

## Rules of the Game.

### Über den formenden Einfluss eines technischen Regelbegriffs

#### *Abstract*

Das Denken in Form von Regeln, wie es Digitalisierung und Algorithmisierung erfordern, ist ein Denken mit spezifischen Exklusionen – was angesichts des Erfolgs der damit hervorgebrachten Technologien leicht vergessen werden kann. Diese Exklusionen verschieben sich auch nicht angesichts des Wandels, den dieses Regeldenken durch das Aufkommen maschineller Lernverfahren erfährt – eine Möglichkeit, die etwa Hubert Dreyfus in seiner KI-Kritik noch zugestand. Der Beitrag diskutiert einen technischen Regelbegriff als Teil einer mit Ernst Cassirer als Form verstandenen Technik, stellt also die Frage, wie dieser Begriff der Regel Auffassungen von im Lichte ihrer Algorithmisierbarkeit betrachteten Prozessen formiert. Nach einer allgemeinen Entfaltung des Begriffs wird diskutiert, was sich durch lernende Systeme gegenüber sogenannter symbolischer KI (nicht) ändert, um schließlich am Beispiel des Sprachverständnisses (das bereits Kern von Dreyfus Kritik war) mit einem Rekurs auf Ludwig Wittgensteins Konzept des Gebrauchs und Thomas Kuhns Paradigmenbegriff herauszustellen, was es ist, das sich der durch den Regelbegriff implizierten Differenzlogik von Regularität und Chaos entzieht: Erst auf Basis der Anerkennung eines Paradigmas können Ableitungsregeln und Daten als deren Relata überhaupt als solche erkannt werden.

Thinking in the form of rules, as digitalization and algorithmization demand it, is always subject to specific exclusions – a fact which tends to be forgotten given the success of the technologies made possible by this form of thinking. These exclusions do not shift, even though rule-based thinking is changed by the advent of machine learning. In his early critique of AI, Hubert Dreyfus still thought such a shift to be possible. The article discusses a technological concept of rule as part of technology taken (with Ernst Cassirer) as a form, asking how such a concept of rule forms the perception of processes in regard of their algorithmization. After generally unfolding the concept of rule, the article discusses what does and what does not change with the advent of learning systems opposed to so-called symbolic AI. Finally, using the example of language comprehension (which was already the core of Dreyfus' critique), recurring to Wittgensteins concept of use and Thomas Kuhns notion of paradigm, it is pointed out what eludes the difference of regularity and chaos implied by the concept of rule: It is only after acknowledging a paradigm that inference rules and data as their relata can be recognized and defined.

Den Fragen, die aktuelle Technik aufwirft, wird man Ernst Cassirer zufolge nicht gerecht werden können, wenn man Technik bloß als Menge von Apparaturen begreift; stattdessen ist Technik aufzufassen als eine Form, mit deren Hilfe der Mensch die Welt erschließt: Wie Sprache, Religion, oder Wissenschaft stellt sie ein symboli-

sches Repertoire bereit, das der Umwelt Bedeutung verleiht.<sup>1</sup> Damit werden bestimmte Sichtweisen ermöglicht, andere – und das kann in Cassirers insgesamt eher optimistisch gestimmtem Aufsatz leichter überlesen werden – verstellt.<sup>2</sup> Prägend für die heutige technologische Situation sind ohne Zweifel Digitalisierung und Algorithmisierung: Immer mehr industrielle, dienstleistende, aber auch kognitive und epistemische Prozesse erweisen sich als algorithmisch besser, gleichwertig oder wenigstens zufriedenstellend substituierbar. Herstellungs- und Kreationsvorgänge jedweder Art zu beschreiben, ihre Funktionalität zu erschließen, sie im Zuge von Digitalisierungsmaßnahmen technologisch nachzubauen, erfordert bestimmte formale Darstellungsfiguren. Diese agieren auf unterschiedlichen Niveaus: Bis ein Prozess visuell mittels Flowcharts, implementiert als Quellcode, oder simpler in Form einer prosaischen Schritt-für-Schritt-Anleitung dargestellt werden kann, muss er bereits mit abstrakteren Mitteln beschrieben worden sein. Eine solche abstrakte Beschreibungsfigur ist die Regel. Im Folgenden wird untersucht, was die Betrachtung von Prozessen unter einem technologisch geprägten Begriff der Regel bedeutet, d.h. es wird punktuell beleuchtet, welche Konsequenzen Technik als Form im cassirerschen Sinne zeitigt. Insbesondere gilt es herauszustellen, was diesem speziellen Betrachtungsinstrument notwendig entgeht.

Die Angabe von Verfahrensregeln eines Prozesses bedeutet seine Demystifikation und damit den entscheidenden Schritt hin zur Ermöglichung von Beschreibungs- und Nachbaumaßnahmen. Das heißt, das Problem der Regelangabe stellt sich nur bei in ihrem Ablauf unbekannten Vorgängen. Mathematische Lösungsstrategien, implementierte Software, automatisierte Industrieprozesse usw. sind bereits als Regelabfolge beschrieben. Ebenso ist ein Spiel wie Schach in diesem Sinne regulär – gewissermaßen ist es durch sein Regelwerk definiert. Anders verhält es sich dagegen mit dem *Schach-Spielen*: Die Regularität der Vorgänge, mit der ein Mensch seine Spielentscheidungen fällt, ist weit weniger offensichtlich, nicht einmal notwendigerweise vorhanden, denn ob der Regelbegriff ein geeignetes Beschreibungsinstrument dieser kognitiven Vorgänge darstellt, ist keineswegs klar (dazu unten).

---

1 Ernst Cassirer: »Form und Technik«, in: Ernst Wolfgang Orth und John Michael Krois (Hg.): *Ernst Cassirer. Symbol, Technik, Sprache. Aufsätze aus den Jahren 1927–1933*, Hamburg 1985, S. 39–91.

2 Trotz der finalen Bestimmung des Sinns der Technik als »Freiheit durch Dienstbarkeit« (Ebd., S. 89) sieht Cassirer aber ein absolutistisches Risiko: »Sie beharrt nicht nur auf ihrer eigenen Norm, sondern sie droht diese Norm absolut zu setzen und sie den anderen Gebieten aufzuzwingen.« (Ebd., S. 78).

## 1. Zum technischen Regelbegriff

Um Konsequenzen der Anwendung des Beschreibungsinstruments der Regel feststellen zu können, muss (es kann ja nicht anders sein) zunächst der Begriff entfaltet werden. Zuerst können Regeln präskriptiver und deskriptiver Natur gesondert werden: Präskriptive Regeln sind bspw. juristische und moralische »Soll«-Vorschriften, aber auch Programmierbefehle, die Verfahrensschritte auf einem Input festlegen. Deskriptive Regeln sind hauptsächlich statistische Aussagen, welche die Erwartbarkeit von Ereignissen beschreiben, also eine Art Befunde: »In der Regel« wird man nicht im Lotto gewinnen, aber im Laufe des Lebens 1,5 Kinder zeugen, eine Räuber-Beute-Population wird sich gemäß einer gewissen Funktionskurve entwickeln, usw. Das bedeutet, dass eine Regel nicht durch den Fall ihrer Nichteinhaltung entkräftet wird, sondern auch Wahrscheinlichkeitsaussagen Regeln sein können. Sie können aber nur dann Regel genannt werden, wenn mehr als bloße statistische Struktur vorliegt, diese Struktur nämlich aus einem (bspw. in einer mathematischen Gleichung) verallgemeinerbarem Muster resultiert. In einem Datensatz werden immer zufällige Korrelationen zwischen Merkmalen auftreten, Häufungspunkte vorhanden sein, er wird sich mehr oder weniger gut durch bestimmte Funktionen beschreiben lassen, etc. Solche Korrelationsmuster können aber nur dann als Regel bezeichnet werden, wenn sie in irgendeiner Weise systematisch bedingt sind, und sich nicht mit Aufnahme eines weiteren Datums als reine Zufallsprodukte entlarven lassen. Das heißt, dass ein Vorgang nur dann als regulär bezeichnet werden kann, wenn er aus einer systematisch bedingten statistischen Struktur resultiert. Die Chancen auf den Lottogewinn müssen nicht empirisch bestimmt und entsprechend je nach gewähltem Datensatz angepasst werden, sondern lassen sich kombinatorisch exakt auf  $\binom{49}{6}$  bestimmen.<sup>3</sup>

Eine solche zugrundeliegende Systematik muss nicht notwendigerweise bekannt sein: Es genügt implizite Vorhandenheit einer Regularität, wobei es praktisch beliebig schwer sein kann, zu entscheiden, ob man es mit impliziter Regularität oder zufälliger Korrelation zu tun hat. Präskriptive und deskriptive Regeln können kreuzklassifikatorisch jeweils implizit und explizit vorliegen.<sup>3</sup> Als Regularität werde ich die Prägung durch beliebige Regeln im Sinne dieses Schemas bezeichnen – sie ist damit prinzipiell graduell denkbar, je nachdem, wie stark die Prägung ist und welche nicht-regulierten Freiheitsgrade vorliegen.

Die Ergebnisse von Herstellungs-/Kreationsvorgängen, allgemein: Prozessen, lassen sich als Datensatz, der notwendig eine statistische Struktur aufweist, beschreiben

---

<sup>3</sup> Am wenigsten explizit mag wohl eine präskriptive, aber nur intuitive Regel sein. Außerhalb des Bereichs des Technischen könnte es sich dabei aber bspw. um nirgends ausformulierte, latente moralische Intuitionen handeln.

ben: Prozess-Input x führt (ggf. mit Wahrscheinlichkeit p) zu Prozess-Output y.<sup>4</sup> Soll ein Prozess automatisiert/digitalisiert/algorithmisiert werden, müssen präskriptive Regeln gebildet werden, welche die impliziten, deskriptiven Regeln, denen die statistische Struktur folgt, abbilden. Mit anderen Worten: Es muss ein Algorithmus formuliert werden, der die gleiche Input-Output-Beziehung wie der zu substituierende Prozess realisiert.<sup>5</sup> Die dafür benötigten präskriptiven Regeln können in einem solchen Algorithmus explizit ausbuchstabiert sein, müssen das aber nicht notwendigerweise: Mit maschinellen Lernverfahren kann die Explikation mithilfe statistischer Optimierung anhand implizit bleibender Regeln umgangen werden (dazu unten). Entscheidend für die unten vorgenommene Kritik wird sein, dass Regeln stets Relationen beschreiben: Sie sind Verbindungen von Stellen. Obwohl die Form dieser Verbindung relativ frei ist, ist die Voraussetzung der Existenz fixierter Stellen keine unschuldige Annahme.<sup>6</sup>

## 2. Prozesse mit Regeln beschreiben – klassische KI und maschinelles Lernen

Die Explikation vermuteter impliziter Regularitäten ist also – abgesehen von Mitteln des *machine learning*, auf die später eingegangen wird – die entscheidende Aufgabe der Formalisierung von Prozessen. So wird beispielsweise immer wieder probiert (freilich noch ohne direkten Bezug zur KI), Regularitäten von Forschungsverfahren der Geisteswissenschaften aufzudecken. Zwei typische Beispiele solcher Versuche sind die Ansätze Max Benses und Friedrich Kittlers. Sie sind im Kontext dieses Aufsatzes deshalb illustrativ, weil sie offenbaren, dass die Festlegung von (expliziten) Regeln eine Reduktion ist, die notwendig etwas ausstreckt. Bense versuchte eine Quantifizierung der Beurteilung des ästhetischen Werts eines Kunstwerks in Form einer deskriptiven, expliziten mathematischen Regel, die er als numerischen Quotienten aus Ordnung und Komplexität,  $M_A : = O/C$ , ausdrückte. Der Ansatz hat allerdings viele Probleme, die vor allem aus teils fragwürdigen Fixierungen resultieren: Höhere Komplexität reduziert das ästhetische Maß, eine Variable der Berechnung ist

- 
- 4 Der Prozessbegriff ist hier absichtlich so weit gefasst, dass damit keineswegs nur industrielle Produktionsvorgänge, sondern auch kognitive Vorgänge oder wissenschaftliche Methoden umfasst werden können. Die Terminologie einer Input-Output-Relation hat hier bereits ganz eigene Voraussetzungen und Konsequenzen, die diskutiert werden sollten. Dieser Text tut das dort, wo diese ebenfalls in Bezug zum Begriff der Regel stehen.
  - 5 Das ist eine rein theoretische Beschreibung dessen, was geschieht, und in der Regel wohl nicht die Überlegung, mittels der eine Prozessautomatisierung praktisch angegangen wird. Darüber soll hier nicht gesprochen werden.
  - 6 Diese noch recht abstrakten Überlegungen zum technischen Regelbegriff genügen als Grundlage der folgenden argumentativen Punkte. Es wird daher darauf verzichtet, darzustellen, in welchen Formen diese Regeltypen in KI-Anwendungen (wie etwa den Inferenzmaschinen wissensbasierter Systeme) konkret verbaut werden.

die ›Freundlichkeit‹ eines Werks, deren Definier- und Quantifizierbarkeit fraglich erscheint, sodass bspw. Frieder Nake Benses Ideen als »hanebüchenen Quatsch«<sup>7</sup> bezeichnete. Ebenso erfolglos blieb Kittlers noch größer angelegtes Projekt der ›Austreibung des Geistes aus den Geisteswissenschaften‹: Shannons Entropieformel der Information, das Modell der universalen Turingmaschine und Theoreme der Psycho-linguistik scheinen weiterhin nicht hinreichend, den auslegungsbedürftigen Gegenstandsbereich der Geisteswissenschaften automatisch verarbeiten zu können und einen menschlichen Interpreten überflüssig zu machen. Kittlers Behauptung, mit Shannons Informationsmaß lasse sich ›der Erfolg der Medien‹ berechnen, steht vor den gleichen Schwierigkeiten wie Benses ästhetisches Maß: Sie versucht, ein kontextabhängiges, polyvalentes Konzept wie ›Erfolg‹ in einer expliziten Berechnungsvorschrift festzulegen, womit produktive Offenheit zugunsten numerischer Darstellbarkeit aufgegeben wird – Kittler hat dann vorab entschieden, was Erfolg bedeutet, und ihn damit seiner Anwendbarkeit in verschiedenen Kontexten beraubt.<sup>8</sup> Ein dem Erkennen solcher Eindeutigkeiten vorgängiges Anerkennen wird dabei verschwiegen (siehe unten).

Die beschriebenen Formalisierungsversuche verfolgen den Ansatz, geschlossene Formeln zu finden, das heißt explizite Regeln zu formulieren. Damit treffen sie die gleichen Vorannahmen wie die frühe KI der 50er/60er Jahre, die davon ausging, Prozesse menschlicher Intelligenz digital nachbauen zu können, wenn nur die richtigen Regeln und Grundeinheiten gefunden werden. Regeln dienen dabei häufig der Desambiguierung einer Situation, also der Reduktion der Anzahl überhaupt in Frage kommender Aktionen. Im Schachspiel erkennt ein Mensch auf einen Blick, dass eine Figur gefährdet ist und weiß, dass das wichtig ist, womit Spielzüge intuitiv priorisiert sind. Im Sprachverständnis kann sich die Phrase ›bleib in meiner Nähe‹ je nach Kontext auf Distanzen von weniger als einem Meter (ein Kind an die Hand nehmen) oder vielen Kilometern (für einen Berufswechsel nicht weit umsiedeln) beziehen. Was ein Mensch anscheinend ohne geistigen Aufwand und ohne explizite Regelkenntnisse entscheidet, muss regelbasierten Algorithmen als Berechnungsvorschrift einprogrammiert werden. So bezeichnet Hubert Dreyfus im Rahmen seiner einflussreichen Kritik ›klassischer‹ KI die Annahme, dass das immer und für alle Leistungen der Intelligenz möglich sei, als die ›metaphysische Hypothese‹ (oft auch

7 Zitat unvollständig ausgewiesen in David Oswald: »Max Bense und die Informationsästhetik«, in: David Oswald u.a. (Hg.): *Rückblicke. Die Abteilung Information an der hfg ulm*, Lemgo 2015, S. 116–122.

8 Freilich tut man Kittler ein wenig unrecht, diese Randbemerkung mit dem gleichen spitzfindigen Argwohn wie Benses detailliert ausgearbeitetes ästhetisches Maß zu betrachten. Es gehört aber zur rhetorischen Strategie Kittlers, provokante und radikale Äußerungen beiläufig einzustreuen (man denke an den ›sogenannten Menschen‹); ihre Analyse darf aufgrund dieser Beiläufigkeit nicht weniger scharf ausfallen.

»ontologische Hypothese«) der »Good old fashioned AI« (GOFAI).<sup>9</sup> Diese konnte ihren (wie sich herausstellte übertriebenen) Optimismus nur entwickeln, indem sie davon ausging, dass sich für intelligente Prozesse explizit angebbare, einem Algorithmus einspeisbare, also letztlich mathematische Regeln formulieren lassen, was Dreyfus einer ausführlichen Kritik unterzog. Für die metaphysische Hypothese zentral war/ist also die Unterstellung fixer Grundelemente bzw. Relata (im Sinne eines logischen Atomismus), deren Zusammenwirken sich in relationalen Strukturen beschreiben lässt – bspw. mit Thomas Kuhns Paradigmenbegriff lässt sich zeigen, dass eine solche Unterstellung problematisch ist (dazu unten).

Trotzdem lassen sich auf diese Weise bekanntlich viele Prozesse so nachbilden, dass das Substitut genauer, schneller, besser arbeitet als das intelligenzbasierte Original: Schachcomputer; in Dreyfus Buch noch ein Beispiel für die verzweifelte Suche nach den richtigen Entscheidungsregeln der Großmeister, sind heute für Menschen nicht mehr systematisch zu schlagen. Auch machen solche regelbasierten Systeme weiterhin Fortschritte, wenn auch häufig – ganz im Einklang mit Dreyfus Diagnosen – die tatsächlichen Leistungen nicht mit den Versprechungen mithalten können: So soll IBMs Watson als wissensbasiertes System, also eine Wissensdatenbank versehen mit einer regelbasierten Abfragelogik, wahlweise medizinische Diagnosen stellen können oder, in der Variante Ross, Juristen mindestens zuarbeiten können; letztlich handelt es sich aber um avancierte Suchmaschinen.

Nun ist es die große Neuerung statistischer Verfahren des maschinellen Lernens, wie künstlicher neuronaler Netze als wohl bekanntestes Beispiel, dass ein solcher Prozessnachbau mittels expliziter Regeln nicht mehr unbedingt nötig ist. Mit der statistischen Struktur eines Trainingsdatensatzes, d.h. Datenpaaren aus einem Input und einem gewünschten Output, passt ein Lernalgorithmus seine numerische Komposition iterativ so an, dass er sodann unbekannte Inputs gemäß dieser in die geforderte Kategorienlogik einsortieren kann.<sup>10</sup> Der stochastische Trainingsprozess, im Falle neuronaler Netze bestehend aus iterativem *Feed Forward* und Fehler-Backpropagation, bildet also eine Zuordnungsfunktion, die für jeden Input einen im Sinne einer Kategorisierung zu interpretierenden Zahlwert generiert. Das ist keine so starke Einschränkung der behandelbaren Probleme, wie man intuitiv meinen könnte: Immer mehr Probleme werden mit *machine learning* bearbeitet, Sprach- und Bilderkennung, die Selektion von Spielzügen, erst recht das Treffen von Entscheidungen (soweit es in einer Auswahl von Möglichkeiten besteht, wie bspw.: Person ist kreditwürdig oder nicht, Bewerber erfüllt die Einstellungsvoraussetzung zu x %, Situation

9 Hubert L. Dreyfus: *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge 1992, S. 62.

10 So zumindest beim überwachten Lernen. Beim unüberwachten Lernen wird auf vorgegebene Outputs im Trainingsdatensatz verzichtet und die Kategorienlogik soll ebenfalls vom Algorithmus anhand der statistischen Struktur gebildet werden. Darauf wird hier nicht näher eingegangen.

erfordert Reduktion der Fahrzeuggeschwindigkeit um y km/h) lassen sich allesamt als Zuordnungsprobleme auffassen. Die im Datenmaterial verhafteten Kriterien für eine Zuordnung sind und bleiben dabei unbekannt: Der Lernalgorithmus benötigt keine Auskunft darüber, welche Merkmale des Inputs er zur Generierung der Outputwerte heranziehen soll, und kann auch im Nachhinein keine Auskunft über die von ihm letztlich verwendeten Merkmale geben.<sup>11</sup> Das bedeutet, Regeln müssen nicht explizit bekannt sein, wohl aber muss das zu lösende Problem impliziter Regularität unterliegen. Die Hoffnung ist dann, dass sich die implizite Regularität, auch wenn nicht explizit bezeichbar, in der statistischen Struktur der Trainingsdaten ausreichend abzeichnet, um vom Lernverfahren abgebildet werden zu können und eine sinnvolle, nicht auf Zufällen beruhende Zuordnung erlaubt. Da es keine letztendliche Gewissheit darüber geben kann, ob implizite Regularität genutzt wird oder doch ein häufig passender Zufall, den das nächste Datum schon als solchen und die Entscheidungsfindung damit als Idiotie entlarven könnte, lassen sich diesbezügliche Restzweifel nie vollständig ausräumen, sondern nur mit geschickt gewählten Kontrolldaten minimieren.

*Machine learning* ist damit von vielen von Dreyfus Kritikpunkten an GOFAI nicht oder nicht in gleichem Maße betroffen. Er selbst gesteht lernenden Verfahren in Einleitungen neuerer Ausgaben seines Buchs höhere Potentiale zu und nimmt sie von seiner Kritik teilweise aus.<sup>12</sup> Insbesondere ist die seiner Meinung nach unhaltbare metaphysische Hypothese der Symbolisierbarkeit aller Geistesvorgänge nicht mehr erforderlich: ML-Verfahren kommen ohne Explikation sämtlicher Regeln aus, ihnen reicht theoretisch ihre implizite Vorhandenheit. Letztere ist aber unumgänglich (ohne sie ist kein erfolgreiches statistisches Training möglich), sodass der Regelbegriff auch für auf ihnen beruhende Systeme fundamental bleibt.

Es wäre ein Fehler, daraus zu schließen, dass ML-Verfahren nicht mehr symbolisch wären. Das legt die eingebürgerte Unterscheidung zwischen symbolischer KI und konnektionistischen Systemen für klassisch-regelbasierte Ansätze und statistische Lernalgorithmen (bzw. enger: künstliche neuronale Netze) nahe. Statt von konnektionistischen wird mitunter von »subsymbolischen« Verfahren, die »subsymbolische« Daten verarbeiten, gesprochen.<sup>13</sup> Das ist insofern irreführend, als man es bei

11 Forschung im Bereich der sogenannten Explainable AI (XAI) beschäftigt sich damit, inwieweit man doch Auskunft über die genutzten Merkmale gewinnen kann. Dafür werden beispielsweise Eigenschaften des Datenmaterials getilgt und beobachtet, wie sich das erlernte Verhalten eines Algorithmus dadurch ändert. Siehe hierzu etwa Alan B. Tickle u.a.: »The truth will come to light: Directions and challenges in extracting the knowledge embedded within trained artificial neural networks«, in: *IEEE Transactions on Neural Networks* 9 (2015), Heft 6, S. 1057–1068.

12 Dreyfus: *What Computers Still Can't Do*, S. XXX.

13 Steve K. Esser u.a.: »Cognitive computing systems: Algorithms and applications for networks of neurosynaptic cores«, in: *The 2013 International Joint Conference on Neural Networks* (2013), S. 1–10.

KI und Algorithmen letztendlich notwendigerweise mit mathematischen Verfahren zu tun hat, und das bedeutet: mit Symbolmanipulation. ML-Systeme betreiben Statistik und Optimierung, was natürlich zwangsläufig auf Basis von Zahlen, also Symbolen, stattfindet. Die technische Implementierung dieser Verfahren erfordert eine symbolische, und zwar numerische Darstellung der Eingabedaten. Überhaupt sind Daten immer schon erhoben und bereits dadurch dargestellt, das heißt symbolisiert – subsymbolische Daten (als eigenständige, ontologisch absolute Größen) kann es nicht geben, wenn man die Begriffe ernst nimmt. Auch ML-Systeme betreiben also, weil es gar nicht anders sein kann, symbolische Operationen auf symbolischen Daten. Hier geht folglich auch Dreyfus einen Schritt zu weit, wenn er behauptet, »neural networks abandon representation altogether«.<sup>14</sup> Repräsentation kann nicht umgangen werden.

### *3. Zwischen Regularität und Chaos – was dem technischen Regelbegriff entgeht*

Damit ergibt sich nun hinsichtlich der algorithmischen/digitalen Modellierung bzw. funktionsäquivalenten Substitution von Prozessen die folgende Systematik: Regelbasierte Systeme sind angewiesen auf explizite, ausformulierte Regeln. ML-Verfahren benötigen keine expliziten Regeln, sondern ihnen genügt subkutanes Vorhandensein von Regularität (soweit deren implizite Regeln sich statistisch ausreichend deutlich niederschlagen, sodass sie vom jeweiligen Lernalgorithmus auch tatsächlich erlernt werden können). Hat man es dagegen mit einem rein chaotischen Zufallsprozess zu tun, so ist dieser unbeherrschbar, was klar und theoretisch unproblematisch ist. Unweigerlich stellt sich nun die Frage, ob damit alle Optionen erschöpft sind, oder ob es noch etwas ganz Anderes, nicht *zwischen* Regularität (implizit oder explizit) und Chaos Liegendes, geben kann, also einen Strukturtyp, der mit dem Begriff der Regel und der aus ihm folgenden Differenz Regularität/Chaos schlicht nicht erfasst werden kann.

Mit einer spontanen Überlegung könnte man zu dem voreiligen Schluss gelangen, dass dem nicht so ist. Erste Kandidaten für weder regulär noch chaotisch zu nennende Bereiche sind sicherlich solche, denen nachgesagt wird, sie erfordern so etwas wie die Erfahrung eines langjährig tätigen Experten, menschliches Einfühlungsvermögen, oder Intuition, Instinkt, gar Genie; Tätigkeiten, für die man ›ein Gefühl‹ entwickeln muss. Attestieren würde man ein solches Gefühl wohl Ärzten oder Therapeuten, die einen Menschen mit ›geschultem Auge‹ auswerten und dadurch irgendwie ›ganzheitlicher‹ erfassen, als eine Maschine das könnte; Intuition und Instinkt findet man bei einem Stürmer im Fußball, der ›einen Riecher‹ für Torchancen hat.

---

14 Vgl. Dreyfus: *What Computers Still Can't Do*, S. XXX.

Genie benötigt der Mathematiker, der einen Beweis *sieht*, den andere nicht gesehen haben, oder auch der Poet, der Lyrik von anscheinend nicht erklärbarer Schönheit generiert. Solche Beschreibungen sind Black-Box-Metaphern für Produktionsprozesse, die nicht systematisch erklärt sind, sondern anscheinend ausschließlich von besonderen Individuen mit besonderen Fähigkeiten hervorgebracht werden können. Damit ist die vermeintliche Regularität fraglos nicht explizit. Nimmt man aber an, diese besonderen Fähigkeiten der besonderen Individuen beruhen nicht auf einer impliziten Regularität des jeweiligen Bereichs, so müssten sie doch wohl reines Raten sein. Denn was genau soll ein Experte, ein Profi, ein Genie denn instinktiv oder durch intuitive Eingabe erkennen, wenn nicht implizite Regelzusammenhänge? Der Mensch muss diese Regeln sicherlich nicht explizieren können, aber man sollte doch hoffen, dass sie vorhanden sind und Ärzte etc. tatsächlich aufgrund antrainierter, also einigermaßen unsystematisch erzeugter Systematik, zu Diagnosen gelangen. Existieren keine solchen Regelzusammenhänge, liegt der Gedanke nahe, dass auch ein Arzt nur raten kann – und das scheint, bei aller berechtigten Kritik an Gesundheitssystemen, nicht der Fall zu sein. Dann gibt es aber keinen prinzipiellen Unterschied mehr zu einem trainierten ML-System, das ebenfalls Zuordnungen aufgrund impliziter statistischer Regularitäten vornimmt. Hinsichtlich der Performanzunterschiede lässt sich freilich immer auf dem Algorithmus noch fehlende Kontextdaten, die der Mensch wahrnimmt, verweisen, aber das ist kein Argument für eine fundamental differente Verfahrensweise. Es scheint dann nur diese Alternativen zu geben: Entweder ein Mensch ist in der Lage, zu erklären, wie er ein Resultat erzeugt hat; ist diese Erklärung plausibel, wurde es offenbar systematisch erzeugt – man wird dann aber kaum von so etwas wie Genie sprechen wollen, schließlich könnte prinzipiell jeder diesen Produktionsprozess nachahmen. Oder ein Mensch kann nicht erklären, wie ein Resultat zustande kommt; dann basiert es entweder auf unbewusst erkannten Regularitäten, oder es ist geraten. Im ersten Fall könnte man von Genie sprechen; schließlich gibt es keine klare Methode, mit deren Befolgung man es ihm gleich tun könnte, sondern der Mensch hat einfach eine besondere Begabung für intuitive Mustererkennung oder ähnliches – dann gäbe es aber keinen prinzipiellen Unterschied zwischen Genie und *machine learning*, was hieße, dass Genie sich durch hochdimensionale, aber nicht übermäßig komplizierte Differentialrechnung modellieren ließe.

Nur, diese spontane Überlegung beinhaltet bereits (als *Petitio Principii*) ihre eigene Hypothese, dass es nichts neben oder zwischen Regularität und Chaos gibt, und hat somit ihren blinden Fleck genau in der Unterstellung des technischen Regelbegriffs. Um nun die dadurch bedingten Schranken aufzuzeigen kann es nicht hinreichen, momentane Verfehlungen der an Regeln orientierten technologischen Systeme aufzuzählen; so geriete man gegenüber technologischer Entwicklung in ein permanentes Rückzugsgefecht. Stattdessen muss, wenn der Regelbegriff mit Cassirer als

Teil der Form der Technik behandelt werden soll, aufgezeigt werden, was beim *Nachdenken über* Vorgänge in einem formal-regelgeleiteten Betrachtungsmodus notwendig nicht gedacht wird.

Besonders deutlich lassen sich die Schwierigkeiten an Sprachverständens-Prozessen festmachen.<sup>15</sup> Während Wilhelm von Humboldt solche noch auf Basis generativer Regeln verstand (wie es Noam Chomsky später erneut tut),<sup>16</sup> die sich also prinzipiell angeben ließen, verhält sich Ludwig Wittgenstein, welcher die Problematik der Regularität von Sprachverständen prominent behandelte, dazu zwiegespalten. Er formuliert ein berühmtes Paradox: »Unser Paradox war dies: eine Regel könnte keine Handlungsweise bestimmen, da jede Handlungsweise mit der Regel in Übereinstimmung zu bringen sei.«<sup>17</sup> Eine Regel, die nur für endlich viele Fälle gelernt werden kann, muss auf weitere Fälle angewendet werden können. Wie kann diese Erweiterung, die ja im Prinzip einer neuen Metaregel, welche die Anwendbarkeit der ersten Regel reguliert, bedarf, funktionieren? »Aber erklärst du ihm wirklich, was du selber verstehst? Lässt du ihn das Wesentliche nicht erraten? Du gibst ihm Beispiele, – er aber muss ihre Tendenz erraten, also deine Absicht«.<sup>18</sup> Eine Regel ist dann wie ein »Wegweiser«,<sup>19</sup> d.h. man kann ihr folgen, muss es aber nicht. Die Regulierung mit Metaregeln, deren Anwendbarkeit ja wieder reguliert werden muss, führt letztlich in einen Regress: »Aber wie hilft mir dann eine Erklärung zum Verständnis, wenn sie doch nicht die letzte ist? Die Erklärung ist dann ja nie beendet; ich verstehe also noch immer nicht, und nie, was er meint!«<sup>20</sup> Diese vieldiskutierten<sup>21</sup> Probleme der Wittgenstein-Interpretation können hier nicht vertieft werden, illustrieren aber eine der Schwierigkeiten, welche die Betrachtung von Sprachverständen als Regelbefolgung mit sich bringt: Eine Regel muss in etwas resultieren – ihre Anwendung benötigt Start- und Zielpunkte. Unter ihrem Diktat stellen sich auch Bedeutungsprozesse als relationale Strukturen dar. Das Problem, das Wittgenstein aufzeigt, lautet: Relationen zwischen was? »Jedes Zeichen scheint allein tot. Was gibt ihm Leben? Im Gebrauch lebt es.«<sup>22</sup> Erst die praktische Verwendung konstituiert das Zeichen, schärfer noch: „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.«<sup>23</sup> Es kann

15 An deren Automatisierung in der gesamten Geschichte der KI großes Interesse geherrscht hat, von noch recht simplen Dialogsystemen wie Joseph Weizenbaums ELIZA bis hin zu heutigen Chatbots und digitalen Assistenten wie Apples Siri und Googles Cortana.

16 Jürgen Habermas: »Vernünftige Freiheit. Spuren des Diskurses über Glauben und Wissen«, in: *Auch eine Geschichte der Philosophie*, Bd. 2, Berlin 2019, S. 455.

17 Ludwig Wittgenstein: *Philosophische Untersuchungen*, Oxford 1953, § 201.

18 Ebd., § 210.

19 Ebd., § 85.

20 Ebd., § 87.

21 Schon ob das Paradox tatsächlich ein Paradox ist, ist umstritten, siehe z.B. Saul A. Kripke: *Wittgenstein on Rules and Private Language*, Cambridge 1982 und als Kritik daran Gordon P. Baker und Peter M. S. Hacker: *Scepticism, Rules and Language*, Oxford 1984.

22 Wittgenstein: *Philosophische Untersuchungen*, § 432.

23 Ebd., § 43.

somit nicht einfach als fixer Input einer Zuordnungsfunktion gedacht werden, sondern existiert erst mit seinem *Gebrauch* (der es ist). Was als algorithmisches Input-datum in Frage käme, wird erst mit einer anerkannten Verwendungspraxis konstituiert. Dieser Vorgang der Konstituierung durch ein praktisches Dispositiv kann mit der Ableitung von Bedingungen für das, was als Datum relevant ist, aus einem Paradigma im Sinne Kuhns gedacht werden (siehe unten).

Der Zweifel an solchen Fixierbarkeiten innerhalb des Sprachverständens wurde durch den Poststrukturalismus ausgearbeitet zu einem radikalen Prozessdenken, das Festlegungen grundsätzlich misstraut. Besonders deutlich tritt dieser Gedanke in Derridas Konzept der Dissemination des Sinns hervor. Wenn Zeichen und ihre Bedeutungen sich erst durch einen laufenden Gebrauchsprozess konstituieren, bedeutet ein Abbruch dieses Prozesses eine Setzung, die notwendig etwas abwürgt: »Denn wenn Sie sagen ›dies ist das‹, ››dies‹ heißt ›das‹‹ [...]», so essentialisiert, substantiellisiert und immobilisiert genau die Form ihres Satzes, das dem Heißen (vouloir-dire) vermahlte ›ist‹ (est), den Text. Seine Bewegung wird damit auf eine Reihe von Stanzen reduziert und seine Schrift auf eine thematische Übung.«<sup>24</sup> Einen Text auf eine bestimmte Weise zu verstehen bedeute immer die Ausstreichung von Differenzen, das Beenden von Bewegungen, also dessen, was laut Derrida den Text-Sinn erst hervorbringt. Diese Reduktion auf ›Stanzen‹ ist aber genau das, was relationale Regeln voraussetzen müssen: Relationale Regeln, ob durch statistisches Training oder ›von Hand‹ gebildet, unterstellen die Vorgängigkeit einander zuordenbarer Elemente (Inputs und Outputs des Algorithmus). Dass die Elemente erst in einem Auslegungsvorgang entstehen, und ihre Fixierung ihre essentielle Natur zerstört, wie es die Dissemination postuliert, kann so nicht gedacht werden.

Die Dissemination ist nur unzureichend als Prozess im oben genannten Sinne eines Produktionsverfahrens zu fassen, da sie nicht mit der Erzeugung eines Ergebnisses endet, sondern in unabgeschlossener Bewegung verbleibt. Das macht sie als Leitmotiv interpretativer Verfahren, die konkrete Aussagen über Textbedeutungen treffen wollen, problematisch.<sup>25</sup> Ungeachtet dessen, ob man Derridas Konzeptionen plausibel findet, sollten sie aber zumindest *gedacht* werden können, und eine strenge Orientierung an einem technisch-relationalen Regelbegriff macht das schwierig – von ihm muss man sich lösen, um die Dissemination in ihrer Radikalität überhaupt gedanklich zuzulassen.

Können Sprachverständens-Prozesse praktisch funktionieren ohne Regeln, die den Verweisungszusammenhang von Begriffen und Bedeutungen regulieren? Einen Hinweis darauf kann der bereits mehrfach angesprochene Paradigmenbegriff Thomas

24 Jacques Derrida: *Dissemination*, Wien 1995, S. 397.

25 Siehe etwa den humoristischen Artikel von Chip Morningstar: »How To Deconstruct Almost Anything. My Postmodern Adventure«, in: *FUDCO*, Juni 2013, <http://www.fudco.com/chip/deconstr.html> (aufgerufen: 27.3.2020).

Kuhns geben. Unter einem neuen Paradigma verändert sich Kuhn zufolge nicht einfach die Interpretation ansonsten stabiler Daten, sondern was Daten sind folgt erst aus dem Paradigma: »What occurs during a scientific revolution is not fully reducible to a reinterpretation of individual and stable data. In the first place, the data are not unequivocally stable. A pendulum is not a falling stone, nor is oxygen dephlogisticated air.«<sup>26</sup> Fixe Daten und Ableitungsregeln zwischen diesen bestehen höchstens innerhalb eines Paradigmas, und zwar *als Folge aus* diesem Paradigma. Das bedeutet, bevor Daten *erkannt* werden, wird ein Paradigma *anerkannt*. Dreyfus führt diese kuhnsche Argumentation für Systeme der GOFAI zwar an, übersieht aber in der Folge, dass sie sehr viel weiter reicht und für ML-Systeme bzw. jedes algorithmische Verfahren nichts von ihrer Gültigkeit verliert. Diese bilden zwar ihre Ableitungsregeln automatisch, aber eben anhand vordefinierter Daten, die wiederum – wenn man Kuhn folgt – aus einer paradigmatischen Konstellation (des Anerkennens) folgen. Eine solche Konstellation funktioniert grundlegend anders als ein anleitendes Set von Regeln:

»They can, that is, agree in their identification of a paradigm without agreeing on, or even attempting to produce, a full interpretation or rationalization of it. Lack of a standard interpretation or of an agreed reduction to rules will not prevent a paradigm from guiding research. Normal science can be determined in part by the direct inspection of paradigms, a process that is often aided by but does not depend upon the formulation of rules and assumptions. Indeed, the existence of a paradigm need not even imply that any full set of rules exists.«<sup>27</sup>

Ein anerkanntes Paradigma funktioniert somit als leitendes Beispiel, ohne dass notwendig Vorgaben existieren, wie sich zu diesem Beispiel zu verhalten ist.<sup>28</sup> Anerkennen eines Paradigmas bedeutet somit keineswegs klare Strukturierung der zu bearbeitenden Fragen und auch nicht Verzicht auf Kritik, sondern lediglich einen gemeinsamen Interessenfokus, vor dessen Hintergrund und unter dessen Orientierung man die Aufgabe des Erkennens, d.h. der Suche nach Regeln, Daten und Fakten, erst beginnen kann.

---

26 Thomas S. Kuhn: *The Structure of Scientific Revolutions*, Chicago, London 2012, S. 121.

27 Ebd., S. 44.

28 Eine ähnliche Konstruktion, die sich direkt dem Problem des Sprachverständens widmet, verwendet Eleanor Roschs Prototypensemantik: Dieser zufolge entscheiden Menschen über kategoriale Zuordnungen von Begriffen weniger anhand festgelegter Kriterien der jeweiligen Kategorie, sondern anhand eines ›besten Beispiels‹ und der nicht durch Regeln definierten, sondern ad hoc verstandenen Abweichung von diesem. So ist ›Flugfähigkeit‹ eine wichtige Eigenschaft von Vögeln, verhindert aber nicht, dass Pinguine zu den Vögeln zählen – sie werden aber als ›weniger vogelhaft‹ aufgefasst als ein Rabe. Siehe Eleanor Rosch: »Cognitive reference points«, in: *Cognitive Psychology* 7 (1975), Heft 4, S. 532–547.

#### 4. Schluss

Natürlich scheint das letzte Problem aus technischer Perspektive damit weiter ungelöst: In welcher Form sonst werden Handlungsdirektiven aus einem Paradigma abgeleitet? In dieser Frage zeigt sich das Potential der Form der Technik: Ihr Ziel der Instrumentalisierung führt zu einem epistemischen Zwang, unter dem man sich nicht mit unscharfen Erklärungen zufriedengeben kann, sondern der auf strenge, funktionale Formalismen angewiesen ist und nicht aufhören kann, diese zu suchen. Unzureichend bleibt sie dort, wo die Existenz von Formalismen nicht gegeben ist bzw. die Annahme ihrer Existenz die Perspektive bereits einengt – Derridas Dissemination ist in technischer Begrifflichkeit nicht vollends zu denken.<sup>29</sup> Die Produktivität liegt gerade in der Unvereinbarkeit solcher Ansätze, die sich so wechselseitig ihre Auslassungen und Verkürzungen offenbaren, die sie aus sich heraus nicht beschreiben können.

Den Sinn des Technischen sieht Cassirer in der »Freiheit durch Dienstbarkeit«.<sup>30</sup> Diesem kann die Technik als Form nur gerecht werden, solange sie eine unter vielen Möglichkeiten ist, mit deren Hilfe sich der Mensch die Welt erschließt. Der technisch verstandene Begriff der Regel eröffnet so einen eigenen Blick auf Phänomene wie das hier beispielhaft andiskutierte Sprachverstehen. Der freiheitliche Aspekt geht aber verloren, wenn die technische Form diskursiv dominant wird. Das ist dann der Fall, wenn bspw. der komplette Mensch in technischen Begriffen gedacht werden soll: als »phyiscal symbol system«<sup>31</sup> oder »information processing system«,<sup>32</sup> oder eben jeder Prozess mittels formaler Regeln zu beschreiben versucht wird, ohne dass die damit bereits anerkannten Vorbedingungen dieser Beschreibungsform wahrgenommen werden.<sup>33</sup> Solche Beschreibungsversuche sind dann produktiv, wenn ihnen Alternativen gegenüberstehen, und sie führen zu perspektivischer Verengung, wenn diese Alternativen marginalisiert werden. Es ist die Aufgabe von Kunst, Geis-

29 Gleiches gilt bspw. für seinen Begriff der Gabe, die sich quasi selbst eliminieren muss, um außerhalb jedes Austauschs stehen zu können und damit wirkliche Gabe zu sein, siehe Matthias Flatscher: »Derridas «coup de don» und Heideggers ‚Es gibt‘. Bemerkungen zur Un-Möglichkeit der Gabe«, in: Peter Zeillinger und Matthias Flatscher (Hg.): *Kreuzungen Jacques Derridas. Geistergespräche zwischen Philosophie und Theologie*, Wien 2004, S. 34–54.

30 Cassirer: »Form und Technik«, in: Orth und Krois (Hg.): *Ernst Cassirer*, S. 89.

31 Terry Winograd zit. in Roger Schank, u.a.: »Panel on Natural Language Processing«, in: *IJCAI* 77 (1976), S. 1007–1013, hier S. 1008.

32 Joseph Weizenbaum: *Computer Power and Human Reason. From Judgment to Calculation*, San Francisco 1976, S. 226.

33 Christoph Hubig arbeitet unter dem Stichwort des Technomorphismus aus, dass Vorstellungen »vom Zustandekommen und der Verfasstheit der Welt« häufig mit solcher Selbstverständlichkeit eine technische Prägung aufweisen (»Vorstellungen vom Schöpfer, vom Weltplan, vom Prozess seiner Realisierung mit seinen internen Prozessen der Verwirklichung von Zuständen und Werken [...]«), dass sie allzu leicht übersehen wird. Siehe Christoph Hubig: *Die Kunst des Möglichen I. Grundlinien einer dialektischen Philosophie der Technik*, Bielefeld 2006, S. 77.

teswissenschaft und Philosophie, solche alternativen Beschreibungen den technischen entgegenzusetzen – wovon dann beide Seiten profitieren können.