rezensionen:kommunikation:medien

Rezensionen aus den Bereichen Kommunikation und Medien

Tarleton Gillespie: Custodians of the Internet

Redaktion · Donnerstag den 23. August 2018

Rezensiert von Stephan Mündges

Das freie Netz, das jeder und jedem die Möglichkeit gibt, sich frei und ohne Furcht vor Repression zu äußern, dieses Netz ist schon lange ein Mythos. In den vergangenen rund 20 Jahren haben sich digitale Strukturen entwickelt, die prägen wie und was das Netz heute ist. Plattformen betrieben von den großen Tech-Riesen aus dem Silicon Valley stellen die Infrastruktur bereit, über die Nutzer in aller Welt miteinander kommunizieren. Sie erfüllen unterschiedliche Funktionen: Facebook, LinkedIn, YouTube, Instagram, Twitter, TripAdvisor oder Yelp – alle haben einen anderen Fokus. So vielfältig Plattformen also sein können, es eint sie, dass sie Kommunikation ermöglichen und strukturieren. Sie kuratieren und empfehlen Inhalte und vor allem: sie moderieren. Das heißt sie entscheiden darüber, was auf den Plattformen gepostet werden darf, was gelöscht wird und welche Nutzer gesperrt werden.

Im Englischen wird diese Tätigkeit Content Moderation genannt. Entscheidungen, die Plattformen dabei treffen, führen immer wieder zu Kontroversen. In den USA wird momentan intensiv darüber debattiert, ob der Verschwörungstheoretiker Alex Jones und die verschiedenen Social Media-Präsenzen seiner Sendung “Info Wars“ gesperrt und gelöscht werden sollten. Facebook-Gründer Mark Zuckerberg sorgte vor kurzem für einigen Wirbel, als er in einem Interview sagte, Posts, in denen der Holocaust geleugnet wird, würden auf Facebook nicht gelöscht. Manche Nutzer wüssten es vielleicht einfach nicht besser.

In Deutschland gibt es seit Jahren Debatten um Hasskommentare vor allem auf Facebook. Auch die Frage, wie Plattformen mit Nacktheit umgehen, hat schon häufiger für Aufregung gesorgt. Fotos stillender Mütter wurden gelöscht, Aktbilder, die viele als Kunst einstufen, wurden von

Plattformen als Pornographie kategorisiert. Die öffentliche Diskussion um Content Moderation verebbt aber zumeist nach jeder Erregungswelle schnell wieder.

Tarleton Gillespie, Forscher im Social Media Collective von Microsoft Research und Professor an der Cornell University, hat nun ein grundlegendes Buch über Content Moderation vorgelegt. Darin geht er über die Betrachtung einzelner Skandale hinaus, erläutert Funktionsweise und Arbeitsabläufe von Content Moderation (soweit sie bekannt sind), analysiert die zahlreichen Probleme und Dilemmata, mit denen Plattformen bei der Content Moderation konfrontiert werden, und macht Vorschläge, wie das institutionelle Gefüge dieser Moderation verändert und demokratisiert werden könnte. Es ist ein differenzierendes, detailliertes Buch und der Autor widersteht erfolgreich der Versuchung, die großen Plattformen als Bösewichte zu portraitieren, die ständig falsch liegen.

Denn – und das ist eine von Gillespies grundlegenden Thesen – Content Moderation ist ein essentieller Teil dessen, was Plattformen ausmacht, und eine der wichtigsten Dienstleistungen, die die Plattformbetreiber für ihre Nutzer erbringen. Da ein großer Teil digitaler Kommunikation mittlerweile auf den Plattformen stattfindet, hat die Fähigkeit, nahezu ungehindert löschen und filtern zu können, die Betreiber der Plattformen aber auch zu mächtigen Gatekeepern gemacht. Mächtiger als es einzelne Medien in der analogen Welt jemals waren. “Platforms may not shape public dicourse by themselves, but they do shape the shape of public discourse. And they know it” (23). Deshalb ist eine genaue Betrachtung, wie die Firmen die Inhalte auf ihren Plattformen moderieren, nach welchen Grundsätzen und mit welchen Mitteln, so wichtig.

Zu Beginn seines Buches skizziert Gillespie den institutionellen Rahmen, in dem sich die Plattformen bei der Content Moderation bewegen. Insbesondere der Blick auf die Gesetzeslage ist fokussiert auf die USA – ein traditionell sehr liberales Land, wenn es um die freie Rede geht. Andere Herangehensweisen, wie sich auch Content Moderation gesetzlich regulieren ließe, kommen nur am Rande vor. Das in Deutschland viel diskutierte und kritisierte Netzwerkdurchsetzungsgesetz (NetzDG), das private Unternehmen für die Durchsetzung deutscher Gesetze auf den Plattformen in die Pflicht nimmt, taucht zum Beispiel überhaupt nicht auf.

In einem nächsten Schritt analysiert der Autor die Regeln, die sich Plattformen selbst auferlegen: Nutzungsbedingungen und Community-Richtlinien. Schon dabei wird ein zentraler Widerspruch der Content Moderation deutlich: die Angst, zu stark einzugreifen und freie Rede zu unterdrücken auf der einen, und die Sorge, durch zu wenig Moderation das Klima auf einer Plattform zu vergiften auf der anderen Seite.

Dieses Feld steckt Gillespie in den ersten drei Kapiteln sehr ausführlich ab. Für die Praxis relevanter wird das Buch in Kapitel vier. Darin beschäftigt sich der Autor mit den drei möglichen Vorgehensweisen bei der Content Moderation. Passenderweise ist das Kapitel mit “Three imperfect solutions“ überschrieben. Denn jede Option hat gewaltige Nachteile: Option 1 ist ‘editorial review‘, also eine redaktionelle Überprüfung der Inhalte vor der Veröffentlichung. Man könnte es auch Vorzensur nennen. Aufgrund der enormen Menge an zu überprüfenden Inhalten ist das für fast keine Plattform eine realistische Option. Lediglich Apple wendet das Verfahren bei seinem App-Store an.

Option 2 ist hingegen bei allen Plattformen im Einsatz: ‘community flagging‘. Nutzer einer Plattform melden Inhalte, die sie als anstößig oder den Community-Regeln widersprechend einstufen. Die gemeldeten Inhalte werden dann von menschlichen Angestellten, den Content

Moderatoren, überprüft. Millionenfach müssen die Entscheidungen treffen, unter Zeitdruck und häufig ohne kulturelle Hintergründe oder den Kontext von Posts zu kennen. Dass dabei immer wieder auch falsche Entscheidungen getroffen werden (von denen die krassesten auch medial thematisiert werden), arbeitet Gillespie überzeugend heraus.

Option 3 ist ‘automatic detection‘, also die automatisierte Überprüfung aller hochgeladenen Inhalte durch Software. Ein Verfahren, dass YouTube und Facebook bereits intensiv anwenden. Die Entscheidung, ob etwas gelöscht wird, liegt bei diesen Systemen bislang aber immer noch bei Menschen. Noch gibt es keine intelligenten automatisierten Systeme, die z. B. Kontexte von Veröffentlichungen bewerten können (ist z. B. ein Bild mit viel nackter Haut durch die Kunstfreiheit gedeckt? Wird ein brutales Kriegsbild in einem journalistischen Zusammenhang benutzt, oder um Voyeurismus zu bedienen?). Content Moderation funktioniert ohne Menschen bislang einfach nicht.

Die Frage, welche Gruppen eigentlich daran beteiligt sind und wer die oftmals anstrengende Arbeit der Content Moderation übernimmt, stellt Gillespie im anschließenden Kapitel. Seine Argumentation, dass die Regeln, nach denen Content Moderation erfolgt, von einer sehr homogenen Gruppe junger privilegierter Menschen im Silicon Valley erlassen werden, ist überzeugend. Leider ist der Teil, in dem er sich mit der großen Armee der Content Moderatoren beschäftigt, die über die Einhaltung der Regeln wachen, zu oberflächlich und bietet wenig Neues. Dass zigtausende Clickworker, die zu einem großen Teil auf den Philippinen und in Indien sitzen, den Dreck von den Plattformen kärchern müssen, ist bekannt. Zuletzt hat der Dokumentarfilm *The Cleaners* das Leben und zum Teil auch Leiden der Content Moderatoren eindrücklich illustriert.

Was allerdings nach wie vor unbekannt ist und was auch Gillespies Buch nicht bieten kann, ist eine Quantifizierung der Content Moderation-Branche. Wer sind die führenden Unternehmen, an die Plattformbetreiber diese Aufgaben outsourcen? Wie viele Menschen arbeiten in der Branche? Zu welchen Löhnen? Content Moderation ist in weiten Teilen eine Blackbox, in die es nur wenige Einblicke gibt. Insbesondere wenn es um solche strukturellen Fragen geht. Zwar hat Gillespie, wie er schreibt, mit zahlreichen Angestellten der Plattformen, die in führenden Positionen im Bereich Content Moderation arbeiten, gesprochen. Antworten auf die oben genannten Fragen kann er allerdings nicht geben.

Spannend wird es im letzten Kapitel des Buches. Darin entwirft der Autor eine Utopie, wie es besser werden könnte und wie die Probleme der Content Moderation zumindest zum Teil behoben werden könnten. Kern seines Vorschlags ist die Institutionalisierung demokratischer Prozesse, wenn es um Content Moderation geht: “their shared responsibility for the public requires that they share that responsibility with the public – not just the labor, but the judgement“ (209). Eine Möglichkeit das umzusetzen: Nutzer stimmen über Regeln ab und haben ein Mitspracherecht, wenn es um zentrale Fragen der Content Moderation geht. Eine solche Utopie lädt natürlich sofort zu Kritik ein: Ist so etwas überhaupt umsetzbar? Diese Frage zu stellen ist sicherlich legitim. Gillespies Verdienst ist aber, überhaupt einmal Alternativen zum Status Quo anzubieten. Denn ständiges Lamentieren über die Fehler der Plattformen wird die Probleme, die es zweifelsohne gibt, nicht lösen. *Custodians of the Internet* ist ein gut fundierter erster Schritt auf dem Weg, Content Moderation im Netz neu zu erfinden.

Eine ungekürzte Fassung des Schlusskapitels findet sich unter http://www.custodiansoftheinternet.org/application/files/5815/3185/4676/Gillespie_-_improving_moderation_-_extended_CH8.pdf

Links:

- Verlagsinformationen zum Buch
- Webpräsenz von Prof. Tarleton Gillespie an der Cornell University
- Webpräsenz von Stephan Mündges an der Technischen Universität Dortmund

Dieser Beitrag wurde publiziert am Donnerstag den 23. August 2018 um 08:00
in der Kategorie: Einzelrezension, Rezensionen.
Kommentare können über den Kommentar (RSS) Feed verfolgt werden.
Kommentare und Pings sind momentan geschlossen.