

Antworten Google

Hassrede

- **Frage:** Bezüglich Standard-Entwicklung: Google gibt an, dass verschiedene Stakeholder in die Entwicklung von Standards einbezogen werden. Wie und nach welchen Kriterien werden diese Stakeholder ausgesucht? Besteht Transparenz darüber wer wie einbezogen wurde?
- **Antwort:** Die benötigte Expertise in der Standardentwicklung ist jeweils sehr unterschiedlich. Aus diesem Grund gibt es auch keine vordefinierte Liste von möglichen Kontaktpunkten. Die zuständigen internen Teams bedienen sich der Expertise von relevanten Trusted Flaggers, aber auch der Wissenschaft und anderen Gruppen. Es ist unmöglich mit jeder Institution in jedem Land zu sprechen, aber Google macht ein gewisses Mapping von besonders wichtigen Institutionen. Dabei werden nicht nur auf amerikanische Stakeholder zurückgegriffen, sondern Organisationen aus aller Welt kontaktiert. Google nutzt dazu spezifische Teams, die sich mit Free Speech und Human Rights befassen. Es soll nicht ein gewisses nationales Recht sondern ein globales Human Rights Charter als Guideline gelten.
- **Frage:** Inwiefern orientieren sich die Community Guidelines in Bezug auf Hassrede an rechtlichen Massstäben? An welchen rechtlichen Massstäben (national, international) orientiert man sich besonders?
- **Antwort:** Google hält sich an alle nationalen Regeln, die Community Guidelines reflektieren aber nicht nationale Richtlinien, sondern sollen weiter gefasst sein – so dass idealerweise Verstöße unter nationalem Recht bereits von den Community Guidelines als Verstöße erkannt werden. Zu diesem Zweck orientiert sich Google auch vor allem an internationale Standards und Instrumente wie die UN Guiding Principles on Business and Human Rights oder den Code of Conduct der EU.
- **Frage:** Bei den entfernten Videos auf Youtube nehmen die automatisierten Entfernungen stetig zu (5.3 Mio in Q4 2019, bzw. 90% aller Entfernung auf 9.5 Mio in Q1 2021 und 95% aller Entfernung). Wie erklärt Google das? Gleichzeitig steigt die Reaktivierungsrate nach Beschwerden (von 19.5% aller Beschwerden in Q4 2019 auf 30.5% in Q1 2021 – mit zwischenzeitlich 50% in Q2 2020). Wie erklärt Google das? Bedeutet dies, dass der Algorithmus vermehrt Fehler macht?
- **Antwort:** Die erhöhten Reaktivierungsraten lassen sich damit erklären, dass Google seine Mitarbeiter:innen während der Pandemie ebenfalls ins Home Office schickte. Gerade im Bereich Hate Speech schaut Google darauf, dass Inhalte nur auf bestimmten Systemen und unter geschützten Bedingungen überprüft werden. Dies ist zum Schutz der Mitarbeiter:innen – bedeutet aber auch dass die Überprüfungskapazität von Google rapide sank und vermehrt auf die automatisierten Entscheidungen zurückgegriffen wurde. Erst jetzt langsam stellt sich bei Google wieder ein Normalbetrieb ein.

Desinformation

- **Frage:** Wie wird die Schädlichkeit von Desinformation definiert? Wie wird die Abstufung zwischen „border-line content“ und erwünscht/unerwünscht bestimmt? Gibt es zusätzliche Abstufungen?
- **Antwort:** Es wird immer zuerst überprüft ob ein Inhalt Community-Guidelines verletzt oder nicht. Sind Guidelines verletzt worden, dann wird der Inhalt entfernt. „Border-line content“ beschreibt Inhalte, welche die Guidelines zwar nicht verletzen, aber entweder nahe dran sind oder dann „spammy“ sind. Dies werden reduziert, d.h. er kommt nicht im Empfehlungen

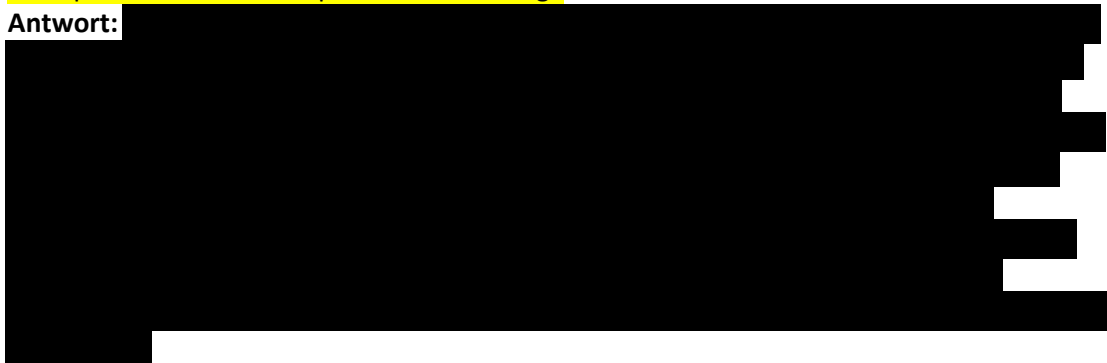
oder News. Google verfolgt die policy, dass alle Inhalte die nicht verletzend oder illegal sind auch gefunden werden sollten, wenn nach diesen gesucht wird. Es gibt aber kein Recht auf privilegierten Zugang und Reichweitenvergrößerung. Google hat in diesem Zusammenhang auch ein neuer Gradmesser: Den Violative View Rate. Die VVR zeigt auf was der Anteil der Inhalte auf Youtube sind, die problematisch sind. Er steht aktuell bei 0.16% und wurde in den letzten Jahren um 70% reduziert. Siehe auch: [Violative View Rate \(blog.youtube\)](#)

- **Frage:** Bezüglich automatisierten Systeme in Zusammenhang mit Desinformation – insbesondere in der Beurteilung von „borderline content“ und nicht erwünschten Inhalten: Wie gross ist insgesamt betrachtet die Genauigkeit der eingesetzten Systeme? Konkret, was ist der Anteil von Inhalten, die zunächst als „borderline“ oder „unerwünscht“ identifiziert werden, aber danach als unproblematisch eingeschätzt werden? Wie oft werden Inhalte als „borderline“ qualifiziert, wo sie unerwünscht sein sollten (und umgekehrt)?
- **Antwort:** Google konnte diese Fragen nicht konkret beantworten. Sie haben aber auf ihre übergeordneten Massnahmen verwiesen, nämlich dass jeder Kreator die Möglichkeit hat bei einer Entfernung eine Beschwerde einzureichen.

Wahlen und Abstimmungen

- **Frage:** Aus welchem Grund wird die Schweiz bezüglich politischer Werbung anders gehandhabt als der Rest von Europa? D.h. warum gibt es in der Schweiz keinen Transparenz-Bericht über politische Werbung?

- **Antwort:**



- **Frage:** Wird politische Werbung in der Schweiz auf Google-Diensten gekennzeichnet?
- **Antwort:** Google kennzeichnet jede Form der bezahlten Werbung mit dem Tag „Ad“ bzw. „Anzeige“. Es wird in der Kennzeichnung jedoch nicht zwischen politischer und nichtpolitischer Werbung unterschieden.

Transparenz

- **Frage:** Werden in der Schweiz alle bezahlten Inhalte als solche gekennzeichnet? Falls nein: welche nicht? Weshalb nicht?
- **Antwort:** Bezahlte Inhalte werden alle als Ad/Anzeige gekennzeichnet. Es wird auch festgehalten und ausgewiesen warum diese Ad gezeigt wird. Was in Search natürlich nicht ausgewiesen ist, ist das Suchergebnis und was dahintersteht. Wenn also jemand ein Suchbegriff eingibt und als Resultat eine Seite mit lauter Werbung kommt, dann ist dies nicht an Google auszuweisen, sondern an dieser Drittseite. Dasselbe gilt auch bei Youtube – wenn eine Partei auf ihrem Kanal ein Video raufstellt mit Werbung für sich selbst (das Video aber nicht mit Bezahlung pusht), dann kann Youtube das nicht beeinflussen. Dasselbe gilt für

Product Placement und Influencer – Kreatoren haben aber die Möglichkeit dies selbst zu deklarieren.

- **Frage:** Ist Google grundsätzlich bereit, in Zukunft CH-Reports zu allen bezahlten Inhalten zu publizieren?
- **Antwort:** Kann so nicht beantwortet werden
- **Frage:** Ist Google grundsätzlich bereit, Berichte zum Umgang mit problematischen Phänomenen (z.B. Hatespeech und Desinformation) spezifisch für die Schweiz zu veröffentlichen?
- **Antwort:** Transparenzberichte von Google sind global ausgerichtet und spiegeln nicht die Aktivitäten im nationalen Kontext. Grund dafür ist, dass das System nicht ausgelegt worden ist um Transparenzberichte zu schreiben, sondern um effizient zu entfernen. Google hat in einem ersten Schritt nun auch Berichte für die 30 bevölkerungsreichsten Ländern erstellt. Die Schweiz ist nicht darunter. Der Wunsch ist Google sehr bewusst und die Schweiz ist auch nicht allein. Die Skalierung ist jedoch schwierig und Google kann nichts versprechen.