Hans Böckler Stiftung
Mitbestimmung · Forschung · Stipendien

STUDY

Study 417 · Februar 2019

CROWDPRODUKTION VON TRAININGSDATEN

Zur Rolle von Online-Arbeit beim Trainieren autonomer Fahrzeuge

Florian Alexander Schmidt

FORSCHUNGS-
FÖRDERUNG

Dieser Band erscheint als 417. Band der Reihe Study der Hans-Böckler-Stiftung. Die Reihe Study führt mit fortlaufender Zählung die Buchreihe „edition Hans-Böckler-Stiftung“ in elektronischer Form weiter.

Hans Böckler Stiftung
Mitbestimmung · Forschung · Stipendien

STUDY

Study 417 · Februar 2019

CROWDPRODUKTION VON TRAININGSDATEN

Zur Rolle von Online-Arbeit beim Trainieren autonomer Fahrzeuge

Florian Alexander Schmidt

FORSCHUNGS-
FÖRDERUNG

Autor

Dr. Florian A. Schmidt ist Professor für Designkonzeption und Medientheorie an der HTW Dresden. Seit 2006 erforscht er die Mechanismen digitaler Plattformen. Sein letztes Buch, „Crowd Design“, erschien 2017 bei Birkhäuser.

Hans-Böckler-Straße 39, 40476 Düsseldorf
www.boeckler.de

Satz: DOPPELPUNKT, Stuttgart

ISBN: 978-3-86593-330-0

INHALT

Abbildungsverzeichnis

Tabellenverzeichnis

ZUSAMMENFASSUNG

Seit 2017 gibt es einen starken Anstieg in der Nachfrage nach hochpräzisen Trainingsdaten für die KI-Modelle der Automobilindustrie. Ohne große Mengen dieser Daten ist das ehrgeizige Ziel des autonomen Fahrens nicht zu erreichen. Damit aus selbstlernenden Algorithmen selbstlenkende Fahrzeuge werden können, braucht es allerdings zunächst viel Handarbeit, die von Crowds auf der ganzen Welt geleistet wird. Sie bringen den lernenden Maschinen das Hören, das Sehen und das umsichtige Fahren bei.

Auf der Kundenseite für crowdproduzierte Trainingsdaten drängen Dutzende gut finanzierte Firmen in den entstehenden Markt für das autonome Fahren. Mit den großen Automarken (OEMs) und deren etablierten Zulieferern konkurrieren nun auch große Hardware- und Softwarefirmen wie Intel und Nvidia, Google und Apple, die bisher wenig mit Autos zu tun hatten; hinzu kommen zahlreiche risikokapitalfinanzierte Start-ups. 2018 verfügten bereits 55 Firmen über eine Lizenz zum Testen eigener autonomer Fahrzeuge in Kalifornien. Tendenz steigend. Sie alle sind für das Funktionieren ihrer Algorithmen auf präzise beschriftete – „annotierte" – Trainingsdaten angewiesen. Sie brauchen Millionen von Fotos aus dem Straßenverkehr, bei denen jedes abgebildete Pixel einer Szene semantisch einem Objekt zugeordnet wurde. Fahrbahnmarkierungen, Fahrzeuge, und Fußgänger müssen trennscharf voneinander abgegrenzt und mit Zusatzinformationen versehen werden, damit im Zuge des maschinellen Lernens daraus Regeln abgeleitet und Softwaremodelle entwickelt werden können.

Durch die deutlich gestiegenen Qualitätsanforderungen dieser neuen Kunden verändert sich die Crowdsourcing-Branche. Zwar findet die Arbeit auch auf herkömmlichen Plattformen wie Amazon Mechanical Turk (MTurk) statt, zugleich haben sich jedoch eine Reihe Plattformen entweder extra zu diesem Zweck neu gegründet oder ihr Angebot und ihre Prozesse komplett auf die Anforderungen der Autoindustrie umgestellt.

Die neuen Plattformen betreiben mehr Qualitätsmanagement, indem sie den Kunden nicht mehr einfach nur Zugang zur Crowd verschaffen, sondern die Prozesse viel stärker orchestrieren. Außerdem müssen sie immer präzisere Werkzeuge für die Durchführung der Arbeit entwickeln und die Crowdarbeiter_innen anlernen. Durch die Investition in das Training der Crowd kennen die Plattformbetreiber ihre Arbeiterschaft besser und haben ein größeres Interesse daran, sie auf der Plattform zu halten. Kunden und Crowdworker_innen kommen hingegen nicht mehr direkt in Kontakt.

Darüber hinaus versucht die Branche, sich in diesem Zuge ein moderneres Image zu geben. Der Begriff der „Crowd" gerät aus der Mode und wird durch den Zusatz „AI" als vielversprechende Formel in den Firmennamen der Plattformen ersetzt. Zwar geht es weiterhin um „humans-in-the-loop", um billige Arbeitskräfte, die hinter den Kulissen die künstlichen Intelligenzen mittels menschlicher Intelligenz trainieren und anleiten. Die Crowdplattformen setzen zugleich jedoch tatsächlich auch selbst mehr KI-Technologien zwecks Prozessoptimierung ein. Teile der Arbeit können bereits von Machine-Learning-Systemen übernommen werden, sodass die Crowdworker_innen partiell nur noch die Ergebnisse der KI überprüfen und korrigieren. Die Arbeitsteilung von Mensch und Maschine wird hier immer komplexer.

Die für die Studie befragten Arbeiter_innen sehen sich von den neuen Plattformen respektvoller behandelt und verlässlicher bezahlt als von herkömmlichen Anbietern. Auch können sie sich besser auf Teilaufgaben spezialisieren, die ihnen besonders liegen, und dadurch innerhalb der Plattform aufsteigen. Insgesamt erscheinen sie zufriedener, klagen allerdings darüber, dass es zu wenig Arbeit dieser Art gäbe und sie immer schlechter bezahlt werde.

Ein wesentlicher Faktor für die schlechte Bezahlung, die bei ein bis zwei Euro die Stunde für qualifizierte Vollzeitarbeitskräfte liegt, ist der Umstand, dass es sich um einen extrem volatilen globalen Arbeitsmarkt handelt, bei dem der Wert der Arbeit permanent aus zwei Richtungen bedroht ist: durch das ständige Wettrennen mit der Automatisierung und dadurch, dass die Arbeit dynamisch zu jenen Menschen auf der Welt fließt, welche die niedrigsten Löhne zu akzeptieren bereit sind – sei es, weil es sich um Hobbyisten handelt oder weil ihre wirtschaftliche Not besonders groß ist.

Zum Zeitpunkt der Recherche kam Venezuela hier eine Schlüsselrolle zu – einem Land mit gut ausgebildeter und gut vernetzter, jedoch von Hyperinflation völlig ausgezehrter Bevölkerung. Für viele Menschen aus Venezuela ist Crowdarbeit zur Devisen bringenden Lebensader geworden. Sie selbst sind heute Teil eines Heers von digitalen Wanderarbeiter_innen, die wie Erntehelfer zwischen den neuen Plattformen hin und her ziehen.

VORGEHENSWEISE UND STRUKTUR

Ausgangspunkt für die vorliegenden Kurzstudie war die Beobachtung Ende 2017 im Umfeld der am Code of Conduct beteiligten deutschen Crowdsourcing-Plattformen (http://crowdsourcing-code.com/) und der IG Metall, dass sich die Nachfrageseite auf dem Markt für Microtasking-Aufgaben durch neue Kunden aus der Automobilindustrie zu verändern begonnen hatte. Offenbar führte die veränderte Nachfrage seitens der Großkunden zu neuen Strukturen und anderen Arbeitsbedingungen für Crowdarbeit. Gestartet Anfang 2018 mit dem Ziel, diese Verschiebungen besser zu verstehen, handelt es sich bei der Kurzstudie um eine erste Branchenanalyse dieses gerade im Entstehen begriffenen, hochdynamischen und globalen Marktes für Trainingsdaten, die ein Schlaglicht auf die Situation der Crowdworker_innen in diesem Bereich wirft.

Die Plattformen, um die es hier geht, sitzen an der Schnittstelle zwischen Automobilindustrie, Crowdsourcing-Branche und Künstliche-Intelligenz-Forschung. Zu allen drei Teilbereichen gibt es zahllose Fachpublikationen, aber die Entwicklungen innerhalb der gemeinsamen Schnittmenge dieser drei Bereiche sind so neu, dass es, als die Studie entstand, noch keine wissenschaftlichen Publikationen zu den Wechselwirkungen von Crowdsourcing und der Produktion von Trainingsdaten für autonome Fahrzeuge gab. Eine akademische Literaturanalyse erschien wegen des als Momentaufnahme einer Branche im Umbruch verstandenen Charakters der Kurzstudie als wenig sinnvoll. Stattdessen sei hier auf die Publikationen 323, 324 (Leimeister et al. 2016) und 391 (Gegenhuber et al. 2018) der Reihe Study der Hans-Böckler-Stiftung sowie auch auf die vorherigen Publikationen des Autors (insbesondere Schmidt 2016, 2017a, Schmidt/Kathmann 2017) verwiesen, in welchen der Crowdsourcing-Diskurs in seiner historischen Tiefe sowie in der Breite der gegenwärtigen Erscheinungsformen ausführlich analysiert wird.

Die meisten Plattformen, um die es im Folgenden geht, wussten Anfang 2017 selbst noch nicht, dass sie sich nur ein Jahr später ganz auf Kunden aus dem Automobilbereich spezialisiert haben würden. Erschwerend für die Analyse kam hinzu, dass viele Unternehmen in diesem Bereich, insbesondere aufseiten der Autoindustrie, größten Wert auf Diskretion legen und sich weder bezüglich ihrer Fortschritte beim autonomen Fahren noch bezüglich ihrer Outsourcing-Praxis in die Karten blicken lassen. Auch ältere Crowdsourcing-Plattformen, die in der Vergangenheit wegen ihrer Arbeitsbedingungen öffentlich in der Kritik standen oder gar schon mit Sammelklagen der Crowd

konfrontiert waren (Schmidt 2015), sind alles andere als auskunftsfreudig, was die Arbeitsbedingungen, die Entlohnung und die Auftragsvolumina angeht. Zahlreiche Interviewanfragen liefen so leider ins Leere.

Die Studie basiert deshalb in erster Linie auf sechs explorativen Interviews mit den CEOs von Mighty AI (Daryn Nakhuda), understand.ai (Marc Mengler), Playment (Siddharth Mall), Hive (Kevin Guo), clickworker (Christian Rozsenich) und Crowd Guru (Hans Speidel) sowie fünf Interviews mit Crowdworker_innen der zu Mighty AI gehörenden Plattform Spare5. Kurze Ausschnitte aus den Interviews mit den CEOs ziehen sich quer durch die Studie. Alle Zitate der CEOs stammen, wenn nicht anders hervorgehoben, aus den 2018 zumeist auf Englisch vom Autor geführten und übersetzten Interviews.

Der Plattform Mighty AI/Spare5 kam bei der Recherche besonderes Gewicht zu, weil sie zu den Ersten gehörte, die sich ganz auf Kunden aus der Automobilbranche spezialisiert hatte. Außerdem waren die Betreiber besonders offen und kooperativ gegenüber der Studie und haben den Kontakt zu ihrer Crowd ermöglicht. Nicht zuletzt war das Phänomen der venezolanischen Crowdworker_innen an dieser Plattform sehr früh zu beobachten, wobei es über Spare5 hinausreicht.

Zum besseren Verständnis der Arbeitsabläufe wurden vom Autor Benutzerkonten auf verschiedenen Plattformen angelegt und stichprobenartig zur teilnehmenden Beobachtung genutzt. Zudem bestand Austausch mit einer professionellen Crowdworkerin auf Mechanical Turk, die dort zum Vergleich die Vergabepraxis von Trainingsdaten-Aufgaben beobachtete.

Weitere Quellen für die Studie sind die Webseiten und Pressemitteilungen der Plattformen, KI-Podcasts, Branchentagungen wie die AutomotiveIT in Berlin, Presseberichte, insbesondere aus der Start-up-Szene und aus technologienahen Magazinen wie Wired und TechCrunch sowie deren Firmendatenbank CrunchBase, auf der sich aktuelle Informationen zu Firmenaufkäufen und Finanzierungsrunden finden. Auch die von Amazon betriebene Analyseseite www.Alexa.com, die den Internet-Traffic von Webseiten aufzeichnet, hat sich als nützliche Quelle erwiesen, insbesondere um die globalen Migrationsbewegungen der digitalen Wanderarbeiter_innen nachzuvollziehen.

EINLEITUNG: CROWD, KI UND AUTOINDUSTRIE

Die hier zu analysierende Entwicklung lässt sich vereinfacht in wenigen aufeinanderfolgenden Stufen darstellen: In den letzten zehn Jahren hat die KI-Forschung beachtliche Fortschritte im Bereich des maschinellen Lernens gemacht und damit völlig neue Produkte ermöglicht.

Um die allerdings zugleich weiterhin bestehenden Unzulänglichkeiten der KI aufzufangen, wird an verschiedenen Stellen im Entwicklungsprozess auf Crowdarbeit zurückgegriffen. Seit 2012 wird das maschinelle Lernen dank des Einsatzes neuronaler Netze so schnell besser, dass am Horizont die Vision des vollautonomen Fahrens Konturen gewinnt und diverse Firmen in ein Wettrennen getreten sind, dieses Ziel als erste zu erreichen.

Die dadurch stark angeheizte Nachfrage der Automobilindustrie nach Trainingsdaten, ohne die das autonome Fahren nicht möglich ist, verändert wiederum die Crowdsourcing-Branche. Neue Plattformen entstehen, mit neuen Arbeitsprozessen, und dies hat wiederum Auswirkungen auf die Arbeitsbedingungen der Crowdworker_innen in diesem Bereich. Zugleich lässt sich deren spezielle Arbeitssituation als eine Vorahnung einer möglichen künftigen Arbeitswelt sehen, in der Menschen gerade nicht – wie häufig behauptet – durch Algorithmen ersetzt werden, sondern mit ihnen in einem komplexen und zugleich prekären Wechselverhältnis zusammenarbeiten.

Die Verknüpfungen zwischen sogenannter „künstlicher Intelligenz“ und Crowdsourcing sind vielgestaltig und reichen bis an die Anfänge der plattformbasierten Arbeitsauslagerung über das Internet zurück. Schon der Name der ersten crowdbasierten Microtasking-Plattform, Amazon Mechanical Turk (MTurk), gegründet 2005, verweist durch Bezug auf den historischen „Schachtürken“ darauf, dass hier künstliche Intelligenz durch in der Maschine verborgene Menschen vorgetäuscht wird. Lange über die Gründung hinaus bewarb Amazon die dort angebotene Dienstleistung folgerichtig auch als „*artificial* artificial intelligence“, also *künstliche* KI.

Schon 2005 konnten Auftraggeber über eine entsprechende Programmier-Schnittstelle (API), in diesem Fall genauer einer „human API“, ihre zu lösende Aufgaben an das Mensch-Maschine-System geben. Die Ergebnisse bekamen sie von einem scheinbar vollautomatischen Computer zurückgespielt. Die Crowd übernahm hinter den Kulissen genau die Aufgaben, zu deren Lösung Software noch nicht in der Lage war.

In Human-Computer-Interaction-Kreisen wird dieses Prinzip auch als „Wizard-of-OZ-Technique“ bezeichnet. Hinter dem Vorhang der scheinbar

magischen Prozesse versteckt sich eine Maschinerie, in der letztlich Menschen die Strippen ziehen. Ursprünglich war diese Technik lediglich ein Verfahren zum kostengünstigen Testen von Software-Prototypen, in welchem Menschen einen Teil der geplanten Funktionalität simulierten. Mit MTurk und diversen Folgeplattformen ist diese händische Übergangslösung selbst zum Produkt geworden. Paradoxerweise ist es heute dieses Gegenteil von Automatisierung, welches die Automatisierung erst möglich macht. Ein Ende der Arbeit ist hier nicht in Sicht.

Inzwischen ist die Entwicklung deutlich weiter vorangeschritten, und die Wechselwirkungen zwischen Crowd und KI sind komplexer geworden. Die Crowdworker_innen trainieren einerseits genau die Maschinen, die sie früher oder später ersetzen werden. Die Menschen werden andererseits jedoch nicht vollständig ersetzt, sondern lediglich Aufgabe für Aufgabe, während sie zugleich für immer neue Tasks benötigt werden, um dort wieder der Maschine auf die Sprünge zu helfen. Ein gutes Beispiel ist das maschinelle Lesen von gedruckten und gescannten Texten. Zwar gibt es schon lange OCR-Software, die dazu prinzipiell in der Lage ist, aber bis vor einigen Jahren wurde dafür dennoch häufig eine Crowd eingesetzt, etwa um Visitenkarten zu transkribieren oder den Text gescannter Bücher zu entziffern. Inzwischen werden hierfür keine Menschen mehr benötigt, weil die Maschinen dank mehr Training gut genug geworden sind.

> „Das ist eine Entwicklung, die sich in unserer Branche ständig wiederholt. Gerade die Erkennung von Objekten und Schriften in Bildern wird inzwischen gut von Algorithmen bewältigt. Solche Fortschritte in der Technik bedeuten für uns, dass Geschäftszweige wegbrechen." (Hans Speidel, CEO von Crowd Guru)

Oftmals ist es auch so, dass Algorithmen zwar eine Aufgabe schon ziemlich gut lösen können, es aber schlichtweg billiger ist, über das Internet auf den Wahrnehmungsapparat von Hilfsarbeiter_innen zuzugreifen. Das Versprechen von Diensten wie MTurk ist es, dass man als Auftraggeber mit der gleichen funktionalen Leichtigkeit auf natürliche neuronale Netze wie auf künstliche neuronale Netze zugreifen können soll. Deswegen ist in dieser Branche nicht nur die Rede von „human intelligence tasks" (HITs) sondern auch von „humans-as-a-service" – menschliche Hirnleistung wird hier wie Prozessorleistung als technische Dienstleistung vermietet. Genau diese Reduktion auf die Rolle von Maschinen bzw. das Verschwinden in der Maschine ist aber wiederum etwas, woran viele Crowdworker_innen leiden (Irani 2013, Irani/Silberman 2013).

Während menschliche Kognition als technische Dienstleistung unter Wert vermietet wird, um künstliche Intelligenz zu produzieren, verirrt sich die öffentliche Diskussion über „künstliche Intelligenz" in der mystisch überhöhten Vorstellung, es handle sich dabei um eine menschenähnliche, generelle Intelligenz, wie wir sie aus zahllosen Science-Fiction-Filmen kennen. Eine derartig imaginierte KI hat ein allgemeines Verständnis von der Welt, und ihr werden Vernunftbegabung sowie eigene, oftmals böswillige Intentionen angedichtet. Von solchen „Skynet-Szenarien" wie in den Terminator-Filmen ist die Realität zum Glück weit entfernt. Tatsächliche KI-Anwendungen sind zwar sehr leistungsstark, aber viel banaler, weil extrem schmalspurig.

Reale Machine-Learning-Produkte wie etwa die „Autokorrektur" verlieren außerdem genau in dem Moment den futuristischen Anklang „künstlicher Intelligenz", in dem sie als stabil funktionierende Standardwerkzeuge in den Alltag integriert und damit quasi unsichtbar werden. Die „KI verwandelt sich dann in IT, in normale Computertechnologie", wie es der Berliner KI-Forscher Raúl Rojas González ausdrückt (Schmidt 2017b). Manche deutsche KI-Expert_innen übersetzen das aufgeladene Kürzel deshalb nüchtern mit „künftige Informatik".

Im gegenwärtigen Hype um die KI werden sowohl die Kreativleistung der Programmierer_innen als auch die Trainingsleistung der Crowd ausgeblendet, um die Technik beeindruckender erscheinen zu lassen. Aus diesem Blickwinkel ist Crowdwork das schmutzige Geheimnis vieler Automatisierungsverfahren, und das Kürzel KI bzw. meist AI verdeckt als verkaufswirksames Etikett die Arbeitsleistung der Crowd. Die amerikanische Plattformkritikerin und Aktivistin Astra Taylor hat für derartige Pseudoautomatisierungen, welche die tatsächlichen Arbeitsbedingungen verschleiern, den treffenden Begriff „Fauxtomation" eingeführt (Taylor, 2018).

Ab welchem Punkt man im Amalgam aus menschlicher und maschineller kognitiver Leistung, von „künstlicher" Intelligenz sprechen kann, bleibt letztlich eine Frage der Definition. Unverfänglicher ist es, von maschinellem Lernen zu sprechen und den menschlichen Anteil bei der Planung, dem Training und der Kontrolle nicht aus dem Blick zu lassen.

Bezeichnenderweise gehen die großen Fortschritte bei der Automatisierung auf einen Paradigmenwechsel in der KI-Forschung zurück, im Zuge dessen man den langjährigen Versuch verworfen hat, der Maschine logisches, deduktives Denken beizubringen. Man versucht der Maschine heute gerade nicht mehr zu erklären, dass ein Auto eine motorisierte Blechkiste mit einem Steuer, vier Rädern, fünf Türen und einer ähnlich endlichen Anzahl von Eigenschaften und Verhaltensweisen ist, sondern lässt einen Algorithmus an-

hand Hunderttausender gut beschrifteter Bilder selbst seine Schlüsse ziehen. Ohne auf einer symbolischen Ebene wirklich zu verstehen, was ein Auto ist, kann die Maschine irgendwann dann auch in neuen, unbeschrifteten Bildern selbstständig Autos erkennen, wobei danach nochmals Menschen zum Einsatz kommen, um die Ergebnisse zu überprüfen.

Menschen produzieren also die Trainingsdaten und prüfen und korrigieren die sogenannten Validierungsdaten. Es handelt sich hierbei um „supervised learning", die Maschine lernt unter Aufsicht, eigene Korrelationen zu entdecken und daraus Regeln abzuleiten, allerdings ohne dass diese dann noch auf für Menschen nachvollziehbare logische Schlüsse und Kausalitäten beruhen müssten. Es handelt sich nicht um ein Verstehen im menschlichen Sinne, sondern um statistische Mustererkennung auf Basis von Wahrscheinlichkeiten, und das funktioniert so außerordentlich gut, dass diese Technologie uns derzeit eine Fülle neuer KI-Produkte und -Dienstleistungen beschert. Potenziell sogar selbstfahrende Autos.

Sowohl bei der Entwicklung dieser Verfahren als auch bei dem darauf aufbauenden Einsatz der Technologie in autonomen Fahrzeugen kommt Google eine Vorreiterrolle zu. Das Google „Brain Team" betreibt seit vielen Jahren Grundlagenforschung auf dem Gebiet des „Deep Learning" für diverse Google-Dienstleistungen, vom Sprachassistenten bis zur Suche. Schon 2012 erzielte man hier den ersten Durchbruch beim Erkennen und geografisch präzisen Verorten von Hausnummernfotos aus Google Street View mittels Deep-Learning-Techniken. Inzwischen hat Google nach eigenen Angaben in den meisten Ländern der Welt über 95 Prozent aller Hausnummern auf diese Weise erfasst und zu Google Maps hinzugefügt (Arnoud 2018).

Der Durchbruch dank Einsatz von Deep-Learning-Verfahren bzw. neuronalen Netzen, die mittels gestaffelter Erkennungsebenen und in iterativen Schleifen stufenweise ein semantisches dreidimensionales Modell von der Umgebung erstellen können, ist Teil des Entwicklungserfolges. Die Algorithmen sind zwar extrem rechenintensiv, bestehen aber aus vergleichsweise wenigen Zeilen Programmiercode. Was sie jedoch benötigen, um funktionieren zu können, sind riesige annotierte Trainingsdatensätze.

2013 musste die oben erwähnte Hausnummernzuordnung laut Google noch mit Milliarden von Bilddaten trainiert werden. Inzwischen kommt man hier mit weniger Daten aus, es werden aber immer noch Millionen von annotierten Bildern für das Training benötigt. Mittels der Software reCAPTCHA lagert Google wiederum einen Teil dieser Arbeit an Milliarden von Internetnutzer_innen aus.

Musste man früher noch verzerrte Wörter aus Googles Buch-Scan-Projekt

entziffern um beim Anlegen eines Onlinekontos zu beweisen, dass man ein Mensch ist, so muss man inzwischen eher Hausnummern, Verkehrsschilder oder Fahrzeuge erkennen. Das reicht für eine erste grobe Annotierung. Inzwischen werden allerdings, wie wir sehen werden, so präzise Beschriftungen gebraucht, dass man dies nicht mehr mit beliebigen Internetnutzer_innen nebenher bewerkstelligen kann, sondern trainierte und hochfokussierte Crowdworker_innen braucht.

Im Dienst des autonomen Fahrens kommen crowdgestützte Machine-Learning-Verfahren inzwischen in vielfältiger Weise zum Einsatz. Es ist eines der weitreichendsten und besonders schnell wachsenden Anwendungsgebiete dieser Technologie. Es zeichnet sich dadurch aus, dass hier ein breites Spektrum an sehr gut finanzierten Unternehmen in ein Wettrennen um die Automatisierung des Autos getreten ist und dass hier der Anspruch an die Qualität der Trainingsdaten außerordentlich hoch ist. Schließlich geht es um Leben und Tod. Nach den ersten fatalen Unfällen mit autonomen Fahrzeugen von Tesla 2016 und von Uber 2018 ist dies kein hypothetisches Szenario mehr.

> „Bei der Onlinesuche kann man schon mit neunzig Prozent Genauigkeit gut arbeiten. Es schadet niemandem, wenn jedes zehnte Ergebnis falsch ist. Für die Unfallgefahr eines Autos sind selbst neunundneunzig Komma neun Prozent Genauigkeit, also ein schwerer Unfall alle tausend Fahrten, absolut inakzeptabel. Im Straßenverkehr geht es schließlich um Menschenleben, und entsprechend hoch ist der Preis, wenn man hier mit seinen algorithmischen Modellen falschliegt." (Siddharth Mall, CEO von Playment)

Die Kontrahenten auf dem entstehenden Markt für autonome Fahrzeuge kommen aus unterschiedlichen Industriezweigen, verfolgen unterschiedliche Strategien, und die Motivationen für die Automatisierungsbestrebungen reichen von Sicherheit über Bequemlichkeit bis hin zu Arbeitsplatzeinsparung. Doch sie alle brauchen Trainingsdaten, und zwar sehr viele und besonders genaue, denn die Qualität der Machine-Learning-Modelle, die den Menschen hinter dem Steuer ersetzen sollen, kann nicht besser sein als die Daten, mit denen sie angelernt werden. Wie es so schön heißt unter Informatiker_innen: „Garbage in, garbage out."

Manche der neuen Zulieferfirmen für KI-Trainingsdaten konzentrieren sich auf die Verarbeitung von Bild- und Sensordaten aus dem Straßenverkehr, andere auf die Kommunikation zwischen Fahrzeug und Passagieren. Die Interaktion mit dem Fahrzeug ist über Sprachassistenzsysteme und Gestensteuerung zunehmend digital vermittelt, und die Maschine behält mittels

Gesichtserkennung zudem den Wachheitsgrad und die Fahrtüchtigkeit der Menschen hinter dem Steuer im Blick. All dies geschieht ebenfalls mit Machine-Learning-Verfahren.

Insbesondere die etablierten Hersteller halten sich sehr bedeckt, wenn es um Einblicke in ihre konkreten Entwicklungsfortschritte bezüglich des autonomen Fahrens geht (Laing 2018). Eine gute Quelle ist jedoch das Department of Motor Vehicles DMV CA, die kalifornische Kfz-Zulassungsbehörde, denn in diesem US-Bundesstaat testen die meisten Firmen ihre autonomen Prototypen im Straßenverkehr (DMV CA, 2017, 2018). Ein weiterer US-Bundesstaat mit ähnlich liberaler Gesetzgebung (und gutem Wetter) ist Arizona, hier testen beispielsweise Uber und Google ihre Fahrzeuge.

Um entsprechende Lizenzen zu erhalten, müssen sich die Firmen nicht nur anmelden, sondern auch im Nachhinein darüber Bericht erstatten, wie häufig Testfahrer_innen ins Steuer greifen mussten. Für 2017 hatten bereits 20 Firmen sogenannte „disengagement reports" eingereicht, darunter die deutschen Firmen BMW, Bosch, Mercedes und VW. Lizenzen zum Testen autonomer Fahrzeuge hatten 2018 bereits 55 Firmen – neben etablierten Fahrzeugherstellern und Zulieferern auch Hardware- und Softwarefirmen wie Apple, der Grafikkartenhersteller NVDIA, die Suchmaschinenfirmen Alphabet (Google, Waymo) und Baidu (aus China) sowie verschiedene Startups aus dem Silicon Valley.

Die klassischen Automobilfirmen befinden sich auf diesem Feld also in unmittelbarer Konkurrenz zu einigen der weltweit wertvollsten Technologieunternehmen, die zudem große Expertise auf dem Gebiet des Machine-Learning vorzuweisen haben und aufgrund ihrer sonstigen Geschäfte (siehe z. B. Google, mit den Anwendungen Maps und Street View) schon sehr große und für das autonome Fahren hochrelevante Datenbanken anlegen konnten.

Es gibt allerdings einen entscheidenden Unterschied in der Strategie: Während die etablierten Hersteller auf eine schrittweise „evolutionäre" Ausweitung verschiedenster Fahrerassistenzsysteme der Stufen drei und vier setzen und davon ausgehen, dass Menschen noch lange Zeit in brenzligen Situationen das Steuer werden übernehmen müssen, setzen Firmen wie Google auf einen „revolutionären" Markteintritt direkt auf Stufe fünf – also der vollständigen Automatisierung ohne Vorhandensein eines Lenkrads oder Erfordernis eines Führerscheins.

Auch Blinde sollen alleine Autofahren können, so die werbewirksame und hoffnungsvolle Vision. Durch Menschen verursachte Unfälle sollen der Vergangenheit angehören. Darüber hinaus geht es sicherlich auch darum, die Aufmerksamkeit der ehemaligen Fahrer_innen, jetzt Passagiere, für mehr

Zeit vor dem Bildschirm freizubekommen (immerhin Googles Kerngeschäft), und natürlich auch um Einsparung von Arbeitsplätzen bei Taxi- und Lastwagenfahrer_innen. Letzteres Ziel wird besonders von Firmen wie Uber, die ebenfalls in den Markt für autonomes Fahren drängen, ganz offen als Hauptmotivation genannt.

Das Rennen ist noch nicht entschieden. Die meisten Patente im Bereich des autonomen Fahrens zwischen 2010 und 2017 wurden bezeichnenderweise von Bosch angemeldet (958), gefolgt mit einigem Abstand von Audi (516) und Continental (439). Auch BMW, VW und Daimler finden sich unter den Top 10. Aus dieser Perspektive betrachtet ist Deutschland also führend auf dem Gebiet, gefolgt von Japan und den USA (Bardt 2017, DPMA 2018).

Betrachtet man jedoch die bereits erfolgreich mit vollautonomen Fahrzeugen zurückgelegten Strecken, ergibt sich ein anderes Bild. Hier liegt Waymo, die zum Google-Konzern Alphabet gehörige Firma für selbstfahrende Autos, mit großem Abstand vorn. Die Testfahrzeuge von Waymo haben inzwischen etwa 6,5 Millionen Kilometer Strecke im öffentlichen Straßenverkehr. Auch bei der Frequenz, mit der noch ein Mensch ins Steuer greifen muss, liegt Waymo deutlich vor der Konkurrenz: fast 6.000 Meilen schafften es die Fahrzeuge der Firma ohne menschliche Intervention; bei General Motors, dem zweitbesten Entwickler in diesem Vergleich, waren es nur 1.300 Meilen. Alle anderen Firmen, die 2017 einen offiziellen Bericht abgaben, lagen weit hinter diesen Werten zurück.

Der große Vorsprung lässt sich teils damit erklären, dass Google bereits 2009 mit der Entwicklung autonomer Fahrzeuge begonnen hatte, sogar bevor die Deep-Learning-Revolution industriell nutzbar wurde. Die Ausgründung von Waymo im Jahre 2016 war für Google schon der Übergang aus der reinen Forschungsphase hin zur Entwicklung eines kommerziellen Endprodukts in industriellem Ausmaß. Seit Ende 2017 wagt es Waymo, die Prototypen ganz ohne menschliche „Safety-Driver“ als Rückversicherung in den Straßenverkehr zu schicken. In Phoenix, Arizona, hat die Firma eine Lizenz dafür erhalten und im Dezember 2018 den Dienst für erste Testfahrgäste geöffnet (Krafcik 2018).

Trotz dieser beachtlichen Fortschritte warnt selbst Waymos Chefentwickler Sacha Arnoud davor, die noch zu lösenden Problem zu unterschätzen. Selbstfahrende Autos seien zwar „fast schon da“, aber eben nur fast. 90 Prozent der Probleme habe man in den letzten acht Jahren gelöst, aber die verbleibenden zehn Prozent würden das Zehnfache des bisherigen Entwicklungsaufwands kosten, um ein unter allen Straßenbedingungen sicheres autonomes Fahrzeug bis zur Marktreife zu entwickeln (Arnoud 2018).

Für ein vollautonomes Fahrzeug reicht es nicht aus, andere Verkehrsteilnehmer_innen im Bilddatenstrom zu erkennen. Die Maschine muss Modelle von deren Verhalten, Absichten und wahrscheinlichen künftigen Positionen in Zeit und Raum bilden und vor allem mit unzähligen Grenz- und Sonderfällen, wie Baustellen, Notfallfahrzeugen, Schneegestöber etc., in Echtzeit so zurechtkommen, dass unter keinen Umständen ein Mensch zu Schaden kommt. Die Anforderungen und der Komplexitätsgrad sind also außerordentlich hoch.

> „Gerade selten auftretende, aber für den Algorithmus verwirrende Ereignisse sind kritisch. Diese Randfälle muss immer ein Mensch validieren – also prüfen, ob der Computer die richtige Einschätzung vorgenommen hat. […] Ohne Menschen wäre diese Arbeit nicht möglich, dafür ist das Leben da draußen zu vieldimensional und komplex. Im Straßenverkehr passieren die verrücktesten Sachen – z. B. Kängurus –, wenn der Algorithmus einen Sonderfall nicht vorab gelernt hat, kann er damit nicht umgehen, und es passieren Unfälle." (Marc Mengler, CEO von understand.ai)

Obwohl die bereits existierenden Prototypen schon sehr viel mehr können, als man noch vor wenigen Jahren für möglich gehalten hatte, sind die zwischenzeitliche Euphorie und der Optimismus, das Problem bald gelöst zu haben, nach aktuellen Rückschlägen wie den ersten tödlichen Unfällen deutlich gedämpft.

Ein heikler Punkt unter vielen bleibt das mitunter erratische oder bewusst regelwidrige Verhalten von Fußgänger_innen, die etwa bei Rot über die Ampel laufen oder an Stellen, wo sie es nicht sollten, die Straße kreuzen. Dies hat den Machine-Learning-Experten Andrew Ng, inzwischen Vizepräsident von Baidu, vorher bei Google, 2018 zu der viel kritisierten Aussage veranlasst, man müsse erst die Fußgänger umerziehen, bevor die autonomen Fahrzeuge funktionieren würden (Kahn 2018). Da hiermit jedoch nicht zu rechnen ist, wird die Autoindustrie auf unbestimmte Zeit weiter darauf angewiesen sein, die Fahrzeuge mithilfe von Crowdworker_innen auf alle nur erdenklichen Sonderfälle hin zu trainieren, z. B. tatsächlich Kängurus.

> „Kängurus sind wegen ihrer unorthodoxen Fortbewegungsweise eines dieser bisher ungelösten Probleme. Wenn sie springen, sind sie für ein paar Momente frei in der Luft, und die kameraabhängigen autonomen Fahrzeuge haben große Schwierigkeiten damit, diese für sie scheinbar schwebenden Objekte in Distanz und Geschwindigkeit richtig einzuschätzen." (Marc Mengler, understand.ai)

ANALYSE: CROWDPRODUKTION VON TRAININGSDATEN

Die Kurzstudie war geleitet von vier Grundfragen:

1. Welche Automobilfirmen nutzen welche Crowdplattformen?
2. Wie wirkt sich deren Nachfrage auf die Crowdsourcing-Branche aus?
3. Wie verändert sich dadurch die Situation der Crowdworker_innen?
4. Handelt es sich bei der Crowdproduktion von Trainingsdaten um ein temporäres Phänomen oder eine langfristige Perspektive für Crowdworker_innen?

Die zweite und dritte Frage werden in diesem Kapitel ausführlich untersucht. Auf die vierte Frage wird in Diskussion und Ausblick eingegangen. Was jedoch die erste Frage angeht, so zeigte sich leider, dass hierzu aufgrund von Complianceregeln und Verschwiegenheitserklärungen keiner der Interviewpartner offizielle Angaben machen durfte. Die Plattformen führen zwar zahlreiche kleinere Start-ups auf ihren Webseiten als Kunden auf, die Namen von etablierten Herstellern sucht man jedoch bis auf wenige Ausnahmen vergeblich. Bei der stichprobenartigen eigenen Arbeit als Crowdworker tauchten immer wieder Bilddatensätze auf, die anhand von Straßenschildern und markanten Gebäuden eindeutig Stammregionen deutscher Autobauer zuzuordnen waren, aber eine aussagekräftige Zuordnung war auf diese Weise natürlich nicht möglich, und so wurde diese Frage schließlich als im Rahmen der Kurzstudie unbeantwortbar verworfen.

Man kann jedoch auch so einige Rückschlüsse auf die Größe des Marktes ziehen. Mindestens die 55 Hersteller, die bereits offiziell eigene autonome Prototypen im Straßenverkehr testen, realistisch betrachtet aber noch deutlich mehr Firmen, wenn man all die Zulieferer ohne vollständig fahrtüchtige eigene Systeme mitbedenkt, brauchen heute millionenfach eigene Trainingsdaten.

Als Alternative zu den Crowdsourcing-Plattformen bleibt den Firmen nur die Möglichkeit, auf eine Mischung aus unspezifischen generischen Daten aus offenen Datenbänken und firmenintern antrainierten Werkstudent_innen zurückzugreifen. Der CEO einer deutschen Zulieferfirma, der hier nicht genannt werden möchte, erklärte dem Autor im Zuge der Recherche, dass man, wenn es um noch in der Entwicklung befindliche Fahrzeugmodelle der Hersteller gehe, auf studentische Kräfte vor Ort angewiesen sei, weil die

sensiblen Daten das Firmengelände nicht verlassen dürften. Eine Auslagerung an die Crowd wäre schon aus Compliancegründen nicht möglich. Andererseits sind der firmeninternen Annotation je nach Datentyp und Umfang deutliche Grenzen gesetzt:

> „Bei Verhandlungen mit Kunden treten wir bisher nicht direkt gegen externe Konkurrenten an, sondern gegen die Überlegung mancher Auftraggeber, die Arbeit lieber im eigenen Betrieb zu erledigen. Deren Forschungsabteilungen würden am liebsten ausschließlich selbst erstellten Datensätzen vertrauen, aber deren hoch bezahlte Datenwissenschaftler können ja schlecht eigenhändig zehntausend Bilder einer semantischen Segmentierung unterziehen. Das skaliert nicht, würde ewig dauern, sich finanziell nicht lohnen und die Forscher von ihrer eigentlichen Arbeit abhalten. Deshalb wenden sich die Kunden an uns. Insbesondere jetzt, wo der Anspruch an Qualität und Expertise sehr viel höher ist als noch vor zwei Jahren." (Daryn Nakhuda, Mighty AI)

Aufschlussreich für die hohe Nachfrage nach Trainingsdaten ist darüber hinaus die extreme Redundanz, mit der diese Arbeit in Auftrag gegeben wird, vor allem weil sich die Auftraggeber in einer ausgeprägten Konkurrenzsituation befinden, dementsprechend ihre Daten nicht mit Wettbewerbern teilen wollen und es aufgrund von Spezialisierungen ab einem gewissen Punkt in der Entwicklung auch nicht mehr können.

Global betrachtet wäre es augenscheinlich viel effizienter, wenn die zahlreichen Kunden für Trainingsdaten einen gemeinsamen Pool an annotierten Daten anlegen würden. Und das gilt insbesondere für die vielen Unternehmen aus der Automobilindustrie. Doch hier siegt – vorerst zumindest – Konkurrenz über Kooperation. Zwar gibt es auch Anbieter vorproduzierter Trainingsdatensätze, z. B. Octant (www.octant.io) und Mapillary (www.mapillary.com), doch noch immer legen die meisten Kunden Wert auf maßgeschneiderte eigene Trainingsdatensätze.

> „An den vorproduzierten Trainingsdaten scheiden sich die Geister. Manager finden die Idee super, denn sie wollen unbedingt eine Lösung einkaufen, die sie versichern bzw. rückversichern können. Deswegen hätten sie gerne ein generisches System. Die meisten Techniker sind jedoch der Auffassung, dass es nicht möglich ist, die notwendigen neunundneunzig Komma neunhundertneunundneunzig Prozent aller Menschen im Straßenverkehr mit einem generischen System zu erkennen." (Marc Mengler, understand.ai)

Darüber hinaus ist es so, dass insbesondere bei den künftig immer wichtiger werdenden Validierungsdatensätzen, bei denen die Crowd rückwirkend prüft, ob ein proprietärer Algorithmus die richtige Entscheidung getroffen

hat, es auch nicht anders gehen kann, denn diese Daten sind firmen- und fallspezifisch und unterliegen besonderer Geheimhaltung. Die Auftraggeber aus der Automobilindustrie brauchen also mindestens mittelfristig weiterhin große Crowdkapazitäten zur individuellen Auswertung ihrer Daten.

> „Bis zu einem gewissen Entwicklungsstand könnten die Kunden auch auf generische Open-Source-Datensätze zurückgreifen. Aber irgendwann haben sie es immer mit Spezialproblemen zu tun. Manche Kunden sind z. B. besonders am Verhalten von Fußgängern interessiert, andere wollen anhand des Gesichtsausdrucks von Menschen deren Grad der Aufmerksamkeit analysieren. Letzteres ist für viele Firmen relevant, aber sie orientieren sich z. B. an verschiedenen Ankerpunkten im Gesicht – die einen schauen mehr auf die Augen, die anderen auf die Augenbrauen. Es kann gut sein, dass Trainingsdaten für allgemeine Lösungen in Zukunft leichter wiederverwendbar sind. Dem steht allerdings der enorme Wettbewerb zwischen den Firmen für autonomes Fahren entgegen. Vorerst will da niemand seine Daten teilen." (Daryn Nakhuda, Mighty AI)

Nach Einschätzung von Plattformbetreibern wie Daryn Nakhuda, CEO von Mighty AI aus Seattle, und Marc Mengler, CEO von understand.ai aus Karlsruhe, ist davon auszugehen, dass derzeit praktisch alle großen Hersteller ihre Trainingsdaten zumindest versuchsweise bei diversen Plattformen parallel in Auftrag geben – auf der Suche nach dem besten Verfahren und dem besten Preis-Leistungs-Verhältnis. Marc Mengler hält es deshalb für wahrscheinlich, dass es bald zu einer Konsolidierung bei den Plattformen für Trainingsdaten kommen wird und sich nur einige wenige hoch spezialisierte Plattformen durchsetzen werden.

Bisher ist die Anzahl der Anbieter allerdings weiter am Wachsen, und die Venture-Capital-Flüsse in Kombination mit der veränderten Selbstdarstellung der Plattformen und den Aussagen der Experten in den Interviews zeichnen durchweg das Bild, dass hier mit vorerst weiter wachsenden Auftragsvolumina durch die Automobilindustrie gerechnet wird.

Die nachfolgende Tabelle gibt einen Überblick über die wichtigsten Plattformen für Microtasking-Crowdwork im Allgemeinen und für Trainingsdaten im Besonderen. Die Unternehmen sind hier sortiert nach ihrem Alexa-Rank, welcher den Traffic misst, der den Plattformen zufließt (www.alexa.com/siteinfo, zu Amazon gehörig, Stand: Mai 2018). Dies ist gegenüber der von den Plattformen selbst angegebenen Größe ihrer Crowd die aussagekräftigere Kennziffer, da zur Crowd in den offiziellen Angaben meist alle jemals registrierten Arbeitskräfte gezählt werden, der Anteil an „Karteileichen" ist also hoch. Der Traffic hingegen wird ständig neu und von außen erfasst.

Wie eingangs erwähnt haben die CEOs von clickworker, Mighty AI, Hive, Playment, Crowd Guru und understand.ai die Studie mit Interviews unterstützt. Amazon Mechanical Turk und Figure Eight, Scale und Samasource waren trotz mehrfacher Anfragen leider nicht für Interviews zu gewinnen.

Über die Alexa-Webtraffic-Analyse lässt sich außerdem nachvollziehen, für welche Länder die jeweiligen Arbeitsplattformen besonders relevant sind. Die Plattformen Spare5 (https://app.spare5.com) und Hive Work (https://hivemicro.com) werden z. B. insbesondere von Menschen aus Venezuela angesteuert (bei Spare5 sind es schon seit Anfang 2018 75 %, bei Hive Work stieg der Anteil von 55 % im Mai 2018 auf 75 % im Dezember 2018); Remotasks (https://www.remotasks.com/) zieht mehr Traffic aus den Philippinen und Indien (jeweils 31 %); Playment (https://playment.io/) fast ausschließlich aus Indien (85 %). Die deutsche Plattform clickworker (https://www.clickworker.com/) zieht ihren Traffic vorwiegend aus den USA (27 %) und Deutschland (18 %). Crowd Guru (https://www.crowdguru.de/) ist mit 93 Prozent ganz auf Deutschland fixiert (zur Sonderrolle deutscher Plattformen siehe den vorletzten Abschnitt der Analyse).

Samasource (https://www.samasource.org/) ist neben dem Auftraggeberland USA (29 %) für Menschen aus Kenia (23 %) und Indien (16 %) interessant; CloudFactory (https://www.cloudfactory.com/) für Menschen aus Kenia (41 %) und Nepal (24 %). Wobei zu beachten ist, dass Samasource und CloudFactory keine offenen Onlineplattformen für an heimischen Computern verstreute Crowds betreiben, sondern als sogenannte BPOs (Business-Process-Outsourcing-Firmen) Menschen in callcenterähnlichen Werkhallen gebündelt an ausgewählten Standorten in Entwicklungsländern beschäftigen. Sie sind hier mit aufgeführt, weil sie inzwischen ebenfalls gezielt Kunden aus der Automobilindustrie mit der Produktion von Trainingsdaten umwerben; seit Ende 2018 führt Samasource z. B. auch VW als Kunden an.

CloudFactory beschäftigt z. B. 3.000 Menschen auf Vertragsbasis, die etwas mehr Sicherheiten genießen als Crowdworker_innen. Die Firma positioniert sich ausdrücklich als Gegenentwurf zu Crowdwork im Stil von MTurk, mit dem Argument, dass externe, anonyme, untrainierte Arbeitskräfte nicht genug Qualität und Sicherheit böten. Mit den Standorten in Entwicklungsländern und dem erklärten Ziel, mit dieser Form der lokal gebündelten Arbeitsauslagerung Entwicklungshilfe zu leisten, unter anderem dadurch, dass zweieinhalbmal mehr als der örtliche Mindestlohn gezahlt werde, ähnelt ihr die Firma Samasource. Beide Plattformen bewerben ihre Dienstleistung mit dem Kunstwort „impact sourcing“, um die positiven Auswirkungen sozial verantwortlichen Outsourcings auf die Regionen zu betonen.

Tabelle 1

Die wichtigsten Microtasking-Plattformen für Trainingsdaten

Firma	Plattform	Sitz (gegr.)	Alexa-Rang	Crowdgröße	Finanzierung
Amazon	MTurk	USA (2005)	5.800	500.000	–
Appen	Appen	AUS (1998)	21.000	1.000.000	(Börse: APX)
Figure Eight	(Diverse)	USA (2007)	30.000	(5.000.000)	$58 Mio.
clickworker	clickworker	BRD (2005)	35.000	1.200.000	$13,7 Mio.
Mighty AI	Spare5	USA (2014)	37.000	500.000	$27,3 Mio.
Hive (AI)	Hive Work	USA (2013)	49.000	300.000	$18 Mio.
Playment	Playment	IND (2015)	168.000	300.000	$2,3 Mio.
Scale	Remotasks	USA (2016)	187.000	–	$4,6 Mio.
CloudFactory	(vor Ort/BPO)	UK (2011)	(334.000)	(3.000)	$13 Mio.
Crowd Guru	Crowd Guru	BRD (2008)	416.000	52.000	–
Samasource	(vor Ort/BPO)	USA (2008)	(815.000)	(7.000)	$1,5 Mio.
Alegion	(Diverse)	USA (2011)	855.000	–	$4,1 Mio.
understand.ai	(vor Ort/BPO)	BRD (2016)	(3.300.000)	–	$2,8 Mio.

Quelle: Eigene Darstellung

Der Mindestlohn in Nepal liegt seit 2016 bei 3,74 US-Dollar pro Tag. In einem Interview von 2014 erklärte Mark Sears, CEO von CloudFactory, die Firma stelle ihren Kunden sechs Cent die Minute bzw. 3,60 US-Dollar die Stunde in Rechnung.

CloudFactory versteht sich als Virtual Production Line bzw. als Fließbandarbeit „in den Wolken" und bezieht sich direkt auf Henry Fords radikale Arbeitsteilung. Diese ermögliche es CloudFactory, auch mit ungelernten Arbeitskräften komplexe Produkte zusammenzusetzen. Eine Strategie, die von vielen Interviewpartnern diese Kurzstudie ebenfalls als essenziell eingeschätzt wird.

Die Plattformen Hive (bzw. Hive AI, Hive Work, Hive Micro, die Schreibweise variiert) und understand.ai nehmen insofern eine Sonderrolle in der Tabelle ein, als beide Gründer ihre Verankerung in der KI-Forschung betonen.

> „Ursprünglich hatten wir Hive bloß gegründet, weil wir für unseren eigenen Bedarf große Datensätze annotieren mussten. Jetzt nutzen wir die Plattform auch für unsere Kunden. […] Wir liefern ihnen nicht nur Trainingsdaten, sondern helfen ihnen auch dabei, eigene KI-Modelle zu entwickeln. Unternehmen aus der Automobilbranche unterstützen wir dabei, Wahrnehmungsmodelle (perception models) für ihre Fahrzeuge zu entwickeln." (Kevin Guo, CEO von Hive)

Sowohl Hive-CEO Kevin Guo als auch understand.ai-CEO Marc Mengler haben ihre Firmen gegründet, weil sie in ihrer Forschungsarbeit an Machine-Learning-Modellen die Schlüsselfunktion von Trainingsdaten erkannt und dann eigene Lösungen für deren Produktion entwickelt haben. Sie sind also nicht aus herkömmlichen eCommerce-Crowdsourcing-Plattformen hervorgegangen, wie es z. B. bei Mighty AI und Playment der Fall ist. Doch während Hive explizit auf massenhafte, redundante Crowdarbeit als Produktionsprinzip setzt, distanziert man sich bei understand.ai von diesem Verfahren:

> „[Was wir anbieten] hat Elemente von Crowdsourcing. Aber für uns arbeiten keine Freizeitjobber, sondern dezidierte Teams in Indien und in Deutschland. Allein mit Menschen aus Indien, die sich mit den Verkehrsregeln, der Beschilderung und dem Aussehen von Straßenszenen in Europa nicht auskennen, könnten wir die hohen Qualitätsauflagen insbesondere der europäischen Autoindustrie nicht erfüllen. Als Zulieferer sind wir zudem zu Konformität mit dem Compliance-Code und den Werten unserer Auftraggeber verpflichtet. Und eine der obersten Regeln ist: Kein Crowdsourcing. Die Firmen wollen

> keine Freizeitkräfte, die sich einfach per App anmelden und dann Kundendaten einsehen können. […] Außerdem ist es fast unmöglich, Tausende Crowdworker so zu trainieren, dass sie einheitlichen Standards folgen.“ (Marc Mengler, understand.ai)

Bei understand.ai arbeitet man deshalb mit einer Mischform aus studentischen Kräften in Deutschland, einem BPO-ähnlichen Ansatz in Indien und vor allem einem hohen Einsatz von maschinellem Lernen, um die Handarbeit auf ein Minimum zu reduzieren. Auf Student_innen setze man deshalb, weil die Arbeit „stark konzentrationsabhängig“ sei und diese Gruppe diese Fähigkeit in der Schule besser gelernt habe als andere Arbeitskräfte. „Unserer Erfahrung nach kann man nicht in Vollzeit fokussiert Bilder annotieren. Deshalb fahren wir ganz gut damit, unsere Leute nur 20 Stunden die Woche einzusetzen“ (Marc Mengler, understand.ai). Im Gegensatz dazu setzt man bei Hive auf massenhaft untrainierte Arbeitskräfte.

> „Manche Plattformen setzen auf einzelne hochtrainierte Individuen, wir verfolgen da den gegensätzlichen Ansatz. Ich glaube an die Weisheit der vielen. Um bei Hive zu einem Ergebnis für einen bestimmten Datenpunkt zu gelangen, setzen wir voraus, dass sich eine ganze Reihe von Arbeitskräften einig ist. Es gibt immer eine Abstimmung, einen Crowdkonsens zwischen ungefähr sechs Leuten.“ (Kevin Guo, Hive AI)

Etablierte Generalisten und neue Spezialisten

Die Tabelle lässt sich des Weiteren unterteilen in vergleichsweise bereits lange am Markt bestehende Plattformen wie MTurk, Appen, Figure Eight und clickworker, die sich durch ein sehr breites Leistungsspektrum auszeichnen. Dem gegenüber stehen mit Mighty AI, Hive AI, Playment, Scale und understand.ai sehr junge und für die kurze Dauer ihres Bestehens bereits auffällig gut finanzierte schnell wachsende Plattformen, die sich fast vollständig auf die Crowdproduktion von Trainingsdaten für das autonome Fahren spezialisiert haben.

> „Für Unternehmen, die herkömmliches Crowdsourcing nutzen, sind Geschwindigkeit und Kosten die wesentlichen Faktoren. Die Qualität muss einigermaßen in Ordnung sein, aber mehr auch nicht. Bei Firmen, die Modelle für maschinelles Lernen bzw. KI-Modelle entwickeln, kommt es hingegen ganz besonders auf Qualität an. Diese Kunden brauchen hochpräzise Ground-Truth-Daten, denn jeder Fehler in den Ausgangsdaten macht die KI-Modelle weniger effektiv.“ (Daryn Nakhuda, Mighty AI)

Relativ neue Marktzugänge wie Playment sprechen von den etablierten Plattformen wie MTurk inzwischen als „legacy crowdsourcing“ und versuchen so den Eindruck zu erwecken, dass die Methoden der alten Plattformen sich bereits überholt hätten oder nur schlechte Qualität lieferten. Auch vergleicht sich Playment auf seiner Webseite mit „traditionellem“ Business Process Outsourcing (BPO), welches aufgrund fester Teamgrößen schlecht skaliere.

Die „althergebrachten“ Plattformen haben sich inzwischen ebenfalls größtenteils auf die Crowdproduktion von Trainingsdaten umgestellt. Einzig MTurk und clickworker werben auf ihrer Startseite nicht zuoberst für eine derartige Spezialisierung oder Neuausrichtung, sondern bleiben ihrem breiten Angebot treu, wobei auch clickworker „AI-Training Data“ als Dienstleistung aufführt (dazu später mehr). Hinter dem neuen Namen Figure Eight verbirgt sich übrigens die lang etablierte Plattform CrowdFlower, die sich inzwischen ebenfalls vollständig auf Machine-Learning-Kunden umgestellt hat.

Bei den neuen Crowd-Anbietern Mighty AI, Hive und Scale ist zudem auffällig, dass die zugehörigen Crowdplattformen (www.Spare5.com, www.hivemicro.com, und www.remotasks.com) getrennt von den jeweils den Kunden zugewandten Firmenwebseiten gehalten werden. Die Firmen treten janusköpfig auf und haben gegenüber der Crowd einen anderen Namen und ein anderes Gesicht als gegenüber den Auftraggebern. Letztere kommen mit der Crowd nicht mehr in Berührung. Aus Marketingsicht macht dieser Schritt viel Sinn, da die Industriekunden über eine andere Ansprache zu erreichen sind als die Crowdworker_innen. Es erlaubt den Plattformen aber auch, gegenüber ihren Kunden und der Öffentlichkeit den inzwischen oft mit schlechten Arbeitsbedingungen assoziierten und im Vergleich fast schon altmodisch erscheinenden Begriff der „Crowd“ hinter sich zu lassen und stattdessen von der Aura des neuen Business-Hypes „AI“ zu profitieren. Am besten, indem man die Formel schon im Namen oder in der Webadresse führt (wie bei www.thehive.ai und www.understand.ai).

Auch das ehemalige CrowdFlower betont insbesondere seit seiner Umbenennung 2018 in Figure Eight, dass es sich heute primär um eine Firma für künstliche Intelligenz handelt. Einerseits hat sich Figure Eight auf Kunden aus diesem Bereich spezialisiert, andererseits setzt die Firma nach eigenen Angaben auch selbst künstliche Intelligenz ein, um die Abwicklung der Aufgaben zu orchestrieren. Nichtsdestotrotz spielt menschliche Arbeit weiterhin eine entscheidende Rolle, sodass sich Figure Eight auch als „human-in-the-loop AI platform for data science & machine learning“ bezeichnet. Aus den Selbstbeschreibungen der anderen Plattformen lässt sich der Wandel in der Crowdsourcing-Branche ebenfalls ablesen: Mighty AI bietet „training data as

a service“; Hive liefert „large scale, quality training data within hours“ mittels einer „manual labeling platform“; Playment sieht sich als „fully managed human intelligence platform“; Scale bietet „human-powered data for AI applications“ und sieht sich als „developer API for human intelligence“, und understand.ai aus Karlsruhe offeriert „training and validating your algorithms with accurate annotations made with German precision“.

Nicht alle hier aufgeführten Plattformen bieten ausschließlich Trainingsdaten für maschinelles Lernen an. Nicht alle Trainingsdaten sind für die Automobilindustrie. Nicht alle Trainingsdaten für die Automobilindustrie basieren auf der Auswertung von Bilddaten. Appen und clickworker sind z. B. eher auf Audiodaten für das Training von Sprachassistenten spezialisiert, Dazu Christian Rozsenich, Geschäftsführer von clickworker:

> „In den letzten zwei Jahren sind die Entwicklungen rund um die künstliche Intelligenz und das Deep Learning zu Haupttreibern der Branche geworden. Im letzten Jahr ist noch das Thema Voice Commerce hinzugekommen, also Spracherkennung wie bei Amazons ‚Alexa‘.“

Die Verschiebung macht sich für clickworker seit 2017 in den Umsatzzahlen bemerkbar, die inzwischen „stark vom KI-Bereich getrieben“ werden, so Rozsenich. Innerhalb dessen nehmen Kunden aus dem Automobilbereich bei dieser Plattform allerdings nur einen relativ kleinen Anteil von 20 bis 30 Prozent ein. „Teils“, so vermutet Rozsenich, „weil die deutsche Autoindustrie aus Compliancegründen vor Crowdsourcing zurückschreckt. Vermutlich gehen diese Aufträge ins Ausland. Wir sehen davon jedenfalls nicht so viel.“

Die Firma Appen stellt ihre Audiospezialisierung wiederum explizit in den Dienst des Automobilbereichs. Das börsennotierte Unternehmen ist bereits seit 20 Jahren am Markt und seit 15 Jahren für die Automobilindustrie tätig. Nach eigenen Angaben arbeitet Appen mit sechs der zehn führenden OEMs zusammen. 2017 hat Appen ein Büro in Detroit eröffnet, um näher an der Autoindustrie zu sein (Lundgren 2017a). Die automobilnahen Bereiche umfassen neben der Beschriftung von Bilddaten die Spracherkennung im Fahrzeug zur Steuerung des Multimediasystems sowie die Erkennung von Gemütszuständen wie Müdigkeit, Wut etc. zur Gefahrenprävention durch Vorhersage von Verhalten.

Ende 2017 hat Appen das Konkurrenzunternehmen Leapforce gekauft und ist dadurch in seinem Bereich nach eigenen Angaben zum Weltmarktführer aufgestiegen (Redrup 2017). Die von Appen selbst antrainierte Crowd umfasste davor bereits 400.000 Menschen. Durch die Übernahme hat die Fir-

ma nun Zugang zu 1,2 Millionen Crowdworker_innen. Appens Aufkauf von Leapforce geschah laut Pressemitteilung ausdrücklich mit dem Ziel, im dynamischen Markt der Trainingsdaten für KI den kompetitiven Vorteil durch Crowdgröße weiter auszubauen.

Mighty AI wurde 2014 in Seattle als Spare5 gegründet und war anfangs noch eine Crowdsourcing-Plattform für E-Commerce-Kunden. Ende 2016 spielten Kunden aus der Automobilindustrie für die Firma noch überhaupt keine Rolle, nur ein Jahr später dann hatte sich die Firma bereits fast vollständig diesen neuen Aufgaben und Kunden verschrieben. Die Fokusverschiebung, die in zwei ausführlichen Podcast-Interviews mit dem Spare5-Management zum Ausdruck kommt (Charrington 2016, 2017), illustriert sehr gut, wie neu und dynamisch dieser Markt ist. 2018 ist Mighty AI einer der wichtigsten Dienstleister für die Erstellung von Trainingsdaten für autonomes Fahren und verfügt auf seiner Plattform Spare5 über etwa eine halbe Million registrierte Crowdworker_innen. Seit Februar 2018 hat Mighty AI eine Kooperation mit der Forschungseinrichtung MCity der Universität Michigan in Zusammenarbeit mit General Motors, Ford, Honda, Toyota, Intel, LG, Verizon. Darüber hinaus hat die Firma aber auch ein Büro in Boston eröffnet, um näher an Europa zu sein:

> „Boston ist für uns der wichtigere Standort, in Detroit haben wir nur eine Person sitzen. Zwar haben wir auch klassische amerikanische Autofirmen als Kunden, aber die europäischen Firmen sind für uns von größerer Bedeutung, und von Boston aus haben wir bei der Kundenbetreuung mit einem geringeren Zeitunterschied zu tun, als dies von Seattle aus der Fall ist." (Daryn Nakhuda, Mighty AI)

Auch die Plattform Playment hat ihre Wurzeln im Crowdsourcing für E-Commerce. Sie wurde 2015 von drei ehemaligen Flipkart-Mitarbeitern (Flipkart ist das indische Pendant zu Amazon) in Bangalore gegründet und hat erst Ende 2017 den Schwerpunkt ihrer Dienstleistungen komplett auf die Erstellung von Trainingsdaten für autonomes Fahren verlagert. Playment ist nach eigenen Angaben für etwa 30 internationale Kunden tätig, darunter Fahrzeughersteller, Zulieferer und Start-ups für autonomes Fahren – 70 Prozent der Kunden haben einen Bezug zur Automobilindustrie. Auf der Firmenwebseite waren Anfang 2018 unter anderem Starsky Robotics, Cyngn und drive.ai als Kunden gelistet. Inzwischen produziert Playment mit 300.000 indischen Crowdworker_innen durchschnittlich eine Million Annotationen/Labels am Tag.

> „Die Automobilindustrie ist für uns extrem wichtig geworden, weil hier der Bedarf an hochqualitativen Trainingsdaten besonders groß ist und immer wieder aktualisierte Daten benötigt werden, um die Sicherheit auf der Straße zu gewährleisten. Deswegen erwarten wir, dass künftig die meisten unserer Aufträge aus der Automobilindustrie kommen werden, und stellen uns entsprechend darauf ein." (Siddharth Mall, Playment)

> „Der Fokus auf das autonome Fahren ist aus dem Team gekommen. Mein Mitgründer Philip Kessler hat vorher bei Mercedes-Benz Research in Sunnyvale, Kalifornien gearbeitet, und auch viele andere aus unserem Team waren vorher bei Firmen wie Bosch, BMW und Mercedes. Dementsprechend war diese Spezialisierung für uns ein logischer Schritt." (Marc Mengler, understand.ai)

Auch wenn man die Webseiten und Pressemitteilungen der in der Tabelle aufgeführten Plattformen miteinander vergleicht, wird schnell deutlich, dass das Hauptgeschäft der neuen Plattformen die Auswertung von Bilddaten für die Automobilindustrie ist. Es gibt diese Aufgaben in unterschiedlichen Komplexitätsgraden, je nachdem, wo im Entwicklungsprozess die Plattform und ihre Kunden jeweils gerade stehen. Das Spektrum reicht vom Setzen loser, rechteckiger Objektbegrenzungen, sogenannter „bounding boxes", über das präzise Nachzeichnen von Objektkanten mit polygonalen Linien bis hin zur dreidimensionalen Verortung von Fahrzeugen und sogar Fußgängern im Bildraum. Hier spricht man auch von „cuboids". Die wohl weitverbreitetste Aufgabe ist die „semantische Segmentation" – hier gilt es, das gesamte „Sichtfeld" des Fahrzeugs im Straßenverkehr Bild für Bild mit sogenannten „semantic segmentation masks" aufzubereiten – d. h. Fotos händisch in Areale mit Objekten unterschiedlicher Bedeutung zu unterteilen und so z. B. Straßenschilder und Fahrbahnmarkierungen vom Hintergrund abzuheben bzw. Umrisslinien um Verkehrsteilnehmer_innen und Hindernisse zu zeichnen und diese dann spezifischen Objektkategorien zuzuweisen. Dies muss vollflächig, lückenlos, pixelgenau und millionenfach geschehen.

Scale und einige andere der neueren Plattformen bieten auf ihren Webseiten Preiskalkulatoren, mit denen man überschlagsweise die Kosten einschätzen kann, wobei bei den großen Auftragsvolumina Mengenrabatte eine Rolle spielen. Bei Scale kosteten zum Zeitpunkt der Recherche neun Annotationen pro Bild einen US-Dollar – bei einer Abnahme von 10.000 Bildern. Die mit Abstand am aufwendigsten umzusetzende semantische Segmentierung kostete pro Bild 6,40 US-Dollar, und wenn man die Bilder mit „Express-Dringlichkeit" bestellt, steigt der Preis sogar auf 16 US-Dollar pro Bild. Wenn man bedenkt, wie viele Firmen diese Daten in hoher Stückzahl brauchen, erscheinen die 50 bis 60 Millionen US-Dollar Risikokapital die den Topanbie-

tern unter den neuen Plattformen in den letzten Jahren zugeflossen sind, eher niedrig. Dabei darf man aber nicht übersehen, wie viel arbeits- und damit kostenintensiver die Crowdproduktion durch die gestiegenen Qualitätsansprüche für die Plattformen geworden ist.

> „Ground-Truth-Daten müssen um mindestens eine Größenordnung präziser sein als die Algorithmen, die dann später im Auto eingesetzt werden. Die Präzision ist absolut kritisch." (Marc Mengler, understand.ai)

Um unter diesen Bedingungen kompetitiv zu bleiben, verfügen die Plattformen über mehrere Stellschrauben, mit denen sie die Qualität steigern und die Kosten senken können:

1. Investition in fortschrittliche, KI-gestützte Präzisionswerkzeuge.
2. Investition in Qualitätsmanagement durch mehrstufige interne Prozessoptimierung, Kontrolle der Ergebnisse und viel mehr Service nach außen.
3. Investition in Training, Community-Management und Gamification-Mechanismen zur Fortbildung und Motivation der Crowd.
4. Zugang zu noch billigeren Arbeitskräften auf dem globalen Markt für Crowdworker_innen, was eine Übersetzung der Aufgaben in die Sprache der Arbeitskräfte erfordert.

Insbesondere die ersten drei Punkte, also Produktionsmittel, Prozessoptimierung und Schulung der Arbeiterschaft, gehen in der Praxis fließend ineinander über, weil Werkzeuge und Training zunehmend zu einer einzigen Softwarelösung verschmelzen.

Maßgeschneiderte Werkzeuge und handverlesene Crowds

Für die vollständige, pixelgenaue Segmentierung eines Bildes braucht ein einzelner Mensch ohne intelligente Softwareunterstützung etwa 60 bis 90 Minuten. Marc Mengler, der vor der Gründung von understand.ai im Rahmen seiner Grundlagenforschung selbst 15.000 Bilder annotiert hat, betont, wie kognitiv anstrengend und langwierig dieser Prozess ist.

Bei einem zuerst hoch erscheinenden Standardendverkaufspreis von 6,40 US-Dollar pro Bild (wie bei Scale) ist klar, dass hier für reine Handarbeit kein Mindestlohn auf dem Level von Industrienationen gezahlt werden kann. Selbst bei einer Bezahlung der Arbeitskräfte auf dem Niveau von Entwicklungsländern sind die Margen der Plattformen begrenzt, wenn sie lediglich am Kostenfaktor Arbeit sparen. Sie müssen deshalb eigene Spezialwerk-

zeuge entwickeln, die den Arbeitskräften zuarbeiten, sie kontrollieren und den Prozess deutlich schneller und präziser machen.

> „Wir haben einiges investiert, um unsere IT-Ausstattung mit Werkzeugen für die Vollbildsegmentierung aufzurüsten, und das speziell für einen Kunden. Dabei haben wir anfangs unterschätzt, wie aufwendig die Arbeit für die einzelnen ‚Gurus', so heißen bei uns die Crowdworker, ist. Das ist nicht in drei Klicks getan. Ein Einzelner braucht für die vollständige semantische Segmentierung eines Bildes in der gewünschten Qualität durchaus schon mal eine Stunde. Mit deutschen Stundenlöhnen ist man da gerade bei großen Auftragsvolumina international nicht konkurrenzfähig, deshalb ist noch unklar, ob sich unsere Investition gelohnt hat." (Hans Speidel, Crowd Guru)

> „Früher hat es gereicht, Rechtecke um Verkehrsschilder zu zeichnen, heute sind die Anforderungen deutlich komplexer. Das erfordert Investitionen in Spezialwerkzeuge, um das pixelgenau umsetzen zu können. Darüber hinaus setzen wir auch selbst KI-Techniken ein – zur Qualitätssicherung und zur Unterstützung der Crowd. Teilweise schon allein deswegen, weil die Aufgaben zu teuer wären, wenn sie in der benötigten Menge komplett von Hand gemacht werden würden." (Christian Rozsenich, clickworker)

Um die immer komplexer werdenden Aufgaben in hohen Stückzahlen preislich kompetitiv anbieten zu können, müssen sie in so viele kleine Arbeitsschritte wie möglich zerteilt werden. Das gilt für die meisten Formen von Crowdsourcing. Hier greifen nach wie vor die klassischen, von Frederic Taylor entwickelten Prinzipien des „Scientific Management", um die maximale Leistung aus den Arbeitskräften herauszuholen, ohne dass diese allzu schnell erschöpfen.

> „Wir versuchen den Menschen die kognitive Überlastung dadurch abzunehmen, dass wir ihnen möglichst einfache Fragen stellen. Etwa: ‚Is this bounding box tight enough?' Oder: ‚Is this a car at all?' Ziel ist es, die Aufgaben in möglichst atomare Einheiten zu zerteilen." (Marc Mengler, understand.ai)

Auch Mighty AI hat das herausgefunden. Anfänglich mussten die Crowdworker_innen 75 verschiedene Objektklassen von Dingen im Straßenverkehr erlernen, um ein Bild entsprechend vollständig zu annotieren. Es hat sich dann aber als viel effizienter erwiesen, die 75 Objektklassen auf ebenso viele Crowdworker_innen aufzuteilen und diese wiederum immer nur ein Objekt markieren zu lassen. Die Arbeitskräfte sitzen also nicht mehr stundenlang vor einem riesigen Puzzle, sondern werden nur noch gefragt: „Siehst du in diesem Bild einen weiteren Lastwagen, der bisher noch nicht markiert wurde?" Erst wenn sie dies bejahen, bekommen sie Zugang zum Zeichenwerk-

zeug und sind kurzfristig hochfokussiert, bevor es für sie weiter zur nächsten Kleinstaufgabe in einem anderen Bild geht.

Auf diese Weise lässt sich auch eine automatisierte Qualitätskontrolle in den Prozess integrieren, berichtet Daryn Nakhuda. Außerdem lassen sich durch die Kleinteiligkeit einzelne Bilder von vielen Menschen parallel bearbeiten und wieder zusammensetzen und die Prozesse so viel schneller und detaillierter steuern. Damit die einzelnen Arbeitskräfte die Tasks überhaupt ausführen dürfen, müssen sie für jeden Aufgabentyp erst ein längeres Training durchlaufen, bei dem der Anspruch an die Präzision bereits sehr hoch ist. Hier wird auch sortiert, wem welche Aufgabe am meisten liegt.

> „Wir vergeben die Aufgaben nach individuellen Fertigkeiten. Manche Menschen sind außerordentlich gut darin, Außenkanten von Objekten detailliert nachzuzeichnen, andere sind viel besser darin, Kategorisierungen vorzunehmen, Metadaten wie etwa Fahrzeugtypen hinzuzufügen oder die Intention von Passanten zu erkennen." (Daryn Nakhuda, CEO von Mighty AI)

Die neuen Plattformen setzen alle auf eine Mischung aus Werkzeugoptimierung, Training und Motivation, wobei die Bereiche immer mehr zusammenfließen, weil die Softwarewerkzeuge, an denen die Arbeitskräfte ausgebildet werden, zugleich auch die Präzision der Arbeitskräfte messen und ihnen diese als quantitatives Feedback gekoppelt mit anfeuernden Sprüchen zurückspiegeln.

Während Mighty AI besonders auf persönliches und positives Feedback sowie auf Community-Building setzt und wie Playment auch Gamification-Elemente einsetzt, mit denen die Arbeit zu einem Spiel mit quantitativem Feedback wird, setzt understand.ai besonders auf die Entwicklung von Softwarelösungen, welche die Arbeit so weit wie möglich automatisch vorab erledigen und die menschlichen Arbeitskräfte dann nur noch jene Stellen überprüfen lassen, bei denen sich der Algorithmus noch unsicher ist. Die eben beschriebene Spezialisierung der Crowdworker_innen bei Mighty AI findet sich bei understand.ai in Bezug auf verschiedene Algorithmen wieder:

> „Bei uns kommt ein Gremium aus Algorithmen zum Einsatz, das über das Ergebnis abstimmt. Ein Algorithmus ist z. B. sehr gut darin, Straßenschilder zu erkennen, ein anderer ist Experte für Fahrzeuge oder Fahrradfahrer. In nur dreihundertzwanzig Millisekunden erhalten wir so eine erste semantische Segmentierung. Nach diesem Durchlauf gibt es immer noch einzelne Elemente, die nicht vollständig erkannt wurden. Das Zwischenergebnis schicken wir dann an unsere darauf spezialisierten Partner in Indien, welche die noch unpräzisen oder fehlerhaften Stellen mit von uns dafür entwickelten Werkzeu-

> gen korrigieren. Normalerweise braucht ein Mensch etwa eineinhalb Stunden, um ein knapp zwei Millionen Pixel großes Bild perfekt zu annotieren bzw. zu segmentieren. Da ist dann auch die Qualitätssicherung schon mit einkalkuliert. Bei uns dauert dieser Prozess dank unseres algorithmischen Vorschlags insgesamt nur noch fünfunddreißig Minuten. Unser Werkzeug weiß übrigens auch, wie sicher oder unsicher sich der Algorithmus mit seinen jeweiligen Vorschlägen ist. Die Aufmerksamkeit unserer ‚Refiner', der Menschen, die von Hand die Nachbesserungen vornehmen, wird dann gezielt auf die Stellen gelenkt, bei denen noch Unsicherheit in der automatischen Erkennung herrscht." (Marc Mengler, understand.ai)

Die zeitlichen Abfolgen dieser gestaffelten Mensch- und Maschinen-Bearbeitung variieren von Plattform zu Plattform, und das Gleiche gilt für die Anzahl der Iterationsstufen und die Redundanz – also wie oft ein Schritt hintereinander und parallel ausgeführt wird, um die Qualität zu sichern. Bei billigen Arbeitskräften rechnet es sich, eine Aufgabe mehrfach zu vergeben und aus den Ergebnissen einen Mittelwert zu bilden. (Was Kevin Guo weiter oben als „Crowdkonsens" bezeichnet, im Gegensatz zum „Gremium aus Algorithmen" bei Marc Mengler.)

Bei teuren Arbeitskräften braucht es andere Kontrollmechanismen, wie etwa dass sich Arbeitskräfte eine quantifizierte Reputation aufbauen können und diese halten wollen oder eine Klasse von Crowdworker_innen (oder ein Algorithmus) die Ergebnisse der anderen kontrolliert.

> „Generell gibt es in unserem Prozess die Rollen ‚Maker' und ‚Checker'. Eine Person erledigt die Aufgabe am Desktop, mehrere andere überprüfen dann das Ergebnis von unterwegs aus auf ihrem Telefon. Auch hier geht für uns Qualität vor Effizienz." (Siddharth Mall, Playment)

> „Die ‚Player' [so heißen die Crowdworker auf der Plattform Playment, A. d. A.] können verschiedene Level und Positionen erlangen, die es ihnen dann erlauben, besser bezahlte Tasks zu erledigen. So haben wir ‚Superplayer', denen wir die schwierigeren Aufgaben geben und die wir dafür besser vergüten. Die Plattform ist wie ein Arbeitsplatz, an dem man auch befördert werden kann. Mit mehr Verantwortung geht eine bessere Entlohnung einher." (Siddharth Mall)

> „Die Player haben natürlich kein Hintergrundwissen über autonome Fahrzeuge, wir müssen ihnen alles beibringen. Dies gelingt nur durch eine sorgfältige Gestaltung der Prozesse auf der Plattform sowie der Aufgaben selbst. Auf interaktive Weise erlernen die Player neue Fertigkeiten zur Bewältigung der jeweiligen Aufgaben. So qualifizieren sie sich nach und nach und steigen im System auf." (Siddharth Mall)

Aus der ursprünglich amorphen Crowd werden auf diese Weise hierarchische Strukturen herausgebildet, und es kann den Arbeitskräften passieren, dass sie sowohl über als auch unter sich in der Hierarchie Algorithmen und Menschen haben, die ihnen zuarbeiten oder sie kontrollieren. Man kann sich als einzelne Arbeitskraft nicht mehr sicher sein, was genau gerade der Fall ist. Außerdem kann man nicht mehr, wie es etwa bei MTurk tendenziell der Fall ist, selbst frei wählen, welche Aufgaben man erledigen will:

> „Wir verteilen Aufgaben gezielt auf kleinere Untergruppen, um eine bessere Qualitätskontrolle zu haben. Wenn es ein paar Hunderttausend Aufgaben zu erledigen gibt und man diese gleichmäßig auf die Community verteilt, kann jeder nur ein paar Aufgaben abschließen. Nach unserem Prinzip können hingegen ein paar Tausend Leute jeweils Hunderte von Aufgaben machen und werden entsprechend schnell viel besser darin." (Daryn Nakhuda, Mighty AI)

Schlagworte, mit denen die Plattformen auf den Wandel aufmerksam machen, sind „trained crowd labour", „curated crowds", „crowd qualification", „known crowds" – die Kunden sollen davon überzeugt werden, dass man es nicht mehr mit einer anonymen und womöglich unfähigen oder unzuverlässigen Masse zu tun hat, sondern mit der Plattform bekannten ausgewählten und trainierten Fachkräften.

Die Crowdworker_innen von Spare5, die für diese Studie interviewt wurden, waren einerseits stolz auf ihre quantifizierten Fertigkeiten in Form von Erfahrungspunkten und Spezialistentiteln. Andererseits waren sie teils frustriert, weil sie Aufgaben einfach nicht zu sehen bekamen, während Kollegen Arbeit hatten. Für diejenigen, die gerade keine Arbeit hatten, war es oft nicht transparent, ob es ihnen jetzt an vom System ermittelter Genauigkeit, an Erfahrungspunkten oder an Spezialisierung mangelte oder ob die Plattform, wie Nakhuda im oben stehenden Zitat schildert, es aus Trainingsgründen vorzog, neue Aufgaben nur an ausgewählte kleine Untergruppen zu verteilen.

> „Wir kennen unsere Crowd gut, wir wissen, wer für welche Aufgaben geeignet ist und wie man solche Projekte aufsetzt. Diesen Service schätzen viele Kunden, aber wenn es um große Volumen geht, wird es für manche Unternehmen dennoch attraktiver, diese Kapazitäten selbst aufzubauen oder zu billigeren Anbietern zu gehen." (Hans Speidel, Crowd Guru)

Auch die Plattform Appen versucht, sich mit „handverlesenen" Crowds von herkömmlichem Crowdsourcing abzuheben, und wirbt dafür mit dem Begriff der „curated crowds" – eigentlich ein Widerspruch in sich, denn die Haupteigenschaft der klassischen Crowd ist, dass sie eben nicht handverlesen ist (Schmidt 2017a). Die „curated crowds" sind nach Angaben von Appen auf

bestimmte Tasks spezialisiert und werden besser bezahlt (Christensen 2015). Zudem gibt es laut Appen weniger Redundanz, dieselbe Aufgabe wird an weniger Crowdworker_innen parallel vergeben, unterliegt dafür aber strengeren Qualitätskontrollen bei der Durchführung, um trotzdem zu verlässlichen Endergebnissen zu gelangen. Qualitäts- und Produktivitätsmanagement gewinnen folglich an Relevanz.

Die Crowdworker_innen verpflichten sich bei Appen außerdem zu Mindestarbeitszeiten pro Tag und pro Woche, für sie handelt es sich oftmals nicht um ein Hobby, sondern um ihre Haupteinnahmequelle. Für die Auftraggeber wird die Arbeit teurer, dafür sind die Ergebnisse von höherer Qualität und lassen sich in einem besser kalkulierbaren, engeren Zeitfenster erzielen, so zumindest das Versprechen Appens (Christensen 2015).

Laut Appen arbeiteten 2017 im Schnitt zu jedem Zeitpunkt etwa 10.000 Menschen gleichzeitig als Crowdworker_innen für die Firma. Viele von ihnen vier Stunden am Tag, fünf Tage die Woche, allerdings nur für die Dauer des jeweiligen Projekts. Mitunter pausieren sie danach so lange, bis wieder eine passende Aufgabe für sie auftaucht (Lundgren 2017b). Wegen der Fokussierung auf Sprache erfolgt die Vergabe von Jobs bei Appen nach Ländern. Man muss sich auf die Jobs bewerben und dafür im jeweiligen Land wohnhaft sein.

Qualitätsmanagement und Scheinselbstständigkeit

Ähnlich wie am Versprechen der „curated crowds“ lässt sich der aktuelle Wandel in der Branche auch an Verheißungen wie „fully managed“, „end-to-end project management“ und „nothing is done by you“ ablesen. Sowohl die Gründer von Mighty AI als auch die von Playment hatten MTurk als Vorbild, aber zugleich das Ziel, eine nutzerfreundlichere Dienstleistung zu entwickeln.

> „Amazon Mechanical Turk war für uns keine taugliche Lösung, weil wir für unsere Anforderungen viel mehr Projekt- und Qualitätsmanagement brauchten. Außerdem kann einem Mechanical Turk nicht garantieren, bis wann die Bearbeitung erfolgt sein wird (turnaround time). […] So entstand die Idee, einen ‚fully managed distributed workplace‘ zu bauen.“ (Siddharth Mall, Playment)

> „Für die Unternehmen und unsere Community funktionieren wir wie eine ‚Black Box‘, die sie vertrauensvoll nutzen können, ohne die genauen Abläufe im Hintergrund kennen zu müssen, weil wir dafür einstehen, dass alles, was durch unsere Kanäle bei ihnen ankommt, auch valide ist.“ (Daryn Nakhuda, Mighty AI)

All dies lässt sich als ein Gegenentwurf zur rudimentäreren Form des Crowdsourcings im Stile von MTurk verstehen, bei dem die Plattform so weit wie möglich versucht, in der Rolle des technischen Infrastruktur-Providers im Hintergrund zu bleiben und die Kunden dafür über die „human API" unmittelbar selbst mit einer Crowd aus ihnen unbekannten Akteuren agieren lässt.

Ganz wesentlich ist hier, dass die Auftraggeber beim MTurk-Schema selbst die Erfahrung mitbringen müssen, wie man eine semianonyme Crowd am besten koordiniert, um gute Ergebnisse zu erhalten. Die Qualitätsprüfung obliegt beim „legacy crowdwork" ebenso den Kunden wie die Programmierung der Werkzeuge, mit denen die Aufgaben umgesetzt werden können. Das Gleiche gilt für ein etwaiges Trainieren oder Aussieben von Arbeitskräften.

Da die Crowdworker_innen auf MTurk und ähnlichen Plattformen folglich bei jedem Auftraggeber andere Werkzeuge erlernen müssen, gibt es hier viel größere Reibungsverluste durch den dafür notwendigen Zeitaufwand sowie durch missverständliche Aufgabenbeschreibungen und fehleranfällige Werkzeuge. Fehleranfällig deshalb, weil die Auftraggeber auf MTurk natürlich signifikant weniger Zeit und Geld in die Entwicklung von Werkzeugen und Aufgabenbeschreibungen investieren und auch in der Regel nicht ansprechbar sind, falls Fragen oder Probleme auftauchen; das Gleiche gilt für MTurk als Plattform. Als Crowdworker_in ist man hier auf sich allein gestellt. Dies zeigte sich auch bei der stichprobenartigen eigenen Arbeit als Crowdworker auf MTurk im Vergleich zu Spare5. Im ersten Fall waren die Werkzeuge sehr viel primitiver und die Aufgaben sehr viel schwerer zu verstehen als im zweiten Fall, wobei man bei Spare5 erst viel Zeit in (leidlich bezahlte) Trainingsaufgaben investieren muss, um im System aufzusteigen.

Unter professionellen Crowdworker_innen auf MTurk, so der Eindruck einer die vorliegende Studie unterstützenden gut vernetzten amerikanischen „Turkerin" mit mehrjähriger Erfahrung, erfreuten sich die Tasks zur Erstellung von Trainingsdaten für das autonome Fahren aus diesen Gründen und wegen der besonders schlechten Bezahlung keiner großen Beliebtheit.

Die befragten Crowdworker_innen von Spare5 waren im Gegensatz dazu im Netz gezielt auf der Suche nach solchen Aufgaben, weil sie diese als besonders attraktiv und gut bezahlt kennengelernt hatten, nur eben nicht auf MTurk. Letzteres schien ihnen keine taugliche Alternative zu Spare5, insbesondere wegen der Abwesenheit von Community-Management und firmenseitigen Ansprechpartnern_innen im Fall von Fragen oder Problemen.

Die neuen, hoch spezialisierten Full-Service-Plattformen ersparen ihren Kunden und den Crowdworker_innen die Reibungsverluste, die ihnen durch

die direkte Interaktion miteinander entstehen. Für die Kunden ist eine solche Lösung selbstverständlich deutlich teurer, als wenn sie zu MTurk gehen.

Die erste Plattform, die von MTurks Prinzip der eher primitiven „Selbstbedienungs-API" abgewichen war, ist übrigens CrowdFlower, heute Figure Eight. Sie wird von manchen auch als Microtasking der zweiten Generation verstanden.

CrowdFlower bzw. Figure Eight operiert dabei bezeichnenderweise ohne eigene Crowd. Stattdessen verteilt die Firma die Aufträge ihrer Kunden als „Metaplattform" auf unterschiedliche Crowds anderer Plattformanbieter, so auch lange auf MTurk. 2018 nutzte Figure Eight unter anderem Crowdplattformen wie IndiVillage und Clixsense, in der Vergangenheit auch Samasource. Das Unternehmen fungiert also als weitere Zwischenstufe bzw. als zusätzlicher Intermediär im Auslagerungsprozess. Die Firma besteht in erster Linie aus genau dem Service-Layer, das bei MTurk vollständig fehlt und das bei Firmen der dritten Generation wie Mighty AI und Playment direkt in die Crowdplattform integriert ist.

Mit seiner frühen Dienstleistungsorientierung hat CrowdFlower die heute gängigen Standards in der Produktion von Trainingsdaten um Jahre vorweggenommen. Zugleich sah sich die Firma auch als Erste mit einem weitreichenden rechtlichen Konflikt konfrontiert: Eine Plattform, die wie MTurk betont primitiv strukturiert ist und sich aus der Interaktion zwischen Kunden und Crowd komplett herauszuhalten versucht, kann relativ glaubwürdig vertreten, dass sie nur ein Infrastruktur-Provider und nicht etwa ein Arbeitgeber ist. Zugespitzt formuliert ist die schlechte Behandlung der Crowd für MTurk rechtlich von Vorteil.

Sobald eine Plattform jedoch das Management der Crowd in den Vordergrund stellt, die Crowd schult, betreut, an sich bindet, als eigene Arbeiterschaft versteht und die Aufgaben an Untergruppen verteilt, so wie das bei Plattformen für Trainingsdaten zunehmend der Fall ist, wächst auch das Risiko von Scheinselbstständigkeitsklagen. Denn Crowdworker_innen – so lässt sich argumentieren –, die viel Zeit auf der Plattform verbringen, und das ohne jeden Kundenkontakt, arbeiten nicht mehr freiberuflich für eine Vielzahl unterschiedlicher Kunden, sondern sind eigentlich scheinselbstständige Angestellte der Plattform.

Eine auf dieser Argumentation aufbauende Sammelklage von 2013 – Otey v. CrowdFlower – konnte die Plattform 2014 noch durch einen Vergleich abwenden (Schmidt 2013, Cherry 2016), und heute läuft Figure Eight durch Rückgriff auf eine Vielzahl unterschiedlicher externer und eigener Plattformen nicht mehr so leicht Gefahr, hier erneut verklagt zu werden. Für

die neuen Anbieter der dritten Generation, bei denen Crowdworker_innen de facto in Vollzeit auf der Plattform arbeiten, ohne dabei je mit den Auftraggebern in Kontakt zu kommen, ist dies durchaus ein rechtliches Risiko.

Das führt zu der paradoxen Situation, dass Plattformen wie Mighty AI bzw. Spare5, die sehr viele Ressourcen in besseres Training und vor allem in besseres Community-Management stecken und bei denen sich die Crowd viel besser, weil als Menschen und nicht bloß als Algorithmen behandelt fühlt, letztlich dadurch eher Gefahr laufen, wegen der Arbeitsbedingungen verklagt zu werden, als MTurk, wo es all diese Investitionen in die Arbeitskräfte nicht gibt. Auch bei den befragten deutschen Plattformen wird übrigens Wert auf Community-Management und ein gutes Verhältnis zur Crowd gelegt. Hier besteht wiederum weniger Risiko bezüglich Scheinselbstständigkeitsklagen, weil die Arbeitskräfte hier meist nur sporadisch und phasenweise, also als Gelegenheitsarbeiter_innen, tätig werden (dazu mehr im übernächsten Abschnitt).

> „Die Crowd ist der zentrale Bestandteil unseres Produkts, und deswegen ist es für uns essenziell, dass die Crowd glücklich ist." (Hans Speidel, Crowd Guru)

> „Unser Credo ist, die Crowd fair zu behandeln, so wie das auch in unserem Code of Conduct steht. Es ist wichtig, den Leuten hinreichend Support und eine Entwicklungsperspektive zu geben und sie aus der Isolation als digitaler Einzelkämpfer zu holen, z. B. über Foren, Chats und andere Möglichkeiten des Austauschs und der Vernetzung untereinander. Das wird zum Garanten für eine loyale Crowd, die dann auch entsprechend gute Ergebnisse liefert." (Christian Rozsenich, clickworker)

Ausgleich von Schwankungen in der Nachfrage

Die vielleicht wichtigste Funktion von Crowdsourcing-Plattformen ist, dass sie für ihre Kunden deren rapide Schwankungen im Bedarf an Arbeitskraft abfedern. Hierin ähneln sie Leiharbeitsfirmen, nur dass die Frequenz und Dynamik, mit denen man große Gruppen mobilisieren und wieder „entlassen" kann, bei Online-Crowdarbeit potenziell noch viel extremer sind. Oder, wie es Lucas Biewald, damaliger CEO von CrowdFlower schon 2014 auf den Punkt gebracht hat:

> „Before the Internet, it would be really difficult to find someone, sit them down for ten minutes and get them to work for you, and then fire them after those ten minutes." (Lucas Biewald, zitiert nach Marvit 2014)

Bei der Produktion von Trainingsdaten kommt diese zentrale Eigenschaft von Crowdsourcing besonders zum Tragen, weil die Kunden hier sehr kurzfristig riesige Datenmengen und damit Arbeitsstunden brauchen und die Tätigkeit zudem ausgesprochen repetitiv und zugleich doch nicht vollständig automatisierbar ist.

Da dieser hohe Bedarf an Arbeitskraft aber bei den einzelnen Kunden schubweise auftritt, je nach der Entwicklungsphase, in der sich Research-and-Development eines bestimmten KI-Produkts gerade befindet, macht der Aufbau einer eigenen Belegschaft eigens für diese Zwecke für die Kunden meist keinen Sinn. Wobei sich dies in den nächsten Jahren noch ändern könnte, falls die von den Interviewpartnern vermutete Verschiebung von kurzfristig benötigten immer neuen Trainingsdaten hin zu einem stetigen Strom von Validierungsdaten eintritt.

Vorerst lässt sich jedenfalls konstatieren, dass die Kunden das Problem der extremen Schwankungen im Bedarf von Hilfsarbeiter_innen dank der Crowdsourcing-Plattformen sehr gut externalisieren können. Theoretisch können Letztere das Problem der Schwankungen ihrerseits wiederum dadurch auffangen, dass sie für verschiedene Kunden die Arbeitskräfte vorhalten und sich deren Schwankungen im Idealfall in der Summe ausgleichen:

> „Anfangs liefen die Auftragsvolumina als einzelne große Wellen durch unser System, mit entsprechenden Schwankungen in der Verfügbarkeit von Arbeit. Inzwischen gleicht sich das durch eine größere Anzahl von Kunden besser aus. Es gibt immer noch Wellen, aber keine Phasen ganz ohne Arbeit mehr. Aus der Perspektive einzelner Arbeitskräfte mag dies jedoch anders erscheinen, weil deren Qualifikation oder Spezialisierung gerade nicht zur verfügbaren Arbeit passt." (Daryn Nakhuda, Mighty AI)

Letztlich ist es so, dass die Plattformen das Problem der Schwankungen an die Crowdarbeiter_innen durchreichen, die sich nie sicher sein können, ob sie zu einem gegebenen Zeitpunkt genug Arbeit finden werden. Für sie bleibt die Verteilung der Arbeitsvolumina komplett intransparent und erratisch (vgl. hierzu insbesondere die Interviews mit den Crowdworker_innen im nächsten Abschnitt).

Da die Plattformen die Crowd bekanntlich nicht anstellen, sind Schwankungen in Richtung einer zu geringen Auslastung auf den ersten Blick kein so großes Problem, die Arbeiter_innen haben dann einfach nichts zu tun, kosten aber auch nichts. In die andere Richtung sind die Schwankungen für die Plattform allerdings schon eine Herausforderung, da sie den Kunden gerne eine möglichst schnelle Bearbeitung zusichern wollen. Um Spitzen in der

Auslastung im Richtigen Moment auffangen zu können, ist es dann doch wichtig, die Crowd nicht durch einen zu geringen Auslastungsgrad zu vergraulen.

> „Wir sind darum bemüht, dass sich die Menge an verfügbaren Jobs mit der verfügbaren Arbeitskraft die Waage hält. Ein Ungleichgewicht in die eine oder andere Richtung ist ein Problem, und beide Situationen treten manchmal auf." (Hans Speidel, Crowd Guru)

> „Das Wichtigste ist, dass man immer Arbeit für die Crowd hat. Sonst schauen sie auf der Plattform vorbei, finden nichts, was sie tun können, und dann erreicht man sie auch nicht mehr." (Christian Rozsenich, clickworker)

Dieses Problem des Aufeinanderabstimmens von extrem schwankender Nachfrage mit einer Kernbelegschaft aktiver Crowdarbeiter_innen und einer schnell mobilisierbaren „Reservearmee" gibt es auch in anderen Bereichen des plattformbasierten Arbeitens, etwa bei Lieferdiensten und der Personenbeförderung (Schmidt 2016, 2017b). Und aus diesem Bereich kommt auch ein Lösungsansatz, das sogenannte „Surge-Pricing" – also die dynamische Preisanpassung, welche zuerst von Uber etabliert wurde und inzwischen auch von der Trainingsdatenplattform Hive angewendet wird. Auch in der Preisgestaltung von Scale.api, bei der es die Option einer mehr als das Doppelte kostenden Expressbearbeitung gibt, lassen sich ähnliche Ansätze ablesen, wobei von außen nicht zu erkennen ist, zu welchem Anteil dieser Aufschlag an die Crowd ausgezahlt wird. Bei Uber können die Arbeitskräfte so jedenfalls in Spitzenzeiten deutlich mehr verdienen, sind also motiviert, die Schwankungen in der Nachfrage aufzufangen. Zugleich erhöht sich für sie aber der Stress, weil sie immerzu auf dem Sprung sein müssen. Unter erfahrenen Uber-Chauffeur_innen hat sich deshalb offenbar der Konsens durchgesetzt, dass es sich nicht lohnt, dem Surge-Price hinterherzujagen (Rosenblat 2018).

Das Problem der Schwankungen nach unten, also der zu geringen Auslastung, ist für Menschen, die diese Arbeit nur gelegentlich als eine Art Hobby mit Zuverdienst ausüben, wenn die Konditionen gerade günstig sind, nicht weiter schwerwiegend. So scheint es bei vielen europäischen Crowdworker_innen der Fall zu sein, die ohnehin nicht allein von dieser Arbeit leben könnten.

Für Menschen hingegen, die sich in Hochphasen der Auslastung in die Abhängigkeit einer prekären „Vollbeschäftigung" auf einer Plattform begeben, die sich und teilweise sogar ihre Familie (siehe übernächster Abschnitt

mit den Crowdworker_innen-Porträts) ausschließlich über eine Tätigkeit finanzieren, die von einem Moment auf den anderen wegbrechen kann, ist die Volatilität dieses Arbeitsmarktes ein großes Risiko und ein ständiger Stressfaktor. In diese Abhängigkeit können natürlich nur Menschen in Ländern geraten, in denen man mit den geringen Ausschüttungen, die man im besten Fall als freie Hilfskraft monatlich von einer Plattform erwarten kann, irgendwie über die Runden kommen kann.

Globale Wanderarbeiter_innen und deutsche Plattformen

Da ein großer Teil der hier besprochenen Arbeit völlig ortsunabhängig erledigt werden kann (mit einigen wichtigen Ausnahmen, dazu gleich mehr), ist es nicht überraschend, dass sich der Marktwert der Arbeit wie in einem System kommunizierender Röhren dem niedrigsten Level von Entlohnung annähert, welches noch von einer hinreichend großen Gruppe von hinreichend gut an das Internet angeschlossenen und verlässlich arbeitenden Menschen akzeptiert wird. Sei es, wie manchmal der Fall, weil es sich um Hobbyisten handelt, die Spaß an der mitunter sogar als Spiel aufbereiteten Freizeitarbeit finden und die Bezahlung als sekundär erachten, sei es, wie weit häufiger der Fall, weil es sich um Menschen aus sehr armen Regionen der Welt und/oder in persönlichen Notlagen handelt.

Auf den hier besprochenen Plattformen für Trainingsdaten lag das monatliche Einkommen der motiviertesten und erfolgreichsten Arbeitskräfte den Interviews mit den CEOs und den Crowdworker_innen zufolge zwischen 200 und maximal 400 US-Dollar im Monat bzw. zwischen 1 bis 2 US-Dollar die Stunde. Da es sich hier nur um die Topverdiener handelt, muss das Durchschnittseinkommen noch deutlich niedriger sein. Die Entlohnung ist den Interviews mit Crowdworker_innen von Mighty AI nach zudem sinkend, offenbar weil die schnell wachsende Anzahl von Jobs in der Produktion von Trainingsdaten ein noch schneller wachsendes Überangebot an Arbeitssuchenden nach sich gezogen hat, die eine sehr niedrige Bezahlung zu akzeptieren bereit sind.

> „Es gibt mit Sicherheit ein Überangebot an Arbeitsstunden bzw. an Supply von Arbeit." (Christian Rozsenich, clickworker)

> „Bei uns verdienen die Arbeitskräfte umgerechnet etwa zweihundertfünfzig bis dreihundert Dollar im Monat. In einem sehr guten Monat und wenn sie besonders aktiv sind, verdienen einzelne ‚Player' auch schon mal vierhundert

> Dollar. Das ist für indische Verhältnisse sehr viel Geld. Die Regierung in Indien definiert Armut als ein Einkommen von unter einem Dollar pro Tag." (Siddharth Mall, Playment)

> „Die Leute verdienen bei uns bis zu hundert Dollar die Woche. […] Wenn die Aufgaben anspruchsvoller sind und entsprechend besser vergütet werden, kann ich mir auch vorstellen, dass etwa mehr Leute in den USA diese Arbeit machen. Auch jetzt ist es schon manchmal so, dass sensible Kundendaten ihr Herkunftsland nicht verlassen dürfen und wir deswegen amerikanische Crowdworker nehmen müssen." (Kevin Guo, Hive)

In vielerlei Hinsicht kennt man die Auslagerung von Arbeit in die sogenannte „Dritte Welt" natürlich aus früheren Phasen des Outsourcings. Vergleichsweise neu ist jedoch, dass ein Produzent von Trainingsdaten seine Produktionsstätten nicht mehr physisch in einem Entwicklungsland aufbauen oder auch nur einen Vertrag mit einer lokal verankerten Outsourcing-Firma schließen oder sich selbst um die Rekrutierung kümmern muss. Stattdessen suchen sich die neuen globalen Wanderarbeiter_innen ihre Arbeit selbst. Untereinander koordiniert über zahlreiche Spezialforen und soziale Netzwerke, warnen sie einander vor schlechten Plattform-Arbeitgebern und strömen genau dahin, wo man gerade etwas verdienen kann. Wie Erntehelfer finden die globalen Crowdworker_innen so den Weg zur Arbeit und ziehen mit den Schwankungen der Auftragslage zwischen den Plattformen hin und her.

Die Crowdsourcing-Plattformen müssen lediglich in die Übersetzung der Aufgaben in jene Sprachen investieren, in denen es besonders viele arbeitssuchende Menschen mit Internetanschluss gibt. Um die Crowd darüber hinaus motiviert zu halten und zu trainieren, hilft es auch, das Community-Management ebenfalls in dieser Sprache anzubieten, so wie es z. B. von Mighty AI/Spare5 sehr erfolgreich gemacht wird (siehe nächster Abschnitt).

Nachdem es in Ländern wie Indien (wo Playment vertreten ist) und den Philippinen (wo Scale/Remotasks stark ist) wegen der Geschichte als ehemalige Kolonien viele Menschen mit guten Englischkenntnissen gibt, weswegen diese Regionen bereits seit Langem für das Outsourcing von Callcentern und ähnlichen Serviceleistungen bekannt sind, kommt auf dem Markt für Trainingsdaten Spanisch als Verkehrssprache und Südamerika als Region wachsende Bedeutung zu – und das hängt auch mit der ökonomischen und humanitären Katastrophe in Venezuela zusammen. Nach Schätzungen des Internationalen Währungsfonds vom Oktober 2018 betrug die Hyperinflation in Venezuela in diesem Jahr unvorstellbare 1,37 Millionen Prozent. Ein Hühnchen kostete zu dem Zeitpunkt in der Landeswährung Bolívar schon

14 Millionen – oder 2 US-Dollar. Der Druck und zugleich die Chance, sich über Onlineplattformen ausländische Devisen zu erarbeiten, ist also enorm.

2018 ließ sich eine sehr starke Fluktuation der Arbeitskräfte zwischen manchen der neuen Plattformen für Trainingsdaten beobachten, insbesondere zwischen Spare5 (vgl. Abbildung 2), Remotasks (vgl. Abbildung 3) und Hive Micro (vgl. Abbildung 1), die alle drei sehr ähnliche Aufgaben jeweils auf Englisch und Spanisch anbieten. Wie eingangs bereits erwähnt, kamen dabei überproportional viele der Arbeitskräfte aus Venezuela. Ende 2018 war die Plattform Spare5 auf Platz 167 der beliebtesten Webseiten in Venezuela, Hive auf Platz 187, jeweils ca. 75 Prozent der Besucher_innen dieser Plattformen kamen zu diesem Zeitpunkt aus Venezuela (www.alexa.com/siteinfo).

Nach eigenen Angaben ist die Crowd von Hive in der ersten Hälfte 2018 von 100.000 auf 300.000 Arbeitskräfte gewachsen, in einer Geschwindigkeit von 3.000 Neuanmeldungen pro Tag.

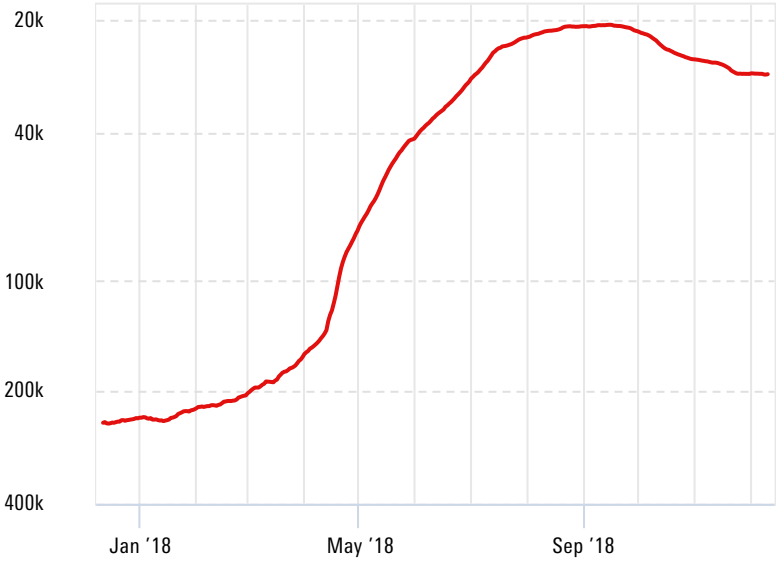
Wenn der von Alexa gemessene Traffic korrekt ist, und die Interviews mit den Crowdworker_innen stützen diese Annahme, dann waren 2018 mindestens 200.000 Menschen aus Venezuela auf der Suche nach Arbeit auf Crowdplattformen für Trainingsdaten. Gegen Mitte 2018 hatte Mighty AI/Spare5 (vgl. Abbildung 2) der Crowd nicht mehr genug Arbeit zu bieten, zeitgleich stieg die Nutzung von Hive Micro (vgl. Abbildung 1) und Scale/Remotasks deutlich an. Da auf Remotasks (vgl. Abbildung 3) aber nur gut 15 Prozent des Traffics aus Venezuela kommen (im Vergleich zu 31,1 % aus den Philippinen), ist davon auszugehen, dass die größten Pendelbewegungen von venezolanischen Crowds zwischen Spare5 und Hive stattfinden. Remotasks fällt auch deswegen aus der Reihe, weil die Plattform ihre Aufgaben nicht auf Spanisch anbietet.

> „Es stimmt, dass Menschen aus Venezuela inzwischen den Großteil unserer Arbeiterschaft bilden. Die Wirtschaft des Landes liegt bekanntlich am Boden. Wir zahlen in US-Dollar, und der ist in Venezuela extrem begehrt. Bei uns sind auch viele andere Länder vertreten, aber Venezuela stellt klar die Mehrheit. Das war natürlich nicht absehbar, denn wir wachsen organisch und betreiben keine Akquise von Arbeitskräften. Die Leute erfahren über uns durch Mund-zu-Mund-Propaganda, meist von ihren Freunden.“ (Kevin Guo, Hive)

> „Angesichts der Herkunft unserer Arbeiterschaft müssen wir natürlich alles ins Spanische übersetzen und sind jetzt zweisprachig. Die Arbeitskräfte sind für die Art der Aufgaben ungelernt, aber mithilfe der Werkzeuge bringen wir ihnen die nötigen Fertigkeiten bei.“ (Kevin Guo, Hive)

Abbildung 1

Alexa-Besucherstatistik für hivemicro.com 2018



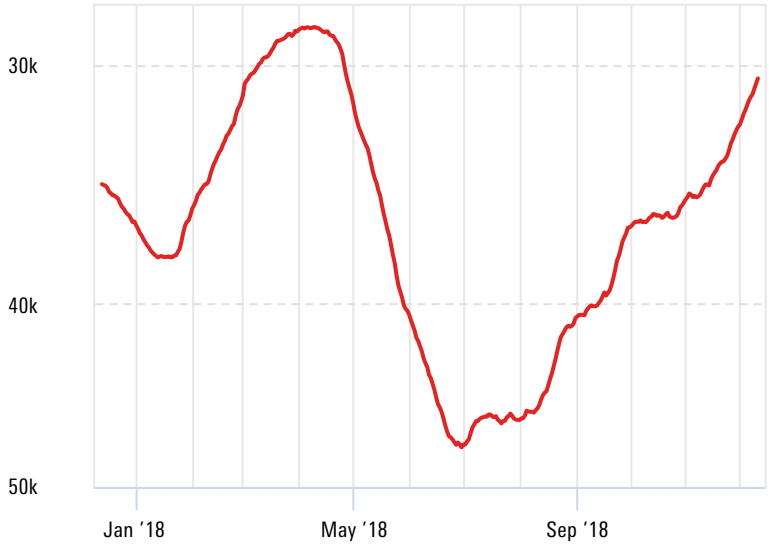
Globaler Rang: 27.941
Rang in Venezuela: 187

Herkunft der Besucher nach Ländern:
Venezuela 75,9%
Indien 6,0%
Vietnam 1,8%
USA 1,7%
Philippinen 1,6%

Quelle: Daten abgerufen am 12.12.2018 https://www.alexa.com/siteinfo/hivemicro.com

Abbildung 2

Alexa-Besucherstatistik für spare5.com 2018



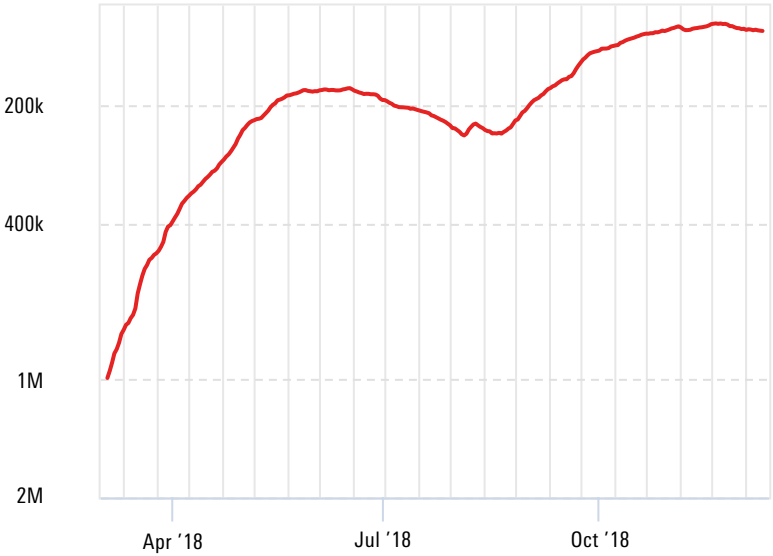
Globaler Rang: 30.531
Rang in Venezuela: 162

Herkunft der Besucher nach Ländern:
Venezuela 76,2%
Kolumbien 4,1%
Mexiko 2,5%
Indien 2,1%
Argentinien 1,4%

Quelle: Daten abgerufen am 12.12.2018 https://www.alexa.com/siteinfo/spare5.com

Abbildung 3

Alexa-Besucherstatistik für remotask.com 2018



Globaler Rang: 128.938
Rang auf den Philippinen: 2.129

Herkunft der Besucher nach Ländern:
Philippinen 31,1%
Venezuela 15,0%
Kenia 10,5%
Nigeria 9,2%
Indien 7,2%

Quelle: Daten abgerufen am 12.12.2018 https://www.alexa.com/siteinfo/remotasks.com

Das extreme globale Gefälle in der Höhe der Entlohnungen führt zu starken Spannungen und Fliehkräften auf dem Markt für Trainingsdaten. Einerseits kann natürlich keine funktionierende Volkswirtschaft mit dem Preisgefälle von Krisenregionen konkurrieren. Andererseits lässt sich aber auch nicht jede Art von Arbeit in jede Region exportieren. Zum einen haben manche Kunden, wie bereits weiter oben erwähnt, Complianceregeln, die aus Sorge um den Datenschutz oder um einen etwaigen Imageschaden durch Assoziation mit Billigarbeit den Export von Arbeit an eine Crowd oder in bestimmte Länder untersagen. Zum anderen sind, selbst wenn es sich um Crowdarbeit handelt, mitunter spezielle kulturelle und sprachliche Kenntnisse erforderlich.

> „Die Bilddaten kommen aus den unterschiedlichsten Ländern, und dort gibt es jeweils ganz eigene Fahrbahnmarkierungen, Zeichensysteme und Fahrzeugtypen. Das bedeutet für uns einerseits, dass wir unsere Leute sehr gut schulen müssen und sie auf jedes Detail ebenso achten müssen wie auf den Kontext. Es bedeutet aber auch, dass bestimmte Bedeutungsebenen nur von Menschen erfasst werden können, die sich in einer Region gut auskennen. Jemand aus Venezuela kann keine Verkehrsschilder aus Japan lesen. Deswegen ist es in unserem Interesse, eine möglichst vielfältige, internationale Arbeiterschaft zu haben." (Daryn Nakhuda, Mighty AI)

Angesichts der riesigen Mengen an benötigten Trainingsdaten schlagen die Kosten für die Arbeitskraft aber so deutlich zu Buche, dass die Anreize, das Lesen fremder Zeichensysteme mit einer Kombination aus Arbeitskräften aus Billiglohnländern, KI-Unterstützung und möglichst wenigen regionalen Expert_innen zu bewältigen, natürlich sehr hoch sind. Schließlich liegt der Mindestlohn in Ländern wie Japan, Deutschland und den USA um das Acht- bis Zehnfache über dem, was hoch motivierte Crowdworker_innen aus Venezuela derzeit als einen guten Verdienst erachten.

> „Die deutschen Automobilzulieferer und auch andere IT-Unternehmen, die KI-Forschung im Bereich Bilderkennung betreiben und große Mengen an Trainingsdaten benötigen, fragen deshalb bei uns schon gar nicht mehr nach, sondern gehen direkt ins Ausland, zu Mighty AI, Mechanical Turk oder CrowdFlower, wo die Arbeitskraft nur ein Zehntel kostet." (Haus Speidel, Crowd Guru)

> „Für uns wird die Wahl der Sprache gerade zu einer wichtigen strategischen Entscheidung über die Weiterentwicklung des Unternehmens, denn der neue Markt für Trainingsdaten ist für uns mit unserer rein deutschen Crowd eine schwer zu bewältigende Herausforderung." (Haus Speidel, Crowd Guru)

> „Es ist für uns auch ein Faktor, dass der deutsche Arbeitsmarkt im Moment sehr positiv dasteht. Das führt dazu, dass weniger Menschen nach Nebenjobs im Internet suchen, und wenn, dann mit entsprechender Erwartung an die Bezahlung. Die Attraktivität von Crowdarbeit ist nicht nur eine Frage der Lebenshaltungskosten, sondern der Jobalternativen." (Hans Speidel, Crowd Guru)

An dieser Stelle muss noch auf einen wichtigen Sonderbereich in der Produktion von Trainingsdaten hingewiesen werden, für den das eben Beschriebene nicht gilt. Während man die Verarbeitung von Bilddaten durch geschickte Zerteilung der Aufgaben zum allergrößten Teil ins Ausland verlagern kann, ist dies mit dem Trainieren von Sprachassistenzsystemen für regionale Dialekte nicht möglich.

In diesem Bereich sind besonders die Plattformen Appen und clickworker aktiv, und hier ist das Interesse an einer kulturell und global möglichst weit verstreuten Crowd nicht bloß ein Lippenbekenntnis, sondern essenziell. Wenn man also als Crowdworker_in der Maschine dabei hilft, besser Plattdeutsch oder Bayrisch zu verstehen, braucht man keine Sorge zu haben, dass einem jemand aus Venezuela den Job wegnimmt – andererseits kann es sich bei so einem Job aber im Gegenzug unweigerlich nur um eine Gelegenheitstätigkeit handeln, die nur ab und zu von Firmen nachgefragt wird. Professionelle Crowdarbeit gibt es in dieser Nische eher nicht. Dafür wandert diese Art der Crowdarbeit interessanterweise gerade in die Regionen, in denen die Wirtschaft floriert, weil die dort zu lernenden Dialekte mit entsprechender Kaufkraft einhergehen.

> „Gerade bei den Sprachaufnahmen geht es darum, eine möglichst große Bandbreite von Originärdaten, also ganz verschiedene Dialekte zum Beispiel oder unterschiedliches Sozialverhalten, zu erfassen, damit die KI künftig diese unterschiedlichen Sonderfälle verstehen kann. Dementsprechend breit muss auch das Spektrum an Clickworkern sein, und es hilft nicht, wenn einer von denen schon seit Jahren nur Sprachaufnahmen macht. Die Masse macht's. Deswegen ist es eher nicht so, dass daraus ein Berufsbild wird." (Christian Rozsenich, clickworker)

> „Auf der Kundenseite sind die USA und Israel [die für uns wichtigsten Länder], was daran liegt, dass die meiste Forschung im Bereich der künstlichen Intelligenz dort stattfindet. Was die Crowdseite angeht, so ist das sehr gemischt. Auch hier sind die USA wichtig für uns, hinzu kommen die großen Märkte in Europa, besonders Großbritannien und Deutschland, aber auch Skandinavi-

en. Die Nachfrage nach Crowdworkern aus bestimmten Ländern orientiert sich an der ökonomischen Stärke der einzelnen Regionen. Wenn ein Unternehmen ein neues Produkt einführt oder seinen E-Commerce-Bereich ausbaut, dann geschieht das erst einmal dort, wo am meisten Wertschöpfung zu erzielen ist. Das beeinflusst wiederum, welche Sprachkenntnisse bei uns nachgefragt werden." (Christian Rozsenich, clickworker)

Fünf „Fives" von Spare5 im Porträt

Die Crowdworker_innen der zu Mighty AI gehörenden Arbeitsplattform Spare5 werden als „Fives" bezeichnet. Die folgenden Zusammenfassungen von per Videochat auf Englisch im Mai 2018 geführten halbstündigen Interviews mit fünf Fives sind natürlich nicht repräsentativ, aber sie geben doch ein konsistentes Stimmungsbild der Plattform zu diesem Zeitpunkt wieder.

Eine erste, vielleicht überraschende Gemeinsamkeit ist, dass sie sich alle fünf mit der Bezeichnung „Fives" identifizieren – sie seien eher „Fives" als „Crowdworker_in", wobei auch der allgemeinere Begriff passe. Trotz großer Unterschiede bezüglich des demografischen Hintergrundes und, damit zusammenhängend, der Motivation, auf dieser Plattform zu arbeiten, waren alle fünf in der Bewertung von Spare5 sehr positiv. Dies mag dem Umstand geschuldet sein, dass vier der fünf Befragten von der Firma selbst als Interviewpartner vorgeschlagene hoch motivierte und verdiente Arbeitskräfte sind, aber ihre Haltung zur Plattform deckt sich mit anderen Stimmen in Crowdwork-Foren im Netz sowie mit den eigenen stichprobenartigen Tätigkeiten als Crowdworker auf Spare5.

Bevor auf die biografischen Besonderheiten der fünf Fives eingegangen wird, vorab zusammengefasst die wichtigsten Gemeinsamkeiten, die sich durch alle Interviews zogen: Alle Interviewpartner_innen haben einen relativ hohen Bildungsgrad. Auf der Plattform waren sie meist gelandet, weil sie sehr gezielt im Internet nach Möglichkeiten gesucht hatten, von zu Hause aus Geld zu verdienen, hatten bereits Erfahrungen mit anderen Crowdwork-Plattformen gesammelt und sich dann bewusst für Spare5 entschieden. Hierfür wurden insbesondere drei Gründe genannt:

Der wichtigste war, dass die Mitarbeiter_innen von Spare5 mit den Crowdworker_innen schnell, verständlich, persönlich und zuverlässig kommunizierten und sich die Crowdworker_innen dadurch ernst genommen fühlten, den Eindruck hatten, mit einer „echten" (so eine häufige Formulierung in Abgrenzung zu Online-Scams) und vertrauenswürdigen Firma als

„Arbeitgeber“ zu tun zu haben und bei Problemen unmittelbar mit Hilfe rechnen zu können.

Als ähnlich wichtig, aber leicht nachgelagert, wurde der Umstand beschrieben, dass die Arbeit besser bezahlt sei als auf ähnlichen Plattformen, und vor allem, dass sehr zuverlässig gezahlt werde. Die wöchentliche Auszahlung per PayPal war für alle sehr attraktiv. Der Umstand allerdings, dass in US-Dollar bezahlt wird, ist nur für diejenigen ein Vorteil, die aus Regionen mit schwächerer Währung kommen. Dort aber ist dieser Vorteil alles entscheidend. Alle Interviewpartner_innen klagten darüber, dass ihre Einkünfte zurückgegangen seien, zum einen, weil weniger Arbeit verfügbar sei, zum anderen, weil pro Aufgabe weniger bezahlt werde als noch einige Monate zuvor. Alle Interviewten hegten die Vermutung, dass dies mit einem Überangebot an Arbeitskraft zusammenhänge, manche brachten dies wiederum insbesondere mit dem Zustrom von Menschen aus Venezuela in Zusammenhang.

Der dritte häufig genannte Grund für die Arbeit bei Spare5 war, dass die Arbeit viel Spaß mache und dass insbesondere die semantische Segmentierung eine intrinsisch befriedigende Tätigkeit sei, bei der man stolz auf die eigenen Fertigkeiten, den wachsenden Grad an Perfektion und das sichtbare Ergebnis der Arbeit sein könne.

Daniela, 55 Jahre, aus Italien

Daniela arbeitet seit Januar 2017 für Spare5, zum Zeitpunkt des Interviews seit fast eineinhalb Jahren. Ihrer Einschätzung nach ist sie ungewöhnlich alt für diese Art von Arbeit. Sie ist sehr aktiv im Community-Forum und hat dort den Eindruck gewonnen, dass die meisten ihrer Kolleginnen und Kollegen zwischen 18 und 25 Jahre alt sind, nur wenige über 30, noch viel weniger über 50.

Daniela hat einen Bachelorabschluss in Modedesign und bis 2008 eine Schmuckfirma geleitet. Aus finanziellen Gründen müsste sie nicht mehr arbeiten, ist aber der Untätigkeit schnell überdrüssig geworden und hat deshalb angefangen, im Internet nach einer Beschäftigung mit Nebenverdienst zu suchen.

Seit 2009 hat sie so zahlreiche Plattformen getestet, darunter Upwork, Microworkers und Mechanical Turk. Die meisten Plattformen gefielen ihr nicht. „Ich habe alle ausprobiert und dann die gewählt, die am verlässlichsten sind und die eindeutigsten Aufgabenbeschreibungen haben.“ Dies seien Spare5, und, mit einigem Abstand, clickworker. Auch für Remotasks und Hive

Work ist sie gelegentlich tätig, betont jedoch, dass Spare5 ihr Favorit ist, weil Remotasks technische Fehler habe und man auf clickworker zu viel Zeit in Umfragen mit ungewisser Bezahlung verliere. (Man werde erst, nachdem man viele Fragen beantwortet habe, als „nicht zur Zielgruppe passend" aussortiert und dann nicht bezahlt – ein Problem über das auch Crowdworker_innen auf anderen Plattformen klagen.)

Daniela konsultiert regelmäßig Foren für Onlinearbeit, z. B. „We Work Remotely" (www.weworkremotely.com) und das Reddit-Subforum „/r/beermoney". Darüber hinaus sucht sie im Internet gezielt nach Schlagworten wie „Trainingsdaten für autonomes Fahren", um mehr Aufgaben wie bei Spare5 zu finden, weil sie die Tätigkeit „liebt" und immer viel positives Feedback für ihre akkurate Arbeitsweise erhält. Sie hat sechs Millionen Erfahrungspunkte auf der Plattform gesammelt, was zwei- bis dreimal so viel ist wie die anderen Interviewpartner_innen.

Sie habe die Arbeit für Spare5 zwar als Hobby begonnen, sei dann aber schnell „süchtig" danach geworden. Daniela beschreibt sich als stark intrinsisch motiviert und von Neugier geleitet. Es falle ihr leicht, sich in neue Werkzeuge und Anforderungen einzuarbeiten, immer mit dem Ziel, perfekte Ergebnisse zu liefern. Die Arbeit mache ihr so viel Spaß, dass es ihr manchmal schwerfalle, damit aufzuhören.

Inzwischen schaue sie täglich gleich nach dem Aufstehen nach neuen Tasks und dem Community-Forum von Spare5. Wenn auf Spare5 alle Tasks abgearbeitet seien, schaue sie noch auf anderen Plattformen vorbei. Sie verwende den größten Teil ihrer verfügbaren Zeit auf diese Tätigkeiten, betont zugleich jedoch, dass man von dieser Arbeit leider nicht leben könne, schon gar nicht in Italien, und dass sie Glück habe, weil sie das ja nicht müsse. In Spitzenzeiten kann sie bis zu 100 Dollar in einer Woche auf Spare5 verdienen, ihr Durchschnitt liegt bei etwa 50 bis 60 Dollar pro Woche.

Seit einem halben Jahr ist sie Community-Moderatorin. Sie war gefragt worden, weil der Firma ihre Hilfsbereitschaft im Forum aufgefallen war. Für diese zusätzliche Arbeit erhält sie gelegentlich eine kleine Kompensation. Daniela fühlt sich mitverantwortlich dafür, dass ihre Kolleg_innen die Aufgaben qualitätsvoll erledigen. Sie testet neue Aufgaben deshalb vorab, um Fragen in der Community nach der korrekten Durchführung gut beantworten zu können. Auf die häufigste aller Fragen im Forum – „Warum sehe ich so wenig Tasks?" – weiß sie allerdings auch keine Antwort.

Es sei nicht immer leicht, ihren Freunden in Italien zu erklären, was sie als Job macht. Meist sagt sie nur, dass sie für eine amerikanische Firma im Internet arbeitet, manchmal: „Ich helfe Computern, zu sehen, was ich sehe."

Ihre Identifikation mit Spare5 ist groß. „Wir sind Fives – ich bin eine Five." Sie glaubt, dass sie auch in einem Jahr noch für Spare5 arbeiten wird. „Für mich ist es ein Traumjob. Aber ich wäre gerne mehr als nur eine „Five" für Spare5! Ich hoffe, sie werden eines Tages ein Büro in Italien eröffnen, sodass ich fest für sie arbeiten kann." Daniela sieht in Crowdwork zwar die Zukunft der Arbeit, wünscht sich aber, sie könne als qualifizierte Fachkraft zur Vorarbeiterin und Trainerin aufsteigen. Schon jetzt sind viele ihrer Tasks Revisionen, bei denen sie die Ergebnisse der anderen kontrolliert. Gern würde sie auf dieser Basis Empfehlungen über die Eignung anderer Arbeiter_innen aussprechen.

Daniela äußert nur einen großen Kritikpunkt: „Ich finde, die Arbeit ist wirklich unterbezahlt, wenn man all die Zeit und Mühe einberechnet, die man hineinstecken muss." Sie wünscht sich, dass sich Qualität und Erfahrung stärker in der Bezahlung niederschlagen. Wie sich die Bezahlung genau zusammensetzt, wann es etwa Boni für besonders akkurate Ergebnisse gibt, weiß sie nicht. Im Gegensatz zu vielen anderen auf der Plattform hat sie aber meist genug Aufgaben zur Verfügung und glaubt, dass das an der Genauigkeit ihrer Arbeitsweise liegt.

Mit Blick auf die Mehrheit der Kolleg_innen aus Venezuela, deren Anteil sie als „gefühlt 90 Prozent" beschreibt, sagt sie, dass die Plattform eigentlich bei der Bezahlung die drastisch unterschiedlichen Lebenshaltungskosten der Arbeitskräfte berücksichtigen müsste. „Aber wenn ich diese Bedingungen akzeptiere, darf ich mich eigentlich auch nicht beschweren. Wahrscheinlich werden sie so nie die Bezahlung erhöhen – warum auch, es ist ja sehr bequem für sie."

Sara, 58 Jahre, aus Brasilien

Sara arbeitet bereits seit Ende 2014 für Spare5 und kennt die Plattform von Anfang an. Eigentlich ist sie Erzieherin und arbeitet phasenweise als Kindergärtnerin und Englischlehrerin für Kinder. Sara spricht Holländisch, Portugiesisch, Spanisch und Englisch jeweils fließend. Einige Monate im Jahr lebt sie in den Niederlanden, wo sie ihre erwachsene Tochter besucht.

Die Arbeit für Spare5 ist für sie einer von vielen Teilzeitjobs. Ab und zu schaut sie auch nach anderen Plattformen, aber ihrer Erfahrung nach sind die meisten Webseiten, die mit der Möglichkeit eines Zuverdiensts im Internet werben, „fake" oder unseriös. Sie schätzt es, bei Spare5 reale Menschen als Ansprechpartner zu haben, die auf E-Mail-Rückfragen schnell und zuverläs-

sig reagieren. Gelegentlich kommuniziert sie auch mit anderen Mitgliedern der Spare5-Community, aber dieser Austausch spielt für sie eine untergeordnete Rolle.

Als langjährige Nutzerin von Spare5 betrachtet Sara den starken Zustrom von Menschen aus Venezuela mit gemischten Gefühlen. Sie äußert Verständnis für deren schwierige politische und ökonomische Situation, die auch im Community-Forum heftig diskutiert werde, zugleich glaubt sie jedoch, dass die Popularität von Spare5 in Venezuela zu einem Preisverfall auf der Plattform geführt hat. Für Aufgaben, die vorher fünf Cent erbracht haben, würden jetzt nur noch zwei Cent gezahlt. Dieser Wandel habe Anfang 2017 begonnen.

Saras Lieblingsaufgabe aus den Anfangsjahren war es, telefonisch Daten bei Hotels zu erfragen. Vor der Spezialisierung auf Trainingsdaten, als es diese laut Sara sehr gut bezahlten Tasks noch gab, hat sie bis zu acht Stunden am Tag für Spare5 gearbeitet und bis zu 150 Dollar die Woche verdient. Insgesamt hat sie seit Ende 2014 etwa 3.700 Dollar verdient, im Durchschnitt also nur 18 Dollar die Woche. Trotzdem freut sie sich über die Gesamtsumme: „Es ist nie verkehrt, ein bisschen Geld dazuzuverdienen."

Inzwischen hat sie große Mengen an Trainingsdaten für das autonome Fahren produziert. Aber auch für diese Aufgaben habe es früher deutlich mehr Geld gegeben. Heute überlege sie zweimal, ob ihr die eigene Zeit die geringe Entlohnung wert ist. Oft sind auch keine Aufgaben für sie verfügbar, und ihr ist nicht klar, welche Faktoren dafür ausschlaggebend sind. Am Grad der Erfahrung könne es wohl kaum liegen, so Sara, denn sie sei ja praktisch von Anfang an dabei.

Sara ist stolz auf ihren versierten Umgang mit den Spare5-Werkzeugen, sie ist sich zugleich jedoch auch bewusst, dass diese Fertigkeiten außerhalb von Spare5 wenig Wert haben. Sie ist sich sicher, dass sie auch in einem Jahr noch gerne für Spare5 arbeiten wird. Am meisten schätzt sie die Flexibilität, zu arbeiten, wann, wo und wie es ihr passt, und keinen Chef über sich zu haben. Für Sara ist die Arbeit für Spare5 zwar kein Vollzeitjob, aber doch mehr als nur ein Hobby. Sie genießt es, die Arbeit gelegentlich auch vom Strand aus ausüben zu können. In Brasilien komme sie mit dem Zusatzeinkommen auch sehr viel weiter als in den Niederlanden.

Ihre Hauptkritikpunkte sind, dass es zu wenig Arbeit gibt und dass sie zu schlecht bezahlt ist. Eine Bezahlung nach aufgewendeter Zeit im Gegensatz zur vorherrschenden Akkordarbeit hält sie nicht für sinnvoll. Paradoxerweise glaubt sie einerseits, dass sie selbst durch ihr Pflichtbewusstsein davon gestresst wäre, weil sie dann das Gefühl hätte, härter für den stündlichen Lohn

arbeiten zu müssen; umgekehrt glaubt sie, dass andere Fives dann nicht mehr genügend oder extralangsam arbeiten würden. Hier wechselt sie in ihrer Einschätzung der Sinnhaftigkeit eines Stundenlohns zwischen Arbeitgeber- und Arbeitnehmerperspektive hin und her.

Isabel, 34 Jahre, aus Venezuela

Isabel ist angehende Zahnärztin, hatte allerdings zum Zeitpunkt des Interviews ihr Studium temporär ausgesetzt. Spare5 hat sie durch gezielte Suche nach Onlinearbeit gefunden. In intensiven Phasen, wenn genug Arbeit verfügbar ist, arbeitet sie zehn bis zwölf Stunden am Tag für die Plattform. Wenn dort keine Tasks verfügbar sind, greift sie auch auf Figure Eight und Remotasks zurück. Wegen der freundlicheren Benutzeroberfläche und der besseren Bezahlung zieht sie aber Spare5 vor. Ihr Einkommen auf der Plattform schwankt stark mit der Verfügbarkeit von Tasks. Wenn es Arbeit gibt, liegt ihr Einkommen bei 30 bis 40 Dollar die Woche – das reiche, um damit ihre Rechnungen zu bezahlen, Essen zu kaufen und Fortbildungskurse zu buchen.

Allein von Spare5 könne sie aber nicht leben, weil die Arbeit nicht verlässlich verfügbar sei. Deswegen gehe sie auch verschiedenen Offlinenebenjobs nach. Aufgrund der ökonomischen Situation in Venezuela ist es für sie essenziell, dass die Plattformarbeit in US-Dollar vergütet wird, da dies eine der wenigen Möglichkeiten ist, an dringend benötigte Devisen zu kommen. In ihrem Umfeld bekommt Isabel für die Plattformarbeit Anerkennung und neugierige Fragen. „Na, dir geht es ja bestens, wie machst du das bloß?“ sei eine typische Reaktion von Freunden. In der Folge lernt Isabel immer neue Freunde auf der Plattform an. Auch ihre beiden Brüder arbeiten inzwischen für Spare5, doch Isabel stellt immer wieder fest, dass nicht jeder für diese Arbeit geeignet sei, da sie sehr viel Geduld und Genauigkeit erfordere.

Isabel hat inzwischen etwa zwei Millionen Experience-Points auf Spare5 gesammelt. Die Arbeit auf Spare5 mache ihr größtenteils viel Spaß, insbesondere bei der semantischen Segmentierung freue sie sich über die Endergebnisse. „Wenn man so ein Bild fertig hat, ist man stolz, etwas geschafft zu haben.“

Dennoch ist die Arbeit nur eine Übergangslösung, bis sie ihr Studium fortsetzen kann. Ihr Hauptkritikpunkt ist, dass es oft zu wenig Arbeit gibt. Besonders schätzt sie den guten Support durch die Mitarbeiter_innen von Spare5 und auch deren regelmäßig veranstaltete Live-Fragesessions. In auf

Spanisch und Englisch durchgeführten Videochats beantworten Spare5-Community-Manager via Facebook und YouTube Fragen und holen das Feedback der Community ein. Dies schaffe eine persönliche Bindung zur Plattform, und Isabel fühlt sich dadurch auf Spare5 deutlich respektvoller behandelt als auf anderen Plattformen.

Isabel geht davon aus, dass der Großteil der Spare5-Community aus Venezuela kommt. Sie weiß von Ärzt_innen, Anwält_innen und Ingenieur_innen aus ihrem Heimatland, die aufgrund der katastrophalen ökonomischen Situation vor Ort inzwischen dazu gezwungen sind. Es gebe eine gut organisierte Community venezolanischer „Fives", die sich über den Telegram-Messenger, WhatsApp, Facebook und das offizielle Forum austauscht.

Isabel schätzt, dass man derzeit als Einzelperson in Venezuela mit 200 Dollar im Monat über die Runden kommt und dass es nur unter diesen Bedingungen überhaupt Sinn macht, für Spare5 zu arbeiten. Sie versteht nicht, warum auch Menschen aus anderen Ländern diese Arbeit machen.

Ob die Arbeit fair entlohnt wird, sei eine in ihrer Community kontrovers diskutierte Frage. Manche Aufgaben seien fair bezahlt, andere nicht. Es hänge stark vom Zeitaufwand und von der verlangten Präzision einzelner Aufgaben ab. In ihrer Community herrsche Uneinigkeit darüber, ob internationale Firmen einen Vorteil daraus schlügen, dass es Venezuela so schlecht geht, und es deshalb ausbeuterisch sei, so wenig zu bezahlen, oder ob im Vordergrund der Beurteilung stehen sollte, dass diese Arbeit für die Menschen in der Region eine große Hilfe sei. Isabel selbst will, dass international bekannt wird, dass die Menschen in Venezuela aus der Not heraus auf den Plattformen arbeiten.

Douglas, 20 Jahre, aus Venezuela

Douglas hat zwei Jahre Maschinenbau studiert, das Studium jedoch nicht abgeschlossen. Für Spare5 arbeitete er zum Zeitpunkt des Interviews seit acht Monaten und hatte bereits hat drei Millionen Erfahrungspunkte gesammelt. Von der Plattform hatte er über einen Freund erfahren, der dort auch nach wie vor arbeitet. Douglas hat einige ähnliche Plattformen ausprobiert (www.hivemicro.com und www.Humanatic.com), war aber von allen anderen Anbietern schnell frustriert, weil die Arbeit hart und die Bezahlung schlecht war.

In seinen ersten Monaten auf Spare5 hatte Douglas noch genug Arbeit vorgefunden, um zehn Stunden am Tag oder länger dort zu arbeiten. Seit Anfang 2018 sei die Menge an Arbeit aber stetig zurückgegangen, sodass er jetzt

höchstens noch Arbeit für zwei Stunden am Tag vorfindet. In den letzten Wochen hatte es gar keine Arbeit mehr für ihn gegeben, weshalb er und seine Kolleg_innen sich gezwungen sahen, auf die weniger populären Plattformen auszuweichen.

Für Douglas ist die Arbeit auf Spare5 seine Haupteinnahmequelle, mit der er seine Mutter und seinen jüngeren Bruder mitfinanziert. In der hochproduktiven Anfangsphase konnte er sogar noch Geld ansparen, mit dem er derzeit die ausbleibende Arbeit überbrücken kann. Als es noch gut lief, habe er etwa 50 Dollar die Woche verdient, und er gibt (im Kontrast zu Isabel) an, dass man in seiner Region auch mit 20 Dollar im Monat überleben könne.

Die Arbeit mache Spaß und sei interessant genug, um sie auch in einem Jahr noch machen zu wollen. Am liebsten mag Douglas die semantische Segmentierung. Innerhalb der letzten Monate habe sich aber sehr viel an den dafür eingesetzten Werkzeugen und am Interface verändert. Die Arbeit sei dadurch anspruchsvoller und komplizierter geworden. Douglas habe eine sehr steile Lernkurve absolviert und viele spezielle Fähigkeiten auf Spare5 erlernt, die ihm jedoch zu seinem Bedauern nur dort zugutekommen.

Auch er lobt den Umgang der Mitarbeiter_innen mit der Crowd: „Ich habe Spare5 schon dreimal wegen irgendwelcher Fragen angeschrieben, und jedes Mal haben sie so freundlich und persönlich geantwortet – das sind echt tolle Leute."

Douglas sagt, dass sich die Bezahlung in den letzten Monaten deutlich verschlechtert habe und dass Aufgaben, für die man zuvor noch fünf Cent erhalten habe, jetzt nur noch mit zwei oder drei Cent vergütet würden. Seine Theorie ist, dass die Arbeit deswegen jetzt schlechter bezahlt werde, weil die Plattform ein und dieselbe Aufgabe an mehr Leute vergeben müsse, um den gestiegenen Qualitätsansprüchen zu genügen. Douglas glaubt, dass Spare5 ausgleichen müsse, dass es viele ungeduldige und unerfahrene neue Kolleg_innen gibt, die möglichst schnell, ohne die Beschreibungen zu lesen und ohne Anspruch auf Genauigkeit, ein paar Dollar verdienen wollen, bevor ihr Account wegen zu schlechter Ergebnisse gesperrt wird und sie dann einfach einen neuen Account eröffnen.

Die Akkordarbeit hält Douglas für ein gutes System, denn wenn nicht nach „Stückzahl" bzw. Tasks, sondern per Zeiteinheit bezahlt würde, könne man ja einfach den Computer laufen lassen, ohne die Arbeit zu erledigen.

Douglas ist in mehreren Facebook-Gruppen, in denen er sich mit anderen über die Tasks auf Spare5 austauscht. Das offizielle Forum erfährt er ebenfalls als sehr hilfreich. Die Community komme zu 75 Prozent aus Venezuela, schätzt er. Im spanischsprachigen Teil des Forums sei die Stimmung nervös

oder aufgebracht, weil die Menschen inzwischen angewiesen sind auf die derzeit ausbleibenden Tasks. Er betont, dass die Schuld dafür nicht bei Spare5 liege. „Vielleicht haben sie selbst zu wenige Aufträge." In letzter Zeit hat er auf der Plattform besonders viele Bilder aus Japan bearbeitet, aber auch aus Deutschland seien immer wieder viele Bilder dabei.

An seiner Arbeit für Spare5 schätzt er die Freiheit und die Unterstützung durch die Firma. Wie Isabel lobt auch er die Live-Fragesessions auf YouTube und Facebook. Als einer der besten „Fives" darf er manchmal neue Werkzeuge vorab testen, wird um seine Meinung gefragt und fühlt sich dadurch geehrt. Es gefällt ihm, direkt mit Mitarbeitern_innen der Plattform zu tun zu haben und Einfluss auf die Produktentwicklung nehmen zu können.

Douglas berichtet, dass in der spanischsprachigen Spare5-Community immer wieder diskutiert werde, ob man überhaupt noch neuen Leuten von der Plattform erzählen solle, um nicht durch ein Überangebot an Arbeitskraft weiter die Preise zu verderben. „Manche Leute wollen die Henne, die die goldenen Eier legt, lieber für sich behalten." Er selbst sieht das anders und hält es für selbstverständlich, dass man andere Landsleute in misslicher Lage auf die Plattform einladen müsse.

Wilson, 22 Jahre, aus Venezuela

Wilson studiert Elektrotechnik und hofft, in einem Jahr mit dem Studium fertig zu sein. Er lebt allein, und die Plattform ist seine einzige Einnahmequelle. Seit einem Jahr arbeitet er für Spare5, und dies mache ihm nicht nur Spaß, sondern sei für ihn auch eine große Hilfe, um mit der schwierigen Situation in Venezuela klarzukommen. Zum Zeitpunkt des Interviews hatte Wilson auf Spare5 knapp eine Million Experience-Points angesammelt und etwa 20.000 Tasks erledigt.

Spare5 hat er über YouTube gefunden, wo er nach Wegen gesucht hatte, um im Internet Geld zu verdienen. Nebenbei ist er auch auf Hive Work und Remotasks sowie auf den breiter aufgestellten Plattformen Clixsense und Neobux aktiv, die unter anderem ebenfalls Trainingsdaten für autonomes Fahren anbieten, Clixsense sogar im Auftrag von Figure Eight.

Wilsons Hauptkritikpunkt an allen Plattformen ist, dass die Instruktionen oft unklar seien. Hier brauche es dringend Verbesserungen. Weil die Bezahlung besser ist und die Aufgaben leichter zu verstehen sind, zieht er Spare5 vor, beobachtet aber eine Wanderbewegung vieler Kolleg_innen hin zu anderen Plattformen, weil es dort wenigstens Arbeit gebe, wenn auch schlechtere.

Je nach Verfügbarkeit von Aufgaben auf Spare5 verdient Wilson dort 23 bis 27 Dollar die Woche, wobei er hauptsächlich am Wochenende arbeitet, und zwar in „Schichten" von fünf bis sechs Stunden, etwa drei Tage die Woche. Er verdient also im Schnitt zwischen 1,50 und 1,80 Dollar die Stunde. Um zu überleben, brauche er monatlich nur etwa 30 Dollar. Angesprochen darauf, dass eine andere Interviewpartnerin aus Venezuela (beide nicht aus Caracas, sondern aus unterschiedlichen Provinzregionen) deutlich mehr Geld im Monat brauche, erklärt er sich diesen Unterschied aus dem Umstand, dass er allein lebt und keine Familienmitglieder versorgen muss.

Wilson glaubt, dass er auch in einem Jahr noch Trainingsdaten für autonome Fahrzeuge erstellen wird: „Weil mir die Arbeit Spaß macht und weil ich darauf angewiesen bin." Er hofft, dass die Situation in Venezuela sich so weit bessert, dass diese Form der Arbeit nur eine temporäre Phase in seinem Leben bleibt. Zugleich glaubt er, dass es sich bei der Erstellung von Trainingsdaten um einen Wachstumsmarkt handelt, der noch lange bestehen wird. Spare5 repräsentiert für ihn „Hoffnung für die Menschen aus Venezuela", und entsprechend häufig empfiehlt er die Plattform weiter.

DISKUSSION UND AUSBLICK

In einer bemerkenswerten Studie zur „Zukunft der Crowdarbeit“ von 2013 stellte eine Gruppe namhafter Crowdsourcing- und Human-Computer-Interaction-Experten um Aniket Kittur (Carnegie Mellon) und Michael Bernstein (Stanford) eine wichtige und weitreichende Leitfrage: „Können wir uns einen Crowd-Arbeitsplatz der Zukunft vorstellen, der so gestaltet ist, dass wir unsere Kinder dort arbeiten sehen wollten?“ (Kittur et al. 2013; dt. Version in: Benner 2015).

Besagte Studie hat durch die aktuelle Entwicklung der hier beschriebenen Crowdproduktion von Trainingsdaten noch an Relevanz gewonnen. Denn einerseits machten die Autoren 2013 eine ganze Reihe von Vorschlägen, wie man Crowdarbeitsplätze weiterentwickeln und verbessern müsste, die sich heute in den neuen Plattformen bereits umgesetzt wiederfinden. Und in mancher Hinsicht ist die Qualität der Arbeit aus Perspektive der Crowdworker_innen wirklich besser geworden. Andererseits sind wir trotz dieser Verbesserungen immer noch weit davon entfernt, Crowdarbeit auf ein Niveau zu heben, welches man seinen Kindern als Arbeitsplatz der Zukunft wünschen wollte.

Was sich verbessert hat, warum dies trotzdem nicht ausreicht, um einen guten Arbeitsplatz zu schaffen, und wie die Perspektiven für diese spezielle Form der Crowdarbeit und die Situation der Arbeiter_innen aussehen, das soll hier abschließend kurz erörtert werden.

Die drei zentralen Verbesserungsvorschläge, die von den Wissenschaftlern um Aniket Kittur 2013 für die Zukunft der Crowdarbeit gemacht wurden, sind:

1. Die Erzeugung von Karriereleitern für Crowdworker_innen innerhalb der Plattformen.
2. Die Verbesserung der Aufgabenstellung durch bessere Kommunikation mit den Crowdworker_innen.
3. Die ständige Weiterbildung der Crowdworker_innen durch KI-Unterstützung.

Alle drei Kernpunkte sowie zahlreiche weitere Visionen für die Zukunft der Crowdarbeit aus der Studie von 2013 sind bei Mighty AI/Spare5 in Gänze und bei den meisten anderen hier im Detail besprochenen Dienstleistern für Trainingsdaten zumindest in Abstufungen in die Plattformen integriert worden.

In den Interviews zeigte sich eine große Zufriedenheit der „Fives“ mit der Tätigkeit selbst, insbesondere ein gesteigertes Selbstwertgefühl durch häufiges quantitatives und qualitatives Feedback. Ersteres durch Gamification-Funktionen wie Erfahrungspunkte für Tätigkeiten, auf die man sich gemäß eigener Vorlieben und Talente spezialisieren und in denen man dann zu besserer Bezahlung, anspruchsvolleren Aufgaben, mehr Verantwortung und vorarbeiterähnlichen Rollen aufsteigen kann. Letzteres durch persönliches und aufwendiges Community-Management, bei dem den Crowdworker_innen immer unmittelbar und persönlich auf ihre Verständnisfragen geantwortet wird und bei dem regelmäßig ein Team von festen Mitarbeiter_innen der Plattform per Videochat via YouTube oder Facebook Live mit der Crowd spricht.

Die Beispiele zeigen, dass all dies dazu führen kann, dass die Crowdproduktion von Trainingsdaten als intrinsisch lohnende, attraktive Tätigkeit wahrgenommen wird, bei der man stolz auf seine Arbeit sein kann, sich von den Plattformbetreibern als Mensch respektiert fühlt und sogar das Gefühl hat, ständig innerhalb der Plattform aufzusteigen. Das ist ein echter Fortschritt gegenüber Amazon Mechanical Turk, den man auch als solchen anerkennen muss.

2013 war all dies noch weitgehend Utopie, und aus diesem Blickwinkel betrachtet hatte die Studie „Zukunft der Crowdarbeit“ geradezu prognostische Qualitäten. Zum Vergleich: Die Crowdworkerin und MTurk-Aktivistin Kristy Milland klagte zu dem Zeitpunkt noch über Amazon: „We have been sold as nothing more than an algorithm“ (in: Benner 2015). Die gleiche, sehr verständliche Klage findet sich auch in der Aktion „We Are Dynamo“, in welcher sich „Turker“ wie Kristy Milland und viele andere in offenen Briefen direkt an Jeff Bezos wandten, um ihn daran zu erinnern, dass sie Menschen aus Fleisch und Blut sind, keine Maschinenteile, und dass sie erwarten, dass man mit ihnen entsprechend respektvoll umgeht (http://www.wearedynamo.org/dearjeffbezos; sowie Salehi et al. 2015).

Das entscheidende Thema der Bezahlung noch einen weiteren Moment außen vor gelassen, könnte man jetzt zu dem Schluss gelangen, dass die Crowdproduktion von Trainingsdaten eine vielversprechende Zukunft für Crowdarbeiter_innen auf der ganzen Welt bieten könnte, weil ein großer Bedarf an dieser Arbeit besteht; weil nicht nur die Crowd die Algorithmen trainiert, sondern auch umgekehrt, und weil eben dieses Training die Crowdworker_innen wertvoll für die Plattformen macht und diese dadurch einen Anreiz haben, ihre Arbeiterschaft gut zu behandeln. Die Arbeiter_innen kämen, so könnte man optimistisch annehmen, durch das Training im System

bzw. die ständigen Fortbildungen, die sie selbst erfahren, aus der Falle der ungelernten Hilfskräfte, der unskilled labour. Sie könnten aufsteigen zu professionellen Crowdworker_innen mit mehr Sicherheit im Job dank beständig großer Nachfrage nach dieser Art von Arbeit. Und allem Anschein nach bleibt diese Nachfrage, soweit man dies vorhersagen kann, weiterhin groß, darüber besteht zumindest unter den Plattformbetreibern kaum Zweifel:

> „Insgesamt wird die Nachfrage nach von Menschen annotierten Originärdaten […] zumindest in den nächsten fünf Jahren noch deutlich steigen, weil noch nicht absehbar ist, ob es überhaupt irgendwann KI-Systeme geben wird, die ohne solche Referenzdaten völlig autonom arbeiten können." (Christian Rozsenich, clickwork)

> „Ich glaube der Bedarf wird weiter wachsen, aber die Art der Aufgaben wird sich ständig ändern. Wir machen heute schon ganz andere Aufgaben als noch vor einem Jahr. […] Wenn autonome Fahrzeuge erst einmal marktreif sind, werden wir weniger den Input als den Output der Maschine-Learning-Modelle unserer Kunden prüfen […]. Das Auftragsvolumen für Validierungsdaten wird sehr viel höher sein als das für Trainingsdaten." (Daryn Nakhuda, Mighty AI)

> „Sobald mehr als nur Prototypen auf der Straße sind, wird die Menge an anfallenden Daten drastisch steigen. Prognosen zu treffen ist immer schwierig, aber ich rechne mit einem exponentiellen Wachstum. Die Art der Aufgaben wird sich aber immerzu ändern. Ich denke, dass man z. B. bei der Auswertung von Unfällen künftig immer ‚humans-in-the-loop' wird haben wollen, um zu verbindlichen Ground-Truth-Daten zu gelangen." (Siddharth Mall, Playment)

Es ist zwar durchaus denkbar, dass die Entwickler vollständig autonomer Fahrzeuge die letzten paar Prozent unbedingt notwendiger Zuverlässigkeit nicht in den Griff bekommen und es deshalb zu weiteren tödlichen Unfällen kommt, die sich dann wiederum ungünstig auf die Akzeptanz der Technologie in der Bevölkerung und vor allem die Rechtsprechung auswirken könnten. Aber selbst wenn wir nicht flächendeckend vollautonome Autos bekommen, werden ja auch schon für die unabhängig davon expandierenden semiautomatischen Fahrerassistenzsysteme weiterhin riesige Mengen annotierter Bilddaten benötigt.

Wir können dieses „best case scenario" zur Crowdproduktion von Trainingsdaten also wie folgt zusammenfassen: Die Arbeit macht vielen Menschen Spaß, die Arbeitskräfte fühlen sich gut behandelt und als Teil einer Community, in der sie das Gefühl haben, sich mit ihren Fähigkeiten gut einbringen zu können, immer besser zu werden und dafür respektiert zu wer-

den, sowohl innerhalb der Plattform also auch, wie einige der Crowdworker_innen aus Venezuela berichteten, außerhalb im persönlichen Umfeld. Ein Ausbleiben der Arbeit ist zudem nicht in Sicht.

Doch selbst in diesem Idealfall, der natürlich längst nicht überall in der Crowdsourcing-Branche vorherrscht oder zu erwarten ist, bleiben zwei fundamentale und, wie es scheint, unlösbare Probleme: das Problem der Schwankungen in der Verfügbarkeit von Arbeit und das Problem der Abwärtsspirale in der Bezahlung. Schon jedes einzelne dieser beiden miteinander verknüpften Probleme – insbesondere aber beide im Verbund – haben zur Folge, dass die Arbeit in der Crowdproduktion von Trainingsdaten für die Crowdworker_innen außerordentlich prekär ist. Zumindest vom in dieser Studie dargelegten Kenntnisstand ausgehend ist keine Stellschraube erkennbar, mit der man der extremen Volatilität in diesem sehr speziellen Arbeitsmarkt auf globaler Ebene nachhaltig begegnen könnte.

Das Problem der Crowd bleibt die Crowd. In den letzten zwei Jahren waren auf den hier beschriebenen neuen Plattformen in der Summe etwa 1 bis 1,5 Millionen Neuregistrierungen von Crowdarbeiter_innen zu verzeichnen teils mit deutlichen Überschneidungen. Hinzu kommen 2,7 Millionen Registrierungen auf MTurk, clickworker und Appen, ebenfalls sicherlich mit Überschneidungen, und angeblich 5 Millionen Crowdworker_innen auf der Metaplattform Figure Eight. Selbst wenn man von starken Überschneidungen und zahlreichen Karteileichen ausgeht, arbeiten inzwischen mehrere Millionen Menschen als Microtasking-Crowdworker_innen. Das mag man viel oder wenig finden. Der entscheidende Punkt ist hier, dass es den neuesten Plattformen innerhalb sehr kurzer Zeit gelungen ist, mehrere Hunderttausend Arbeitskräfte zu rekrutieren. Genauer gesagt rekrutieren sich diese Menschen selbst, nachdem sie sich auch selbst diese virtuelle Arbeitsmöglichkeit im Internet gesucht haben.

Die Crowd ist und bleibt ein Massenphänomen, in dem das Individuum per Definition austauschbar ist. Es ist offenbar global immer einfacher, große Gruppen zu rekrutieren und dann per Software, also automatisiert und folglich skalierbar, zu trainieren, sodass die Verhandlungsmacht der Crowdworker_innen, was die Entlohnung angeht, gegen null geht.

Hierin liegt ein zentrales Paradox der Crowdproduktion von Trainingsdaten. Das Training der Algorithmen lässt sich nicht direkt automatisieren, aber das Training der Crowd, welches Ersteres wiederum indirekt ermöglicht, lässt sich offenbar sehr gut automatisieren, wenn man als Plattform bereit ist, fortlaufend in Werkzeuge zu investieren, in denen Crowd und KI eine kybernetische Einheit bilden. Ein Apparat, in dem die Massen die Ma-

schine trainieren und umgekehrt. Ein Apparat, in dem die Massen die Maschine kontrollieren und umgekehrt. All dies mit dem Zweck, wiederum andere Automatisierungssysteme zu optimieren.

Das Besondere an den Trainingsdaten ist, dass sie so plötzlich in so riesigen Mengen benötigt werden, und das von Kunden mit sehr viel verfügbarem Kapital. Sei es, weil es sich um schon lange am Markt befindliche Großkonzerne handelt oder um mit viel Risikokapital ausgestattete Start-ups. Die kapitalgetriebene massenhafte Nachfrage nach gleichförmigen Trainingsdaten, die nur durch ebenso massenhafte repetitive Crowdarbeit hergestellt werden können, erzeugt sich so ihre eigene Crowd. Die Prinzipien dahinter sind nichts Neues, aber am Beispiel der Trainingsdaten lassen sich die wirkenden Kräfte wie im Zeitraffer beobachten. Die austauschbaren Tasks erzeugen die austauschbare Crowd.

Alle drei Gruppen von Beteiligten, also Technologiehersteller/Plattformkunden, Crowdplattformen/Intermediäre und Crowdworker_innen/Hilfs- und Wanderarbeiter_innen, haben es auf ihrer Ebene jeweils mit einem enormen globalen Preisdruck zu tun. So wäre es verkürzt, etwa die Auftraggeber aus der Autoindustrie oder die Plattformbetreiber isoliert als rücksichtslose Ausbeuter darzustellen. Zugleich wäre es aber auch falsch, zu bestreiten, dass dieses aktuelle Wettrennen um das autonome Fahren aufseiten der Crowd Ausbeutung zur Folge hat, selbst wenn man wie z. B. Mighty AI bemüht ist, die Crowd so gut wie möglich zu behandeln.

Wie wir an den Aussagen von Crowd-Guru-CEO Hans Speidel gesehen haben, ist eine Plattform, die sich um die großen Trainingsdaten-Budgets der Automobilkunden bewirbt, ohne auf eine „Drittwelt-Crowd“ zurückgreifen zu können, angesichts einer um den Faktor zehn teureren Arbeitskraft schlichtweg nicht konkurrenzfähig.

Die Crowdsourcing-Branche als solche kann man ohnehin schon als Triumph der Intermediäre verstehen. Ist in anderen Bereichen „to cut out the middle man“ eine Businessstrategie, um Kosten zu sparen, scheint sich das Crowdgeschäft in dieser Hinsicht in die andere Richtung zu verkomplizieren, was man auch schon am Prinzip der Metaplattform CrowdFlower/Figure Eight beobachten kann.

In dem Moment, in dem große Budgets global auf der Suche nach billiger, gleichförmiger Arbeitskraft sind, eröffnet dies offenbar Möglichkeiten für neue Schichten von Intermediären, die das Problem der Schwankungen in der Arbeitskraft abpuffern, der nächsthöheren Ebene ein Mindestmaß an Qualität sichern und vor allem das Training und die Orchestrierung einer fremdsprachigen Crowd übernehmen.

> „Ein Phänomen, das uns neuerdings verschiedentlich begegnet, ist das Auftauchen kleinerer Crowd-Subunternehmer in afrikanischen Ländern, in Asien, aber auch in der Ukraine, die Crowdworker vor Ort rekrutieren, in Trainingscentern sowohl sprachlich als auch für spezifische Aufgaben und Plattformen schulen, um diese Arbeitskraft dann Plattformbetreibern wie uns im Paket anzubieten. Bisher stehen wir solchen Angeboten sehr skeptisch gegenüber, nicht nur, weil unklar ist, zu welchen Bedingungen diese Crowdworker dann am Ende arbeiten, sondern schon allein, weil wir große Probleme hätten, solche externen Crowds mit unserer eigenen Community zu vereinen. Das wäre für uns auf der Plattform politisch sehr heikel, wenn es plötzlich eine indische Sub-Crowd gäbe, die pro Stunde nicht mehr neun Euro sondern nur noch einen Euro bekäme." (Hans Speidel, Crowd Guru)

Was für Crowd Guru nicht möglich wäre, weil man sich über Jahre eine vergleichsweise homogene, einsprachige Crowd aufgebaut hat, ist für neue internationale Plattformen oder Metaplattformen wie Figure Eight natürlich kein so großes Problem.

Wenn wir abschließend noch einmal einen Schritt zurücktreten, um die aktuelle Crowdsourcing-Landschaft abstrakter zu betrachten, eröffnet sich anhand der Untersuchung der Crowdproduktion von Trainingsdaten der Blick auf ein verschachteltes System aus Schichten und Stapeln („Stacks"). Nicht nur bauen die Plattformen selbst potenziell aufeinander auf, ähnlich der Verkettungen von Outsourcing-Operationen in anderen Bereichen. In sich wiederum bestehen die Plattformen aus parallel prozessierenden Schichten von Menschen und Maschinen, Schichten von Crowds und KIs. (Eine Makrostruktur, die der Mikrostruktur der Datenverarbeitung innerhalb der „neuronalen Netze" nicht unähnlich ist.)

Auch wenn die Nachfrage nach crowdproduzierten Trainings- und Validierungsdaten wie prognostiziert weiter ansteigt, können sich die einzelnen Hilfsarbeiter_innen innerhalb eines dieser „Stacks" niemals sicher sein, ob nicht eine Fähigkeit, die sie gerade erst mühsam erlernt haben, aus Kostengründen von heute auf morgen entweder doch in einer höheren KI-Ebene teilautomatisiert oder auf eine externe Sub-Crowd ausgelagert wird.

So wie 2018 plötzlich viel Crowdarbeit unerwartet nach Venezuela geflossen ist, kann selbst diesen Menschen noch die Arbeit durch eine andere Schichtung der Prozesse innerhalb des Stapels oder durch eine neue Sub-Crowd in einer anderen Krisenregion streitig gemacht werden, mit der sich innerhalb der angestammten Plattform plötzlich die Preise senken lassen. Oder die Auftraggeber entscheiden sich für eine komplett andere Plattform, eine andere Form der Stapelung von Billigarbeitskräften, Fachkräften und KI, weil sich diese vielleicht als effizienter oder präziser erweist.

Noch ist unklar, ob sich die eher KI-intensiven Ansätze mit vielen gestaffelten Iterationsstufen bzw. Verarbeitungsschichten oder die auf massenhafte billige Arbeitskräfte in paralleler und redundanter „Schaltung" setzenden Plattformen durchsetzen werden.

Klar ist jedoch, dass der größere Grad an Training, den die Crowdarbeiter_innen in der Produktion von Trainingsdaten selbst durchlaufen müssen, zwar einerseits dazu führt, dass die individuellen Arbeitskräfte für die Plattformen ein bisschen mehr wert sind und besser behandelt werden, weil man sie gerne halten möchte. Da dieses Training der Arbeiterschaft aber für Aufgaben durchgeführt wird, die nach kurzer Zeit immer wieder automatisiert werden, hat diese Form der „skilled labour" eine sehr kurze Halbwertszeit und ist nur für die jeweilige Plattform, auf der man gerade arbeitet, und deren maßgeschneiderte Spezialwerkzeuge nützlich.

Die Crowdworker_innen sind also durch das Training stärker an die Plattform gebunden, als die Plattformen an die Crowdarbeiter_innen gebunden sind, weil sich das Training der Menschen wie gesagt besser automatisieren lässt als das Training der Maschinen.

Das Problem der Bindung an die Plattform wird durch die Rolle von Gamification-Elementen wie Erfahrungspunkten, Rängen, Expertentiteln für die Crowdworker_innen natürlich noch einmal deutlich verstärkt und führt zu einem Lock-in-Effekt. Zwar können sie als Wanderarbeiter_innen global von Plattform zu Plattform ziehen, aber sie müssen dann immer neue Werkzeuge lernen und sich immer neu ihre Reputation aufbauen.

Schlussendlich – und das ist auch der Grund, warum man nicht wollte, dass die eigenen Kinder so arbeiten müssen – bleibt auf allen Ebenen der Arbeit im „Stack" die Abhängigkeit vom Algorithmus, der einem die Arbeit zuteilt, sie bewertet, die Entlohnung bestimmt oder die Aufgabe gleich selbst übernimmt. All dies sind natürlich immer auch menschliche Managemententscheidungen, aber aus der Perspektive der Crowdworker_innen kann man unter diesen Bedingungen im „Stack" nicht zwischen menschlichen und algorithmischen Entscheidungen unterscheiden, und man kann man keine Antwort erwarten auf die häufigste im Forum gestellte Frage: „Warum sehe ich so wenige Tasks?"

LITERATUR

Arnoud, Sacha (2018): The rise of ML in self-driving cars. Guest lecture for course 6.S094: Deep Learning for Self-Driving Cars. MIT Boston. Veröffentlicht am 16.02.2018. Online unter: https://youtu.be/LSX3qdy0dFg (Abruf am 07.12.2018).

Bardt, Hubertus (2017): Deutschland hält Führungsrolle bei Patenten für autonome Autos. IW-Kurzberichte 61.2017. Köln: Institut der deutschen Wirtschaft. Online unter: https://www.iwkoeln.de/fileadmin/publikationen/2017/356331/IW-Kurzbericht_55_2017_Patente_autonomeAutos.pdf (Abruf: 24.01.2019).

Benner, Christiane (Hg.) (2015): Crowdwork – zurück in die Zukunft? Perspektiven digitaler Arbeit. Frankfurt am Main: Bund-Verlag.

Bordoli, Robin (2018): Focused on the Future with a New Name. In: Figure Eight (ehemals Crowdflower), 27.02.2018. Online unter: https://www.figure-eight.com/focused-future-new-name/ (Abruf am 07.12.2018).

Charrington, Sam (2016): TWiML Talk #6 – Angie Hugeback – Generating Labeled Training Data for Your ML/AI Models. In: TWiML & AI, 29.09.2016. Online unter: https://twimlai.com/twiml-talk-6-angie-hugeback-generating-labeled-training-data-mlai-models/ (Abruf am 24.01.2019).

Charrington, Sam (2017): Training Data for Autonomous Vehicles with Daryn Nakhuda. In: TWiML & AI, 23.10.2017. Online unter: https://twimlai.com/twiml-talk-057-training-data-autonomous-vehicles-daryn-nakhuda/ (Abruf am 24.01.2019).

Chen, Philipp, et al. (2017): Smart Moves Required – Der Weg zu künstlicher Intelligenz im Mobilitätssektor. McKinsey Center for Future Mobility. Online unter: https://www.mckinsey.com/de/^/media/McKinsey/Locations/Europe%20and%20Middle%20East/Deutschland/News/Presse/2017/2017-09-14%20Kunstliche/2017_smart_moves_required_de.ashx (Abruf am 24.01.2019).

Cherry, Miriam A. (2016): Beyond Misclassification: The Digital Transformation of Work (18.02.2016). In: Comparative Labor Law & Policy Journal, Forthcoming; Saint Louis U. Legal Studies Research Paper No. 2016-2. Online unter: http://papers.ssrn.com/abstract=2734288 (Abruf am 07.12.2018).

Christensen, Ben (2015): Crowdsourced Data: When to Use Curated Crowds vs. Crowdsourcing. In: Appen Blog, 17.11.2015. Online unter: https://appen.com/curated-crowds-vs-crowdsourcing/ (Abruf am 07.12.2018).

Corporaal, Greetje F./Lehdonvirta, Vili (2017): Platform Sourcing – How Fortune 500 Firms are Adopting Online Freelancing Platforms. University of Oxford: Oxford Internet Institute (OII). Online unter: https://www.oii.ox.ac.uk/publications/platform-sourcing.pdf (Abruf am 24.01.2019).

Davies, Alex (2017): Nissan's Path to Self-Driving Cars? Humans in Call Centers. In: Wired, 05.01.2017. Online unter: https://www.wired.com/2017/01/nissans-self-driving-teleoperation/ (Abruf am 24.01.2019).

Davies, Alex (2018): The Unavoidable Folly of Making Humans Train Self-Driving Cars. In: Wired, 22.06.2018. Online unter: https://www.wired.com/story/uber-crash-arizona-human-train-self-driving-cars/ (Abruf am 24.01.2019).

DMV CA (2017): Autonomous Vehicle Disengagement Reports 2017. Online unter: https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/disengagement_report_2017 (Abruf am 24.01.2019).

DMV CA (2018): Autonomous Vehicle Permit Holders 2018. Online unter: https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/permit (Abruf am 24.01.2019).

DPMA (2018): Autonomes Fahren, Teil 3: Zahlen und Fakten. München: Deutsches Patent- und Markenamt. Online unter: https://www.dpma.de/dpma/veroeffentlichungen/hintergrund/autonomesfahren-technikteil1/autonomesfahren-zahlenteil3/index.html (Abruf am 07.12.2018).

Figure Eight (2018): What is Human-in-the-Loop Machine Learning? In: figure eight resources. Online unter: https://www.figure-eight.com/resources/human-in-the-loop/ (Abruf am 07.12.2018).

Gegenhuber, Thomas/Ellmer, Markus/Scheba, Claudia (2018): Partizipation von Crowdworkerinnen auf Crowdsourcing-Plattformen: Bestandsaufnahme und Ausblick. Study der Hans-Böckler-Stiftung, Nr. 391. Düsseldorf: Hans-Böckler-Stiftung (HBS Study 391). Online unter: https://www.boeckler.de/5248.htm?produkt=HBS-006913 (Abruf am 17.01.2019)

Graham, Mark/Hjorth, Isis/Lehdonvirta, Vili (2017): Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods. In: Transfer: European Review of Labour and Research 23, S. 135–162. DOI: https://doi.org/10.1177%2F1024258916687250.

Gray, Mary L./Suri, Siddharth (2017): The Humans Working Behind the AI Curtain. In: Harvard Business Review, 09.01.2017. Online unter: https://hbr.org/2017/01/the-humans-working-behind-the-ai-curtain (Abruf am 24.01.2019).

Irani, Lilly (2013): The cultural work of microwork. In: New Media & Society 17, S. 720–739. DOI: https://doi.org/10.1177%2F1461444813511926.

Irani, Lilly C./Silberman, M. Six (2013): Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In: CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, Frankreich – 27.04.–02.05.2013. New York: ACM, S. 611–620.

Jürgens, Kerstin/Hoffmann, Reiner/Schildmann, Christina (2017): Arbeit transformieren! Denkanstöße der Kommission „Arbeit der Zukunft". Forschung aus der Hans-Böckler-Stiftung, Arbeit, Beschäftigung, Bildung, Bd. 189. Bielefeld: transcript Verlag. Online unter: https://www.boeckler.de/6299.htm?produkt=HBS-006612 (Abruf am 24.01.2019).

Kahn, Jeremy (2018): To Get Ready for Robot Driving, Some Want to Reprogram Pedestrians. In: Bloomberg, 16.08.2018. Online unter: https://www.bloomberg.com/news/articles/2018-08-16/to-get-ready-for-robot-driving-some-want-to-reprogram-pedestrians (Abruf am 24.01.2019).

Kässer, Matthias, et al. (2018): Artificial intelligence as auto companies' new engine of value. McKinsey & Company, Januar 2018. Online unter: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/artificial-intelligence-as-auto-companies-new-engine-of-value (Abruf am 24.01.2019).

Kässer, Matthias/Müller, Thibaut/Tschiesner, Andreas (2017): Analyzing start-up and investment trends in the mobility ecosystem. McKinsey & Company, November 2017. Online unter: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/analyzing-start-up-and-investment-trends-in-the-mobility-ecosystem (Abruf am 24.01.2019).

Kässi, Otto/Lehdonvirta, Vili (2018): Online Labour Index: Measuring the Online Gig Economy for Policy and Research (21.08.2018). In: Technological Forecasting and Social Change, Forthcoming. Online unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3236285 (Abruf am 24.01.2019).

Katz, Miranda (2017): Amazon's Turker Crowd Has Had Enough. In: Wired, 23.08.2017. Online unter: https://www.wired.com/story/amazons-turker-crowd-has-had-enough/ (Abruf am 24.01.2019).

Kittur, Aniket, et al. (2013): The Future of Crowd Work. In: CSCW '13 Proceedings of the 2013 conference on Computer supported cooperative work, San Antonio, Texas, USA – 23.–27.02.2013. New York: ACM, S. 1301–1318.

Krafcik, John (2018): Waymo One: The next step on our self-driving journey. In: Medium, 05.12.2018. Online unter: https://medium.com/waymo/waymo-one-the-next-step-on-our-self-driving-journey-6d0c075b0e9b (Abruf am 07.12.2018).

Laing, Keith (2018): Few carmakers submit self-driving car safety reports. In: The Detroit News, 10.09.2018. Online unter: https://eu.detroitnews.com/story/business/autos/mobility/2018/09/10/few-carmakers-submit-self-driving-safety-assessments/1076691002/ (Abruf am 24.01.2019).

Lehdonvirta, Vili (2017): Where are online workers located? The international division of digital gig work. In: The iLabour Project, 11.07.2017. Online unter: http://ilabour.oii.ox.ac.uk/where-are-online-workers-located-the-international-division-of-digital-gig-work/ (Abruf am 24.01.2019).

Leimeister, Jan Marco, et al. (2016): Systematisierung und Analyse von Crowd-Sourcing-Anbietern und Crowd-Work-Projekten. Study der Hans-Böckler-Stiftung, Nr. 324. Düsseldorf: Hans-Böckler-Stiftung. Online unter: https://www.boeckler.de/6299.htm?produkt=HBS-006379 (Abruf am 17.01.2019).

Leimeister, Jan Marco/Durward, David/Zogaj, Shkodran (2016): Crowd Worker in Deutschland: eine empirische Studie zum Arbeitsumfeld auf externen Crowdsourcing-Plattformen. Study der Hans-Böckler-Stiftung, Nr. 323. Düsseldorf: Hans-Böckler-Stiftung. Online unter: https://www.boeckler.de/6299.htm?produkt=HBS-006421 (Abruf am 17.01.2019).

Levin, Sam/Wong, Julia Carrie (2018): Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian. In: The Guardian, 19.03.2018. Online unter: https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe (Abruf am 24.01.2019).

Lundgren, Carole (2017a): Embracing New Opportunities in the Automotive Industry: Appen Establishes Detroit Office. In: Appen Blog, 05.06.2017. Online unter: https://appen.com/new-opportunities-automotive-industry-appen-establishes-detroit-office/ (Abruf am 24.01.2019).

Lundgren, Carole (2017b): AI Requires a Human Touch: Improve Technology with Crowd Recruiting. In: Appen Blog, 31.08.2017. Online unter: https://appen.com/ai-requires-human-touch-appen-crowd-recruiting/ (Abruf am 24.01.2019).

Marshall, Aarian (2018): Why Self-Driving Cars Will Never Be „Ready". In: Wired, 08.04.2018. Online unter: https://www.wired.com/story/when-will-self-driving-cars-ready/ (Abruf am 24.01.2019).

Marvit, Moshe Z. (2014): How Crowdworkers Became the Ghosts in the Digital Machine. In: The Nation, 05.02.2014. Online unter: https://www.thenation.com/article/how-crowdworkers-became-ghosts-digital-machine/ (Abruf am 24.01.2019).

Redrup, Yolanda (2017): Appen to become global leader after $105 million Leapforce acquisition. In: Financial Review, 29.11.2017. Online unter: www.afr.com/technology/appen-to-become-global-leader-after-105-million-leapforce-acquisition-20171129-gzv2bd (Abruf am 24.01.2019).

Reese, Hope (2016): Is 'data labeling' the new blue-collar job of the AI era? In: TechRepublic, 10.03.2016. Online unter: https://www.techrepublic.com/article/is-data-labeling-the-new-blue-collar-job-of-the-ai-era/ (Abruf am 24.01.2019).

Rosenblat, Alex (2018): When Your Boss Is an Algorithm. In: The New York Times, 12.10.2018. Online unter: https://www.nytimes.com/2018/10/12/opinion/sunday/uber-driver-life.html (Abruf am 24.01.2019).

Ross, Philip E. (2017): 2017: The Year In Robocars. In: IEEE Spectrum: Technology, Engineering, and Science News, 20.12.2017. Online unter: https://spectrum.ieee.org/cars-that-think/transportation/self-driving/the-year-in-robocars (Abruf am 24.01.2019).

Rowley, Jason (2018): The well-funded startups driven to own the autonomous vehicle stack. In: TechCrunch, 18.04.2018. Online unter: https://techcrunch.com/2018/05/27/the-well-funded-startups-driven-to-own-the-autonomous-vehicle-stack/ (Abruf am 24.01.2019).

Salehi, Niloufar, et al. (2015): We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers. In: CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea – 18.–23.04.2015. New York: ACM, S. 1621–1630.

Schmidt, Florian A. (2013): For a Few Dollars More: Class Action Against Crowdsourcing. In: APRJA peer reviewed open access journal. Online unter: http://www.aprja.net/?p=836 (Abruf am 07.12.2018).

Schmidt, Florian A. (2015): The Good, the Bad and the Ugly. Warum Crowdsourcing eine Frage der Ethik ist. In: Benner, Christiane (Hrsg.): Crowdwork – zurück in die Zukunft? Perspektiven digitaler Arbeit. Frankfurt am Main: Bund-Verlag, S. 367–385.

Schmidt, Florian A. (2016): Arbeitsmärkte in der Plattformökonomie: zur Funktionsweise und den Herausforderungen von Crowdwork und Gigwork. Bonn: Friedrich-Ebert-Stiftung, Abteilung Wirtschafts- und Sozialpolitik. Online unter: http://library.fes.de/pdf-files/wiso/12826.pdf (Abruf am 24.01.2019).

Schmidt, Florian A. (2017a): Crowd Design: From Tools for Empowerment to Platform Capitalism. Basel: Birkhäuser.

Schmidt, Florian A. (2017b): „Gehirne in die Cloud laden? Absurd!" Interview mit Raul Rojas. In: t3n Magazin, 28.04.2017. Online unter: https://t3n.de/magazin/kuenstliche-intelligenz-gehirne-cloud-241810/ (Abruf am 24.01.2019).

Schmidt, Florian A./Kathmann, Ute (2017): Der Job als Gig – digital vermittelte Dienstleistungen in Berlin. Berlin: ArbeitGestalten Beratungsgesellschaft und Senatsverwaltung für Integration, Arbeit und Soziales. Online unter: https://www.arbeitgestaltengmbh.de/assets/projekte/Joboption-Berlin/Der-Job-als-Gig-Expertise-Digital-November-2017.pdf (Abruf am 24.01.2019).

Solon, Olivia (2018): The rise of 'pseudo-AI': how tech firms quietly use humans to do bots' work. In: The Guardian, 06.07.2018. Online unter: https://www.theguardian.com/technology/2018/jul/06/artificial-intelligence-ai-humans-bots-tech-companies (Abruf am 24.01.2019).

Somers, James (2017): Is AI Riding a One-Trick Pony? In: MIT Technology Review, 29.09.2017. Online unter: https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/ (Abruf am 24.01.2019).

Stewart, Jack (2017a): The Human Army Using Phones to Teach AI to Drive. In: Wired, 09.07.2017. Online unter: https://www.wired.com/story/mighty-ai-training-self-driving-cars/ (Abruf am 24.01.2019).

Taylor, Astra (2018): The Automation Charade. In: Logic Magazine, 02.10.2018. Online unter: https://logicmag.io/05-the-automation-charade/ (Abruf am 24.01.2019).

Venkatesh, Mothi (2017): What's the difference between Playment and Traditional Crowdsourcing? In: Playment Blog, veröffentlicht 01.12.2017, upgedated 24.09.2018. Online unter: https://blog.playment.io/whats-difference-playment-traditional-crowdsourcing/ (Abruf am 07.12.2018).

Seit 2017 gibt es einen sprunghaften Anstieg in der Nachfrage nach hochpräzisen Trainingsdaten für die Automobilindustrie. Ohne diese Daten ist das ehrgeizige Ziel des autonomen Fahrens nicht zu erreichen. Um den selbstlernenden Algorithmen die Steuerung der selbstlenkenden Fahrzeuge anvertrauen zu können, braucht es viel menschliche Handarbeit, die von Crowdworker_innen auf der ganzen Welt erledigt wird. Sie bringen den Maschinen das Hören und Sehen bei, während sich die Crowdsourcing-Industrie rasant verändert, um den neuen Qualitätsanforderungen gerecht zu werden.

WWW.BOECKLER.DE

ISBN 978-3-86593-330-0