

4 Interdisziplinäre Anknüpfungspunkte

Ein Anliegen dieser Analyse des maschinellen Lernens besteht darin zu prüfen, welche Beiträge die Informatik und die Technikphilosophie für eine interdisziplinäre Diskussion des maschinellen Lernens bieten können. Zu diesem Zweck wurden wiederholt Anknüpfungspunkte und möglicherweise verwandte Technikbereiche aufgezeigt um zu identifizieren, was die nächsten Stationen einer Analyse sein können.

Einige dieser offenen Fragen und Anknüpfungspunkte werden im Folgenden etwas detaillierter ausformuliert.

4.1 VERALLGEMEINERBARKEIT DER SUCHE NACH WELTTECHNIK

Eine bereits mehrfach aufgetretene Frage war diejenige nach der Verallgemeinerbarkeit der Beobachtungen und Beschreibungen des zweiten Hauptteils auf andere Technikbereiche. Zwar basiert etwa speziell die im zweiten Hauptteil entwickelte Vorstellung einer Technik, die den Nutzer bei der Erstellung von Weltbezügen unterstützt, auf der interdisziplinären Aufarbeitung und Darstellung des maschinellen Lernens im ersten Hauptteil, dennoch wurde die so gefasste Welttechnik nicht auf die Informatik begrenzt gedacht. In diesem Zusammenhang entsteht die Frage, welche Aspekte von Welttechnik spezifisch sind für den technischen Kontext, aus dem sie gewonnen wurden, und inwieweit bestimmte Aspekte verallgemeinerbar sind – etwa für neugierige MLA wie nicht vorstrukturierte künstliche neuronale Netze. Eine allgemeine Form von Welttechnik würde weiter auf komplexen iterativen Verfahren basieren, die sich imitativ oder selbstverstärkend verhalten. Solche iterativen Verfahren wiederum resultieren in Strukturen, die

aufbauend auf Sensordaten bestimmte Eigenschaften entwickeln und dabei zunehmend undurchsichtig beziehungsweise intransparent werden. Diese Art von Design oder Konstruktionsverfahren, die im Rahmen des maschinellen Lernens technisch realisiert ist, kann durchaus auch in anderen Ingenieurszusammenhängen auftreten¹. Die nächste Aufgabe besteht entsprechend darin, konkrete iterative Verfahren der genannten Art zu identifizieren und sie auf ihre Vergleichbarkeit mit maschinellem Lernen hin zu betrachten.

Ein Beispiel für einen Bereich von Technik, in dem die Grundoperationen der Steuer- und Regelungsprozesse ebenfalls nicht immer realisiert werden können, bildet die Nanotechnik. Hier werden ebenfalls mitunter mittels medialer Steuerung Umgebungen inszeniert um zu prüfen, welche Phänomene und Strukturen unter speziellen Rahmenbedingungen entstehen können. Der Hilfsbegriff zur Beschreibung dieser Vorgehensweise ist der der Inszenierung, der nicht zufällig auch bei Rheinberger Verwendung findet. Zwar wäre ein großer Schritt getan, wenn die Nanotechnik sich als Welttechnik verorten ließe, allerdings steht dem ein erhebliches Problem im Weg. *Die Nanotechnik* ist eine gleichermaßen unpräzise Beschreibung wie die Rede von *dem maschinellen Lernen* oder *gar der Informatik*. Hier fehlt eine systematische, interdisziplinäre Darstellung der technischen Grundlagen in Hinblick auf Selbstorganisation, wie sie diese Arbeit für das maschinelle Lernen darstellt. Dasselbe gilt auch für andere Technikbereiche von denen vermutet werden könnte, dass dort eine Suche nach Welttechnik sinnvoll sein könnte. Eine interdisziplinäre Darstellung dieser Technikbereiche kann nicht von den jeweiligen Disziplinen selbst verlangt werden, da aus einer rein disziplinären Perspektive die interdisziplinären und technikphilosophischen Fragestellungen nicht sehr präsent sind. Soll dementsprechend die Suche nach Welttechnik und die Diskussion und Analyse von Selbstorganisation nicht zurückgestellt werden, bis ein interdisziplinäres Projekt sich mit der relevanten Facette eines weiteren Technikbereiches beschäftigt, so müssen als Suchraum technikphilosophische Perspektiven genutzt werden, die einen größeren Allgemeinheitsanspruch und eine entspre-

1 Tatsächlich ist gerade das Auftreten solch eines Designs in der Informatik eher kontraintuitiv, da das Auftreten unvorhergesehener Strukturen beim Einsatz von zufallsfreien Algorithmen zunächst überrascht.

chend verringerte Auflösungskraft aufweisen². Den Vorzug können hierbei interdisziplinär geprägte, technikphilosophische Perspektiven erhalten, die zwar einerseits gegebenenfalls nicht ohne Weiteres auf maschinelles Lernen anwendbar sind, aber andererseits auch keine allgemeine Technikphilosophie darstellen, sondern primär Technikbereiche betrachten, die auch im Rahmen der Diskussion von Selbstorganisation betrachtet werden. Ein Beispiel für solch eine Perspektive, das bereits in der vorangegangenen Herleitung von Welttechnik betrachtet wurde, ist Rheinbergers Darstellung epistemischer Dinge. Im Weiteren werden in Kürze zwei weitere Perspektiven benannt, die die genannten Kriterien erfüllen und für eine weiterführende Analyse des maschinellen Lernens herangezogen werden könnten.

4.1.1 Technoscience als möglicher Suchraum

Eine noch nicht genannte Perspektive, die einen besonders vielversprechenden Suchraum für weitere Beispiele von Welttechnik darstellt, bietet die **TECHNOSCIENCE**. Die Technoscience betrachtet Dinge explizit als an der Erzeugung von Möglichkeiten und Weltbezügen beteiligt.

»Things demand intimacy, things are partners that afford possibilities, create successful connections to the world.«

(Nordmann et al. 2011, S. 2)

Gerade wenn anerkannt wird, dass spezielle Teile des maschinellen Lernens sich deutlich von der klassischen Technik unterscheiden, könnte versucht werden die Überlegungen zum Status von Nanotechnik als einer Technoscience auch auf MLA und auf Welttechnik abzubilden. In der Perspektive der Technoscience auf Dinge – wie eben auch Strukturvorschläge – ist insbesondere die Notwendigkeit des In-ein-Verhältnis-Setzens mit den Dingen mitgedacht.

2 Eine Steigerung der Komplexität der technischen Seite der Betrachtung, etwa ein Vergleich von Nanotechnik oder Genetik mit maschinellem Lernen, würde noch höhere Ansprüche stellen als eine direkte Analyse der Nanotechnik. Diese Steigerung der Komplexität wird schon daraus deutlich, dass nicht klar ist, was ein dem maschinellen Lernen als Teilbereich der Informatik vergleichbarer Teilbereich der Nanotechnik wäre.

»Affordances – a relational property that becomes salient only in respect to human purposes and material agencies.«

(Nordmann et al. 2011, S. 2)

Zwar kreieren mit Hilfe von Welttechnik erstellte Strukturvorschläge keine Verbindungen zur Welt, sondern basieren auf dem nutzerseitigen Wunsch nach einer Ersterstellung von Weltbezügen, aber diese Differenz scheint überbrückbar. Ein grundsätzlicheres Problem und gleichzeitig implizit ein konstruktiver Umgang mit dem Problembegriff der Selbstorganisation findet sich in der folgenden Kurzdefinition von Technoscience.

»Technoscience – the knowledge-production of homo faber that often uses scientific representations (e.g. theories, models, diagrams) to make things work.«

(Nordmann et al. 2011, S. 1)

In Hinblick auf Welttechnik ist fraglich, wie notwendig es ist, dass gerade einem neugierigen Umgang mit Dingen wirklich wissenschaftliche Repräsentationen zugrunde liegen.

Eine Aufarbeitung der Potenziale und Schwierigkeiten der Suche nach Welttechnik mit Hilfe der Perspektive der Technoscience kann und soll im Rahmen dieser Analyse nur angedeutet werden. Gleichwohl kann festgehalten werden, dass die Diskussion der Technoscience stark interdisziplinär geprägt ist und den Dialog mit techniknahen Experten sucht. Diese Ausrichtung und die techniknahe Argumentation lassen die Technoscience als einen besonders vielversprechenden Suchraum für die Identifikation von Welttechnik erscheinen.

4.1.2 Abduktionsstufen als Mittel zur Suche

Eine zweite Möglichkeit, nach dem Auftreten von Welttechnik zu suchen, basiert auf der Beschreibung des Agierens von Artefakten als der Erstellung logischer Schlüsse und besteht in der Analyse und Einordnung der jeweiligen Schlussweisen.

»This study embeds abductive search for new ideas within a broader framework concerning human cognitive activity than just reason-

ing as such. [...] Instead of conceptualizing human beings as processing information just in their heads inquirers are seen as fundamentally connected both to other inquirers (cultural resources, communities) and nature (world, indexical object of inquiry) through developing mediating artifacts and shared objects of activity collaboratively.«

(Paavola 2006, S. 70)

Die im ersten Hauptteil vorgenommene Darstellung von Konzepten zur Erstellung von Autoadaptionsprozessen reicht aus, um die Möglichkeit einer pauschalen Antwort auf die Frage nach der Schlussweise von Autoadaptionsprozessen als unwahrscheinlich und die Frage somit als durchaus interessant erscheinen zu lassen. Im Rahmen der Darstellung des analytischen Lernens wurde bereits angedeutet, dass die Frage nach der logischen Schlussweise von MLA auch innerhalb der Informatik gestellt wird. Die Begriffsverwendungen sind jedoch sehr uneinheitlich und zum Teil widersprüchlich³ (Flach et Kakas 2000; Kaminski et Harrach 2010).

Der Klärung dieser Begriffsverwendungen und der anschließenden Nutzung der Perspektive bei der Suche nach Welttechnik steht zunächst eine Gefahr im Wege, die ähnlich gelagert ist wie die Schwierigkeiten bei einer verstärkten Nutzung des Begriffes des Lernens. In beiden Fällen scheinen sich sofort starke Parallelen zwischen MLA und Menschen aufzudrängen.

»Der Versuch, den Denkprozeß in Analogie zum Verarbeitungsprozeß des Computers zu setzen, führt fast zwangsläufig zu der Frage nach der möglichen Funktion der Abduktion innerhalb automatischer Informationsverarbeitung.«

(Wirth 1995, S. 14)

Auch bei einer Betrachtung der Rede von Abduktion muss die Analyse sich dementsprechend dagegen absichern, Artefakten implizit menschliches

3 Eine Verbindung der Perspektive von Peirce und der Nutzung des Abduktionsbegriffes (Peirce 1878; Peirce et Walther 1991) bieten Costa und Wirth (Costa 2008; Wirth 2003, S. 19). Einen Versuch der Darstellung der Konsequenzen der unscharfen Begriffsbildung bietet Gottschalk (Gottschalk 1999).

Verhalten zu unterstellen. Die in der Welttechnik angelegte Trennung der Beiträge des unterstützenden Artefaktes und des Nutzers – der Präsentation und der Interpretation – bei der Erstellung von Weltbezügen lässt sich bei der Betrachtung des Autoadaptionsprozesses als eines logischen Schlusses nicht ohne Weiteres aufrechterhalten. Die Erstellung eines Strukturvorschlages, der eine gewisse Fähigkeit entwickelt, stellt nicht unbedingt einen vollständigen abduktiven Schluss gleich welcher Stufe dar. Die Beschreibung, dass die Eigenschaften von Strukturvorschlägen genutzt werden können, um menschliche abduktive Schlüsse zu ermöglichen und zu unterstützen⁴, kann als Reformulierung eines zentralen Aspektes von Welttechnik verstanden werden, muss jedoch nicht dazu führen, den noch nicht interpretierten Strukturvorschlag als das Ergebnis eines logischen Schlusses anzusehen. Dies gilt vor allem, wenn die zugrunde liegende Lernstrategie sehr stark neugierig geprägt ist und eine stark autoadaptive Vorstruktur nutzt, wodurch der Strukturvorschlag gerade keine klare Menge von Entscheidungsregeln darstellt.

Dennoch lassen sich die Vorgehensweisen von MLA bei einer impliziten Mitbetrachtung des menschlichen Anteils durchaus sehr produktiv als verschiedene Stufen abduktiver Schlüsse beschreiben (Kaminski et Harrach 2010). Insbesondere erleichtert es diese Perspektive, den eigentlichen Autoadaptionsprozess nicht zu Gunsten der Betrachtung des Strukturvorschlages zu vernachlässigen. Die Rede von einem abduktiven Schluss bezieht sich immer auf den Prozess der Erstellung des Strukturvorschlages und berücksichtigt insbesondere die Vorstruktur als Modell für die technischen Formalisierungen der jeweils eingesetzten Selbstorganisationsprinzipien. Gerade die Unterscheidung zwischen neugierigen und zielorientierten Artefakten profitiert von dem Fokus auf die Vorstruktur und der Diskussion, welcher STUFE ein abduktiver Schluss zugeordnet werden kann. Abduktive Schlüsse höherer Stufe zeichnen sich dadurch aus, dass im Vorfeld ein geringeres Maß an Informationen bekannt ist.

4 Menschliche Schlüsse können jedoch weiterhin nur in sehr optimierungsnahen beziehungsweise zielorientierten Ausnahmefällen vollständig ersetzt werden.

»All other forms [of abductive reasoning] are derived thereafter by removing the elements from the set of information given in the basic form one at a time.«

(Kaminski et Harrach 2010, S. 2)

Das bedeutet, zielorientierte Autoadaptionsprozesse würden am ehesten einem abduktiven Schluss entsprechen, der unter einer Reihe von denkbaren Regeln die im konkreten Kontext optimale Regel identifiziert und zum Einsatz bringt. Neugierige Autoadaptionsprozesse hingegen kommen in erster Linie in sehr vage bekannten Kontexten zum Einsatz, in denen eine mathematische Modellierung und Optimierung noch nicht möglich ist und eine autoadaptive Vorstruktur mit geringen Vorgaben zu Vorwissen erstellt werden muss.

4.1.3 Mehrwert einer weiteren Verallgemeinerung

Die hier angedeutete Diskussion von Methoden, die bei der weiteren Suche nach Welttechnik hilfreich sein könnten, zielte darauf ab, eine Begriffsbildung oder Perspektive zu finden, die einerseits ein größeres Feld von Technik betrachtet als nur die Informatik und die andererseits zumindest einige der wesentlichen Teilaspekte der Idee von Welttechnik oder neugieriger Artefakte thematisiert. Eine Vermutung bestand darin, dass eine solche Begriffsbildung aufgrund ihres allgemeineren Anspruches kein ausreichend großes Auflösungsvermögen aufweist, um bestimmen zu können, welche Teile des maschinellen Lernens als Welttechnik identifiziert werden können, dass jedoch der Gewinn von neuen oder präziser gefassten Suchräumen diese Schwäche aufwiegt. Die der Formulierung dieser Schwäche entgegengesetzte Denkrichtung wurde bisher nur implizit mitbetrachtet. Gemeint ist die Frage, ob die Möglichkeit, eine Diskussion maschinellen Lernens führen zu können, nicht generell einen optionalen Charakter hat, wenn andere Technikbereiche betrachtet werden sollen und bereits festgestellt wurde, wie maschinelles Lernen zu beschreiben ist. Weitergedacht kann sogar gefragt werden, ob schon die in dieser Analyse gewonnene Perspektive auf Welttechnik überhaupt noch in der Lage ist, zwischen den im ersten Hauptteil vorgestellten Lernstrategien zu unterscheiden. Tatsächlich ist die Perspektive der Welttechnik hierzu *nicht* in der Lage, denn die Betrachtung der Welttechnik setzt andere Schwerpunkte. Die Unterscheidung zwi-

schen zielorientierten und neugierigen Artefakten etwa hängt nur lose mit der technischen Unterscheidung zwischen einzelnen Lernstrategien zusammen.

Der Analyseschwerpunkt im zweiten Hauptteil wurde bewusst von der Klassifizierung des ersten Hauptteils gelöst und diese Flexibilität stellt auch kein Problem dar, denn der erste Teil sollte gerade einen interdisziplinären und technikphilosophischen Diskurs darüber ermöglichen, was an maschinellem Lernen interessant sein könnte. Entsprechend kann festgehalten werden, dass die Freiheiten bei einer Übertragung der Rede von Welttechnik auf andere Kontexte größer ist, als es im Kontext dieser techniknahen Grundlagenarbeit zu maschinellem Lernen zunächst wirkt. Die Fähigkeit zur Auflösung von Algorithmenklassen stellt in anderen Kontexten kein notwendiges Erfolgskriterium dar. Die Darstellung einer interdisziplinären Grundlagenforschung etwa zur Nanotechnik kann und sollte die technischen Konzepte und Intuitionen, die in dieser Analyse der erste Hauptteil vermittelt, komplett neu bestimmen.

Unabhängig davon wäre es speziell für eine Diskussion des maschinellen Lernens außerordentlich hilfreich, bestimmen zu können, was genau die nutzbaren Lernstrategien voneinander unterscheidet und vor allem, welche weiteren, noch nicht entdeckten Möglichkeiten es geben kann, auf Selbstorganisationsprozessen basierende Vorstrukturen zu entwerfen. Die Konzepte der Informatik zum Einsatz von Selbstorganisation spiegeln sich in den Kriterien zur Bewertung der Autoadaptionsprozesse oder Strukturvorschläge wieder. Die Messung von Performanz, Beschreibungslänge, Fitness, Signalstärke, Abständen, Wahrscheinlichkeiten, Aussagenlogik oder formalen mathematischen Fehlertermen setzt jeweils einen speziellen Zugang zu Selbstorganisationsprozessen voraus. Weiter besteht kaum Veranlassung zu vermuten, dass diese Liste und die korrespondierenden Lernstrategien einen Anspruch auf Vollständigkeit besitzen. Denkanstöße zur Erweiterung dieser Liste könnten durchaus aus einem interdisziplinären Diskurs mit Beteiligung der Technikphilosophie gewonnen werden. Hierzu wäre es erforderlich, andere Kontexte von Welttechnik zu identifizieren und zu beschreiben oder die Arbeit an den Problembegriffen der Selbstorganisation oder der Neugier weiter voran zu bringen und die Ergebnisse anschließend in den Kontext des maschinellen Lernen zurück zu übersetzten beziehungsweise zu übertragen.

4.2 TECHNIKFERNE EINORDNUNG VON MLA

In den vorangegangenen Betrachtungen wurde Technik stark isoliert von ihrem gesellschaftlichen oder sozialen Kontext betrachtet. Beispielsweise wurde bei der Positionierung zum Begriff des Zweckes danach gefragt, was mit dem Einsatz des MLA erreicht werden soll, ohne zu berücksichtigen, dass Zwecke mehrstufig angeordnet sein können⁵ oder dass eine Vielzahl praktischer Erwägungen und externer Einflüsse den Einsatz eines MLA motivieren können. Die Nichtberücksichtigung von Einflüssen dieser Art im Rahmen einer Grundlegung zum maschinellen Lernen soll an dieser Stelle nicht in Frage gestellt werden. Stattdessen soll eine technisch angemessene, interdisziplinäre Diskussion technikferner Themen durch diese Arbeit gerade ermöglicht werden.

Beispielsweise besteht ein Aspekt der Debatte über den sozialen Charakter von Technik, der in verschiedenen Beispielen dieser Arbeit eine Rolle spielte, in der Frage, inwiefern die Nutzer von Last.fm oder die Kunden eines Kaufhauses durch maschinelles Lernen geformt werden. Diese Debatte greift eine besonders interessante Fragestellung auf, da argumentiert werden kann, dass nicht nur MLA in Hinblick darauf entwickelt werden, vorweg zu nehmen was Menschen als potenzielle Nutzer benötigen. Unabhängig davon liegt es in Hinblick auf MLA nahe, explizit die ›Formung‹ der Nutzer beim Einsatz von zielorientierten, vorstrukturierten MLA und beim Einsatz von neugierigen, präsentierenden Artefakten zu vergleichen. Auch und gerade die Konzeption neugieriger MLA kann es mangels Zielen oder Zwecken mit sich bringen, dass die Nutzerwünsche erst aus einer Interpretation des Strukturvorschlages entstehen und damit aus Sicht des Nutzers in gewisser Weise vorgegeben werden – unabhängig davon, wie das Artefakt die konzeptionierte Aufgabe konkret löst. Auch wenn zielorientierte MLA eingesetzt werden, spielt es eine zentrale Rolle, zu entscheiden welche Konzepte von den Nutzern vorgegeben werden und Teil der Vorstruktur sein sollen und welche Ziele durch die Berücksichtigung der Rohdaten erreicht werden sollen. Die Übergabe von Rohdaten an ein MLA ist dementsprechend noch in viel allgemeinerer Hinsicht problematisierbar als bisher beschrieben. Im Fall der Betrachtung von Last.fm ist es etwa von In-

5 Dies wurde jedoch mitbetrachtet, als der MLA als ein gestiftetes Worumwillen beschrieben wurde.

teresse, welche Rohdaten von dem Nutzer vorgegeben wurden, der den Strukturvorschlag betrachtet und welche Eingaben von anderen Nutzern oder den Programmierern des MLA stammen.

Wie bereits betont wurde, sollen die Darstellungen der vorliegenden Analyse interdisziplinäre Diskurse des maschinellen Lernens ermöglichen und motivieren. Die entsprechenden Fragestellungen sollen auf Basis dieser Grundlagenforschung unter Vermeidung von technikfernen Pauschalisierungen und komplexitätsreduzierenden Verkürzungen diskutiert werden können.

4.3 MASCHINELLES LERNEN OHNE COMPUTER

Abschließend wird die bisherige Perspektive auf die im maschinellen Lernen zum Einsatz kommenden Formen von Selbstorganisation noch einmal erweitert. Es wird die These vertreten, dass die Diskussion der Konzepte hinter dem maschinellen Lernens nicht erst mit Hilfe der Informatik möglich wurde. Die Implementierung maschinell lernender Algorithmen durch die Informatik hat lediglich die Geschwindigkeit erhöht, mit der Autoadaptionsprozesse umgesetzt werden können.

Diese These basiert auf der Beobachtung, dass gewisse Formen maschinellen Lernens auch als menschliche Praxis ohne Nutzung eines Computers oder eines anderen technischen Systems möglich sind. Ein Beispiel für evolutionäres Lernen *ohne* Nutzung von Computern bildet die Weiterentwicklung einer Düse zur Herstellung von Seifenpulver in einer Fabrik bei Liverpool.

»Da [eine mathematische Optimierung] nicht klappte, versuchten [die Ingenieure] es auf andere Weise [...] günstige Abweichungen wurden beibehalten, ungünstige verworfen. Mit anderen Worten, es handelt sich um natürliche Selektion. [...] Man nehme eine Düse, die schon ganz gut funktioniert, und stelle Kopien davon her, die jeweils nach dem Zufallsprinzip ein wenig abgewandelt sind. Dann prüfe man, wie gut sie das Pulver herstellen. Anschließend zwingt man ihnen einen Kampf ums Dasein auf, indem man dafür sorgt, daß nicht alle überleben können. Viele der abgewandelten Apparate sind nicht besser (sondern oft sogar schlechter) als die Ausgangsform. Sie werden verworfen, und nur die wenigen, die bei der Erfül-

lung ihrer Aufgabe überlegen sind, dürfen sich fortpflanzen. Im Laufe der Generationen schält sich wie von Zauberhand ein neues, hervorragend funktionierendes Rohr heraus, das eine komplizierte, überraschende Form besitzt.«

(Jones 2002, S. 119)

Sofern der Selektionsprozess nicht gezielt beeinflusst wurde, entsprechen die von den Ingenieuren auf diese Weise erstellten Düsen neugierig erstellten Strukturvorschlägen. Die Aussagen, dass eine mathematische Optimierung im Vorfeld nicht möglich war⁶ und dass die Ergebnisse eine Überraschung darstellten, deuten darauf hin, dass solche Beeinflussungen tatsächlich nicht vorgenommen wurden. Die Rede von einer überraschenden Form legt zwar nahe, dass bezüglich des Aussehens der Düse Erwartungen bestanden, aber es kann vermutet werden, dass diese Erwartungen den Adaptionsprozess nicht beeinflusst haben.

Ein Beispiel für den Einsatz von Data Mining ohne Computer bildet das MUSIC GENOME PROJECT und, mit Einschränkungen, dessen kommerzielle Nutzung im Rahmen des Last.fm-Konkurrenten PANDORA. Das Music Genome Project stellt eine Datenbank dar, die Musiktiteln mit 150 bis 400 bewerteten Attributen eine vergleichsweise große Zahl von Eigenschaften zuordnet. Das Besondere an dieser Datenbank ist, dass die Attribute vollständig von menschlichen Experten zugeordnet und bewertet werden.

6 Die Düse diente dem Zerstäuben einer Flüssigkeit und die Berechnung des Verhaltens von Flüssigkeiten ist mathematisch extrem anspruchsvoll, wenngleich bei der Modellierung und Simulation inzwischen Fortschritte gemacht werden (Klöppel et al. 2011).

»Every time a new song comes out, someone on Pandora's staff — a specially trained musician or musicologist — goes through a list of possible attributes and assigns the song a numerical rating for each one. Analyzing a song takes about 20 minutes.

The people at Pandora — no relation to the alien planet — analyze 10,000 songs a month. They've been doing it for 10 years now, and so far they've amassed a database containing detailed profiles of 740,000 different songs. Westergren calls this database the Music Genome Project.«

(Grossmann 2010)

Die von den Mitarbeitern von Pandora erstellten Strukturvorschläge, in Kombination mit einer analytisch im Vorfeld bestimmten mathematischen Modellierung des Zusammenhangs der Attribute, stellen eine typische Anwendung maschinellen Lernens im Data Mining dar, verwenden jedoch erst im kommerziellen Online-Angebot von Pandora MLA. Pandora stellt den registrierten Kunden zusätzlich zu den Experteneinschätzungen und analog zu Last.fm auch Strukturvorschläge basierend auf Empfehlungen von anderen Nutzern zur Verfügung. Welche Attributausprägung ein Nutzer konkret bevorzugt, muss im Einzelfall festgestellt werden, aber diese Analyse entspricht einerseits einem nachgelagerten Schritt und kann andererseits relativ einfach über die Beantwortung eines umfangreichen Fragebogens festgestellt werden. Die kommerzielle Natur von Pandora verlangt hier aber nahe- liegenderweise nach einer komfortableren Lösung, die in Form von mathematischer Modellierung oder eben maschinellern Lernen zur Verfügung gestellt wird.

Die Vorgehensweise des Music Genome Projects stellt einen extrem zielorientierten Prozess dar, der sehr langfristig geplant werden musste, um die Verfügbarkeit von ausreichend vielen Experten sicherzustellen. Das Music Genome Project lässt sich sehr gut von der Webseite Pandora trennen⁷ und mit dem isoliert betrachteten Music Genome Project liegt ein beeindruckendes Beispiel für Data Mining ohne Einsatz eines maschinell lernenden Artefaktes vor. Weiter stellt das Projekt durch seine Kopplung an Pandora eine besonders transparente Veranschaulichung dafür dar, dass

7 Wohl auch, damit gezielt Marketing mit dem scheinbar unkommerziellen Music Genome Project betrieben werden kann.

häufig verschiedene Arten maschinellen Lernens im Zusammenspiel mit anderer Technik eingesetzt werden.

Der Mehrwert, den die Informatik in solchen Methodenkompositionen bietet, liegt darin, dass Selbstorganisationsprozesse und Datenbankerstellung automatisch, extrem schnell und mit klar benennbaren systematischen Fehlern durchgeführt werden⁸. Der Einsatz von leistungsstarken Computern erlaubt das Entstehen von Welttechnik, deren NEUGIERIGE STRUKTURVORSCHLÄGE unüberschaubare Datenmeere erst wahrnehmbar werden lassen.

8 Im Gegensatz zu den weniger transparenten systematischen Fehlern, die ein menschlicher Experte macht.

