

Digitale Nachhaltigkeit Zusatzserie

Artikel 02 (Facebook knows Instagram is toxic for Teen Girls, Company documents show)

1.) Kurze Zusammenfassung des Artikels

Der Artikel dreht sich um verschiedene Probleme von Instagram und um Facebooks (heute Meta) Bewusstsein der Probleme, welches durch Firmeninterne Dokumente welche an die Öffentlichkeit gelangten, ans Licht kamen. Im Artikel werden mehrere Beispiele von Personen genannt, welche verschiedenste Probleme nach der Nutzung von Instagram haben. Einige Berichten von Essstörungen, andere von Selbstmordgedanken, wiederum andere berichten von unter Depressionen zu leiden. Alle führen zum selben gemeinsamen Nenner, nämlich dem falschen Körper-/Schönheitsideal welches von Influencern auf der Plattform vermittelt wird. Facebook hat seinerseits interne Nachforschungen zur Nutzung, Verhalten und Auswirkungen von Instagram auf Jugendliche gemacht. Diese internen Ergebnisse werde ich später erläutern, wichtig zu wissen ist jedoch noch dass Facebook bis die internen Dokumente ans Licht kamen, stets in der Öffentlichkeit verneinte, dass Instagram schädliche Einflüsse auf die Jugendlichen haben könnte und behaupteten sogar dass ihre eigenen Studien nur positive Einflüsse und Vorteile für Jugendliche . In der internen Untersuchung welche von Mitarbeitenden aus den Abteilungen Data Science, Marketing und Product Developement geleitet wurde, kam Facebook selbst zum Schluss dass die Anschuldigungen wahr sind, nur leugneten sie dies weiterhin in der Öffentlichkeit. Im Dokument erwähnten sie unter anderem dass Instagram selbst das Problem sei und nicht die sozialen Medien im allgemeinen Sinne und das Problem in der Kernidee von Instagram stecke. Damit meinten sie die Tendenz Nur die allerbesten Momente zu poste, damit möglichst viele Likes erzielt werden könnten, was wiederum mehr Druck auf andere auslöst, perfekt auszusehen. Weitere Untersuchungen ergaben, dass 32% von Jugendlichen Mädchen welche sich generell schon etwas schlechter fühlten, sich aufgrund von Instagram zusätzlich nochmals schlechter fühlten. In einer Slide einer Powerpoint welche Mark Zuckerberg vorgeführt wurde, steht: « We make body image issues worse for one in three teenage girls» also dass sie für jedes dritte jugendliche Mädchen das Bild ihres eigenen Körpers verschlechterten. Auf einer weiteren stand dass sie Beweise dazu gefunden haben, dass Instagram selbst einen Anteil von Selbstmordgedanken der Mädchen stärkt. Ein weiteres Problem welches im Artikel erwähnt wird, ist dass die Explorepage von Instagram den Betroffenen aufgrund des Algorithmus nur noch mehr ungesunden Content anzeigt. Aus den Hauseigenen Untersuchungen kam auch raus, dass nicht nur Mädchen von den erwähnten Problemen betroffen sind sondern, dass auch Jungen diesen Problemen gegenüberstehen.

2.) Welche Problematik sehen sie darin?

Ich selbst finde es problematisch dass Facebook diese Informationen nicht mit der Öffentlichkeit geteilt hat. Natürlich sind es interne Forschungen und Daten welche sie nicht direkt mitteilen möchten und sich auch nicht positiv auf ihren Ruf auswirken, jedoch finde ich es trotzdem schlimm, dass sie solche Ergebnisse nicht teilen. Natürlich herrschen solche Probleme auch in anderen Sozialen Medien, hätte Facebook die Ergebnisse geteilt hätten andere Entwickler aus meiner Sicht vielleicht Rücksicht darauf genommen und so ihre Algorithmen oder ihre Nutzungsbedingungen angepasst, sodass solche Probleme gar hätten verhindert werden können. Auch dass nachdem diese Informationen Publik wurden, wurde keine grosse Veränderung vorgenommen und das Thema ist noch heute stark in den Medien vertreten

Artikel 10 (Facebook says AI will clean up the platform. Its own Engineers have doubts)**1.) Kurze Zusammenfassung des Artikels**

Das Thema dieses Artikels handelt um die AI von Facebook welche dazu genutzt werden soll um Hassrede, Gewalt und der Zugang von Minderjährigen auf dieser Plattform zu minimieren. Das derzeitige Problem (Stand Veröffentlichung des Artikels) liegt darin, dass der Algorithmus zwar einige Posts als unsicher einstufen kann und diese danach entfernt, jedoch bleibt der Account mit welchem der Post veröffentlicht wurde bleibt unbestraft. 2019 begann Facebook damit Menschen welche sich anstössige Posts ansehen musste, durch AI zu ersetzen. Dieses vorgehen erfordert natürlich verstärkt Forschung im Bereich AI was zusätzliche Mittel fordert. Ein Entwickler der AI sagte dazu, dass es nie ein Modell geben wird welches alle Verstösse aufdeckt, das schwierigste dabei wäre laut seiner Aussage sensible Themen. Er schätzt, dass nur 2% der von der AI angeschauten Hassrede wirklich entfernt wird und ohne eine grosse Innovation wird es schwierig diesen Prozentsatz gar auf 10-20% zu erhöhen. Die auswärtigen Statistiken widersprechen der Aussage von Mark Zuckerberg, dass die AI bis ende 2019 den grössten Teil von anstössigen Content erkennen und entfernen wird. Der Head of Integrity von Facebook sagte in einem Interview, dass nicht die Häufigkeit gefundener anstössiger Posts sondern dieser welche nicht gefunden wurden, die wichtigste Kennzahl zur Messung wie effizient die AI ist. Es gibt Beispiele in welcher die AI, Videos in welchen Personen erschossen wurden oder Gewalt an transgender Kindern angedroht wurde, nicht erkannt hatte und demnach auch nicht entfernt wurde. Es wurde gar eine Massenschiesserei Live übertragen und nicht von der AI entdeckt. Zu den Merkmalen welche die AI verwendet um Posts als unangebracht zu identifizieren, sagen die Entwickler nur dass sie zu naiv sind und Posts eher eingrenzen können als eindeutig zu identifizieren.

Der Sprecher von Facebook gab bekannt dass 75% der Ausgaben zur Erkennung von Hassrede, zur Einstellung von neuem Personal genutzt wird, welches sie Posts manuell überprüfen muss. Facebook selbst teilt nicht den genauen Prozentsatz von Hassrede welcher durch die AI selbst entfernt wird, oder wie viel durch manuelle Überprüfungen entfernt wurden. Ein weiteres schwieriges Problem der AI seien Hahnenkämpfe, bei denen könne die AI nicht korrekt zuordnen ob dabei wirklich Hühner verletzt werden oder ob nichts darauf zu sehen ist. Zum Teil werden diese so falsch eingeschätzt, dass sie als Autounfälle gekennzeichnet werden von der AI. Ein weiteres Problem sind Fremdsprachen mit verschiedenen und seltenen Dialekten, diese können auch nicht richtig erkannt werden von der AI.

2.) Welche Problematik sehen sie darin?

Meiner Meinung nach ist es schwierig eine «korrekte» AI zu entwickeln, welche alle anstössigen Posts fachgerecht erkennen kann. Schon nur das angesprochene Problem mit den Dialekten sehe ich als ein eher unüberwindbares. Dazu kommt, dass in gewissen Kreisen manchmal auch «harmlose» Worte verwendet werden um einer oder mehreren Personen übel nachzureden. Diese Posts zu filtern und zu erkennen wird aus meiner Sicht nie möglich sein. Mit der heutigen Technologie sollte es aus meiner Sicht jedoch möglich sein, dass Videos von Schiesserein oder anderen Gewalttaten bereits beim Upload herausgefiltert werden und gar nicht erst veröffentlicht werden können. Facebook sollte mehr vermehrt auf Machine Learning setzen und mit den bereits entfernten Inhalte ihre AI versorgen, sodass zumindest ähnliche Posts nicht mehr gepostet werden können und besser erkannt werden.

Artikel 13 (Facebook services are used to spread religious hatred in India, internal documents show)**1.) Kurze Zusammenfassung des Artikels**

In diesem Artikel dreht sich alles um Aufstände und Hassrede in Indien. Aufgrund von Gerüchten und Aufrufen zur Gewalt, welche sich über Whatsapp verbreiteten, kam es sogar zu tödlichen Aufständen in Indien. Indien ist ein wichtiger Markt für Facebook da es mit den mehreren hundert Millionen Nutzern von Facebook und Whatsapp der Grösste Markt des Unternehmens ist. Es kam immer wieder zu Konflikten zwischen Hinduisten und Muslimen aufgrund von Gewalt und Hassrede über Whatsapp. Unter anderem wurden Muslime beschuldigt sie wären der Hauptgrund der Ausbreitung von Covid-19 oder auch, dass sie hinduistische Frauen heiraten um eine Art der Übernahme der Muslimen zu bewerkstelligen. Es gibt viele Facebook Gruppen gleichgesinnter und diese Gruppen werden dann mit aufhetzenden Nachrichten und Hassrede gefüllt. Dazu hat Facebook Nachforschungen veranlasst und dutzende Nutzer interviewt. Das Ergebnis war dass viele Hassreden im Umlauf sind, dass Hindus in Gefahr seien und die Muslime getötet werden sollten. Ein Muslim berichtet aufgrund dieser Hassrede nun Angst um sein Leben zu haben. Die Leute in Indien sind davon überzeugt, dass die Lösung dieses Problems die Aufgabe von Facebook sei, und sie Facebook und Whatsapp davon befreien müssen. Mehrere Nationalistische Hindugruppen posten anstössige Nachrichten mit anti-muslimischem Inhalt auf den verschiedenen Plattformen. Forscher fordern, dass diese von den Plattformen entfernt werden, aufgrund Verstösse gegen die Nutzungsbedingungen. Diese Accounts sind jedoch immer noch aktiv. Berichten zufolge weiss Facebook klar von den aufhetzenden Nachrichten bescheid, Nutzer berichten jedoch dass Facebook dagegen nichts unternimmt. Von anderen Gruppen werden Muslime als Schweine und Hunde bezeichnet, doch auch diese Gruppe bleibt aktiv auf Facebook. Facebook entfernt die besagten Gruppen aufgrund politischer Sensitivität nicht. Der Sprecher von Facebook sagte dazu, dass solche Gruppen erst nach gründlicher Untersuchungen und ohne politische Empfehlungen deaktiviert werden können. Ausserdem erwähnt er dass Hassreden gegen Muslime immer mehr der Fall sei und sie Vorkehrungen treffen um dies zu unterbinden. Indiens Regierung droht derweil Mitarbeitern von Facebook, Whatsapp und Twitter mit Gefängnisstrafen wenn diese nicht auf ihre Forderungen eingehen und den genannten Content entfernen. RSS eine politisch rechts orientierte Gruppe parlamentarisch freiwillige Organisation wird auch nicht von Facebook entfernt. Ein Sprecher der RSS sagt dazu, dass sie keine Facebook Services nutzen und sie offen wären und Facebook jederzeit auf sie zukommen könnte falls es Verstösse gäbe.

2.) Welche Problematik sehen sie darin?

Das Problem hierbei ist das Recht auf eine freie Meinungsäusserung. Facebook selbst ist keine Plattform auf der freie Meinungsäusserung herrscht. Es gibt klare Nutzungsbedingungen und darin wird auch erklärt, dass wer sich nicht daran hält davon ausgeschlossen werden kann. Dadurch erstaunt es mich um so mehr dass Facebook ein solches Verhalten weiterhin toleriert und nicht unterbindet. Ein weiteres Problem erkenne ich darin, dass Indien ein ziemlich grosser Markt ist und mithilfe solcher Gruppen und Posts eine Vielzahl von Menschen erreicht wird. Solche Hassreden und Verbreitung von Falschinformationen sollte sofort unterbunden werden von Seiten Facebooks. Das Argument von politischer Sensitivität finde ich nichtig, da sie bereits andere politische Personen von der Nutzung ihrer Dienste ausgeschlossen haben, als Beispiel den Ex-Präsidenten der USA Donald Trump. Es sollte also durchaus für Facebook auch möglich sein solche Gruppen von der Nutzung ihrer Dienste auszuschliessen.

