Florian Muhle / Josef Wehner

Algorithmus und Entscheidung. Anspruch und Empirie personalisierter Medienangebote im Internet

Die Verbreitung von Computerprogrammen in nahezu alle privaten und beruflichen Bereiche wird in vielen Kommentaren mit der These verbunden, dass Menschen immer weniger, Algorithmen dafür umso mehr Entscheidungsbefugnisse übernehmen und Entscheidungen treffen. Mit dieser Erwartung einer Delegierbarkeit menschlicher Kernkompetenzen an die Technik eng verbunden ist die Idee, dass Algorithmen im Vergleich zum Menschen Entscheidungen anders, um nicht zu sagen: besser und meist auch ›rationaler‹ treffen können (vgl. Rademacher 2015). Diese Idee ist nicht neu, sie begleitet die Computerentwicklung bereits seit ihrem Beginn – wie etwa die Debatten zur Informationsgesellschaft in den 1970er Jahren oder zur Künstlichen Intelligenz in den 1990er Jahren zeigen. Heute scheint sie sich geradezu aufzudrängen, wie etwa in jenen Diskussionen über die ›Macht der Algorithmen‹ in Bereichen der globalen Börsen, sicherheitsrelevanten und geheimdienstlichen Zusammenhängen oder in den Bereichen des Online-Konsums. Demnach werden heute Entscheidungen, Wertpapiere zu (ver-)kaufen, Personen als (un-)verdächtig einzustufen oder Produkte online zu bestellen mit steigender Tendenz von Computerprogrammen unterstützt oder von diesen sogar ganz ohne menschliches Dazutun getroffen (vgl. die Beiträge in Reichert 2014).

Eng verbunden mit dieser Vorstellung, Computerprogramme würden bereits heute und in naher Zukunft in immer größerem Umfang den Menschen Entscheidungen abnehmen, ist die Erwartung, dass solche Entscheidungen nicht gegen, wie häufig befürchtet wird, sondern im Interesse von einzelnen Personen getroffen werden. Diese Erwartung verbindet sich vor allem mit den rasant wachsenden Möglichkeiten des Online-Konsums. Wer im Netz etwas bestellen will, soll nur noch solche Angebote finden, die seinen persönlichen Interessen und Gewohnheiten entsprechen. Die Entscheidungen darüber, ob etwas unserem Geschmack entspricht oder nicht, treffen dabei immer häufiger sogenannte ›Empfehlungs-Algorithmen‹. Dabei handelt es sich um Computerprogramme, die vorgängig unsere Aktivitäten auf entsprechenden Plattformen erfassen, auswerten und vergleichen, um auf dieser Grundlage

›wahrscheinliche‹ Produktvorlieben zu berechnen und dann entsprechende persönliche Konsumvorschläge zu machen.

In unserem Beitrag möchten wir am Beispiel von Online-Musikanbietern Anspruch und Empirie solcher algorithmengesteuerten persönlichen Empfehlungen etwas genauer untersuchen. Dazu werden wir eine Fallstudie zu einem Anbieter vorstellen, der seinen HörerInnen verspricht, dass sie sich dank eines Algorithmus nicht länger mit (massenmedial verbreiteter) ›Konfektionsware‹ zufrieden geben müssen, sondern zum einen nur noch das hören, was ihren persönlichen musikalischen Neigungen entspricht, und zum anderen auf diesen Neigungen beruhend neue Musik entdecken können. Die Erwartung, dass Computerprogramme nicht nur entscheiden können, sondern dies außerdem auch noch im Sinne persönlicher Interessen und Gewohnheiten tun, gelangt hier also in der Delegation von Musikempfehlungen an einen Algorithmus zum Ausdruck.

Bevor wir auf diesen Fall eingehen, werden wir in einem ersten Schritt zunächst den Versuch der Personalisierung der Medienkommunikation mediensoziologisch einordnen. Dazu werden wir darlegen, welche Versprechen für Medienanbieter, UserInnen und ›Interessierte Dritte‹ mit der Implementation von Algorithmen zur Publikumsbeobachtung und Angebotssteuerung verbunden sind, um auf diese Weise den Hintergrund für die Untersuchung zu bereiten und deren Relevanz zu verdeutlichen (1). Daran anschließend werden wir am Beispiel des Online-Musikanbieters das Personalisierungsversprechen auf der Grundlage empirischer Daten problematisieren. Hierbei wollen wir zeigen, dass und wie der ›im Betrieb befindliche‹ Algorithmus den an ihn gestellten Ansprüchen immer nur annäherungsweise oder auch gar nicht gerecht wird. Es soll deutlich werden, wie einerseits im Zuge der Entwicklungsarbeiten Entscheidungskompetenzen an den Algorithmus delegiert werden, andererseits jedoch damit immer wieder auch neue Interpretations- und Koordinationsbedarfe und zusätzliche Entscheidungszumutungen erzeugt werden. Entgegen der Vision einer sich immer weiter verselbständigenden Entscheidungsfindung erfordert der Algorithmus komplementäre, von ihm selbst nicht adaptierbare entscheidungsabhängige Vorgaben, ohne die er nicht arbeiten und funktionieren würde (2). So gesehen bilden Algorithmen immer auch den Gegenstand von Aus- und Verhandlungen, die allerdings ihrerseits Vorgaben durch das jeweils erreichte Niveau algorithmisierter Entscheidungsabläufe gesetzt bekommen. An unserem Fallbeispiel wird dies besonders deutlich, insofern als hier die Online-Plattform mit ihren Sortier-, Filter- und Empfehlungsmöglichkeiten zwar einerseits immer größere Datenmengen verarbeiten und differenziertere Auswertungen vornehmen kann, jedoch immer wieder auch neu zu lösende Pro-

bleme erzeugt. Wir gehen davon aus, dass dieses Dilemma nicht zufällig, sondern typisch für die Entwicklung und Anwendung von Algorithmen ist, die der Beobachtung und Analyse von Nutzeraktivitäten dienen (3).

Personalisierung der Medienkommunikation aus mediensoziologischer Perspektive

Einmal abgesehen von der individuellen Kommunikation wie sie per Telefon, Chat oder Email erfolgt, ist Medienkommunikation traditionell generalisierte Kommunikation (vgl. Esposito 1995). Dies klingt deutlich in den Begriffen der ›Massenkommunikation‹ oder des englischen ›Broadcasting‹ an. Medienkommunikation ist Kommunikation, die in die Breite geht und ihr heterogenes Publikum mit einem identischen Angebot versorgt. Zugleich bedeutet dies, dass Massenkommunikation eine Kommunikation im ›Blindflug‹ ist, die keinen direkten Kontakt zu ihren AdressatInnen hat. Deshalb wissen Medienanbieter nur wenig über ihr Publikum und dessen Vorlieben und Gewohnheiten. Entscheidungen darüber, welche Sendungen ausgestrahlt werden sollen oder welche Werbung zu welchem Zeitpunkt auf welchem Sender geschaltet werden soll, sind entsprechend stets Entscheidungen unter Bedingungen großer Unsicherheit.

Genau aus diesem Grund hat sich parallel bereits mit der Etablierung des Rundfunks die ›administrative‹ bzw. angewandte Publikumsforschung herausgebildet, die heute fester Bestandteil des Mediensystems ist, was sich nicht zuletzt daran zeigt, dass die ›Quote‹ als Währung von Funk und Fernsehen gilt (vgl. Schenk 2007). Das Publikum wird in der Regel in nach Merkmalen wie Alter, Geschlecht, Haushaltsstand oder Bildung klassifizier- und vermessbare Zielgruppen eingeteilt und es interessiert vornehmlich, wie lange ZuschauerInnen vor dem Bildschirm verharren oder HörerInnen das Radio eingeschaltet lassen, und welche Programme und Sendungen sie dabei sehen bzw. hören. Als Individuen mit persönlichen Vorlieben kommen sie hingegen kaum in Betracht (vgl. Wehner 2010, 10ff.). Die Quote mit ihren Hinweisen auf statistisch ermittelte Reichweiten der verschiedenen Programmangebote erzeugt ebenso wie zusätzlich in Auftrag gegebene Umfragen deshalb zwar nur ein indirektes und notwendig grobes Bild vom Publikum, aber dieses scheint die beteiligten Medien ausreichend zu informieren und eine Steuerung des Angebotes zu ermöglichen (vgl. Ang 2001).

Vor diesem Hintergrund versprechen algorithmengesteuerte Internettechnologien eine Veränderung und ›Verbesserung‹ der Medienkommunikation.

Sie sollen die Online-Aktivitäten von UserInnen automatisch erfassen, speichern und auswerten und daraus personalisierte Profile erstellen, die es erlauben, das (Medien-)Angebot an individuelle Gewohnheiten anzupassen und auf diese Weise zu optimieren.◂1 Jeder Klick, jeder Besuch einer Plattform, jeder Kaufakt soll genutzt werden, um persönliche Aktivitätsmuster von InternetnutzerInnen zu erkennen, von diesen auf Vorlieben und Geschmäcker zu schließen und dann darauf ausgerichtete personalisierte Angebote zu unterbreiten (vgl. schon früh hierzu Esposito 1995, 247). Auf diese Weise sollen solche Empfehlungen, die den persönlichen Geschmack der UserInnen nicht treffen, nach und nach aus dem Angebot verschwinden, während diejenigen, die positiv aufgenommen werden, expandieren (vgl. Wehner 2008, 206).

Einfache Beispiele hierfür sind Recommender-Systeme, wie sie sich beim Online-Händler Amazon finden. Hier werden ehemalige Kaufentscheidungen gespeichert und mit Kaufentscheidungen anderer NutzerInnen verglichen, um auf diese Weise angepasste Verkaufsvorschläge unterbreiten zu können, die dann etwa Produkte empfehlen, »weil Sie [XYZ] gekauft haben« oder mitteilen, »Kunden, die diesen Artikel gekauft haben, kauften auch [XYZ]«. Die Technologie, hier der Empfehlungsalgorithmus, trifft somit Entscheidungen darüber, welche Produkte den jeweiligen KundInnen angeboten werden, und diese Entscheidungen unterscheiden sich abhängig davon, wer gerade kauft und was er oder sie (aber auch andere) zuvor gekauft haben. Entsprechend handelt es »sich dabei um einen Mechanismus, der Popularität als Zeichen von Affinität nutzt« (Esposito 2014, 241) und aus der Analyse bereits erfolgter Kaufentscheidungen Prognosen über zukünftige Entscheidungen ableitet.

Während also Kaufempfehlungen im Fernsehen für alle RezipientInnen gleich sind, wird mit den Möglichkeiten der Online-Verdatung die Erwartung verbunden, die Gewohnheiten und Interessen einzelner Kunden oder Gruppen von Kunden immer besser kennenzulernen und bedienen zu können. Das Ziel ist dabei letztlich eine »Massenwerbung ohne Streuverlust« (Greve/Hopf/Bauer 2011, 8) verbreiten zu können. Ähnlich wie der hier skizzierte Bereich des Konsums sollen auch andere Formen der personalisierten Medienkommunikation funktionieren, die sich deutlich in Konkurrenz zur traditionellen Massenkommunikation positionieren. Ein aktuell viel diskutiertes Beispiel hierfür ist der Video-Streamingdienst Netflix, der seinen NutzerInnen Filme und Serien anbietet und hierfür ebenfalls einen Empfehlungsalgorithmus nutzt, der registriert, welche Filme und Serien geschaut und gemocht werden und auf dieser Basis und unter Einbeziehung der Nutzungsweisen anderer UserInnen – angeblich sehr erfolgreich – Vorschläge für neue Filme und Serien unterbreitet.◂2

Durch die Verwendung entsprechender Algorithmen erhoffen sich die *Betreiber* der Plattformen neues und vor allem präziseres Wissen über ihr Publikum, welches sie dazu verwenden können, die Angebote genauer an die unterschiedlichen Gewohnheiten und Vorlieben ihres Publikums anzupassen. Für die *NutzerInnen* soll damit der Vorteil einhergehen, ein auf ihre Gewohnheiten abgestimmtes personalisiertes Medienangebot zu erhalten, indem der Algorithmus entscheidet, welche Produkte zu ihnen ›passen‹. Schließlich soll das durch den Algorithmus erzeugte Wissen über die UserInnen und ihre Vorlieben auch *interessierten Dritten* – etwa der Werbeindustrie – helfen, zu besseren Entscheidungen in Fragen der Investition oder der Platzierung von Werbung zu kommen. Anbieter, NutzerInnen und interessierte Dritte sollen also gleichermaßen von der durch Algorithmen ermöglichten Personalisierung von Medienangeboten profitieren, die es ermöglicht, die mit der traditionellen einseitigen Medienkommunikation verbundene, in der Vergangenheit immer wieder kritisierte Orientierung am »Massengeschmack« zu überwinden.

Wie die entsprechenden Algorithmen entwickelt werden und wie sie arbeiten, weiß kaum jemand. Sie bilden ein wichtiges Betriebskapital der Medienanbieter, so dass deren Funktionsweise in der Regel geheim gehalten wird. Allerdings wird diese von den meisten NutzerInnen auch nicht hinterfragt, so lange sie ausreichend gute Ergebnisse produziert.◂3 Das heißt, Algorithmen selbst und die Grundlagen auf denen sie Entscheidungen treffen, werden in der Regel als eine Art *Black Box* behandelt, die funktioniert und als ›Medium der Entscheidung‹ die konstatierten Probleme für NutzerInnen, Medienanbieter und Interessierte Dritte lösen kann (vgl. Hallinan/Striphas 2014, 1f.).

Dieses Algorithmenverständnis beschränkt sich nicht auf die Praxis. Auch sozialwissenschaftliche BeobachterInnen behandeln Entscheidungsalgorithmen insoweit als Black Boxes, als dass sie diese erst zum Zeitpunkt ihrer Fertigstellung und Inbetriebnahme, also als bereits fertig gestellte und funktionierende Systeme in den Blick nehmen und ihnen starke Wirkmächte unterstellen. Entsprechend wird auch hier (wenn auch durchaus mit kritischem Unterton) die Delegation von Entscheidungen an Algorithmen in der Regel als eine Art Erfolgsstory beschrieben (vgl. Muhle i.E.). So wird die Personalisierung der Medienkommunikation bisher vornehmlich als erfolgreiche Kommodifizierung des Publikums begriffen und davon ausgegangen, dass die algorithmenbasierte Beobachtung und Klassifizierung von Online-Aktivitäten es erlaube, »an exact picture of the interests and activities of users« (Fuchs 2014, 108) zu produzieren, wodurch die Online-Verdatung zu einer »key force of production in an information economy« (Fisher 2015, 62) werde. In dieser Perspektive geraten die Medienproduzenten als Profiteure, die UserInnen als Objekte von Marketing

und Kommerzialisierungsstrategien und die Algorithmen als Instrumente der Ausbeutung in den Blick.

Vor diesem Hintergrund erscheinen uns die Ergebnisse unserer Fallstudie interessant. Sie bezieht sich auf den Versuch eines Online-Musikanbieters, für seine HörerInnen einen Empfehlungsalgorithmus ›zum Laufen zu bringen‹. Das Projekt der Entscheidungsdelegation befindet sich hier noch im Stadium der Entwicklung und experimentellen Erprobung. Mit Blick auf dieses Stadium können wir zeigen, dass und wie der verwendete Algorithmus (in dieser Phase) permanent definitorische Aufwände, Verhandlungen und Entscheidungsbedarfe erzeugt, da immer wieder neue Probleme – in unserem Fall: der Genauigkeit, Zuverlässigkeit, des Analysevermögen etc. – entstehen. Für diese Probleme müssen laufend Lösungen gefunden werden, die immer wieder auch Korrekturen an den Zielen, Konzepten und Funktionsweisen des Algorithmus erforderlich machen – ohne dass offenbar erwartet werden kann, die Arbeiten finalisieren bzw. den Algorithmus perfektionieren zu können. Hierdurch wird deutlich, dass Algorithmen keineswegs ›automatisch‹ erfolgreich funktionieren. Dennoch wäre es vorschnell, ihnen ihre Wirkkraft abzusprechen. Diese liegt aber nicht unbedingt darin, dass sie Entscheidungen besser als Menschen treffen, sondern wohl nicht zuletzt darin, dass sie neue Verhandlungs- und Entscheidungsbedarfe erzeugen, die ohne sie nicht existiert hätten.◂4

Der Fall: Personalisiertes Online-Radio mit algorithmenbasierter Vorschlagstechnologie

Zu den Leitideen des hier untersuchten Dienstes gehört bei dessen Start, ein personalisiertes Online-Radio anzubieten, bei dem ein Algorithmus die Entscheidung über die Programmauswahl trifft. Es sollen also nicht länger beispielsweise Musikredaktionen oder ModeratorInnen entscheiden, welche Lieder bzw. Interpreten gespielt werden; stattdessen soll eine technische Schnittstelle autonom diese Aufgabe übernehmen. Das Konzept des Dienstes passt sich somit in den Trend der ›Googlization‹ ein (vgl. Vaidhyanathan 2012), indem »die Rolle der Experten und Redakteure […] allmählich durch Algorithmen ersetzt« (Esposito 2014, 240) wird. Darüber hinaus sollen die musikalischen Empfehlungen nicht für alle NutzerInnen des Dienstes identisch sein (wie beim klassischen Radio), sondern jeweils auf den individuellen Geschmack einzelner HörerInnen zugeschnitten und entsprechend personalisiert werden. Wie die folgenden Ausführungen allerdings zeigen, werden diese Erwartungen enttäuscht. Der Algorithmus liefert zwar – immerhin – Ergebnisse,

das heißt, er nimmt Nutzungsdaten auf, wertet diese aus und gibt auch Empfehlungen; aus Sicht der HörerInnen wie auch des Online-Betreibers erweisen sich diese jedoch häufig als unpassend und deshalb eher als (dauerhaftes) Problem denn als Lösung von Problemen.

Zielsetzungen, erste Lösungsversuche und Enttäuschungen

Konzeptionell gesehen beruhen die Empfehlungen des Algorithmus auf zwei miteinander kombinierten Methoden und zwar der Berechnung der Vorlieben und Aktivitäten der UserInnen und der statistisch ermittelten Ähnlichkeiten zwischen Liedern bzw. Künstlern.◂5 Die Plattform funktioniert dem Konzept nach so, dass die UserInnen nach dem Login bestimmte Genres oder Künstler auswählen und dann durch den Algorithmus generierte Playlisten erhalten, die zum Genre passende Musik oder »ähnliche Künstler« abspielen. Die User haben dann die Möglichkeit, Songs zu überspringen (»skip«), sie zu bestätigen (»love«) oder sie aus dem persönlichen Angebot auszuschließen (»ban«). Auf der Grundlage entsprechender Daten soll der Algorithmus den Musikgeschmack der einzelnen TeilnehmerInnen immer besser kennenlernen, zunehmend auf die persönlichen Vorlieben ausgerichtete Musik spielen und neue Vorschläge unterbreiten, die auf Ähnlichkeiten mit den erkannten Vorlieben basieren. Es geht also darum, automatisiert, ohne menschliche Einmischung, eine Musikauswahl zu treffen, die für jede/n angemeldete/n Hörer/in unterschiedlich sein kann. Dies macht der technische Leiter des Dienstes im Interview deutlich:

»der aller erste Ansatz war […] rein technisch und sehr vom Glauben inspiriert, die Maschine löst das alles, ich krieg ähnliche Titel ich krieg Stimmung, ich krieg Genre und ich krieg auch noch ähnliche Künstler hat [die Entwicklerfirma] versprochen« (Interview Technischer Leiter).

Die technische Realisierung der Musikauswahl beruht dabei auf der Berechnung von Ähnlichkeiten und Unterschieden in der Musik. Hierzu werden »alle möglichen Eigenschaften aus der Musik rausgezogen die irgendwie mathematisch statistisch verwertbar sind, das sind bis an die knapp zwohundert Stück, die […] raus gezogen werden« (Interview Technischer Leiter). Diese Eigenschaften werden dann durch den Algorithmus

»zueinander ins Verhältnis gesetzt, das ergibt dann sogenannte middle level feature die wiederum dann am Schluss kombiniert werden in etwas was man wirklich der Musik zuordnen kann nämlich so was wie Rhythmus Melodie Melodieinstrument ähm Tonalitäten Perkussivität und solche Geschichten« (Interview Technischer Leiter).

Die Berechnung der Eigenschaften einzelner Lieder dient nun einerseits dazu, diese bestimmten Genres zuzuordnen und andererseits hieraus individuelle Künstlerprofile zu erzeugen, indem aus der Summe der Titel einzelner KünstlerInnen Profile gebildet werden. Auf dieser Grundlage werden dann klangliche Ähnlichkeiten zwischen Songs und KünstlerInnen berechnet und mit einem eindeutigen Wert versehen, der in Form einer Zahl zwischen 0 und 1 ausgegeben wurde, wobei 1 für hohe Übereinstimmung steht und 0 für gar keine Übereinstimmung. Entsprechend integriert der Algorithmus dann Lieder mit einem hohen Ähnlichkeitswert in eine sogenannte Playlist (genre- oder künstlerbasiert), während Lieder mit niedriger Übereinstimmung ausgeschlossen werden. Die »skips«, »loves« und »bans«, die von Usern ausgeführt werden, werden protokolliert und in den individuellen Profilen zugerechneten Datenbanken abgelegt, damit der Algorithmus ›weiß‹, welche Songs gemocht und welche abgelehnt werden, um zukünftige Empfehlungen daraufhin einzustellen.
Zusätzlich sollen Titel bzw. Interpreten empfohlen werden, von denen (wiederum auf statistischer Grundlage) angenommen werden kann, dass sie dem Hörer/der Hörerin unbekannt sind, aber zum individuellen Musikgeschmack passen könnten. Der Dienst soll also den NutzerInnen auch das Entdecken neuer Musik ermöglichen (»dieses Musikdiscoveryding, das in der Firmenidee mit drin steckt«, Interview Musikredakteur). Spätestens hier wird deutlich, dass die UserInnen nicht selbst ihr Programm zusammenstellen (sollen), sondern weiter Radio hören können, bei dem Musik für sie ausgewählt wird. In den Worten des Technischen Leiters:

»unsere Vision is eben halt wir sind Radio, mit Radio kannst du Musik entdecken ähm du kannst deine persönliche Musik hören spielen und immer mehr entdecken, und wir sind […] dein persönlicher Stream« (Interview Technischer Leiter).

Genau in dieser nicht nur für Zwecke der Werbung nach außen, sondern auch der internen Orientierung und Koordinierung in das Unternehmen hinein kommunizierten Plattformperformanz, bei der persönliche musikalische Empfehlungen automatisiert erzeugt werden, wird der entscheidende Vorteil einer algorithmisierten Musikauswahl gegenüber menschlichen Entscheidungen – sowohl denen der UserInnen selbst als auch denen einer menschlichen Redaktion – gesehen. Interessanterweise soll der Algorithmus nicht nur – wie ein Musikredakteur – in der Lage sein zu entscheiden, welche Songs bzw. Interpreten dem persönlichen Geschmack eines Hörers bzw. einer Hörerin entsprechen; er soll sogar besser entscheiden können, insofern als er bei seiner Zusammenstellung das immer begrenzte menschliche Fassungs- und Erinnerungsvermögen

übersteigen soll. Dies wird an der folgenden Auskunft des verantwortlichen Musikredakteurs deutlich:

»wenn dieses Ding mal irgendwann richtig funktioniert, […] dann wäre dieser Algorithmus trotzdem ne Möglichkeit, ähm vielleicht also dieses Musicdiscovery-Ding […] besser umzusetzen sogar noch, als wenn das jetzt irgendein Mensch macht. Weil wir sind alle blöde Szene-Nerds, die irgendwie ähm wenn wenn mir einer meine Lieblingsband sagt, dann fallen mir bestimmt aus dem aus demselben Genre fünfzig Bands ein, die so ähnlich klingen, aber vielleicht gibt's da ganz andere Kombinationen. Vielleicht gibt's da auch ganz andere Gemeinsamkeiten zwischen meiner Lieblingsband und irgendwie was weiß ich, und Stockhausen. Die mir noch nie aufgefallen sind, die aber vielleicht der Computer ausspuckt. Also wenn, wenn der Algorithmus irgendwie richtig funktionieren würde, dann kann ich mir schon vorstellen, dass der Algorithmus auch echt was was bringt. Dass der auch noch mal ne andere ähm ne andere Qualität in das Ganze reinbringt« (Interview Musikredakteur).

Das Problem der Musikauswahl durch eine Redaktion – das gleiche gilt sicher auch für die Auswahl durch die NutzerInnen selbst – wird hier in deren Festlegungen auf bestimmte Geschmacksrichtungen und Hörgewohnheiten gesehen. Selbst professionelle Musikfachleute kennen als »Szene-Nerds« eben nur das, was sie kennen und ohnehin gerne hören. Verblüffende Ähnlichkeiten, die nur durch ›einen Blick über den eigenen Tellerrand hinaus‹ auffallen können, erschließen sich ihnen eher selten. Dies soll bei einem funktionierenden Algorithmus, der über eine stetig wachsende Datenbank mit Musik unterschiedlichster Genres verfügt und fortlaufend nach Ähnlichkeiten und Unterschieden zwischen bereits archivierten und neu dazu kommenden Titeln sucht, anders werden. Er soll, im Unterschied zum menschlichen Personal, aufgrund der Auswertung der Musik-Eigenschaften auch überraschende Zusammenhänge erkennen und entsprechende Empfehlungen generieren können, um die persönlichen Chancen des Entdeckens neuer Musik zu steigern.
Zum Zeitpunkt des Interviews wird jedoch offensichtlich, dass der Algorithmus (noch) nicht in der Lage ist, die Entscheidungen eines Redakteurs zu übernehmen. Der erhoffte Vorteil, anstelle einer Musikredaktion den Algorithmus entscheiden zu lassen, passende Musikstücke bzw. Interpreten zu empfehlen, wird von der *tatsächlichen* Arbeitsweise des Algorithmus nicht bestätigt. So schränkt der Musikredakteur im Interview gleich im Anschluss an die oben angeführte Passage ein, dass »in dem Stadium, in dem das jetzt ist, […] der Algorithmus oft einfach nur ein Ärgernis [ist]« (Interview Musikredakteur). Warum dies so ist, findet sich an anderer Stelle im Interview:

»Nein, der Algorithmus hat einfach viel Scheiße ausgespuckt und ähm das war auch ne zeitlang war das immer ein großes Beschwerdethema, dass wenn jemand ne Artist Station angemacht hat, dass da auf einmal Schlager gelaufen ist oder sonst irgendwas, also dass die Dinge wirklich so absurd aneinandergereiht waren, dass es halt auch überhaupt keine Ähnlichkeit hatte. […] und dann sind ist irgendwann ist da eingeschritten worden sozusagen« (Interview Musikredakteur).

Zwar funktioniert die algorithmenbasierte Ähnlichkeitsermittlung in dem (programmiertechnischen) Sinne, dass sie Berechnungen über klangliche Nähen und Unterschiede durchführt und auch Empfehlungen gibt. Aber die Ergebnisse können weder die HörerInnen noch die musikredaktionelle Leitung des Plattformbetreibers überzeugen. Die zwischen Titeln bzw. Interpreten errechneten klanglichen Ähnlichkeiten und darauf aufbauende Titelzusammenstellungen ignorieren in zu vielen Fällen die eingeschliffenen Genre- und andere musikalische Ein- und Ausschlusskriterien. Der Algorithmus verfehlt so gesehen bereits das Pflichtziel, Entscheidungen zu treffen, die auch eine Musikredaktion hätte treffen können. Erst recht aber verfehlt er das zusätzliche Premiumziel, auch solche Zusammenhänge zwischen Tracks und Künstlern zu erkennen und zu berücksichtigen, die über die partikularen Erfahrungs- und Wissensstände der MusikredakteurInnen hinausgehen. Auch diese Entscheidungen müssen die Redaktion bzw. die HörerInnen ja nicht nur überraschen, sondern auch überzeugen, was sie aber nicht tun. Es gelingt dem Algorithmus also weder in der limitierten, auf bereits bekannte musikalische Ähnlichkeiten, und erst recht nicht in der erweiterten, auf bislang unbekannte Ähnlichkeiten ausgerichteten Variante, an die Erfahrungen der Redaktion und des Publikums anzuschließen. Er kann die ihm zugedachte Funktion, Musikempfehlungen nicht nur anders, sondern auch besser zu tätigen als eine menschliche Redaktion, offensichtlich nicht erfüllen – mit der Konsequenz, dass in dem Unternehmen die Idee einer Algorithmisierung musikredaktioneller Entscheidungen einer folgenreichen Revision unterzogen wird.

Korrekturen, angepasste Ziele, neue Herausforderungen

Die Revision der Entscheidungsbefugnisse des Algorithmus bezieht sich sowohl auf die Berechnung ähnlicher Künstler als auch auf die Zuordnung der Songs zu bestimmten Genres und damit auf den Kernbereich seines Aufgabengebietes. Dem Algorithmus wird nach den ersten Erfahrungen nicht länger zugetraut, hier zu überzeugenden Empfehlungen zu gelangen. Die Annahmen und Erwartungen, die noch in der Anfangsphase des Dienstes die Arbeiten an der Plattform beflügelten (»vom Glauben inspiriert, die Maschine löst das alles«, Inter-

view Technischer Leiter), weichen einer Perspektive, in der nach Gründen des Scheiterns gesucht wird, um zu einer angemessenen Konzeption des Dienstes zu gelangen. Vor allem das dem Algorithmus zugrundeliegende Konzept zur Berechnung der Ähnlichkeiten zwischen Künstlern wird jetzt hinterfragt und auf seine Schwachstellen hin analysiert.

So sah die im Algorithmus implementierte Idee zur Berechnung von Künstlerähnlichkeiten ursprünglich folgendermaßen aus: »ne Summe von ähnlichen titeln ergibt n Profil von nem Künstler also ne Summe von Titeln ergibt n Profil und diese Profile kannste wieder nebeneinander setzen […]« (Interview Technischer Leiter). Wie sich herausstellt, lassen sich auf diese Weise zwar klare Profile erstellen, die mit eindeutigen Werten versehen und somit leicht automatisch verglichen werden können. Gleichzeitig liegt genau in der mit diesem Verfahren verbundenen Reduktion von Komplexität aber das entscheidende Problem. Denn mit der Synthetisierung *eines* Profils geht zwangsläufig einher, dass die mögliche Heterogenität in den Werken einzelner Interpreten nicht berücksichtigt wird. Dies, so die Erkenntnis, sei jedoch notwendig, schließlich seien

> »Künstler auf nem album schon irgendwie so heterogen und wenn de da noch irgendwie ne Zeitspanne nen Horizont von zwanzig dreizig Jahren dann mal nimmst ähm dann haben die einfach schon Werke die sitzen innerhalb der Ähnlichkeitsmatrix soweit auseinander, dass eigentlich ein Querschnitt von sowas immer nur in die Hose gehen kann. Und genau das passiert wenn ich nämlich auf grund so nes Querschnitts einfach Künstler vergleiche« (Interview Technischer Leiter).

Was passiert, wenn die automatische Ermittlung der Künstlerprofile ›in die Hose‹ geht, liegt auf der Hand. Zwar kommen dann bei der Zusammenstellung ähnlicher Künstler möglicherweise unerwartete Ergebnisse zustande, aber »diese unerwarteten Dinge die zerhauen dir […] im Prinzip die ganze Reputation« (Interview Technischer Leiter). Spätestens ein solcher Verlust der Reputation stellt aber für den Dienst, der in (kommerzieller) Konkurrenz zu anderen Versuchen steht, die auch personalisiertes Radio anbieten, ein gravierendes Problem dar.

Angesichts dieser Schwierigkeiten wird entschieden, auf eine editoriale Zusammenstellung ähnlicher Künstler umzustellen. Die Hauptverantwortung für die Zusammenstellung ähnlicher Künstler wird also vom Algorithmus wieder zurück auf das menschliche Musikmanagement übertragen und das Konzept der automatisierten Zusammenstellung auf redaktionelle Arbeit umgestellt. Allerdings wird der Algorithmus nicht einfach komplett ersetzt. Vielmehr sieht die manuelle Bearbeitung der Listen ›Ähnlicher Künstler‹ vor, dass sich eine Per-

son an die zuvor vom Algorithmus erzeugte Ähnlichkeitsliste setzt, diese kontrolliert und bei Bedarf ändert:

»Das sieht so aus, dass neue Künstler zu der Liste hinzugefügt werden bzw. die automatisch gelieferte Liste komplett gelöscht wird und durch eine handgemachte ersetzt wird. Zudem werden dann eigene Zahlenwerte hinzugefügt, die dann die neue Ähnlichkeits-Liste strukturieren und hierarchisieren. Der Maßstab hierfür ist vor allem das Fachwissen und das Bauchgefühl der Redakteure« (Duhr 2013, 42).

Die Ergebnisse des Algorithmus werden jetzt nachträglich gesichtet, evaluiert und auch korrigiert. Sie durchlaufen einen ›Controlling-Prozess‹, in dem nicht mehr die durch den Algorithmus errechneten Eigenschaften der Musik als Qualitätsmaßstab dienen, sondern genau jener komplexe, durch mitunter Jahrzehnte lange Praxis eingeübte Mix unterschiedlicher Ein- und Ausschlusskriterien der MusikredakteurInnen, der ursprünglich durch die Arbeit des Algorithmus überwunden werden sollte. Die Entscheidungs-Kompetenz wird somit wieder an die ›Szene-Nerds‹ zurück delegiert. Allerdings müssen diese ihr »Fachwissen und Bauchgefühl« in maschinenlesbare Form übersetzen und entsprechend in eindeutige Zahlenwerte überführen. Denn der Algorithmus bleibt weiterhin in die Musikauswahl eingebunden und spielt nach wie vor ähnlich klingende Künstler ab. Er entscheidet also nicht mehr über Ähnlichkeiten zwischen Künstlern, bleibt aber dennoch für die Auswahl und das Abspielen der KünstlerInnen verantwortlich.
Auch die Genre-Einteilung wird nun von einem automatischen auf einen manuellen Bearbeitungsmodus umgestellt. Bestand der anfängliche Ansatz auch hier darin, dass der Algorithmus auf Grundlage seiner Analyse bestimmt, welche Titel zu welchem Genre gehören, wird schnell deutlich, dass dies nicht funktioniert. Wie der technische Leiter des Dienstes verdeutlicht, war dies sogar »die erste Arbeit, die wir eigentlich gemacht haben, selber zu definieren [...] was ist wie behandeln wir Genres« (Interview Technischer Leiter). Entsprechend wird in der Funktionsweise des Algorithmus »das Genre ausgegrenzt weil n automatisches Genre extrahieren das funktioniert nicht« (Interview Technischer Leiter). Stattdessen werden nun durch die MitarbeiterInnen des Dienstes insgesamt elf unveränderliche Genres (mit variablen Untergenres) festgelegt, die von Hand der jeweiligen Musik zugeordnet werden. Auf Grundlage dieser manuellen Genreeinteilung sollen dann die Musik-Empfehlungen des Algorithmus verbessert werden, indem sie als Filter für das weitere Auswahlverfahren eingesetzt wird. So berechnet der Algorithmus zwar weiterhin Ähnlichkeiten zwischen Musiktiteln. Bevor diese aber in eine Playlist aufgenommen werden, müssen sie nun den nachträglich eingebauten Filter durch-

laufen und die damit berücksichtigten Bewertungskriterien erfüllen. Auf diese Weise soll sichergestellt werden, dass unterschiedliche Titel auch ein und demselben – durch die Musikredaktion hinzugefügten – Genre angehören:

»der eine Titel ist ähnlich dem anderen das verschiebt sich durch die Tätigkeit der Musikmanager erstmal nicht ähm was als Ergebnis rausgegeben wird ist ne Veränderung die durch nen anderen Layer passiert nämlich n Filter […] bevor wir sagen der und der ist ähnlich müssen wir zusätzlich noch das Genre vergleichen weil die sind jetzt zwar irgendwie jetzt so ähnlich ganz nah beieinander in dieser Matrix aber die können wir niemals zusammen in einer Playliste spielen, weil kulturell gesehen ist der Typ Schlager und der Typ ist Pop, weil der ist aus UK und der ist aus Deutschland« (Interview Technischer Leiter).

Als weitere Modifikation kommt noch das manuelle Erstellen sogenannter Referenzlisten durch die insgesamt in ihrer Bedeutung deutlich aufgewertete Musikredaktion hinzu. Auch diese Referenzlisten sollen das negative ›Überraschungspotential‹ des Algorithmus einhegen. Sie werden vor dem Hintergrund der Erfahrung eingeführt, dass bestimmte Gruppen von NutzerInnen des Dienstes weniger an der Entdeckung neuer als vielmehr am Hören bekannter Musik interessiert sind. Probleme werden deshalb nicht nur in der vom Algorithmus abweichenden Praxis des Urteilens über Ähnlichkeiten und Unterschiede zwischen Musiktiteln und Interpreten in der Redaktion und beim Publikum gesehen, sondern auch in den unterschiedlichen, mit den jeweiligen Hörgewohnheiten und Geschmacksrichtungen variierenden Toleranzen der HörerInnen, überhaupt Empfehlungen unbekannter Musik anzunehmen. So gibt es HörerInnen bestimmter Genre-Playlists, die »gesacht ham öh passt […] nicht« und entsprechend ist die »Zufriedenheitsrate« mit den einzelnen Listen »extrem unterschiedlich« (Interview Technischer Leiter). Vor allem Rock/Pop-HörerInnen scheinen stärker am Massengeschmack bzw. an den massenmedial bereits geförderten Hitlisten eines Genres orientiert zu sein, als zuvor gedacht:

»ja das Wichtige ist ja, dass wir das System erstmal dahingehend umgestellt haben, weil wir gemerkt haben öh öhm, das ist halt auch ne Erfahrung die man halt mit der zeit macht, natürlich ist es toll Musik zu entdecken aber für jemanden der halt sacht oh da is ein neuer Service da geh ich jetzt mal hin äh und auf Rock drückt und dann irgendetwas kommt wo er sacht das kenn ich gar nicht, was is das denn? also so eher sich halt mit dem Kopf schüttelt […] da is uns wenig geholfen […] dafür haben wir die Referenzlisten erstellt, um zu sagen wenn einer eben auf Rock […] drückt […] dann sollte zumindest irgendwas Relevantes kommen was in dieser Referenzliste irgendwie vorhanden ist. Weil das ist das Verständnis des Nutzers: da ist nicht der B-Track von der fünften Single, die halt irgendwie am Schlechtesten verkauft worden ist, sondern es ist referenziert. Das heißt also, es ist ne Gewichtung und diese Gewichtung kann natürlich auch nicht der

Algorithmus bringen, sondern ne Gewichtung kann entweder der Mensch geben oder wir hätten ne Datenbank wo eben halt drinnen steht […] weil der Nutzer eben nicht, zum allergrößten Teil wenn wir übern Mainstream reden äh nicht erwartet, etwas ständig um die Ohren gehauen zu bekommen, was er überhaupt nicht kennt« (Interview Technischer Leiter).

Um ausgehend von diesen Erfahrungen sicherzustellen, dass die HörerInnen verlässlich Musikvorschläge erhalten, die ihren Einstellungen und Gewohnheiten entsprechen, werden von der Musikredaktion genreabhängige Referenzlisten erstellt. Solche Listen enthalten jeweils ca. 300 Songs, aus welchen dann einzelne ausgewählt werden, wenn UserInnen eine Station starten. Diese Referenzlisten werden monatlich aktualisiert und sollen sicherstellen, dass auf jeden Fall bei Einstieg in das Musikprogramm die ›richtigen‹, und das heißt erwartbare und bekannte, Titel abgespielt werden – was vor allem in der anfänglichen Phase des Kennenlernens und Sammelns von Erfahrungen darüber entscheidet, ob die HörerInnen weiterhin auf der Plattform Musik hören oder wieder abspringen. Für die Erstellung der Listen wird den Redakteuren mit an die Hand gegeben, dass Sie sich einerseits in ›den typischen Hörer‹ des jeweiligen Genres hineinversetzen sollen, und anderseits darauf achten, dass das Verhältnis zwischen Hits und unbekannten Songs mindestens 3:1 betragen sollte. Zudem ist es ihre Aufgabe, die Listen nach der Fertigstellung durchzuhören und zu überprüfen, »ob die Lieder zusammenpassen, wie der Gesamteindruck ist [und] ob es auffällige ›Querschläger‹ gibt« (Duhr 2013, 42). Ähnlich wie bei der Editierung der ähnlichen Künstler wird bei der Genre-Zusammenstellung also nach den ersten Erfahrungen mit den Leistungen und Entscheidungen des Algorithmus wieder auf menschliche Urteilskraft umgestellt und die Expertise, Erfahrung und Einschätzung der menschlichen Redaktion als unerlässlich eingestuft. Die RedakteurInnen konstruieren aufgrund ihres (informellen, jedoch geteilten) Vorwissens genrebezogene Hörergruppen und übersetzen die jeweils unterstellten Vorlieben und ›Toleranzen‹ für ähnliche, aber bislang unbekannte Titel in entsprechende Listen, die wichtige Vorgaben für das Auswahl- und Empfehlungssystem darstellen.

Ein weiteres, im Laufe der anfänglichen Einsätze des Algorithmus auftauchendes Problem betrifft die Entdeckung, dass nicht nur über die Zusammensetzungen der Referenzlisten entschieden werden muss, sondern auch, für welche Genres bzw. HörerInnen Referenzlisten überhaupt eine sinnvolle Vorgabe darstellen. Denn während die ›typischen‹ HörerInnen von Pop-/Rockmusik nur geringes Interesse für unbekannte Songs aufweisen, gibt es andere User-gruppen, die eine größere Bereitschaft zeigen, sich auf musikalisches Neuland einzulassen und – wie bei Gründung des Dienstes intendiert – das Angebot des

personalisierten Radios zu nutzen. Folglich wären hier Referenzlisten kontraproduktiv. Offensichtlich sind die HörerInnen verschiedener Genres also in ganz unterschiedlicher Weise offen für bislang unbekannte Musikempfehlungen und wollen entsprechend mehr oder weniger neue Musik empfohlen bekommen. Deshalb erscheinen genrespezifische Lösungen sinnvoller, um mit dem Problem unterschiedlicher Zufriedenheitswerte auf Seiten des Publikums angemessen umzugehen. In diese Richtung soll das Konzept des Dienstes auch weiter entwickelt werden:

»ist halt dann die nächste Frage: Wo fangen wir dann in Zukunft an Stellschrauben zu machen? Ne also was Jazz ist, kriegt im Prinzip äh zwar grundsätzlich content based wenig Metadaten-Filter, weil die Leute offener sind. Die wollen auch mal was anderes haben. Die Popleute sind alles, denen kannste eigentlich Top Ten geben am besten. Da machste irgendwie Referenzlisten sind die alle total zufrieden ähm und dann gibts noch n paar Freaks, die hören alles Mögliche. Da kannst Du auch so irgendwie fast schon (content based) eingeben und dann noch bisschen social kram keine Ahnung also da bewegt es sich halt in Zukunft« (Interview Technischer Leiter).

In dem Zitat klingt eine geplante Publikumssegmentierung an, in welcher der Algorithmus in sehr unterschiedlichem Maße zum Zuge kommen soll. Sinnvoll erscheint sein Einsatz in speziellen Segmenten, die sich durch besondere Offenheit für Neues und Überraschung auszeichnen: »Jazz-Fans« und »Freaks« (Interview Technischer Leiter). Nachteilig wirkt sich sein Einsatz hingegen beim typischen Publikum der Kulturindustrie aus, also jenen Segmenten, die lieber Mainstream und Altbekanntes hören möchten. Offenbar muss vorher entschieden werden, in welchen Fällen der Algorithmus die Datenbank mit ihren klanglich verwandten Titeln bzw. Interpreten ausschöpfen darf, und in welchen Fällen die Überraschungsqualität seiner Musikauswahl durch Referenzlisten gefiltert und damit eingeschränkt werden soll. Diese Entscheidung wird dem Algorithmus nicht zugetraut, sie soll deshalb nur von der Musikredaktion getroffen werden.

Unabschließbares Aushandlungsprojekt

Im Zuge der anfänglichen Misserfolge wird der Algorithmus als Gegenstand ständiger Korrektur- und Entscheidungsbedarfe erfahren. Zugleich geht damit eine Aufwertung der Funktion der Musikredaktion einher. Wurden bei Start des Dienstes »die allerersten Musikmanager […] nicht als Musikmanager angestellt […], sondern als Ripper« (Interview Technischer Leiter), das heißt als Personen, die lediglich Musikinformationen in die Datenbank des Dienstes eingeben, wird die Bedeutung musikalischer Expertise zunehmend als unentbehrlich für das Funktionieren des Dienstes erkannt. Auf sich alleine gestellt

wäre der Algorithmus überfordert, weshalb er durch zusätzlichen von außen hinzukommenden Input (Metadaten, Referenzlisten) und das Urteil der MusikredakteurInnen mit ihren Hörerfahrungen und ihrem (Vor-)Wissen über unterschiedliche Rezeptionskulturen ergänzt werden muss. Erst in dieser Verbindung kann der Dienst offensichtlich den Erwartungen des Publikums gerecht werden und zufriedenstellende Ergebnisse für Anbieter und UserInnen liefern. Es kann deshalb auch nicht überraschen, wenn die Arbeit am Empfehlungsalgorithmus als ein vermutlich nie abzuschließendes Projekt verstanden wird. Denn auch wenn »der Traum, dass die Maschine das in irgendeiner Form hundertprozentig knacken kann« (Interview Technischer Leiter) trotz der notwendigen und offenbar nicht algorithmisierbaren menschlichen Unterstützungs- und Kontrollarbeiten nicht aufgegeben wird, ist den Beteiligten bewusst, dass für die Verbesserung des Algorithmus sich kein einzig bester Weg, geschweige denn eine Finalisierung vorstellen ließe. Stattdessen macht sich nach den bisherigen Erfahrungen die Einsicht breit, dass es *die eine* Lösung nicht geben wird und eine Verknüpfung unterschiedlicher Ansätze erforderlich ist – mit ungewissem Ausgang. Vor allem die unterschiedliche Performance der Empfehlungstechnologie bei verschiedenen Genres scheint diese Einschätzung nachhaltig zu bestätigen:

»ich habe die Erfahrung auch schon mal früher gemacht, dass bestimmte Komponenten von so einer Empfehlungstechnologie für bestimmte Genres auch unterschiedlich funktionieren […] wenn du da Rocksachen haben wolltest, dann hat das fast perfekt funktioniert. Da waren ganz ganz wenig Ausreißer drinne, aber irgendwie ist die Maschine auch auf dieses Thema hin trainiert worden« (Interview Musikredakteur).

Sind die Plattform und die darauf installierten Verdatungs- und Analyseprogramme erst einmal in Betrieb genommen, bedeutet dies also nicht ein Ende (vorgängig) zu leistender Absprachen, Entscheidungen und korrigierender Eingriffe. Das Gegenteil scheint der Fall zu sein. In diesem Kontext erweist sich Andrew Pickerings (2007) Konzept einer ›Mangel der Praxis‹ als treffende Metapher, um den Prozess der Einbindung, Problematisierung und Anpassung des Algorithmus und seiner ›Rolle‹ bei der Herstellung personalisierter Medienangebote zu charakterisieren. Der Wissenschafts- und Technikforscher Pickering hat diese Metapher eingeführt, um darauf hinzuweisen, dass und wie in Innovationsprozessen in einem Spiel von Widerstand und Anpassung menschliche und materiale Wirkmacht (im vorliegenden Kontext ließe sich auch sagen: Entscheidungsbefugnisse) laufend neu verteilt werden. Grundlegend hierfür ist die Vorstellung, dass »die Konturen materieller Wirkungsmacht […] nie im vor-

hinein genau bekannt [sind]« (Pickering 2007, 24), sondern vielmehr »in der Praxis gemangelt werden« (ebd., 28).◂6
Dies trifft auch auf den Algorithmus im untersuchten Fall zu. Zunächst werden durch seine Einführung die Kompetenzen zwischen Menschen und Algorithmus neu verteilt und der Algorithmus mit Entscheidungskompetenz ausgestattet. Dann stellt sich jedoch heraus, dass die technische Schnittstelle dieser Aufgabe nicht gerecht wird und obendrein auf Dauer gestellte Gewährleistungsarbeiten verlangt. Entsprechend werden die Konturen ihrer ›Wirkmacht‹ verschoben und angepasst. Neben den technischen Arbeiten am Algorithmus zeigt sich dies vor allem am sich verändernden Stellenwert und Aufgabenbereich der Musikredaktion. Sie erhält neben ihrer aufgewerteten redaktionellen Arbeit die Aufgabe einer ›Controlling-Instanz‹ und muss durch kontinuierliches Beobachten und Bewerten des Outputs des Algorithmus für ein angemessenes qualitatives Niveau der Empfehlungen und damit auch für eine Verbesserung der Akzeptanz und Reichweite der Plattform sorgen.

Schlussfolgerungen

Wie bereits einleitend angemerkt, hat der in unserem Fall eingesetzte Algorithmus die in ihn gesetzten Hoffnungen und Erwartungen nicht erfüllt. Stattdessen wurden in der alltäglichen Praxis des Dienstes die Unzulänglichkeiten und Probleme einer algorithmengesteuerten Personalisierung des Musikangebotes deutlich, denen mit einer Re-Konfiguration der Entscheidungsbefugnisse begegnet wurde. Die damit einhergehende wechselseitige Anpassung von Mensch, Organisation und Algorithmus haben wir mit Andrew Pickering als ›Mangel der Praxis‹ beschrieben, um deutlich zu machen, dass dieser keineswegs so reibungslos funktioniert, wie dies gegenwärtig in der Diskussion um Algorithmen gerne behauptet wird. Algorithmen scheinen vielmehr immer wieder mit neuen Verhandlungs- und Entscheidungsbedarfen zu konfrontieren, auf die mit angepassten Zielen, korrigierten Konzepten und veränderten technischen Lösungen geantwortet wird. Der von uns untersuchte Online-Dienst scheint da kein Sonderfall zu sein. Auch bei als sehr erfolgreich geltenden Unternehmen zeigt sich, dass Algorithmen in der Praxis nicht einfach reibungslos arbeiten, sondern im Zuge von Anpassungsmaßnahmen ›gemangelt‹ werden. So deutet etwa der Netflix Prize (vgl. Hallinan/Striphas 2014) darauf hin, dass der bei Netflix verwendete Empfehlungsalgorithmus zwar funktioniert, sich aber offensichtlich weiterhin ›under construction‹ befindet – wofür sogar organisationsexterne Hilfen in Anspruch genommen werden.

Auch hier scheint die Delegation von Entscheidungen an Algorithmen nicht gradlinig zu immer besseren und rationaleren Entscheidungen zu führen, sondern aus sich heraus auch neue Problemlagen (die vorher nicht existiert haben) zu erzeugen, für die Lösungen gefunden werden müssen. Insofern üben Algorithmen auch nicht einfach Macht aus oder verleihen diese denjenigen, die sie programmiert haben und anwenden. Solche oder ähnliche Vorstellungen erweisen sich vor dem Hintergrund der bisherigen Argumentation als retrospektive Schönfärberei, die vor allem der Legitimation gegenüber interessierten Dritten wie der Werbeindustrie oder Geldgebern von Innovationsprojekten dienen dürfte. Die – um noch einmal mit Pickering (2007, 32) zu sprechen – zirkulär voranschreitende Dynamik von niemals vollständig gelingenden, da immer partiell scheiternden und daher widerständigen Lösungsversuchen einerseits und darauf bezogenen Anpassungen und Verbesserungen andererseits bleibt jedenfalls im Dunkeln, wenn die *Black Box* des Algorithmus nicht geöffnet wird.◂07

Anmerkungen

01▸ Wir verwenden den Begriff der ›Personalisierung‹ hier so, wie er im Feld selbst auch gebraucht wird. Mit Bezug auf Überlegungen von Elena Esposito (1995, 245ff) wäre zu überlegen, ob es soziologisch nicht angemessener wäre, von der »Spezialisierung eines unpersönlichen Mediums« zu sprechen.

02▸ So basieren angeblich 75% dessen, was die NutzerInnen sich ansehen, auf den Empfehlungen des Systems [http://www.taz.de/!5028276/].

03▸ Dies schließt natürlich nicht aus, dass an der Optimierung der Algorithmen gearbeitet wird, wie der Wettbewerb um den Netflix Prize zeigt. Allerdings geht es bei diesem Wettbewerb nicht um die Offenlegung des Empfehlungs-Algorithmus ›CineMatch‹ selbst, sondern um die Bewertung der Qualität seiner Empfehlungen.

04▸ Im Folgenden beziehen wir uns vor allem auf zwei Leitfadeninterviews mit verantwortlichen Mitarbeitern eines Unternehmens, welches personalisiertes Radio im Internet anbietet und einen Algorithmus verwendet, der die Musikauswahl automatisch steuern soll (vgl. Wehner/Passoth/Sutter 2014). Die Interviews stammen aus dem DFG-Forschungsprojekt ›Numerische Inklusion. Medien, Messungen und gesellschaftlicher Wandel‹ unter der Leitung von Jan-Hendrik Passoth, Tilmann Sutter und Josef Wehner [http://gepris.dfg.de/gepris/projekt/182070052]. Es handelt sich hierbei um Interviews, die mit dem technischen Leiter und einem Musikredakteur des Unternehmens geführt wurden.

05► Damit grenzt sich die Plattform ab gegen konkurrierende Konzepte, die eine Personalisierung entweder »user-generated« (Last.FM) oder mithilfe menschlicher Expertise (Pandora) zu erreichen versuchen.

06► Vorbild ist hier die Wäschemangel, welche aus der nassen Kleidung, wenn sie ›in die Mangel genommen wird‹, die Feuchtigkeit herauspresst und so deren Zustand verändert.

07► Auch der technische Leiter des Dienstes verweist im Interview auf diese ›rhetorische Kraft‹, wenn er anspricht, dass der Algorithmus insbesondere gegenüber Investoren auch eine »Kommunikationsaufgabe« übernähme. So soll die Verwendung der Empfehlungstechnologie mögliche GeldgeberInnen von einer Investition in den Dienst überzeugen, was scheinbar auch in zufriedenstellender Weise gelingt.

Literatur

Ang, Ien (2001) Zuschauer, verzweifelt gesucht. In: Ralf Adelmann, et al. (Hg.): Grundlagentexte zur Fernsehwissenschaft. Konstanz: UVK, S. 454-483.

Duhr, Roman (2013) Projekt „Numerische Inklusion – Medien, Messungen und gesellschaftlicher Wandel“ (Unveröffentl. Forschungsbericht).

Esposito, Elena (1995) Interaktion, Interaktivität und die Personalisierung der Massenmedien. In: Soziale Systeme 2, S. 225-260.

Esposito, Elena (2014) Algorithmische Kontingenz. Der Umgang mit Unsicherheit im Web. In: Alberto Cevolini (Hg.): Die Ordnung des Kontingenten. Wiesbaden: Springer Fachmedien Wiesbaden (Innovation und Gesellschaft), S. 233-249.

Fisher, Eran (2015) ›You Media‹: audiencing as marketing in social media. In: Media, Culture & Society 37,1, S. 50-67.

Fuchs, Christian (2014) Social media. A critical introduction. Los Angeles/London/New Delhi/Singapore/Washington, D.C: SAGE Publications.

Greve, Goetz / Hopf, Gregor / Bauer, Christoph (2011) Einführung in das Online Targeting. In: dies. (Hg.): Online Targeting und Controlling. Wiesbaden: Gabler, S. 3-21.

Hallinan, Blake / Striphas, Ted (2014): Recommended for you: The Netflix Prize and the production of algorithmic culture. In: New Media & Society, S. 1-21.

Latour, Bruno (2004) Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. In: Critical Inquiry 30,2, S. 225-248.

Muhle, Florian (i.E.) Stochastically Modelling the User: Systemtheoretische Überlegungen zur ›Personalisierung‹ durch Algorithmen. In: Thorben Mämecke / Jan-Hendrik Passoth / Josef Wehner (Hg.): Bedeutende Daten. Modelle, Verfahren und Praxis der Vermessung und Verdatung im Netz. Wiesbaden: Springer VS-Verlag.

Pickering, Andrew (2007) Die Mangel der Praxis. In: Ders.: Kybernetik und neue Ontologien. Berlin: Merve Verlag, S. 17-61.

Radermacher, Franz Josef (2015) Algorithmen, maschinelle Intelligenz, Big Data: Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz. Bundesgesundheitsbl. 58,8, S. 859-865.

Reichert, Ramón (2014) (Hg.) Big Data: Analyse zum Wandel von Macht, Wissen und Ökonomie (Digitale Gesellschaft). Bielefeld: transcript.

Schenk, Michael (2007) Publikums- und Gratifikationsforschung. In: Ders.: Medienwirkungsforschung (3. Aufl.). Tübingen: Mohr Siebeck.

Vaidhyanathan, Siva (2012) The Googlization of everything. (and why we should worry). Updated ed. Berkeley, Calif.: University of California Press.

Wehner, Josef (2008) »Social Web« – Zu den Rezeptions- und Produktionsstrukturen im Internet. In: Michael Jäckel / Manfred Mai (Hg.): Medienmacht und Gesellschaft. Zum Wandel öffentlicher Kommunikation. New York/Frankfurt a. M.: Campus, S. 197-218.

Wehner, Josef (2010) Numerische Inklusion – Medien, Messungen und Modernisierung. In: Tilmann Sutter / Alexander Mehler (Hg.): Medienwandel als Wandel von Interaktionsformen. Wiesbaden: Springer VS-Verlag, S. 183-210.

Wehner, Josef / Passoth, Jan-Hendrik / Sutter, Tilmann (2014) The Quantified Listener: Reshaping Providers and Audiences with Calculated Measurements. In: Andreas Hepp / Friedrich Krotz (Hg.): Mediatized Worlds. Culture and Society in a Media Age. Basingstoke/ New York: Palgrave Macmillan, S. 271-287.

Internetquellen:

http://**gepris.dfg.de**/gepris/projekt/182070052
http://www.**taz**.de/!5028276/