

**Bislang ist wenig darüber bekannt, wie die sozialen Netzwerke sich selbst regulieren und Inhalte ihrer Nutzer zensieren. Diese Intransparenz ist beabsichtigt. Sana Ahmad forscht als Doktorandin an der Freien Universität Berlin über Content-Moderation für die bekannten Intermediären und Social-Media-Plattformen durch indische Drittfirmen.**

# „Content-Moderation muss transparenter werden!“

**Christina Heinen im Gespräch mit Sana Ahmad**

## **Wie sind Sie auf das Thema gekommen?**

Ich habe vor einigen Jahren für Greenpeace gearbeitet, in Bangalore, und dort die Social-Media-Auftritte betreut. Immer wieder wurde ich „moderiert“, mein Profil wurde also für eine gewisse Zeit gesperrt mit der Begründung, dass ich gegen die Community-Standards verstoßen hätte.

## **Wurde das näher verargumentiert?**

Nein. Meist kam nur ein Pop-up-Fenster mit der Nachricht, ich sei für eine Woche gesperrt, weil ich gegen die Gemeinschaftsstandards verstoßen hätte. Manchmal wurde immerhin der Post genannt, welcher gegen die Regeln verstieß. Heute bekommt man etwas mehr Information, aber auch nicht viel. Es ist etwas transparenter geworden, aber nicht transparent genug.

Das hat mich dazu gebracht, über Meinungsfreiheit im Social Web nachzudenken. Außerdem interessiert mich die Situation der Arbeiter, der Content-Moderatoren: Wer macht diesen Job, wie läuft das ab, wie sind die Arbeitsbedingungen? Ich habe zu dieser Zeit eine BBC-Dokumentation über Content-Moderatoren in Bangalore gesehen:

Gleich um die Ecke von meinem Büro bei Greenpeace haben sie zensiert, was ich gerade gepostet hatte. Und das alles wird kontrolliert von einem Plattformanbieter im Silicon Valley, am anderen Ende der Welt.

## **Man weiß nur wenig darüber, wie Plattformanbieter den Content ihrer User moderieren...**

Ja, insbesondere über Indien wurde bislang noch kaum etwas geschrieben; abgesehen von einigen Journalisten hat sich noch kaum jemand damit beschäftigt. Die Plattformanbieter wollen keine Informationen dazu herausgeben. Sie verpflichten die Arbeiter der Drittfirmen zur Geheimhaltung, diese dürfen manchmal überhaupt nicht, in einigen Fällen erst neun bis zwölf Monate nach dem Ende ihrer Tätigkeit darüber sprechen oder auf ihrem LinkedIn-Profil veröffentlichen, für wen sie gearbeitet haben.

## **Warum diese Geheimhaltung?**

Die Plattformanbieter selbst sagen, sie wollen die Content-Moderatoren schützen, ihre Identität.

**Das ist nachvollziehbar. Aber es wäre doch möglich, deutlich transparenter zu sein, was die Abläufe anbelangt.**

Sicher. Tarleton Gillespie, der für Microsoft forscht, vertritt die Ansicht, dass Content-Management und -Moderation die zentralen Aufgaben der sozialen Netzwerke sind. Der Content kommt von den Nutzern selbst, die Plattformen bieten dem Nutzer als ihre Leistung genau das: abgemilderten, maßgeschneiderten Content. Es ist also ihr Kerngeschäft, das eigentliche Produkt.

**Niemand möchte vollkommen ungefiltert sehen, was die Menschen so posten ...**

Eben. Deshalb moderieren soziale Netzwerke ihren Content. Der Siegeszug der sozialen Netzwerke begann um 2005 herum, und auch damals gab es meines Wissens schon Content-Moderation. Ich habe einen „New York Times“ Artikel von 2010 gefunden über Content-Moderation durch Billiglohnfirmen in den USA. 2010 wurde diese unbeliebte, schlecht angesehene Tätigkeit also auch schon ausgelagert. Was in den letzten Jahren zugenommen hat, ist die Verlagerung an Überseestandorte, auf die Philippinen, nach Indien. Indien hat mit dem massenhaften Aufkommen der Callcenter in den 1990er-Jahren eine Dienstleistungsökonomie entwickelt. Die indische Wirtschaft konnte viele Erfahrungen damit sammeln, ganze Geschäftsprozesse zu übernehmen, die Firmen im Globalen Norden dorthin auslagern. Die Arbeit im Callcenter und in der Content-Moderation weist viele Ähnlichkeiten auf.

**Inwiefern?**

Die Arbeit ist schlecht bezahlt, prekäre Arbeitsbedingungen, relativ gering qualifizierte Arbeiter, meist ohne Berufserfahrung. Eine Hilfsarbeit, mit sehr standardisierten Aufgaben. Man lernt fast nichts im Job, hat keine Aufstiegschancen.

**Aber man muss doch ständig wichtige Entscheidungen treffen, letztlich werden durch Content-Moderation die Grenzen der Meinungsfreiheit immer wieder neu gezogen. Das kommt mir nicht wie eine unqualifizierte Tätigkeit vor.**

Ich sage ja auch nicht, dass es so sein sollte. Aber das ist die Realität, so wird Content-Moderation durchgeführt. Die Entscheidungsspielräume sind sehr gering, es gibt für jeden Bereich – Kindesmissbrauch, Terrorismus, Hate Speech, Nacktheit etc. – eigene, sehr detaillierte Richtlinien und Regelwerke. Die Social-Media-Firmen senden ihre Policy an die indischen Firmen, und die schulen ihre Arbeiter entsprechend. Man arbeitet im Akkord, mit einer klaren Zielvorgabe, wie viele Posts, Videos, Bilder man innerhalb einer bestimmten Zeit abarbeiten muss. Da bleibt keine Zeit, über Entscheidungen nachzudenken oder gar mit Kollegen darüber zu sprechen. Die Inhalte kommen ohne Kontext, man hat 3 bis 5 Sekunden, um zu entscheiden, bei Livestreaming etwas länger. Jeder bearbeitet immer nur einen Bereich, nur Kindesmissbrauch, nur Terrorismus. Die Fluktuation in diesen Jobs ist hoch – im Schnitt bleiben die Arbeiter nur sechs Monate, maximal drei Jahre. Es sind Einsteigerjobs für Hochschulabsolventen, die meisten sind sehr jung, zwischen 20 und 30 Jahre alt, und sie hoffen, dass das ein Sprungbrett ist für die attraktiven Tätigkeiten bei den großen Plattformanbietern. Allerdings arbeiten sie nicht unmittelbar für bekannte Social-Media-Anbieter oder Intermediäre, sondern bei Drittfirmen, die deren Content moderieren.

**»Die Social-Media-Firmen senden ihre Policy an die indischen Firmen, und die schulen ihre Arbeiter entsprechend. Man arbeitet im Akkord, mit einer klaren Zielvorgabe, wie viele Posts, Videos, Bilder man innerhalb einer bestimmten Zeit abarbeiten muss.«**

**Wie erfolgt die Auswahl der Posts, Videos oder Livestreamings, die zu den Content-Moderatoren weitergeleitet werden?**

Es handelt sich um Inhalte, über die sich jemand beschwert hat oder die durch Algorithmen herausgefiltert wurden. Kindesmissbrauch z. B., danach suchen Algorithmen, aber dann lässt man meist doch noch Menschen draufgucken, denn zu viel zu blockieren bzw. zu zensieren, ist schlecht fürs Geschäft. Menschliche Urteilskraft ist immer noch wichtig, man kann diese Arbeit nicht vollständig automatisieren. Man muss kulturelle Wertmaßstäbe, Kontexte und Symboliken kennen und einordnen können, und man muss sich manchmal auch auf sein Gedächtnis verlassen. Eine Content-Moderatorin aus dem Bereich „Terrorismus“ erzählte mir, dass sie einen indischen Terroristen wiedererkannte, dessen Konterfei noch nicht in den Richtlinien enthalten war. Ausnahmsweise durfte sie sich etwas mehr Zeit nehmen, das zu bearbeiten. Sie durfte einen Vorschlag an die US-amerikanische Firma schicken, diesen Terroristen in die Richtlinien aufzunehmen – und so am Fortschreiben des Regelwerkes mitwirken. Das ist aber die Ausnahme, normalerweise geht es nur darum, die Regeln zu befolgen, und die sind sehr strikt, da gibt es kaum Ermessensspielräume. Die Regelwerke ändern sich ständig, alle drei bis vier Monate. So z. B., als das deutsche Netzwerkdurchsetzungsgesetz (NetzDG) geschaffen wurde.

**Indische Firmen setzen das deutsche NetzDG um und identifizieren z. B. volksverhetzende Inhalte, die dann schnell gelöscht werden müssen?**

Ja. Die Arbeiter bekommen ein entsprechendes Training, das sind seltene Momente, in denen sie etwas Interessantes, Neues im Job lernen. Das geschieht aber eher bei den großen indischen Drittfirmen.

**»Menschliche Urteils-kraft ist immer noch wichtig, man kann diese Arbeit nicht vollständig automatisieren.«**

**Ist Content-Moderation ein wachsender Dienstleistungssektor, weil die Plattformen immer zahlreicher und größer werden? Oder liegt es daran, dass der Druck auf Anbieter wächst, sich selbst zu regulieren?**

Beides spielt eine Rolle, aber vor allem liegt es an der international immer lauter werdenden Kritik an den sozialen Netzwerken. In zahlreichen Ländern wurden neue Gesetze geschaffen, die dem Recht auch im Internet zur Durchsetzung verhelfen sollen, z. B. das NetzDG in Deutschland oder in Indien der Draft „Intermediary Liability Regulation“. Sogar in den USA, wo bis jetzt der Telecommunications Act (Section 230) einen „sicheren Hafen“ („safe harbour“) für Social-Media-Firmen garantierte – sie konnten sich darauf verlassen, nicht für den eventuell rechtswidrigen Content ihrer Nutzer zur Rechenschaft gezogen zu werden –, finden jetzt Anhörungen vor dem Kongress statt, und die Social-Media-Anbieter müssen sich rechtfertigen bzw. es wird darüber nachgedacht, sie in irgendeiner Weise zur Verantwortung zu ziehen.

**Wie stark berücksichtigen die großen Plattformen lokale Gesetze in ihren Richtlinien und Gemeinschaftsstandards?**

Das ist sehr unterschiedlich. Maßgeblich für das grundsätzliche Regelwerk, für die Policy, sind US-amerikanische Wertmaßstäbe und kulturelle Gepflogenheiten, eventuell noch entsprechende Richtlinien der UN. Mitunter besteht gerade bei den großen Unternehmen eine krasse Unwissenheit über die lokalen Gegebenheiten, z. B. über das Kastenwesen in Indien und entsprechende Antidiskriminierungsgesetze. Andererseits gibt es eine mit Blick auf die Meinungsfreiheit erschreckend hohe Bereitschaft der Unternehmen, auch politischen Forderungen zu entsprechen, um nicht den Zugang zu einem Markt zu verlieren. Dazu wird meines Wissens auch Geoblocking eingesetzt – indische Nutzer können bestimmten Content nicht sehen, den die rechtsgerichtete indische Regierung verbannen möchte. Geblockt wird wohl anhand der IP-Adressen, die ja Auskunft über den Standort des Nutzers geben. Wie das im Einzelnen funktioniert, konnte ich bislang aber leider nicht herausfinden.

**Social-Media-Plattformen betonen sehr, dass sie nur Plattformen sind und entsprechend nicht verantwortlich für den Content, den ihre Nutzer veröffentlichen. Das würde sie grundsätzlich unterscheiden von anderen Medienunternehmen wie TV-Sendern oder Zeitungen. Was halten Sie von diesem Argument der Plattformen?**

Die Frage, was eine Plattform aus organisations- und arbeitssoziologischer Sicht ausmacht, ist sehr interessant und schwer zu beantworten. In den Anfangsjahren haben die Intermediären gebetsmühlenartig wiederholt, Technologieplattformen zu sein. Heute, wo der Druck steigt und Verantwortung in Meinungsbildungsprozessen nicht mehr so leicht zurückzuweisen ist, ist man da etwas zurückhaltender geworden. Mark Zuckerberg nannte Facebook in einer Kongress-Anhörung kürzlich eine Hybridplattform – was auch immer das sein soll. Nennen sie sich so gern Plattformen, weil man dann jede Verantwortung für Inhalte und Wirkungen zurückweisen kann? Wenn der gesamte Prozess der Content-Moderation mehr publik würde, wäre es schwieriger das Plattformargument aufrechtzuerhalten, denn dann würden die Parallelen zu anderen Medienunternehmen deutlicher. Dass sie in irgendeiner Form Verantwortung für die Inhalte auf ihrer Plattform tragen, wäre dann nicht mehr so leicht abzustreiten. Sie könnten für Entscheidungen, etwas zuzulassen oder zu blockieren, verantwortlich gemacht werden, müssten sich eventuell rechtfertigen oder könnten gar haftbar gemacht werden.

**Wie könnte man den Prozess der Content-Moderation verbessern? Brauchen wir mehr Gesetze, vielleicht gar staatliche Instanzen für die Regulierung der Plattformen?**

Vor allem müsste Content-Moderation sehr viel transparenter werden. Das ist die Grundvoraussetzung dafür, dass sich etwas ändern kann. Ich würde mir wünschen, dass sich die Situation der Arbeiter in Indien verbessert. Einer der Arbeiter sagte zu mir: „Wir versuchen, die ganze Welt zu beschützen, aber wir selbst müssen diese furchtbaren Dinge alle sehen.“ Die Content-Moderatoren müssten besser geschützt werden, arbeitsrechtlich und auch durch psychologische Unterstützung. Von noch mehr Eingriffen in die Meinungsfreiheit durch staatliche Instanzen halte ich nicht so viel. Das beschneidet meist diejenigen in ihrer Redefreiheit, die ohnehin keine laute Stimme haben.

»Von noch mehr Eingriffen in die Meinungsfreiheit durch staatliche Instanzen halte ich nicht so viel. Das beschneidet meist diejenigen in ihrer Redefreiheit, die ohnehin keine laute Stimme haben.«

© Privat

