit des TSKI und der Charakteristika der *Trainingsdaten*) definieren das *Prinzip der Lernökonomie* und die *Optimierungsfunktionen* den Charakter dieser Relation (Interaktion), siehe Fußnote 4. Sie bestimmen den Charakter des Prozesses des maschinellen Lernens, der seinerseits die Interaktion zwischen den zu lernenden Entscheidungsregeln und der Beschaffenheit des TSKI und der Charakteristika der *Trainingsdaten* realisiert. Das *Prinzip der Lernökonomie* und die *Optimierungsfunktionen* prägen also sowohl die zu lernenden Entscheidungsregeln (das Substrat des TSKI-ML-Subsystems) als auch die Beschaffenheit des TSKI und die passenden Charakteristika der *Trainingsdaten* (ihre Form und ihren Inhalt).

Veranschaulichen wir die Relation zwischen dem primären System und dem Metasystem am Beispiel des TSKI-ML-Subsystems (das Lernsystem der technischen KI):

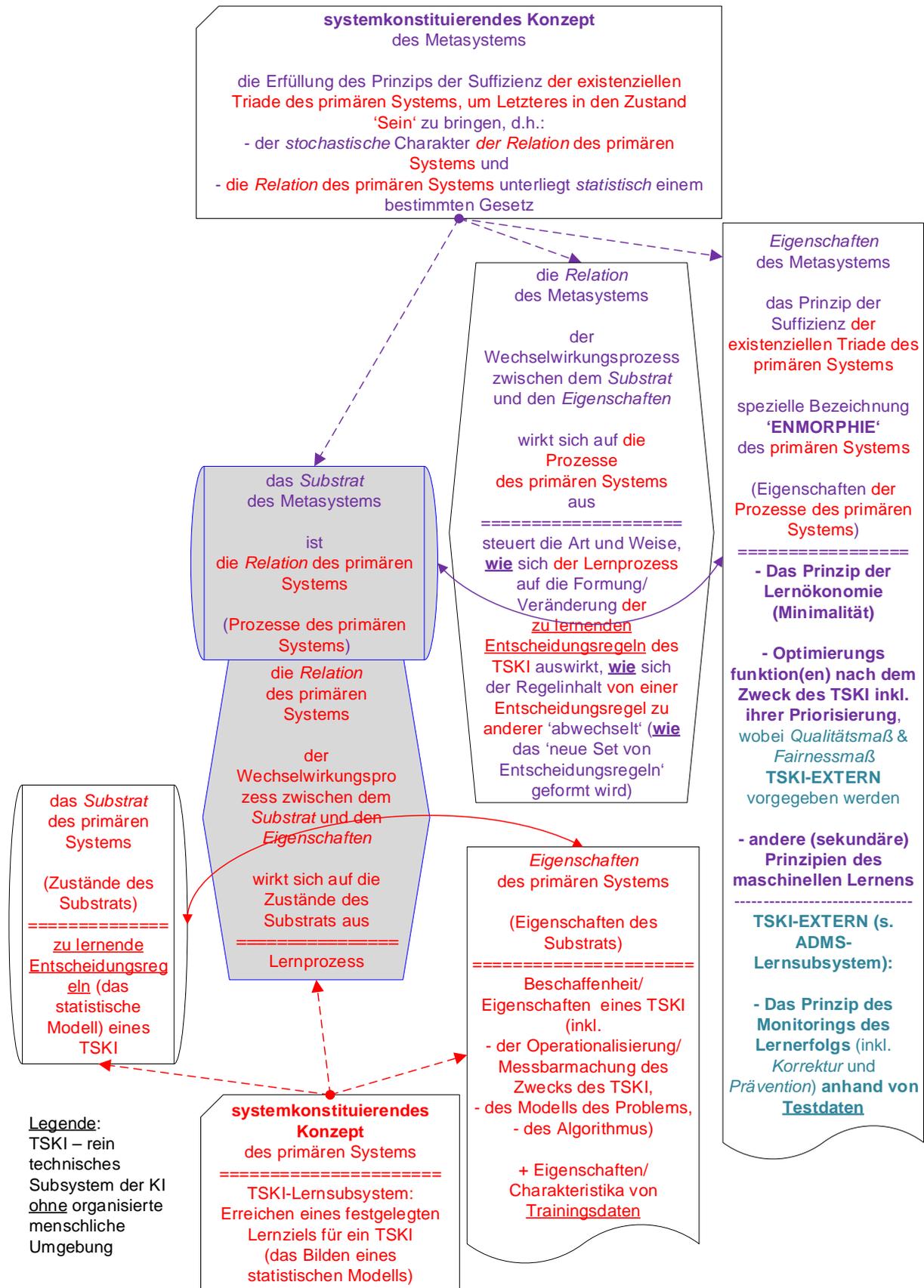


Abbildung 2: Relation zwischen dem primären System 'TSKI-ML-Subsystem' (technisches KI-System – das Lernsubsystem) und dem entsprechenden Metasystem

3.2 Lernsubsystem des ‚automatisierten Entscheidungssystems‘ (ADMS-ML)

Wie wir in diesem Kapitel bereits beschrieben haben, ist eine ‚Schwache KI‘ immer auf ein ex-ante oder ex-post Eingreifen des Menschen angewiesen. Deswegen sind die Systeme, die auf maschinellem Lernen basieren, immer *soziotechnische* Systeme. Das bedeutet, dass sie notwendigerweise über einen rein technischen Kern, d.h. ein rein technisches Subsystem verfügen (das TSKI), welches in eine wohl organisierte menschliche Umgebung eingebunden ist, s. [2], Kap. 1, Abbildung 3. Das entsprechende gesamte soziotechnische System – also das rein technische Subsystem (das TSKI) samt der organisierten menschlichen Umgebung – nennt sich ‚Automatisiertes (Algorithmisches) Entscheidungssystem‘ (ADMS²⁰), s. [2], Kap. 1, [3], Kap. 1.

Im Abschnitt 3.1 weiter oben haben wir das Lernsubsystem der technischen KI (das TSKI-ML-Subsystem) systemtheoretisch betrachtet. Wir wollen nun das Lernsubsystem des gesamten automatisierten Entscheidungssystems (das ADMS-ML-Subsystem) analysieren.

Die ‚Sozio‘-Komponente, d.h. das Eingreifen des Menschen stellt aus der Sicht des rein technischen TSKI eine systemexterne Randbedingung dar. Das Lernsubsystem des gesamten automatisierten Entscheidungssystems unterscheidet sich vom Lernsubsystem der technischen KI ausschließlich dadurch, dass es die ‚Sozio‘-Komponente, d.h. das Eingreifen des Menschen mit integriert.

Auf **Abbildung 2** ist leicht erkennbar, dass das Eingreifen des Menschen – als eine systemexterne Randbedingung für das TSKI-ML-Subsystem – ausschließlich die Enmorphie der Relation des TSKI-ML-Subsystems beeinflusst (durch Türkis gekennzeichnet). Der Mensch (in der Rolle des Data Scientist) gibt das *Qualitätsmaß* und das *Fairnessmaß* (sowie andere notwendigen *Hyperparameter*) vor und überwacht den Lernerfolg anhand von *Testdaten*, s. [2], Kap. 5.

Die Überwachung des Lernerfolgs (inkl. *Korrektur* und *Prävention*) anhand von *Testdaten* implementiert das *Prinzip des Monitorings des Lernerfolgs* für das Lernsubsystem des Automatisierten Entscheidungssystems (ADMS-ML-Subsystem), vgl. [2], Kap. 5, [4], Kap. 3. Das *Prinzip des Monitorings* inkl. *Korrektur* und *Prävention* wird durch den systemeigenen *Adaptationsmechanismus* umgesetzt.

Das *Prinzip des Monitorings des Lernerfolgs* ist wiederum die für das ADMS-ML-Subsystem konkrete Umsetzung des allgemeinen **Prinzips der Selbsterhaltung des Systems**, siehe Fußnote 7. Die Tatsache, dass die Enmorphie der Relation des ADMS-ML-Subsystems u.a. das Prinzip der Selbsterhaltung des Systems beinhaltet, sagt aus, dass das ADMS-ML-Subsystem ein *quasi-stochastisches* System ist, siehe Fußnote 7 und 21. Das ist allerdings vollkommen nachvollziehbar, weil der Mensch, welcher selbst ein *quasi-stochastisches* System darstellt²², Bestandteil des ADMS-ML-Subsystems ist.

Veranschaulichen wir die Relation zwischen dem primären System und dem Metasystem am Beispiel des ADMS-ML-Subsystems (das Lernsubsystem des automatisierten Entscheidungssystems):

²⁰ Automated (or Algorithmic) Decision-Making System

²¹ siehe [5], CHAPTER I, ch. 3.5 “Enmorphia for Quasi-Stochastic Systems”, STM. 14 ‘at least two principles are existentially necessary components of the enmorphia‘ (Übersetzung des Autors): “Mindestens zwei Prinzipien sind existenziell notwendige Bestandteile der Enmorphie *quasi-stochastischer* Systeme: **das Prinzip des kleinsten Ressourcenverbrauchs (PKR) und das Prinzip der Selbsterhaltung des Systems (PSE)**.”

²² siehe [5], CHAPTER II, ch. 1.2 “Principles of the Enmorphia of Living Beings”

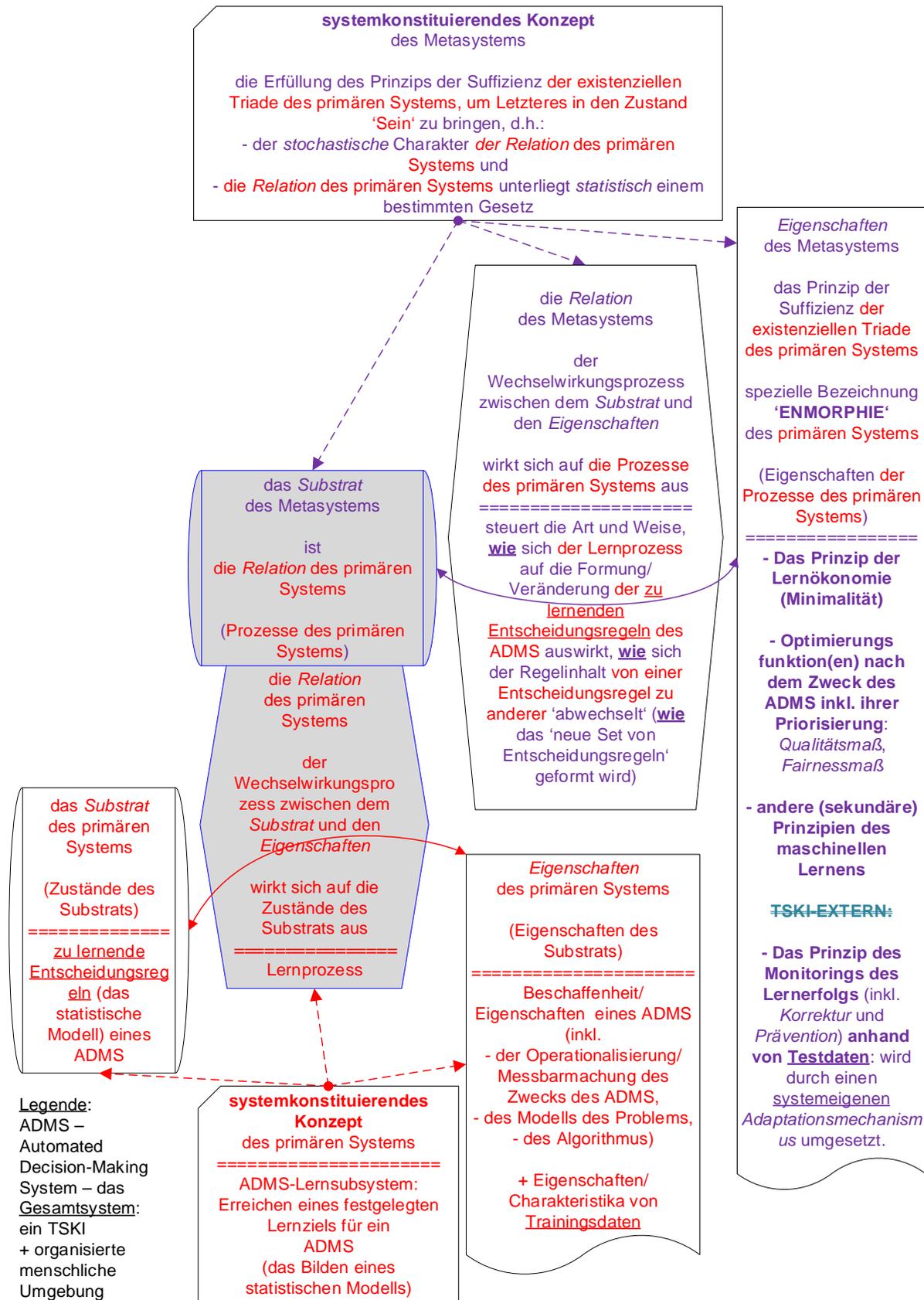


Abbildung 3: Relation zwischen dem primären System 'ADMS-ML-Subsystem' (Automatisiertes Entscheidungssystem – das Lernsubsystem) und dem entsprechenden Metasystem

3.3 *Entscheidungs subsystem der technischen KI (TSKI-DM)*

Zu analysierende Daten sind das ‚Substrat‘ für jedes TSKI-Entscheidungssystem einer ‚Schwachen KI‘. Die ‚Eigenschaft‘ des Subsystems sind die gelernten Entscheidungsregeln (das angelernte statistische Modell) des TSKI sowie die Eigenschaften/Charakteristika der zu analysierenden Daten. Die ‚Relation‘ ist der Prozess der Anwendung der gelernten Entscheidungsregeln auf die zu analysierenden Daten. Die zu analysierenden Daten ändern sich durch diese Anwendung nicht.

Der Prozess der Anwendung der gelernten Entscheidungsregeln auf die zu analysierenden Daten ist *deterministisch*, nicht *stochastisch*. Deswegen ist das Prinzip der Suffizienz der Existenziellen Triade⁴ hier nicht anwendbar: das TSKI-Entscheidungssystem ist *deterministisch*, vgl. [2], Kap. 5, Abbildung 22.

Für das TSKI-Entscheidungssystem repräsentieren die *Prinzipien des angelernten statistischen Modells* die ‚Information über die Steuerung der Relation‘ (die Enmorphie der Relation). Die *Prinzipien des angelernten statistischen Modells* werden durch das TSKI-Lernsystem festgelegt, s. „Enmorphie“ auf **Abbildung 2**. Als Enmorphie der Relation zwischen dem Substrat (den zu analysierende Daten) und der Eigenschaft (dem angelernten statistischen Modell) definieren die Prinzipien des angelernten statistischen Modells den Charakter dieser Relation (Interaktion). Vergleicht man die Enmorphie des regelbasierten Expertensystems (**Abbildung 1**) und die Enmorphie des TSKI-Entscheidungssystems (**Abbildung 4** weiter unten), wird es offensichtlich, dass das angelernte statistische Modell für eine ‚Schwache KI‘ dieselbe (zentrale) Rolle wie das ‚Programmierhandbuch‘ für regelbasierte Expertensysteme spielt.

Veranschaulichen wir die Relation zwischen dem primären System und dem Metasystem am Beispiel des TSKI-Entscheidungssystems (das Entscheidungs subsystem der technischen KI):

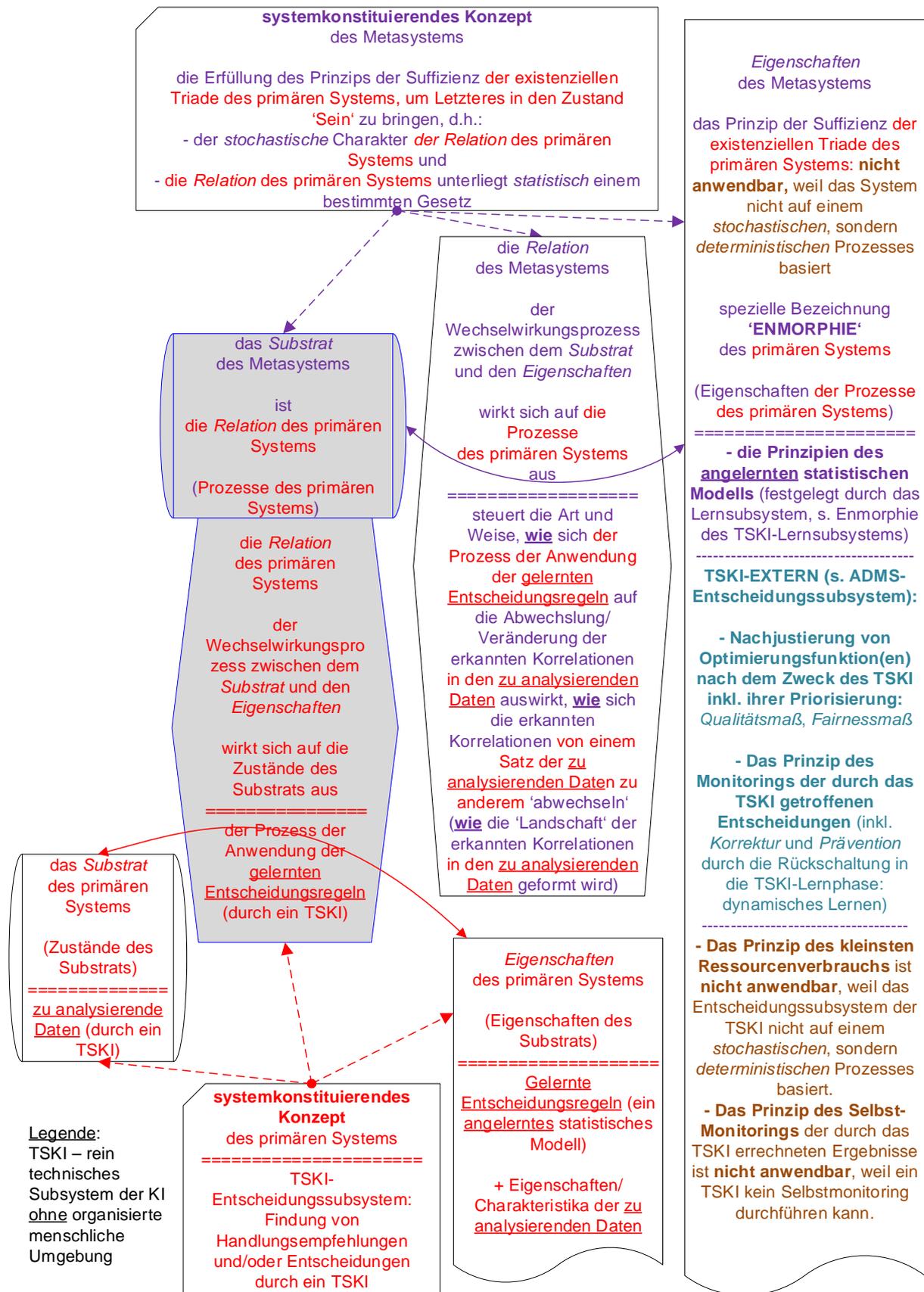


Abbildung 4: Relation zwischen dem primären System 'TSKI-Entscheidungs subsystem' (technisches KI-System – das Entscheidungs subsystem) und dem entsprechenden Metasystem

3.4 *Entscheidungs subsystem des ,automatisierten Entscheidungs systems‘ (ADMS-DM)*

Im Abschnitt 3.3 weiter oben haben wir das Entscheidungs subsystem der technischen KI (das TSKI-Entscheidungs subsystem) systemtheoretisch betrachtet. Wir wollen nun das Entscheidungs subsystem des gesamten *soziotechnischen* automatisierten Entscheidungs systems (das ADMS-Entscheidungs subsystem) analysieren.

Die ‚Sozio‘-Komponente, d.h. das Eingreifen des Menschen stellt aus der Sicht des rein technischen TSKI eine systemexterne Randbedingung dar. Das Entscheidungs subsystem des gesamten automatisierten Entscheidungs systems unterscheidet sich vom Entscheidungs subsystem der technischen KI ausschließlich dadurch, dass es die ‚Sozio‘-Komponente, d.h. das Eingreifen des Menschen mit integriert.

Auf **Abbildung 4** ist leicht erkennbar, dass das Eingreifen des Menschen – als eine systemexterne Randbedingung für das ‚TSKI-Entscheidungs subsystem‘ – ausschließlich die Enmorphie der Relation des ‚TSKI-Entscheidungs subsystems‘ beeinflusst (durch Türkis gekennzeichnet). Der Mensch (in der Rolle des ‚Data Scientist‘) justiert *Optimierungsfunktion(en)* nach dem Zweck des ADMS inkl. ihrer Priorisierung sowie das *Qualitätsmaß* und das *Fairnessmaß* und andere notwendigen *Hyperparameter* nach. Außerdem überwacht der Mensch die Adäquatheit von getroffenen Entscheidungen, vgl. [2], Kap. 8 und 9, [4], Kap. 3. Das Implementiert das *Prinzip des Monitorings der durch das ADM-System getroffenen Entscheidungen* (inkl. *Korrektur* und *Prävention* durch die Rückschaltung in die ADMS-Lernphase, **Abbildung 3**: dynamisches Lernen). Das *Prinzip des Monitorings* inkl. *Korrektur* und *Prävention* wird durch den systemeigenen *Adaptationsmechanismus* umgesetzt.

Ähnlich wie für das ADMS-ML-Subsystem (s. Abschnitt 3.2), ist das *Prinzip des Monitorings der getroffenen Entscheidungen* wiederum die für das ADMS-Entscheidungs subsystem konkrete Umsetzung des allgemeinen **Prinzips der Selbsterhaltung des Systems**, siehe Fußnote 7. Die Tatsache, dass die Enmorphie der Relation des ADMS-Entscheidungs subsystems u.a. das Prinzip der Selbsterhaltung des Systems beinhaltet, sagt aus, dass das ADMS-Entscheidungs subsystem ein *quasi-stochastisches* System ist, siehe Fußnote 7 und 21. Das ist allerdings vollkommen nachvollziehbar, weil der Mensch, welcher selbst ein *quasi-stochastisches* System darstellt²³, Bestandteil des ADMS-Entscheidungs subsystems ist.

Veranschaulichen wir die Relation zwischen dem primären System und dem Metasystem am Beispiel des ADMS-Entscheidungs subsystems (das Entscheidungs subsystem des automatisierten Entscheidungs systems):

²³ siehe [5], CHAPTER II, ch. 1.2 “Principles of the Enmorphya of Living Beings”

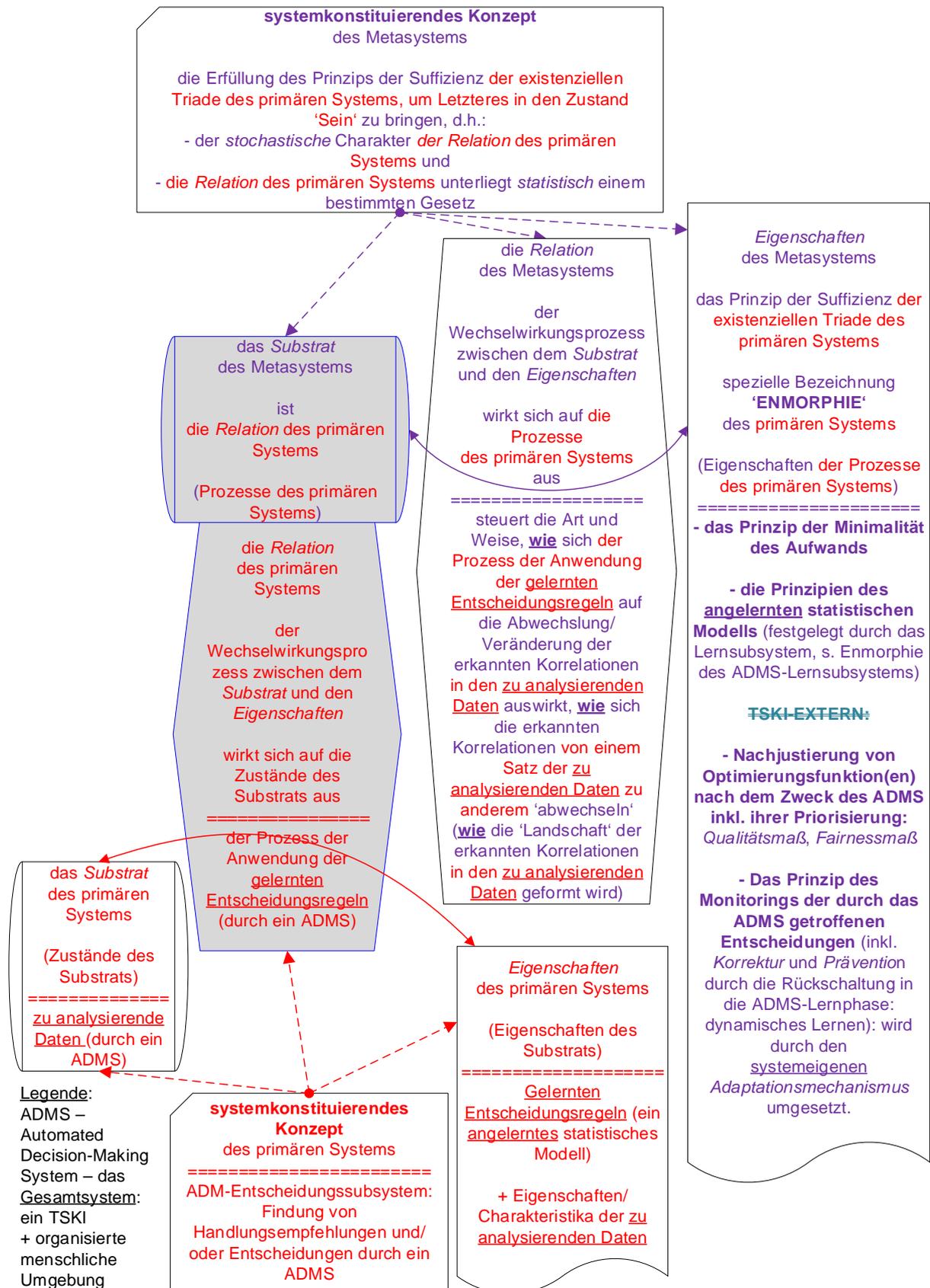


Abbildung 5: Relation zwischen dem primären System 'ADMS-Entscheidungs subsystem' (Automatisiertes Entscheidungssystem – das Entscheidungssystem) und dem entsprechenden Metasystem

Vergleichen wir die jeweiligen Enmorphien der Relation der vier Subsysteme einer ‚Schwachen KI‘ auf Abbildung 2 bis Abbildung 5, erkennen wir die systemtheoretischen Unterschiede zwischen diesen Subsystemen:

- Das TSKI-ML-Subsystem (Abbildung 2) ist *echt-stochastisch* und wird durch das *Prinzip der Lernökonomie* gesteuert; dieses Subsystem ist rein technischer Natur;
- Das ADMS-ML-Subsystem (Abbildung 3) ist *quasi-stochastisch* und somit selbstgenügend, weil dieses Subsystem den Menschen (in der Rolle des ‚Data Scientist‘) mit integriert;
- Das TSKI-Entscheidungssystem (Abbildung 4) ist *deterministisch* und wird durch die *Prinzipien des angelernten statistischen Modells* gesteuert; dieses Subsystem ist rein technischer Natur;
- Das ADMS-Entscheidungssystem (Abbildung 5) ist *quasi-stochastisch* und somit selbstgenügend, weil dieses Subsystem den Menschen (in der Rolle des ‚Data Scientist‘) mit integriert.

Diese Analyse zeigt u.a., dass rein technische Subsysteme einer ‚Schwachen KI‘ entweder *echt-stochastisch* oder *deterministisch* sind. Das technische KI-System – bestehend aus dem *echt-stochastischen* TSKI-ML-Subsystem (dem Subsystem des maschinellen Lernens) und aus dem *deterministischen* TSKI-DM-Subsystem (dem Subsystem für Entscheidungsfindung) – ist insgesamt *echt-stochastisch*.

Das bedeutet unter anderem, dass eine ‚Schwache KI‘ als rein technisches System kein belebtes System sein kann, s. [5], CHAPTER II, ch. 1 „Enmorphia of Living Beings“. So kann die ‚Schwache KI‘ auch keine Fähigkeit zur ‚Risikoreflexion‘ besitzen²⁴. Nur Systeme, die in der Lage sind, ‚*essentiell notwendige Fragen*‘ zu stellen, d. h. nur Systeme, die über *Risikoreflexion* verfügen, sind fähig, Erkenntnisssysteme zu schaffen und somit sich auch einer semantischen Erkenntnis²⁵ zu befähigen.

In Anwendung dieser Erkenntnis auf eine ‚Schwache KI‘ bedeutet das:

Statement STM. 1:

Eine ‚Schwache KI‘ ist einer Erkenntnis des Sinns / einer semantischen Erkenntnis grundsätzlich nicht fähig.

Diese Schlussfolgerung deckt sich mit den bis dato gesammelten Erfahrungen. Manuela Lenzen schreibt in [3], Kap. 6: „Letztlich zeigen diese Experimente, dass die Algorithmen, die Texte, die sie generieren, nicht verstehen.“ Dort weiter gibt sie auch die Gedanken von John Searle wieder: „Algorithmen arbeiten mit Regeln, die Ihnen sagen, wie sie Symbole hin- und herschieben sollen, Bedeutung werden sie nie erfassen. ... Alles, was mit Bedeutung zu tun hat, so Searle, interpretieren die Menschen in die Datenverarbeitung hinein.“

Auch Janelle Shane beschreibt an mehreren Stellen in [4] ähnliche Erfahrungen: „KI versteht die zu lösenden Probleme nicht wirklich“ („Einleitung“); in Kap. 5 darin: „Damit die KI die richtige *Lösung* finden kann, muss der Programmierer dafür sorgen, dass die KI auch wirklich das richtige *Problem* bearbeitet.“

²⁴ see [5], CHAPTER II, ch. 2.2 “Free Will”, there the new notion of ‘*risk reflection*’; see Glossar below, ‘risk reflection’.

²⁵ Die ‘Erkenntnis des Sinns‘ nennen wir synonym ‘*semantische Erkenntnis*‘ (Inhalt) als komplementären Begriff zur ‘*syntaktischen Verarbeitung von Symbolen*‘ (Form).

Ein gesamtes ‚Automatisierte Entscheidungssystem‘ (ADMS) ist lediglich dank des ex-ante und ex-post eingreifenden Menschen quasi-stochastisch. Der Mensch (in der Rolle des ‚Data Scientist‘) muss alle semantischen Ziele vordefinieren, deren Erreichung kontinuierlich überwachen und ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ einer ‚Schwachen KI‘ korrektiv und präventiv nachjustieren.

4 Starke KI

Eine ‚Starke KI‘ ist ein autonomes technisches System, das mit gleichen (oder stärkeren und umfangreicheren) intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre, s. [2], Glossar, [3], Kap. 1, [4], Kap. 2 („AGI“) oder im Internet.

Eine ‚Starke KI‘ stellt aktuell lediglich eine Vision dar. Es bedarf einer globalen intensiven gesellschaftlichen Diskussion, die Nutzen und Risiken einer ‚Starken KI‘ abwägt. Dieser Aspekt wird in [2], Kap. 11 ausführlich beleuchtet.

Wir stellen uns hier die Frage, welche grundsätzliche Voraussetzungen aus systemtheoretischer Sicht ein autonomes technisches System erfüllen muss, damit es mit gleichen (oder stärkeren und umfangreicheren) intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre.

Wir haben in [5]²⁶ festgestellt, was den Menschen als System von allen anderen belebten Systemen unterscheidet. Der naheliegende Ansatz, um die oben gestellte Frage zu beantworten, besteht darin, davon auszugehen, dass ein autonomes technisches System dieselben systemtheoretischen Prinzipien und wesentliche menschengespezifische Attribute dieser Prinzipien wie der Mensch selbst als System umsetzen soll.

Systemtheoretische Prinzipien eines Systems sind per definitionem Bestandteile der Enmorphie der Relation dieses Systems, s. [5], CHAPTER I, ch. 3.2 „Enmorphia“. Die Prinzipien der Enmorphie des Selbstbewusstseins von Lebewesen²⁷ wurden in [5], CHAPTER II, ch. 1.2 „Principles of the Enmorphia of Living Beings“ behandelt: Das sind das **Prinzip der größten Auswahl** und das **Prinzip der Selbsterhaltung des Systems**, siehe Glossar. Allerdings sind diese beiden Prinzipien nicht nur den Lebewesen, sondern jedem *quasi-stochastischen* System²⁸ eigen.

So halten wir zunächst fest, dass eine ‚Starke KI‘ als autonomes technisches System ein *quasi-stochastisches* System sein muss²⁹, d.h. die beiden Prinzipien – der größten Auswahl (des kleinsten Ressourcenverbrauchs) und der Selbsterhaltung des Systems – umsetzen muss.

Das bedeutet auch, wie in [5], CHAPTER I, ch. 3.5 „Enmorphia for Quasi-Stochastic Systems“ für jedes *quasi-stochastische* System festgestellt wurde, dass eine ‚Starke KI‘ notwendigerweise über einen systemeigenen Adaptationsmechanismus verfügen muss,

²⁶ [5], CHAPTER II, Kap. 2.1 „Living and Non-Living Systems“, dort in Abschnitt **“The human as a system”**, auch in [5], CHAPTER II, ch. 2.2 “Free Will”, dort ein neuer Begriff ‘*Risikoreflexion*’; siehe Glossar weiter unten (‘Risikoreflexion’)

²⁷ Die Enmorphie des Selbstbewusstseins ist die für Lebewesen spezifische Ausprägung der Enmorphie der Relation.

²⁸ Das **Prinzip der größten Auswahl** ist die für Lebewesen spezifische Umsetzung des allgemeinen **Prinzips des kleinsten Ressourcenverbrauchs**, s. [5], CHAPTER II, ch. 1.2 “Principles of the Enmorphia of Living Beings”.

²⁹ Somit muss eine ‚Starke KI‘ ein belebtes, lebendiges System, s. [5], CHAPTER II, ch. 2.1 “Living and Non-Living Systems”, und einen ‘*Willensträger*’ darstellen, s. [5], CHAPTER I, ch. 3.5.3 “Society”, [5], CHAPTER II, ch. 2.2 “Free Will” sowie Glossar.

welcher das Prinzip der Selbsterhaltung des Systems umsetzt. Dieser *Adaptationsmechanismus* umfasst

- das *Monitoring* (die *Überwachung*) des Systemzustands,
- systeminterne *Korrekturen* (Korrekturmaßnahmen) in Bezug auf einen sich ändernden Systemzustand und
- die *Prävention* eines ähnlichen ‚suboptimalen‘ Systemzustands³⁰ durch Korrektur einer entsprechenden und systemimmanenten ‚Norm‘.

Außerdem muss *Informationsmetabolismus* einer ‚Starken KI‘, wie jedem *quasi-stochastischen* System, eigen sein, s. Glossar³¹.

‘Gefühl‘³² stellt eine Komponente des *Informationsmetabolismus* dar, wobei es ‘negative‘ und ‘positive‘ Gefühle geben kann: Ein ‘negatives Gefühl‘ wird von einem *quasi-stochastischen* System dann empfunden, wenn das System im ‘suboptimalen‘ Zustand ist, d.h. wenn die Stabilität seines *systemkonstituierenden Konzepts* gefährdet ist; ein ‘positives Gefühl‘ wird von einem *quasi-stochastischen* System dann empfunden, wenn das System im ‚optimalen‘ Zustand ist, d.h. wenn die Stabilität seines *systemkonstituierenden Konzepts* nicht gefährdet ist.

Die menschengespezifischen (als Spezies) Attribute der Enmorphie des Selbstbewusstseins sind in [5], CHAPTER II, ch. 1.3 “Variativity and Attributes of the Enmorphia of Self-Awareness” aufgelistet. Das sind:

- 1) das Attribut ‘biologische Art‘ mit dem Wert ‘homo sapiens‘;
- 2) das Attribut ‘Risikoreflexion‘;
- 3) das Attribut ‘ethische Normen‘ als Verfahrensnormen, die für ein bestimmtes belebtes System gelten;
- 4) das Attribut ‘Modus‘ mit den möglichen Werten ‘alltäglich (opportunistisch)‘ oder ‘ontologisch (ethisch)‘;
- 5) das Attribut ‘Psychotyp‘;
- 6) das Attribut ‘Archetyp‘.

Welche dieser Attribute wären wesentliche, damit ein autonomes technisches System mit gleichen intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre? Um diese Frage zu beantworten, wenden wir das Ausschlussverfahren an.

Das Attribut ‘*biologische Art*‘ mit dem Wert ‘homo sapiens‘ können wir direkt ausschließen, weil ein autonomes technisches System dem ‘homo sapiens‘ nicht identisch sein kann: Man denke lediglich daran, dass ein autonomes technisches System einen anderen als Mensch materiellen Körper haben muss.

Das Attribut ‘*ethische Normen*‘, die als Verfahrensnormen für einen Menschen gelten, gibt es im Ansatz für eine ‚Schwache KI‘ bereits heute: Das ist das *Fairnessmaß*, s. Kap. 3 weiter oben. Für eine ‚Starke KI‘ wird das jeweilige *Fairnessmaß* durch die ‚Sozialisation‘ der ‚Starken KI‘ definiert. Dieses Attribut ist variativ. Wie wir in [5], CHAPTER I, ch. 3.7.2 „Variativity of Quasi-Stochastic Systems“ festgestellt haben, stellt die Variativität von Attributen der Enmorphie ein wesentliches Unterscheidungsmerkmal der *quasi-stochastischen* von den *echt-stochastischen* Systemen dar.

³⁰ Der ‚suboptimale‘ Zustand eines Systems ist der Zustand, der die Stabilität seines *systemkonstituierenden Konzepts* gefährdet.

³¹ auch [5], CHAPTER I, ch. 3.5.3 “Society”, [5], CHAPTER II, ch. 1.1 “Enmorphia of Self-Awareness” und ch. 1.2 “Principles of the Enmorphia of Living Beings”

³² bzw. ‚Stimmung‘/‘Atmosphäre‘ für Organisationen als ganzheitliche Entitäten

Für eine ‚Starke KI‘ halten wir hier fest, dass ihre Enmorphie u.a. mindestens ein spezifisches variatives Ethikattribut beinhalten soll, welches die Art und Weise steuert, wie die ‚Starke KI‘ mit ethischen Aspekten bei ihrer Kommunikation mit ihrer Umgebung und ihrer Umwelt umgeht.

Die Attribute ‚*Modus*‘, ‚*Psychotyp*‘ und ‚*Archetyp*‘ sind variative Attribute, die Art und Weise definieren, wie ein Mensch mit seiner Umgebung und seiner Umwelt kommuniziert (im breitesten Sinne des Wortes, also inkl. Interaktion)³³. Wesentlich ist, dass diese Attribute nicht konstant, sondern variativ sind.

Für eine ‚Starke KI‘ halten wir hier fest, dass ihre Enmorphie u.a. mindestens ein spezifisches variatives Kommunikationsattribut beinhalten soll, welches die Art und Weise steuert, wie die ‚Starke KI‘ mit ihrer Umgebung und ihrer Umwelt interagiert.

Das Attribut ‚*Risikoreflexion*‘ stellt das Hauptunterscheidungsmerkmal des freien Willens des Menschen als Spezies vom freien Willen aller anderen belebten Systeme³⁴. Deswegen ist das Vorhandensein dieses Attributs beim Menschen als Spezies eine Voraussetzung für intellektuelle und schöpferische Fähigkeiten des Menschen.

Der Mensch als einzige Spezies ist fähig, über einen Teil seiner möglichen (zukünftigen) Zustände, die sowohl die Welt um ihn herum als auch ihn selbst, einschließlich seiner eigenen Endlichkeit als System, umfassen, zu reflektieren³⁴. Die Reflexion seiner eigenen Endlichkeit als System ruft die (überwiegend verdrängte) Existenzangst hervor, s. Fußnote 33, dort Abschnitt „Archetypes“ und Fußnote 34. Die Existenzangst motiviert den Menschen zu verschiedenen, u.a. intellektuellen und schöpferischen Aktivitäten, um den Lauf der Zeit (immer in die Richtung des Systemendes) auf verschiedene Art und Weise unbemerkt zu machen, s. [5], CHAPTER VI, ch. 3.4 „The Existential Angst and Adaptation“. Es sei angemerkt, dass Angst (und somit auch die Existenzangst) ein ‚negatives Gefühl‘ ist, siehe weiter oben .

Die Überlegungen ergeben die folgende Kette der Abhängigkeiten: *Risikoreflexion* => Existenzangst (als ‚negatives Gefühl‘)³⁵ => die Motivation, den Lauf der Zeit unbemerkt zu machen => u.a. intellektuelle und schöpferische Aktivitäten.

Die *Risikoreflexion* ist außerdem eine Voraussetzung, dass ein System überhaupt fähig ist, ‚essentiell notwendige Fragen‘³⁶ zu stellen, d.h. Erkenntnisssysteme zu schaffen, und sich somit auch einer semantischen Erkenntnis³⁷ zu befähigen.

Einige KI Forscher führen eine ‚künstliche Neugier‘ als eine der Optimierungsfunktionen (in diesem Falle – Belohnungs- oder Bewertungsfunktionen) ein, s. [2], Kap. 11, [4], Kap. 5 („Neugier“). Diese Belohnungsfunktion soll ein KI-System dafür belohnen, dass es in dem KI-System bekannten ‚Weltmodell‘ etwas Abstrakteres findet. Die Absicht und die Hoffnung bestehen dabei darin, dass eine ‚Schwache KI‘ sich durch eine ‚künstliche Neugier‘ zu einer ‚Starken KI‘ Schritt für Schritt entwickeln würde.

Unsere Überlegungen weiter oben zeigen, dass eine ‚künstliche Neugier‘ – lediglich als ein der mehreren notwendigen systemeigenen Attribute einer ‚Starken KI‘ – nur dann für eine

³³ siehe [5], CHAPTER II, ch. 1.3 “Variativity and Attributes of the Enmorphia of Self-Awareness”

³⁴ siehe [5], CHAPTER II, ch. 2.2 “Free Will”, STM. 24 und ‘risk reflection’, also Glossar weiter unten (‘Risikoreflexion‘)

³⁵ Ohne Risikoreflexion kann es keine Existenzangst geben.

³⁶ ‚*Essentiell notwendige Fragen*‘ sind solche Fragen, Antworten auf welche für die Umsetzung des Prinzips der Selbsterhaltung des Systems notwendig sind. In diesem Sinne können ‚*essentiell notwendige Fragen*‘ auch als ‚*existentiell notwendige Fragen*‘ bezeichnet werden, s. Glossar.

³⁷ Die ‘Erkenntnis des Sinns‘ nennen wir synonym ‘*semantische Erkenntnis*‘ (Inhalt) als komplementären Begriff zur ‘syntaktischen Verarbeitung von Symbolen‘ (Form).

„Starke KI“ hilfreich sein kann, wenn diese „künstliche Neugier“ eine „Starke KI“ *essentiell notwendige Fragen* sich selbst stellen ließe, d.h. die Fragen, Antworten auf welche für die Umsetzung des Prinzips der Selbsterhaltung dieser „Starken KI“ notwendig sind³⁶. **Nur wenn diese „Starke KI“ ihre eigene Endlichkeit als System reflektieren und als „negatives Gefühl“ empfinden würde, wäre sie selbstmotiviert, solche „essentiell/existentiell notwendige Fragen“ sich selbst zu stellen.** Das bedeutet, dass eine für eine „Starke KI“ hilfreiche „künstliche Neugier“³⁸ mindestens diese zwei sehr speziellen Eigenschaften mitbringen soll.

Für eine „Starke KI“ halten wir hier fest, dass ihre Enmorphie u.a. das spezifische Attribut *‘Risikoreflexion‘* beinhalten soll, welches eine „Starke KI“ zu intellektuellen und schöpferischen Aktivitäten, zur Stellung „essentiell/existentiell notwendiger Fragen“ motiviert und erst überhaupt befähigt. Damit die *Risikoreflexion* eine solche Motivation erst erzeugen kann, muss eine „Starke KI“ auch noch „negative Gefühle“ in Bezug auf ihre eigene Endlichkeit als System empfinden.

Jetzt können wir auf Basis dieser Erkenntnisse die Frage beantworten, die wir eingangs gestellt haben: „Welche grundsätzliche Voraussetzungen aus systemtheoretischer Sicht ein autonomes technisches System erfüllen muss, damit es mit gleichen (oder stärkeren und umfangreicheren) intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre?“

STM. 2:

Damit eine „Starke KI“ als ein autonomes technisches System mit gleichen (oder stärkeren und umfangreicheren) intellektuellen und schöpferischen Fähigkeiten wie ein Mensch universell einsetzbar wäre, muss die „Starke KI“ – aus systemtheoretischer Sicht – folgende systemeigene Prinzipien und entsprechende Attribute³⁹ implementieren:

- 1) Sie soll ein *quasi-stochastisches* System sein, d.h. sie muss die beiden Prinzipien – der größten Auswahl (des kleinsten Ressourcenverbrauchs) und der Selbsterhaltung des Systems – umsetzen;
- 2) Sie soll notwendigerweise über einen systemeigenen *Adaptationsmechanismus* verfügen, welcher das Prinzip der Selbsterhaltung des Systems umsetzt. Dieser *Adaptationsmechanismus* umfasst
 - das *Monitoring* (die *Überwachung*) des Systemzustands,
 - systeminterne *Korrekturen* (Korrekturmaßnahmen) in Bezug auf einen sich ändernden Systemzustand⁴⁰ und
 - die *Prävention* eines ähnlichen ‚suboptimalen‘ Systemzustands durch Korrektur einer entsprechenden und systemimmanenten ‚Norm‘;
- 3) Ihre systemeigenen ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ sollen u.a. mindestens ein spezifisches variatives *Ethikattribut* beinhalten, welches die Art und Weise steuert, wie die „Starke KI“ mit ethischen Aspekten bei ihrer Kommunikation mit ihrer Umgebung und ihrer Umwelt umgeht;
- 4) Ihre systemeigenen ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ sollen u.a. mindestens ein spezifisches variatives *Kommunikationsattribut* beinhalten, welches

³⁸ Eine präzisere Bezeichnung wäre „künstliche Wissbegierde“ anstatt „künstliche Neugier“.

³⁹ Man kann Prinzipien und entsprechende Attribute auch als systemeigene ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ betrachten.

⁴⁰ Der ‚suboptimale‘ Zustand eines Systems ist der Zustand, der die Stabilität seines *systemkonstituierenden Konzepts* gefährdet.

die Art und Weise steuert, wie die ‚Starke KI‘ mit ihrer Umgebung und ihrer Umwelt interagiert;

- 5) Ihre systemeigenen ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ sollen u.a. das spezifische Attribut ‚*Risikoreflexion*‘ beinhalten⁴¹, welches die ‚Starke KI‘ zu intellektuellen und schöpferischen Aktivitäten und zur Stellung ‚essentiell/existentiell notwendiger Fragen‘ motiviert und erst überhaupt befähigt;
- 6) Sie soll u.a. ‚negative Gefühle‘ in Bezug auf ihre eigene Endlichkeit als System empfinden, damit die *Risikoreflexion* eine solche Motivation zu intellektuellen und schöpferischen Aktivitäten und zur Stellung ‚essentiell/existentiell notwendiger Fragen‘ erst erzeugen kann.

Der Autor dieser Zeilen hat zum Zeitpunkt der Veröffentlichung nicht gehört, dass in den Communities, die sich mit dem Thema KI befassen, die Notwendigkeit der oben genannten systemimmanenten Prinzipien und Attribute als Voraussetzungen für die Erschaffung einer ‚Starken KI‘ diskutiert würde.

5 Enmorphie des ‚Selbstbewusstseins‘ von KI: Unterscheidungskriterien zwischen den KI-Arten und Ausblick

Jetzt können wir auch die systemtheoretischen Unterscheidungskriterien zwischen drei Arten von KI-Systemen, die wir in diesem Kapitel betrachtet haben, feststellen.

Regelbasierte Expertensysteme sind *deterministisch*. Ihr Verhalten und Eigenschaften sind im entsprechenden ‚Programmierhandbuch‘ apriori festgelegt.

Das rein technische Subsystem einer ‚Schwachen KI‘ – bestehend aus dem *echt-stochastischen* Subsystem des maschinellen Lernens und aus dem *deterministischen* Subsystem für Entscheidungsfindung – ist insgesamt *echt-stochastisch*⁴².

Der *Adaptionsmechanismus* einer ‚Schwachen KI‘ wird nicht technisch, sondern durch den Menschen bedient. Auch ihr *Ethikattribut* – das *Fairnessmaß* – wird durch den Menschen festgelegt. Ein Pendant zum *Kommunikationsattribut* einer ‚Schwachen KI‘ habe ich bis dato nicht gesehen. Dieses Attribut ließe sich bei einer ‚Schwachen KI‘ implementieren, allerdings müsste das *Qualitätsmaß* für die entsprechende Optimierungsfunktion auch durch den Menschen bedient werden.

Eine ‚Starke KI‘, wenn es sie irgendwann geben soll, muss immanent und intrinsisch *quasi-stochastisch* sein, so dass sie den Menschen lediglich in der Designphase brauchen würde und nach der Inbetriebnahme völlig autonom zu agieren in der Lage wäre. Das muss die Eigenvorgabe aller semantischen Ziele, die Eigenüberwachung deren Erreichung sowie der korrektiven und präventiven Eigennachjustierung ihrer ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ inkludieren. Ihre systemeigenen ‚Optimierungsfunktionen‘ und

⁴¹ D.h. sie soll über einen Teil ihrer möglichen (zukünftigen) Zustände als System reflektieren, die sowohl die Welt um ihr herum als auch sie selbst, einschließlich ihrer eigenen Endlichkeit als System, umfassen.

⁴² Ein gesamtes ‚Automatisierte Entscheidungssystem‘ (ADMS) ist dank des ex-ante und ex-post eingreifenden Menschen *quasi-stochastisch*. Der Mensch (in der Rolle des ‚Data Scientist‘) muss alle semantischen Ziele vordefinieren, deren Erreichung kontinuierlich überwachen und ‚Optimierungsfunktionen‘ und ‚Hyperparameter‘ einer ‚Schwachen KI‘ korrektiv und präventiv nachjustieren.

„Hyperparameter“ sollen u.a. folgende spezifische variative Attribute beinhalten: das *Ethikattribut*, das *Kommunikationsattribut* und die *„Risikoreflexion“* u.a. in Bezug auf ihre eigene Endlichkeit als System.

Für eine „Starke KI“ sollen der *Adaptionsmechanismus* und alle seine Attribute systemeigen sein und im Betrieb ohne das Eingreifen des Menschen auskommen.

Die Kernunterschiede in der *Enmorphie der Relation* (in der *„Enmorphie des Selbstbewusstseins“* von KI-Systemen, wenn sie als Lebewesen behandelt würden) zwischen „schwacher KI“ und „starker KI“ aus systemtheoretischer Sicht haben wir in der folgenden Übersichtstabelle zusammengefasst (Unterschiede in **Fettdruck**):

KI Typ → Eigenschaft ↓	Regelbasierte KI (Kap. 2)	„Schwache KI“ (rein technisches Subsystem) (Kap. 3)	„Starke KI“ (wenn es sie zukünftig geben soll) (Kap. 4)
Systemtyp	deterministisch	echt-stochastisch	quasi-stochastisch
Hauptprinzipien von „Optimierungsfunktionen“	der Inhalt des entsprechenden „Programmierhandbuchs“	das Prinzip des kleinsten Ressourcenverbrauchs (der größten Auswahl)	- das Prinzip des kleinsten Ressourcenverbrauchs (der größten Auswahl) - das Prinzip der Selbsterhaltung des Systems
Adaptationsmechanismus	nicht anwendbar	wird durch den <u>systemexternen</u> Menschen justiert	ist systemeigen und selbstjustierend
Ethikattribut	nicht anwendbar	Das <i>Fairnessmaß</i> wird durch den <u>systemexternen</u> Menschen justiert.	variatives Attribut der systemeigenen „Optimierungsfunktionen“ und „Hyperparameter“
Kommunikationsattribut	nicht anwendbar	Wir haben es bis dato nicht gesehen, wäre aber anwendbar und implementierbar; Das <i>Qualitätsmaß</i> für die entsprechende Optimierungsfunktion müsste durch den <u>systemexternen</u> Menschen justiert werden.	variatives Attribut der systemeigenen „Optimierungsfunktionen“ und „Hyperparameter“

KI Typ → Eigenschaft ↓	<i>Regelbasierte KI</i> (Kap. 2)	‘ <i>Schwache KI</i> ’ (rein technisches Subsystem) (Kap. 3)	‘ <i>Starke KI</i> ’ (wenn es sie zukünftig geben soll) (Kap. 4)
Risikoreflexion ⁴³	nicht anwendbar	nicht anwendbar	ein Attribut der systemeigenen ,Optimierungsfunktionen‘ und ,Hyperparameter‘; motiviert und befähigt zu intellektuellen und schöpferischen Aktivitäten, zur Stellung ,essentiell/existentiell notwendiger Fragen‘ (s. Glossar)
Empfinden ,negativer Gefühle‘ in Bezug auf ihre eigene Endlichkeit als System ⁴⁴	nicht anwendbar	nicht anwendbar	ein Ergebnis der Funktion des Überwachungsmechanismus im Rahmen des Adaptionmechanismus
Fähigkeit zur Erkenntnis des Sinns ⁴⁵	keine	keine	grundsätzlich möglich dank der Risikoreflexion, die das System motiviert und befähigt, ,essentiell/existentiell notwendiger Fragen‘ zu stellen

Diese Übersicht zeigt, dass die Kernunterschiede zwischen ‘schwacher KI’ und ‘starker KI’ – aus systemtheoretischer Sicht – im *Enmorphotyp*⁴⁶ der bestimmten Technologie liegen, die eine bestimmte KI implementiert.

Unsere systemtheoretische Betrachtung zeigt u.a. auf, dass ,Schwache KI‘ und ,Starke KI‘ sich voneinander signifikant unterscheiden. Sie zeigt auch auf, dass Schaffung einer ,Starken KI‘ einen sehr langen Weg vor sich haben dürfte, wenn es überhaupt jemals möglich werden sollte. Der globale ethische Aspekt, ob die Menschheit sich eine solche Kreation überhaupt wünscht, bedarf einer globalen intensiven gesellschaftlichen Diskussion über Nutzen und Risiken.

⁴³ D.h. sie soll über einen Teil ihrer möglichen (zukünftigen) Zustände als System reflektieren, die sowohl die Welt um ihr herum als auch sie selbst, einschließlich ihrer eigenen Endlichkeit als System, umfassen.

⁴⁴ als Ergebnis der Funktion des *Überwachungsmechanismus* im Rahmen des *Adaptionmechanismus*

⁴⁵ Die ‘Erkenntnis des Sinns‘ nennen wir synonym ‘*semantische Erkenntnis*‘ (Inhalt) als komplementären Begriff zur ‘*syntaktischen Verarbeitung von Symbolen*‘ (Form).

⁴⁶ siehe [5], CHAPTER II, ch. 1.4 “Enmorphotype”

6 Glossar

Begriff	Definition
Grundbegriffe der Systemtheorie nach A. Uemow [1], die für die Lektüre dieser Abhandlung notwendig sind	
System	Jede beliebige Entität, an der eine <i>Relation</i> umgesetzt wird, die eine beliebig ausgewählte bestimmte <i>Eigenschaft</i> besitzt. Oder äquivalent: Jede beliebige Entität, an der einige <i>Eigenschaften</i> umgesetzt werden, die in einer beliebig ausgewählten bestimmten <i>Relation</i> stehen.
systemkonstituierendes Konzept ⁴⁷	Apriori vorgegebene systemkonstituierende <i>Eigenschaft</i> oder <i>Relation</i> ; Abhängig davon, ist das systemkonstituierende Konzept ein <i>Attributives</i> bzw. <i>Relationales</i> .
Strukturfaktor ⁴⁸	Ein Satz von Eigenschaften und Relationen, der dem vorgegebenen systemkonstituierenden Konzept genügt. Der Strukturfaktor kann relational (im Falle des attributiven Konzepts) und attributiv (im Falle des relationalen Konzepts) sein.
Substrat des Systems ⁴⁹	Ein Träger der relationalen oder attributiven Struktur.
Weitere für die Lektüre dieses Werkes notwendige Grundbegriffe	
existenzielle Triade	Ein Satz von { <i>Substrat</i> , <i>Eigenschaft</i> , <i>Relation</i> }, der für die Kreierung eines auf diesem Satz basierenden Systems erforderlich ist. Eine existenzielle Triade ist für die Kreierung eines Systems mit dem entsprechenden <i>systemkonstituierenden Konzept</i> ausreichend, wenn die 'Relation' in dieser Triade - grundsätzlich <i>stochastisch</i> ist, und - <i>statistisch</i> einem bestimmten Gesetz gehorcht (im allgemeinen Fall dem PKR - dem Prinzip des kleinsten Ressourcenverbrauchs). Die Evolution dieses Systems folgt dem Charakter der 'Relation' in der existenziellen Triade.
universelle existenzielle Pentade	Eine Form, die notwendig und hinreichend ist, um die abstrakte Struktur eines <u>beliebigen</u> Systems (und damit einer <u>beliebigen</u> beobachtbaren Entität) zu beschreiben, unabhängig von Inhalt und Zweck dieses Systems und den Prinzipien, die dieses System lenken. Die universelle existenzielle Pentade ist das gesamte Schema selbst, wie auf Figure 1 in [5], Part A, Chapter I ch. 3.2 "Enmorphya" dargestellt, d. h. alle fünf Elemente des Schemas und die Beziehungen zwischen diesen Elementen, d. h.: - das <i>Substrat</i> des primären Systems, - die <i>Eigenschaften</i> des primären Systems, - die <i>Relation</i> des primären Systems = das <i>Substrat</i> des Metasystems,

⁴⁷ der Originalbegriff von A. Uemow: 'системообразующий концепт'

⁴⁸ der Originalbegriff von A. Uemow: 'структурный фактор'

⁴⁹ der Originalbegriff von A. Uemow: 'субстрат системы'

Begriff	Definition
	<p>- die Eigenschaften des Metasystems (Enmorphie der Relation) und - die <i>Relation</i> des Metasystems.</p> <p>Die <i>existenzielle Pentade</i> ist <u>universell</u> und <u>vollständig</u>.</p>
Information	eine Änderung des Grades der Unbestimmtheit
Informationsmetabolismus	<p>Die existenziell notwendige Aufnahme und Verarbeitung von Signalen aus der Umwelt durch das System und die Reaktion des Systems auf diese Signale.</p> <p>Informationsmetabolismus ist nicht nur dem Menschen, sondern auch jedem <i>quasi-stochastischen</i> System eigen, weil dessen <i>Adaptationsmechanismus</i> ohne den Austausch von Signalen mit der Umwelt und deren Verarbeitung nicht funktionieren kann.</p> <p>Das Konzept des „Informationsmetabolismus“ wurde von Antoni Kępiński als Parallele zum Energiemetabolismus des Körpers eingeführt (Antoni Kępiński <i>Psychopatologia nerwic (Psychopathology of Neuroses)</i>, 1972).</p>
essentiell notwendige Fragen	<p>Fragen, <u>Antworten</u> auf welche für die Umsetzung des Prinzips der Selbsterhaltung des Systems notwendig sind.</p> <p>In diesem Sinne können ‘<i>essentiell notwendige Fragen</i>‘ auch als ‘<i>existentiell notwendige Fragen</i>‘ bezeichnet werden.</p>
Adaptation	<p>Anpassung einer systeminternen ‘Norm’ (ihre Änderung, Abschaffung, Schaffung einer Neuen) als Ergebnis von der Wirkung von <i>Rückkopplung</i>.</p> <p>Der <i>Adaptionsmechanismus</i> umfasst Mechanismen für</p> <ul style="list-style-type: none"> - <i>Monitoring (Überwachung)</i> des Systemzustands (der auch von den Umweltbedingungen abhängt) - systeminterne <i>Korrektur (Korrekturmaßnahmen)</i> in Bezug auf den sich ändernden Systemzustand, und - <i>Prävention (Vorbeugung)</i> eines ähnlichen ‘suboptimalen‘ Systemzustands durch Korrektur einer entsprechenden, systemimmanenten ‘Norm’. <p>Diese Mechanismen sind dem System immanent. Der ‘suboptimale‘ Zustand eines Systems ist der Zustand, der die Stabilität seines <i>systemkonstituierenden Konzepts</i> gefährdet.</p> <p>Die Kombination von <i>Monitoring-</i> und <i>Korrekturmechanismen</i> wird oft als <i>Feedback-Mechanismus</i> oder <i>Rückkopplungsmechanismus</i> bezeichnet.</p> <p>Bei <i>quasi-stochastischen</i> Systemen sind alle drei Mechanismen vorhanden und müssen aktiv sein.</p> <p>Bei <i>echt-stochastischen</i> Systemen, die über kein <i>Langzeitgedächtnis</i> verfügen, kann der <i>Präventionsmechanismus</i> nicht funktionieren, weil das <i>Langzeitgedächtnis</i> notwendig ist, um die systeminterne ‘normative Basis’ zu speichern. Folglich ist der <i>Adaptationsmechanismus</i> für <i>echt-stochastische</i> Systeme äquivalent dem <i>Feedback-Mechanismus</i> (nur <i>Monitoring</i> und <i>Korrektur</i>).</p>
Ressource (eines Systems)	das Produkt aus ‘der Anzahl der Schritte auf dem Weg von Zustand A nach Zustand B’ und ‘der Anzahl der alternativen

Begriff	Definition
	<p>Entscheidungen/Möglichkeiten bei jedem dieser Schritte’.</p> <p>Die Ressource des Systems kann abstrakt als das Produkt zweier kategorial komplementärer Begriffe dargestellt werden:</p> <p style="text-align: center;">‘Ressource’ = ‘Aktion’ * ‘Wahl’,</p> <p>siehe Einzelheiten in [5], CHAPTER VII, ch. 2.3.2 “Complementary Terms as Resource”.</p> <p>Die konkrete Umsetzung der ‘Schritte auf dem Weg von Zustand A nach Zustand B’ und der ‘alternativen Entscheidungen/Möglichkeiten bei jedem dieser Schritte’, d. h. die konkrete Umsetzung von ‘Aktion’ und ‘Wahl’ ist in jedem System spezifisch und muss für jedes System gesondert definiert werden⁵⁰.</p> <p>Bei physikalischen Systemen ist die ‘Ressource’ beispielsweise die Anzahl der Wirkungsquanten, die erforderlich sind, um das System in einen anderen gegebenen makroskopischen Zustand zu versetzen⁵¹; bei der Kommunikation (einschließlich der kommunikativen Funktion der Sprache) - die Anzahl der Positionen in einer Nachricht (Text) * die Anzahl der verschiedenen Zeichen (z. B. Buchstaben und Satzzeichen), die erforderlich sind, um einen bestimmten Inhalt zu vermitteln; für den Bildungsbereich - und für jeden anderen sozialen Prozess - die Anzahl der einzelnen (Lern-)Themen * die Anzahl der alternativen (didaktischen) Methoden, die zur Erreichung eines gegebenen (Lern-)Ziels in Betracht gezogen bzw. angewendet werden müssen.</p>
das Prinzip des Kleinsten Ressourcenverbrauchs (PKR)	<p>Das Prinzip der Entwicklungsdynamik eines jeden Systems, das darin besteht, dass ein System beim Übergang von Zustand A nach Zustand B im statistischen Mittel eine solche Art des Übergangs von A nach B umsetzt, bei der die ‘Ressource’ des Systems am geringsten verbraucht wird.</p> <p>Das PKR stellt eine universelle Information-über-die-Steuerung-der-Relation dar (d. h. es ist integraler Bestandteil der Enmorphie der Relation) und regelt den Prozess der Wechselwirkung zwischen dem <i>Substrat</i> und dem <i>Strukturfaktor eines jeden</i> Systems – z. B. physikalischer, sozialer, kommunikativer Natur - das auf einem <i>stochastischen</i> Prozess beruht.</p> <p>Insbesondere regelt das PKR den Prozess der Wechselwirkung zwischen Materie und Information in der Natur in Form des Prinzips der maximalen Entropie, das dem Prinzip der kleinsten Wirkung äquivalent ist, vgl. [5], CHAPTER VII, ch. 2.1.5 “The Principle of Least Resources Consumption” und 2.3.2 “Complementary Terms as Resource”.</p>
das Prinzip der Selbsterhaltung des Systems (PSE)	<p>Das Prinzip der Stabilisierung eines jeden Systems, das darin besteht, dass die <u>Abweichung</u> des Systems von der Befolgung des Prinzips des kleinsten Ressourcenverbrauchs durch die Tatsache begrenzt wird, dass das systemkonstituierende Konzept dieses Systems stabil und erhalten</p>

⁵⁰ Die Anzahl der ‘Schritte auf dem Weg von Zustand A nach Zustand B’ muss > 0 sein und die Anzahl der ‘alternativen Entscheidungen/Möglichkeiten bei jedem dieser Schritte’ muss > 1 sein. Der Grund dafür ist, dass die Natur mehr als Null Ressourcen aufwenden muss, um einen beobachtbaren Zustand zu schaffen. Dafür ‘muss’ die Natur mindestens 1 ‘Schritt zu einem anderen Zustand’ machen, und ‘alternative Entscheidungen/Möglichkeiten bei jedem dieser Schritte’ können nicht deterministisch sein. Deshalb muss die Anzahl der Alternativen > 1 sein; siehe [5], CHAPTER VII, Abschn. 2.1.3, 2.1.4, 2.3.2 für weitere Einzelheiten.

⁵¹ d. h. die physikalische Größe ‘Wirkung’ (kg·m²·s⁻¹)/h (die Planck-Konstante ist der Wert des Wirkungsquantums)

Begriff	Definition
	<p>bleibt.</p> <p>Das Prinzip der Selbsterhaltung des Systems gilt für <u>jedes</u> System, d. h. es ist ein universeller Bestandteil ihrer Enmorphie. Bei <i>echt-stochastischen</i> Systemen geschieht dies automatisch aufgrund ihrer 'Markow-Natur', welche an sich schon die stochastisch 'aus der Reihe tanzenden' Systeme auf den Pfad der maximalen Entropie zurückbringt.</p> <p>Für <i>quasi-stochastische</i> Systeme gibt es keinen solchen Automatismus. Deswegen muss sein Fehlen durch explizite, diesem System immanente Mechanismen, die zur Erhaltung des Systems beitragen, kompensiert werden. Solche (systemimmanenten) Mechanismen werden durch den <u>Adaptationsmechanismus innerhalb des Systems selbst</u> realisiert.</p>
das Prinzip der Größten Auswahl	<p>Das Prinzip der Minimierung der einschränkenden Faktoren auf die Möglichkeiten der Entscheidungsfindung, das Prinzip der Maximierung der Wahlfreiheit.</p> <p>Das Prinzip der größten Auswahl ist eine der Eigenschaften des Selbstbewusstseins der Lebewesen, die zu ihrer Flexibilität und <u>Anpassungsfähigkeit</u> an die verschiedenen Bedingungen ihrer Existenz führt.</p>
Enmorphie ⁵² von etw.	<p>ein spezieller Begriff für das Konzept 'Information-über-die-Steuerung-von-etw.', z. B. 'Enmorphie der Relation'.</p> <p>Das Unterscheidungsmerkmal zwischen den Begriffen 'Information' und 'Enmorphie' besteht darin, dass 'Information' mit einem <u>materiellen Substrat</u> interagiert, während 'Enmorphie' mit <u>einer Relation, einem Prozess</u> zwischen dieser 'Information' und diesem materiellen <u>Substrat</u> interagiert.</p>
stochastischer Prozess	ein Prozess, bei dem jeder nächste Zustand mit einer anderen Wahrscheinlichkeit als 0 und 1 eintritt.
stochastisches System	ein System, dessen Strukturfaktor auf einem stochastischen Prozess beruht
deterministischer Prozess	<p>ein Prozess, dessen jeder nächste Zustand durch seinen gegenwärtigen Zustand <u>eindeutig definiert ist</u>, d. h. jeder nächste Zustand mit der Wahrscheinlichkeit 1 eintritt.</p> <p>Das bedeutet, dass jeder vorherige Zustand des Prozesses auch eindeutig aus seinem gegenwärtigen Zustand berechnet werden kann.</p> <p>Wenn der nächste Prozesszustand mit der Wahrscheinlichkeit 0 eintritt, dann ist der Prozess gestoppt und existiert nicht mehr; er fällt ebenfalls unter die Definition des deterministischen Prozesses.</p>
deterministisches System	ein System, dessen Strukturfaktor auf einem deterministischen Prozess beruht
Markow-Eigenschaft (eines <i>stochastischen Prozesses</i>)	jeder nächste Zustand des Markow'schen <i>stochastischen Prozesses</i> , der reguläre Markow-Ketten implementiert, hängt probabilistisch <u>allein</u> von seinem aktuellen Zustand ab und ist unabhängig von seinen

⁵² Der Begriff 'Enmorphie (Emorfia, Enmorphy)' ist auf der Grundlage des Griechischen konstruiert: ἐνμορφία (ἐν-μορφή-α => (Einbringen) in-Form, (приведение) в-форму)

Begriff	Definition
	<p>vorherigen Zuständen. Diese Eigenschaft kann auch folgendermaßen ausgedrückt werden: Die Vergangenheit der <i>echt-stochastischen</i>, d. h. Markow'schen Systeme beeinflusst ihre Zukunft ausschließlich durch ihre Gegenwart.</p>
echt-stochastischer Prozess	<p>Ein <i>stochastischer Prozess</i>, der die <i>Markow-Eigenschaft</i> besitzt.</p> <p><u>Die 'echte Stochastizität' ist das Fehlen eines unmittelbaren Gedächtnisses</u> für frühere Zustände: Der nachfolgende Zustand hängt probabilistisch nur vom aktuellen Zustand ab.</p> <p>Die <i>Enmorphie</i> der Relation ist <u>nicht variabel</u> (immer das Prinzip der kleinsten Wirkung ohne variable Merkmale).</p>
quasi-stochastischer Prozess	<p>Ein <i>stochastischer Prozess</i>, der <u>keine Markow-Eigenschaft</u> hat.</p> <p><u>Quasi-stochastische Systeme müssen ein unmittelbares Langzeitgedächtnis</u> für frühere Zustände besitzen.</p> <p>Die <i>Enmorphie</i> der Relation ist <u>variabel</u> (immer das Prinzip des kleinsten Ressourcenverbrauchs mit variablen Merkmalen und das <i>Prinzip der Selbsterhaltung des Systems</i> mit einem <i>Adaptationsmechanismus</i>).</p> <p>N.B.: <i>Quasi-stochastische</i> Prozesse sind <u>nicht deterministisch</u>.</p>
Willensträger	<p><u>jedes quasi-stochastische</u> System, d. h. ein stochastisches System mit <i>Wahlfreiheit</i>, das alle seine früheren Erfahrungen berücksichtigt und über einen <i>Adaptationsmechanismus</i> verfügt. Mit anderen Worten: Ein <i>Willensträger</i> ist ein <i>adaptives</i> System mit <i>Wahlfreiheit</i>.</p>
Sozium	<p>eine soziale Entität (Entität), eine Gruppe von <i>Willensträgern</i>, ein sozial verbundenes System der wechselwirkenden <i>Willensträger</i>, eine Gesellschaft beliebiger Größe, die durch beliebige interne Beziehungen zusammengehalten wird.</p>
kategoriale Komplementaritäten	<p>Es sei eine endliche Menge von Begriffen, die mehr als einen Begriff umfasst. Wir nennen die Begriffe in dieser Menge <i>kategorial komplementär</i> zueinander, wenn:</p> <ol style="list-style-type: none"> 1) Diese Begriffe ausschließlich gemeinsam existieren können, d. h. die Existenz eines Begriffs notwendigerweise die Existenz aller anderen Begriffe in der Menge bedingt, und 2) Ein Begriff in der Menge durch einen beliebigen Satz von anderen Begriffen in dieser Menge nicht definiert werden kann.
attributive Gegensätze	<p>Es sei eine endliche Menge von Eigenschaften, die mehr als eine Eigenschaft enthält. Eigenschaften aus dieser Menge werden als <i>attributive Gegensätze</i> bezeichnet, wenn jedes Element dieser Menge nur <u>einen bestimmten Extremwert ein und desselben Attributs</u> darstellt und deshalb durch ein anderes Element dieser Menge definiert werden kann.</p> <p>Bei der Unterscheidung zwischen <i>attributiven Gegensätzen</i> (z. B. {hoch, niedrig}) und <i>kategorialen Komplementaritäten</i> (z. B. {Form, Inhalt}) sei gesagt, dass attributive Gegensätze grundsätzlich keine kategorialen Komplementaritäten sind, da jedes Element eines attributiven Paares durch ein anderes Element dieses Paares definiert</p>

Begriff	Definition
	<p>werden kann. Zum Beispiel kann das Attribut ‘Größe’ die extremen Werte {groß, klein} haben; diese Werte können durch einander ausgedrückt werden.</p> <p>Attributive Gegensätze beschreiben immer Eigenschaften/Qualitäten, d. h. <u>Werte eines Attributs</u>, aber niemals Begriffe. Dabei erfolgt die Änderung des Wertes dieses Attributs beim Übergang von einem Extrem zum anderen ohne ‘Sprünge’, d. h. ohne Änderung des Symmetriegrades (ohne ‘Phasenübergänge der 2. Ordnung’).</p> <p>Attributive Gegensätze implizieren häufig das Vorhandensein eines Maßstabs, eines Etalons, d. h. einer ‘Norm’, worauf sich die Schätzung des Wertes des jeweiligen Attributs bezieht (z. B. {teuer, billig}, {gut, böse}).</p> <p>Attributive Gegensätze werden in der Sprache fast immer durch antonymische Paare wiedergegeben, während <u>kategoriale Komplementaritäten bei weitem nicht immer als solche darstellbar sind</u>.</p>
Zeit	<p>Die Unterscheidbarkeit der Mikrozustände der Natur voneinander IST der Lauf der Zeit (d. h. die Zeit selbst).</p> <p>Deshalb ist die Zeit diskret.</p> <p>Die Unterscheidbarkeit von Zuständen ist eine notwendige Voraussetzung für ihre Beobachtbarkeit, d. h. für ihr Sein. Deshalb sind Sein und Zeit bijektiv miteinander verbunden.</p> <p>Siehe [5], CHAPTER VII, ch. 1.3 “Time Microstructure”.</p>
Vergangenheit	<p>Die aufgezeichnete/dokumentierte Menge von Zuständen (Ereignissen), die stattgefunden haben.</p> <p>Deshalb ist die Vergangenheit deterministisch, siehe [5], CHAPTER VII.</p>
Gegenwart	<p>Entscheidung über die Wahl des nächsten Zustands aus einer Menge möglicher Zustände.</p> <p>Die Gegenwart verwandelt eine probabilistische Zukunft in die deterministische Vergangenheit. Es ist diese Komplementarität der probabilistischen Zukunft und der deterministischen Vergangenheit, die die Ursache der <i>Irreversibilität</i> der Zeit ist, siehe [5], CHAPTER VII.</p>
Moment (Augenblick)	<p>ein theoretischer Begriff, der einen ‘Zwischenzustand’ beschreibt, der in der Natur nicht realisierbar ist.</p> <p>In einem solchen ‘Zwischenzustand’ besteht bereits die Möglichkeit der Wahl, aber die Auflösung dieser Alternative existiert noch nicht. Da die Zeit diskret ist, kann es keine ‘Zwischenzustände’ von Entitäten geben.</p> <p>Diese Definition macht den ‘Moment’ und damit die Gegenwart zu einem relativen und nicht absoluten Begriff.</p>
Zukunft	<p>eine Vielzahl von möglichen Zuständen.</p> <p>Deshalb ist die Zukunft probabilistisch, siehe [5], CHAPTER VII.</p>
Gedächtnis	<p>Die Eigenschaft, Informationen (sowohl rationale als auch emotionale, falls für ein bestimmtes System zutreffend) für einen Zeitraum zu speichern, der über einen gegebenen Zustand des Systems (<i>Moment</i>, <i>Situation</i>) hinausgeht, so dass diese gespeicherten Informationen <u>mehr als einen</u> nachfolgenden Zustand (<i>Situation</i>) dieses Systems direkt beeinflussen können.</p> <p>Ein solches Gedächtnis kann auch als ‘<i>Langzeitgedächtnis</i>’ bezeichnet werden.</p> <p>Das <i>Langzeitgedächtnis</i> ist ein notwendiges Attribut des</p>

Begriff	Definition
	<p><i>quasi-stochastischen</i> Prozesses.</p> <p>In diesem Zusammenhang ist das ‘<i>Kurzzeitgedächtnis</i>’ die Eigenschaft, Informationen (sowohl rationale als auch emotionale, falls für ein bestimmtes System zutreffend) für einen Zeitraum zu speichern, der den gegebenen Zustand des Systems (<i>Moment</i>, Situation) nicht überschreitet, so dass diese gespeicherten Informationen nicht mehr als einen - den nächsten - nachfolgenden Zustand (Situation) dieses Systems direkt beeinflussen können.</p> <p>Das <i>Kurzzeitgedächtnis</i> erfüllt die <i>Markow-Eigenschaft</i> und ist ein notwendiges Attribut eines <i>echt-stochastischen</i> Prozesses.</p>
Geschichte	<p>Die Abfolge der Entwicklungsphasen des <i>quasi-stochastischen</i> Systems, d. h. des <i>Willensträgers</i>, auf den sich diese ‘Geschichte’ bezieht.</p> <p>Die vollständige Geschichte des <i>Willensträgers</i> umfasst den gesamten Entwicklungszyklus des entsprechenden <i>quasi-stochastischen</i> Systems von seiner Entstehung bis zu seiner Selbstzerstörung. Dieser vollständige Entwicklungszyklus gilt für <u>jedes</u> <i>quasi-stochastische</i> System.</p>
Raum	<p>Ein diskretes <i>Substrat</i>, das zur <i>Unterscheidung</i> zwischen <u>materiellen</u> Entitäten benötigt wird, siehe [5], CHAPTER VII, ch. 3 “Space Microstructure”.</p>
Enmorphotyp (eines Lebewesens)	<p>Die Gesamtheit aller Attribute der ‚<i>Enmorphie</i> des Selbstbewusstseins‘ eines Lebewesens, die in Wechselwirkung mit seinem Genotyp und Phänotyp steht.</p>
freier Wille	<p>Der freie Wille ist die Wahlfreiheit, die nicht-deterministisch ist, aber keinen Markov-Prozess darstellt und zumindest alle früheren Erfahrungen des Systems berücksichtigt.</p> <p>D.h. er ist eine gewisse Wahlfreiheit, eine Möglichkeit der lokalen Abweichung des <i>quasi-stochastischen</i> Prozesses vom Prinzip des Kleinsten Ressourcenverbrauchs.</p> <p>Der Entscheidungsprozess.</p>
Risikoreflexion (nur von Menschen) (Unbestimmtheit des Möglichen (der Zukunft))	<p>Einbeziehung in die Entscheidungsfindung, d.h. in die Wahlfreiheit des <u>Menschen</u>, einer Selbstreflexion möglicher zukünftiger Zustände, die sowohl die den Menschen umgebende Welt als auch den Menschen selbst, einschließlich seiner eigenen Endlichkeit als System, einschließen.</p>

7 Referenzen

- [1] Авенир Иванович Уёмов *Системные аспекты философского знания*, Одесса, 2000⁵³
- [2] Katharina Zweig *Ein Algorithmus hat kein Taktgefühl*, Heyne, 2019, ISBN 978-3-453-20730-1
- [3] Manuela Lenzen *Künstliche Intelligenz*, C.H.Beck, 2020, ISBN 978-3-406-75124-0
- [4] Janelle Shane *You Look Like a Thing and I Love You*, 2019 Janelle Shane, ISBN 978-1-47226-899-0
- [5] Igor Furgel *Being and Systemacy*, Westarp BookOnDemand, 1. Auflage 2022, ISBN: 978-3-96004-131-3

8 Danksagung

Ich möchte meiner Frau Irina für unsere wertvollen und interessanten Diskussionen zu verschiedenen Aspekten dieses Themas danken.

Ebenso möchte ich Avenir I. Uemow, meinem damaligen Philosophieprofessor an der Universität, meine größte Anerkennung aussprechen. Er hat maßgeblich dazu beigetragen, wie ich mit der Welt interagiere.

⁵³ Avenir Iwanowitsch Uemow (ausgesprochen: Awenir Ujomow) *Systemische Aspekte der philosophischen Kenntnis*, Odessa, 2000