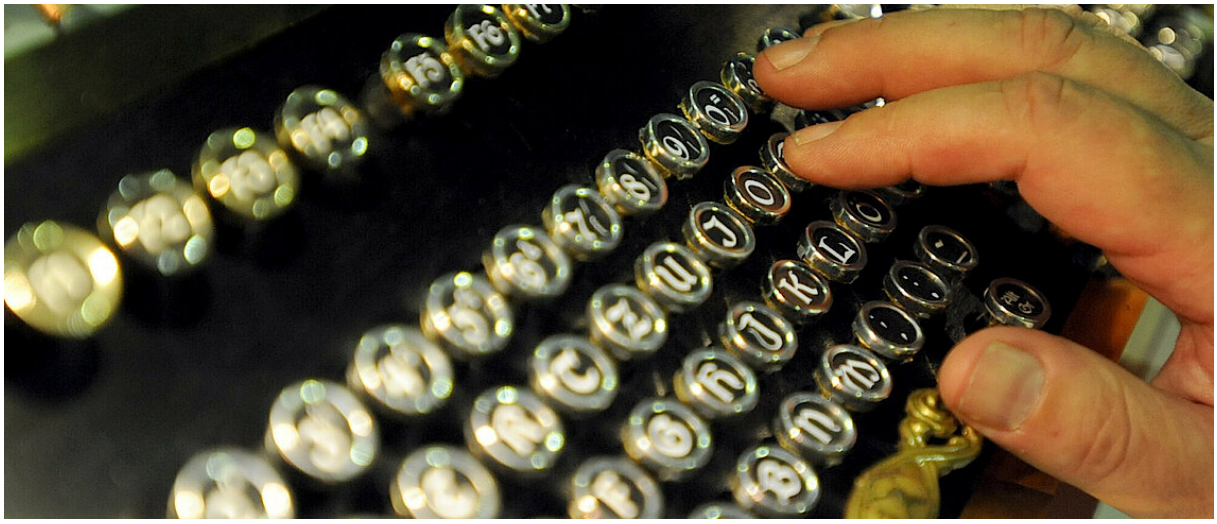
IPG
Internationale Politik
und Gesellschaft

Technik, die begeistert

Von Christoph P. Mohr | 08.22.2016

Wie Algorithmen dschihadistischer Propaganda Einhalt gebieten können.



Das Löschen extremistischer Inhalte könnte in Zukunft vollautomatisch erfolgen, sobald sie als solche eingestuft wurden.

Junge Männer sitzen im Grünen und unterhalten sich darüber, wie schön ihr Leben doch sei. Sie schwärmen vom „Urlaub“, in dem sie sich gerade befänden. Es scheint die Sonne, die Landschaft ist malerisch. Die Idylle wird nur durch ihre kugelsicheren Westen und schwere Bewaffnung getrübt. Dieses Beispiel eines professionell produzierten Rekrutierungs- und Propagandavideos des „Islamischen Staates“ (IS) ist bis heute im Internet zu finden. Es zeigt den Österreicher „Mohamed Mahmoud“ und den Deutschen „Abu Omar al-Almani“ wie sie über ihr Leben im Kalifat sprechen und Drohungen aussprechen – auch gegen Deutschland. Am Ende töten die beiden Soldaten der syrischen Armee.

Der IS produziert solche Botschaften am Fließband und verbreitet sie über soziale Medien wie Twitter, Youtube und Facebook. Die große Reichweite des Internets hat dafür gesorgt, dass solch verheerende Botschaften es bis in die Wohnzimmer deutscher Bundesbürger geschafft haben.

Der einfache Zugang zu solchen Videos kann dazu beitragen, dass das entsteht, was in den

Medien unter dem Begriff Selbstradikalisierung diskutiert wird. Selbstradikalisierung bedeutet hier vor allem, dass junge Menschen, ohne vorher als besonders religiös aufgefallen zu sein, sich dem Radikalismus oder Extremismus zuwenden und zum Beispiel Abu Bakr al-Baghdadi, dem selbsternannten Kalifen und Führer des IS, ihre Treue schwören. Die Islamisierung im Zuge einer Selbstradikalisierung ist allerdings nicht selten der Endpunkt ihrer persönlichen Entwicklung und nicht der Ausgangspunkt. Zuvor haben sie häufig Erfahrungen mit Ausgrenzung, Gewalt, Diskriminierung und Haft gesammelt. Diese Erlebnisse und die damit verbundene inhärente Ablehnung ihrer eigenen Lebensrealität werden exponentiell verstärkt durch online abrufbare extremistische Botschaften – verpackt in Videos, Audio-Dateien oder Texten, irgendwo zwischen Popkultur und Propaganda. „Internethetze“ trägt also dazu bei, dass empfängliche junge Menschen Taten im Namen einer Organisation begehen, mit der sie häufig nie in direktem Kontakt standen.

> Propagandavideos sind mächtige Werkzeuge des IS und müssen als solche bekämpft werden.

Propagandavideos sind mächtige Werkzeuge des IS und müssen als solche bekämpft werden. Ein Ansatzpunkt ist die starke Beschränkung derer Zugänglichkeit und Verbreitung unter Zuhilfenahme modernster technischer Mittel. Wie schnell und wie weit sich die diese Inhalte im Internet verbreiten, liegt nicht nur an der schieren Größe des Internets, sondern vor allem an der sehr langsamen Reaktion der jeweiligen Seitenbetreiber und an der Schwierigkeit, die Materie zu bewerten.

Dabei treten vor allem zwei Probleme auf: Zum einen hinken die technischen Möglichkeiten und Kapazitäten der Seitenbetreiber der Verbreitungsgeschwindigkeit im Internet hinterher. Nutzer müssen ein Video dem Betreiber melden. Dieser prüft den Inhalt, schätzt diesen ein, löscht ihn gegebenenfalls oder sperrt den Benutzer-Account. Das bedeutet, dass sich entsprechende Inhalte bereits vielfach multipliziert haben können bevor sie gelöscht werden. Entsprechende Videos können also mit einfachsten technischen Mitteln gespeichert, kopiert und – ggf. mit einer anderen Beschreibung – neu hochladen werden. Für den Seitenbetreiber beginnt mit erneuter Meldung des gleichen Inhalts der Prüfungsprozess von neuem. Dies mündet in einer Spirale der Redundanzen.

Das zweite Problem liegt darin, dass nicht klar definiert, abgrenzbar oder eindeutig ist, was terroristischer Inhalt ist. Es gibt also keine gängige Lesart, was Terrorismus, was Propaganda oder was schlicht eine Meinungsäußerung ist.

> Die „PhotoDNA“-Technologie ermöglicht es, einmal überprüfte Gewaltvideos zu erkennen und von einer Verbreitung im Internet auszuschließen.

Das erste der beiden Probleme lässt sich beheben, denn die technischen Möglichkeiten zur Eingrenzung der Reichweite von Inhalten haben sich verbessert. Anstatt durch die

mehrfache manuelle Überprüfung von Videomaterial kostbare Zeit zu verlieren, setzt der „PhotoDNA“-Ansatz auf eine Technologie, die es ermöglicht, einmal überprüfte Medien anhand spezifischer Merkmale zu erkennen und von einer Verbreitung im Internet auszuschließen. Diese Merkmale sind in jedem Bild, Video- oder Tonmaterial zu finden und entsprechen einer eindeutigen Signatur. „PhotoDNA“ lässt sich auffinden, selbst wenn ein Video verändert, manipuliert oder geschnitten wurde. Ist ein Video von einer Plattform gelöscht worden, kann ein erneutes Hochladen dadurch verhindert werden, dass die individuelle Signatur des Videos und nicht das Video an sich auf der jeweiligen Plattform gesperrt wird.

Mithilfe eines Algorithmus soll dies auf allen gängigen Plattformen möglich sein und so die Verbreitung von Inhalten stark einschränken. Diese Technologie wurde maßgeblich von Hany Farid entwickelt. Farid arbeitet für ein US-amerikanisches gemeinnütziges Counterterrorism-Projekt und ist Professor für Informatik am Dartmouth College. Bereits seit sechs Jahren wird seine Technik erfolgreich genutzt, um die Verbreitung von Kinderpornographie einzudämmen. Das Löschen oder Sperren von extremistischen Inhalten würde in Zukunft vollautomatisch erfolgen, sobald sie als solche eingestuft worden sind. Nötig hierfür wäre, diesen Algorithmus in alle gängigen Plattformen zu integrieren und entsprechende Videos früh zu erkennen.

Durch die Gründung von zivilgesellschaftlichen und staatlichen Institutionen zur Bestimmung von extremistischen Inhalten können entsprechende Inhalte schneller identifiziert und gelöscht werden. Analysten sichten Videos, ob diese Hass- oder Propagandabotschaften enthalten. Auch das Nutzen von „Crowdsourcing“-Kampagnen, also Kampagnen, an denen sich prinzipiell jeder beteiligen kann, wie beispielsweise das Hashtag „#CEPDigitalDisruption“ auf Twitter, können dazu beitragen, dass Videos schneller markiert und identifiziert werden. Die Reaktionsfähigkeit der Plattformen würde durch die Kombination des Algorithmus, der für eine nachhaltige Sperrung sorgt, mit der frühzeitigen Bewertung von Inhalten durch weitere Akteure, enorm gesteigert werden und so den Zugang zu extremistischen Inhalten stark einschränken.

> Möglichkeiten zur „digital disruption“ können einen neuen Standard für den Umgang mit propagandistischen Inhalten im Internet setzen.

Diese Möglichkeiten zur „digital disruption“ könnten so einen neuen Standard für den Umgang mit extremistischen, gewaltdarstellenden oder propagandistischen Inhalten im Internet setzen, der ganz im Sinne einer neuen „Internet Governance“, auf gemeinsamen Grundsätzen, Normen und Regeln aufbaut. Diese Vorgehensweise muss auf einer Übereinkunft aller Akteure beruhen, dass wir als Gesellschaft(en) Videos ablehnen, die zu Gewalt gegen Unschuldige, Andersgläubige oder Minderheiten aufrufen und schlimmste Gewalttaten zur Unterstützung einer inhumanen Ideologie zeigen. Eine „rote Linie“ muss dort gezogen werden, wo Videos zu Hass, Gewalt und Mord aufrufen. Regierungen, dem Privatsektor und der Zivilgesellschaft muss bewusst sein, dass sich die zugrundeliegenden

Definitionen von Terrorismus, Propaganda oder Extremismus – je nach Akteur –zwar teils fundamental unterscheiden können, es aber Inhalte gibt, die von einem sehr großen Teil der Menschheit als nicht verbreitungswürdig angesehen werden. Ein Beispiel hierfür ist Kinderpornographie.

Die latente Gefahr einer weitergehenden Zensur bei der Umsetzung solcher Maßnahmen zur Eindämmung spezifischer Inhalte kollidiert mit der Tatsache, dass sich die Probleme der Verbreitung von extremistischen Inhalten nicht länger aufschieben lassen.

Die größte Chance besteht darin, dass die internationale Gemeinschaft der erwähnten „roten Linie“ dahingehend folgt, dass einschlägiges Material mit Gewaltdarstellungen oder direktem Aufruf zu Gewalt dauerhaft – im Sinne der Nutzerbedingungen – gesperrt werden würde, ohne gleichzeitig eine massivere Einschränkung von Inhalten zu forcieren. So könnte zumindest zu einem Teil die Anziehungskraft von Gewaltinhalten und vermeintlichen Lebens- oder Siegesrealitäten im IS verhindert werden. Hierfür müssten allerdings die großen Internetanbieter mit Regierungen und zivilgesellschaftlichen Vertretern endlich für Lösungen sorgen und die entsprechenden internationalen Foren besser nutzen; denn bereits seit geraumer Zeit ist die fehlende Kooperation eines der Kernprobleme des Internets.

> Die Suche nach islamistischer Propaganda darf nicht einfacher sein als die Suche nach einem nicht durch die GEMA gesperrten Video.

Technische Neuerungen, wie der Einsatz von Algorithmen, sind – sofern mit Vorsicht eingesetzt – ein gutes Mittel, um Regeln und Normen im Internet durchzusetzen. Getreu dem Motto: Wir können zwar nicht kontrollieren, was ihr hochladet, aber sehr wohl, was sich wie verbreitet. Selbstradikalisierung und der Konsum einschlägiger Medien würde dadurch zwar nicht gänzlich gestoppt, aber deutlich reduziert. Die Suche nach islamistischer Propaganda, Hetze und Gewaltdarstellungen darf nicht einfacher sein als die Suche nach einem nicht durch die GEMA gesperrten Video des südkoreanischen Musikers PSY.