

Resiliente Kommunikation durch Social Media Monitoring und Validierung

Data Week Leipzig, 15. April 2024

Rolf Fricke, Jan Thomsen, Ali Sarioglu, Marie Sophie Funk (Condat AG)

Social Media Monitoring zur Gewinnung zusätzlicher, validierter Informationen in Krisensituationen

- Erkennung von Gefahren
- Betroffene Gebiete
- Zusätzliche Bilder und Videos
- Stimmungslage
- Meinungen aus der Bevölkerung

Zielgruppen

- Feuerwehr
 - TV / Sender
- perspektivisch:
- THW
 - Polizei
 - weitere BOS ...



Quellen

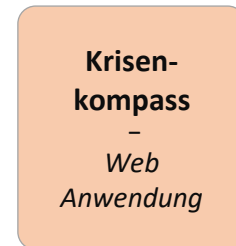
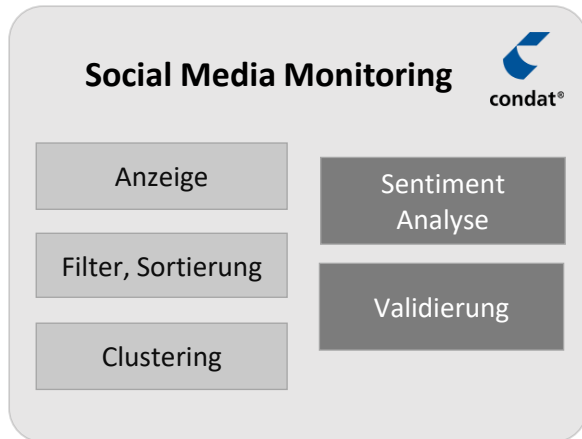


...

kritisches
Ereignis



Suche



weitere
Distributions-
Kanäle
(Broadcast,
Newsroom,
App ...)



- Leitstelle
- Redaktion



Endnutzer

Social Media – Monitoring Modus

Social Media Monitoring



Monitoring

Einstellungen

Überblick

Stromausfall Charité 03.05.2023

Karneval der Kulturen

Großbrand Neukölln 31.05.2023

Gaza Konflikt (Berlin)

Großbrand Grunewald

Feuer/Brand

https://twitter.com/Berliner_Fw
Jun 8, 2023, 10:31:29 AM

Die Einsatzstelle ist unter Kontrolle. Es brannte ein Kellerverschlag in einem 4-geschossigen Wohngebäude. Mehrere Personen wurden mit Fluchthäuben aus dem Gebäude in Sicherheit gebracht. 5 Personen wurden bei dem Brand leicht verletzt, konnten jedoch vor Ort verbleiben.
<https://t.co/k8LSgt0Gs>



Sentiment
W3C Signale
Glaubwürdigkeit

https://twitter.com/Berliner_Fw
Jun 8, 2023, 9:49:33 AM

In der #Emdenzelle in #Spandau brennt es in einem Wohngebäude. #Brandbekämpfung und Personenrettung sind eingeleitet. Zur Zeit sind wir

Stromausfall/Blackout

<https://twitter.com/polizeiBerlin>
Jun 8, 2023, 4:23:57 PM

Die Plakate und die Farbschmierereien in #Neukölln wurden durch unsere Kolleg. vom #A54 und @BerlinNkl entfernt. Unser #Staatschutz ermittelt wegen Volksverhetzung und Sachbeschädigung.
*sm <https://t.co/gltrObSv9N>

Sentiment
W3C Signale
Glaubwürdigkeit
Neukölln A54 Staatschutz

<https://twitter.com/polizeiBerlin>
Jun 8, 2023, 2:17:40 PM

#Fahndung Am 24.3.22 nahmen Kolleg. einen 23-Jährigen fest, der gerade eine Bank überfallen hatte. Unsere #Kripo sucht nun mit Fotos nach dem Mann, der ihn zu dem Bank: einem weiteren Überfall angestiftet haben soll: <https://t.co/xYkT2Dalm0> (030) 4664 – 574116 / – 571100
*sm <https://t.co/E9lwReRQS>



Sentiment
W3C Signale

Glaubwürdigkeit

Verkehr (Berlin)

<https://twitter.com/MichaelFillers1>
Jun 8, 2023, 10:00:59 PM

@BrueggemannRo @joerg_spengler Ich kann wunderbar zur nächsten S-Bahn laufen und dann in die Stadt rein. Auf dem Land kann man wunderbar mit dem Bus auch Vorwerks kommen. Nur Städter glauben dass wir im Mittelalter leben

Sentiment
W3C Signale
Glaubwürdigkeit

<https://twitter.com/MyTransitBerlin>
Jun 8, 2023, 9:48:01 PM

Fern- und Regionalbahn-Gleise der Berliner Stadtbahn auf Freitag gesperrt - S-Bahn nicht betroffen! r024 #Berlin #ÖPNV #BVG #VBB #SBahn #UBahn <https://t.co/Cp2XcRqAKH>

W3C Signale
Glaubwürdigkeit
Berlin ÖPNV BVG VBB SBahn UBahn

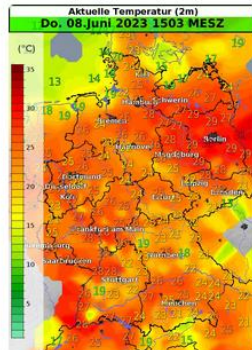
<https://twitter.com/hessenpfalzer>
Jun 8, 2023, 9:46:56 PM

Nach defektem Zug vor uns Kamen - Hamm, BU-Störung Immenhausen und Zugfolgen Kassel Rbf und Kassel: Höhe mit +43 angekommen. Nur Tram + Bus nach Hause. Verkehrsreiche zwei Tage gehen zu Ende. Und dabei sind nicht mal die Tram- und Metrofahrten in Amsterdam erfasst.

Wetter (Berlin)

https://twitter.com/DWD_presse
Jun 8, 2023, 3:30:35 PM

Es ist passiert! Erstmals in diesem Jahr wurde in Deutschland die 30-Grad-Marke geknackt. Die aktuellen Top 3: 30,8°C: Manschnow (Brandenburg) 30,8°C: Berlin-Buch (Berlin) 30,5°C: Potsdam (Brandenburg) /V <https://t.co/BRX57BZAIV>



W3C Signale
Glaubwürdigkeit

https://twitter.com/DWD_presse
Jun 8, 2023, 2:10:58 PM

Immer mehr #Gewitter schließen in den Himmel. Wenn Sie am #Nachmittag oder #Abend draußen unterwegs sind, lohnt es sich, den Himmel im Blick

Sport Ereignisse

<https://twitter.com/polizeiBerlin>
Jun 8, 2023, 4:23:57 PM

Die Plakate und die Farbschmierereien in #Neukölln wurden durch unsere Kolleg. vom #A54 und @BerlinNkl entfernt. Unser #Staatschutz ermittelt wegen Volksverhetzung und Sachbeschädigung.
*sm <https://t.co/gltrObSv9N>

Sentiment
W3C Signale
Glaubwürdigkeit
Neukölln A54 Staatschutz

<https://twitter.com/polizeiBerlin>
Jun 8, 2023, 2:17:40 PM

#Fahndung Am 24.3.22 nahmen Kolleg. einen 23-Jährigen fest, der gerade eine Bank überfallen hatte. Unsere #Kripo sucht nun mit Fotos nach dem Mann, der ihn zu dem Bank: einem weiteren Überfall angestiftet haben soll: <https://t.co/xYkT2Dalm0> (030) 4664 – 574116 / – 571100
*sm <https://t.co/E9lwReRQS>



Sentiment
W3C Signale

Glaubwürdigkeit

Demonstrationen

https://twitter.com/VIZ_Berlin
Jun 8, 2023, 8:01:00 PM

Feierabend für heute. Bis morgen!

W3C Signale
Glaubwürdigkeit

https://twitter.com/VIZ_Berlin
Jun 8, 2023, 7:48:00 PM

Letzte Sperrung des #Tiertgartentunnels in diesem Wartungszyklus: Von 21:00-05:00 Uhr ist der Tunnel in Richtung Nord (Moabit) #GESPERRT!
<https://t.co/g205Xktnl> <https://t.co/q1G1k1k1c>



Sentiment
W3C Signale
Glaubwürdigkeit
Tiertgartentunnel GESPERRT

https://twitter.com/VIZ_Berlin
Jun 8, 2023, 7:47:26 PM

Auf der #A115 (AVUS) laufen von 20:00-ca. 05:00

- Tweet Importer
- Rss Importer
- Sentiment Service
- Region Service

Sortierung nach Datum, Relevanz, Glaubwürdigkeit, Quellen

Filter

Relevanz

Emotionen

Stimmungs-
barometer

ResKriVer
Kognitive Versorgungsnetze

Social Media Monitoring

Überblick Stromausfall Charité 03.05.2023 Karneval der Kulturen **Großbrand Neukölln 31.05.2023** Gaza Konflikt (Berlin) Großbrand Grunewald

Kanäle Details Keywords: großbrand brand grossbrand feuer neukölln nemondis Kriterium 1: Kriterium 2:

Tweets Bilder Videos Quellen

1. Tweet: Bis in das 19. Jahrhundert hinein existierte die Stadt innerhalb ihrer historisch gewachsenen Form. Ein Brand im Jahr 1813 zerstörte die Stadt nahezu vollständig. Das Feuer entstand bei der Erstürmung der damals von den Franzosen besetzten Stadt durch die britische Armee.
Sentiment: [Slider]
W3C Signale: [Slider]
Glaubwürdigkeit: [5 Stars]

2. Bild: Feuerwehreinheiten bei einem Brand.
https://twitter.com/CityReport Jun 7, 2023, 1:39:04 PM

3. Video: KARLS ERLEBNIS-DORF MEHRERE VERLETZTE BEI BRAND IN FREIZEITPARK
https://twitter.com/feuernwehren Jun 5, 2023, 6:57:53 PM

4. Quelle: Name: 1FachMimon_ Beschreibung: LVL 5000 Troll • Medienmogul • Gorkong im DrachenGame • Bildreporter/Pressefoto • SUV Fahrer • Klimaaktivist • KuchenKönig • Satiriker • Spätkunzt
Ort: Berlin, Deutschland
Name: 20PercentBerlin Beschreibung: A Berlin newsletter for the 20 percent without a German passport by veteran Berlin journalists. @20percentberlin.bsky.social Ort: Berlin, Deutschland
Name: 2ftv4de
Name: 2ftv4de
Name: AdvocatDiabol2 Beschreibung: Kognitiv induzierte Verzerrungen in der stereotypen Unterbildung wirken der kognitiven Dissonanz nur zeitweilig entgegen. Ort: Wiedland
Name: Anifa_TKS Beschreibung: Anifas im Raum Teltow, Kleinmachnow & Stehendorf (TKS) - Book in der Region aktiv zu werden oder Nazistat entdeckt? Schreib gern 'ne PN! Ort: TKS
Name: ArgoviaToday Beschreibung: Das Newsportal fürs Mittelalt. Es heißt die Redaktion. #argoviatoday #argovialand Ort: Argovia
Name: BERLINER_KURIER Beschreibung: Berliner Kurier, die Berliner Zeitung und die Berliner Morgenpost

Monitoring Einstellungen

Tweet Importer
Rss Importer
Sentiment Service
Region Service

Ereignisse

Keywords
zum
Ereignis

Meta
Daten

Quellen

Analyse und Bewertung eines Posts

Statistikdaten



Stimmungs-
barometer



Emotionen



Vertrauens-
würdigkeit der
Quelle



Schlüsselwörter
zur Relevanz-
bestimmung



Vertrauens-
würdigkeits-
Signale aus dem
Post



Bewertung der
Glaubwürdigkeit
durch Redakteure



Einrichtung und Konfiguration durch Redaktionen

**ResKriVer**
Resiliente Versorgungsnetze

Social Media Monitoring

Monitoring

Einstellungen

EreignisseKanäleQuellen


Ereignisse

+ Neu

Stromausfall Charité 03.05.2023 - 2023.05.03
Keywords blackout,charite,stromausfall
Kanäle

Karneval der Kulturen - 2023.05.28
Keywords panik,notfall,menschenmassen,karnevalderkulturen,kdk,kdk23,kdk2023,tumult,karneval,demonstration,demo,verletzte,kreuzberg,#KarnevalderKulturen
Kanäle

Großbrand Neukölln 31.05.2023 - 2023.05.31
Keywords groß,brand,brand,grossbrand,feuer,neukölln,remondis
Kanäle
• Feuer/Brand

Gaza Konflikt (Berlin) - 2023.10.13
Keywords palästina,gaza,hamas,freepalestine
Kanäle

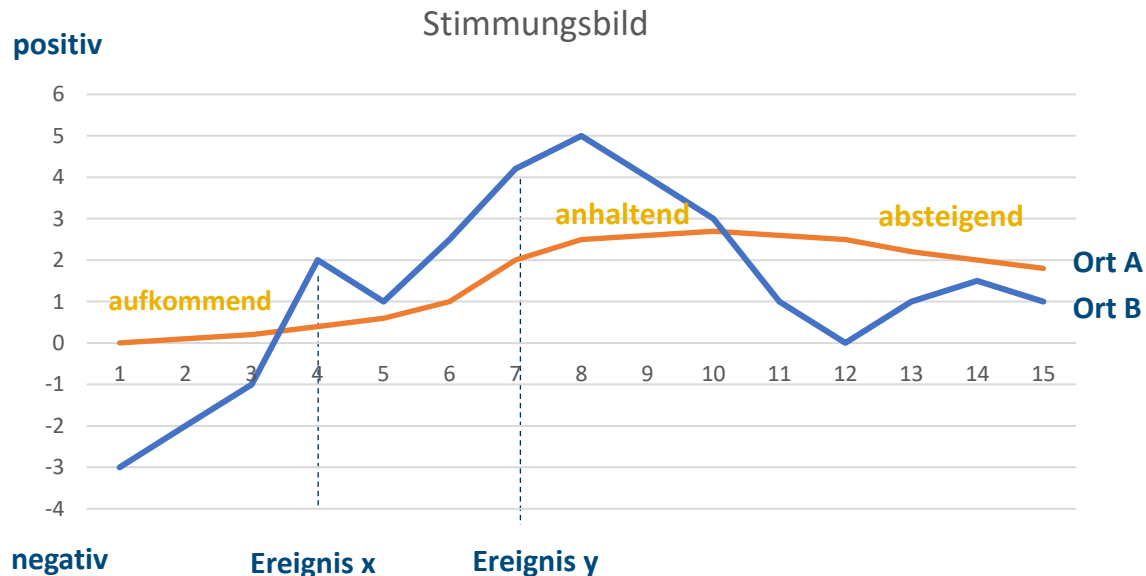
Tweet Importer

Rss Importer

Sentiment Service

Region Service

- Ermittlung Stimmungsbild durch emotional positiv / negativ bewerteten Wörter / Ausdrücke
- Veränderung der Stimmungslage
- Korrelation zu Ereignissen
- Klassifikation nach: Trauer, Wut, Furcht, Freude, Überraschung, Ekel, Verachtung
- Shitstorm-Ermittlung, orts- und zeitbasiert



Sentiment Analyse

Wörterbuchbasiert

„Feierabend.
Ich will nach Hause, Zug, fällt aus, das Stellwerk ist kaputt.
Bus auch schlecht, da schwerer Unfall und alles gesperrt ist.
Wünscht mir Glück!“

| Wort | Sentiment |
|-----------|-----------|
| abbauen | -0,2667 |
| Abendrot | 0,8 |
| Fähigkeit | 0,5392 |
| fallen | -0,3529 |
| Glück | 0,8725 |
| Tisch | -0,7667 |
| schlecht | -0,7833 |
| Unfall | 0,2549 |
| wollen | 0,23 |

Sentiment Analyse

Wörterbuchbasiert

„Feierabend.

Ich **will** nach Hause, Zug, **fällt** aus, das Stellwerk ist kaputt.

Bus auch **schlecht**, da schwerer **Unfall** und alles gesperrt ist.

Wünscht mir **Glück!**“

Gesamtsentiment:

$$\frac{0,2549 + (-0,3529) + (-0,7667) + (-0,7833) + 0,8725}{5} = -0,1551$$

| Wort | Sentiment |
|-----------------|----------------|
| abbauen | -0,2667 |
| Abendrot | 0,8 |
| Fähigkeit | 0,5392 |
| fallen | -0,3529 |
| Glück | 0,8725 |
| Tisch | 0,8725 |
| schlecht | -0,7667 |
| Unfall | -0,7833 |
| wollen | 0,2549 |

Sentiment Analyse - Emotionen

Wörterbuchbasiert

„Feierabend.
Ich will nach Hause, Zug, fällt aus,
das Stellwerk ist kaputt.
Bus auch schlecht, da schwerer
Unfall und alles gesperrt ist.
Wünscht mir Glück!“

| Ekel | Freude | Furcht | Trauer | Überraschung | Verachtung | Wut |
|---------|----------|----------|----------|--------------|--------------|---------|
| Bazille | Applaus | bangen | Abschied | aufschrecken | Abklatsch | Angriff |
| brechen | betören | grauen | Ende | Entsetzen | charakterlos | Gebrüll |
| Fäule | Glück | höllisch | freudlos | Knall | Dulli | hassen |
| igitt | Jubel | Krebs | mies | mitreißen | kaputt | pissig |
| speien | prima | scheu | Schmerz | radikal | schlecht | sauer |
| würgen | zujubeln | Unfall | Unfall | Unfall | Trottel | Vorwurf |

Sentiment Analyse - Emotionen

Wörterbuchbasiert

„Feierabend.
Ich will nach Hause, Zug, fällt aus,
das Stellwerk ist kaputt.
Bus auch schlecht, da schwerer
Unfall und alles gesperrt ist.
Wünscht mir Glück!“

Emotionen:
Freude, Furcht, Trauer,
Überraschung, Verachtung

| Ekel | Freude | Furcht | Trauer | Überraschung | Verachtung | Wut |
|---------|----------|----------|----------|--------------|--------------|---------|
| Bazille | Applaus | bangen | Abschied | aufschrecken | Abklatsch | Angriff |
| brechen | betören | grauen | Ende | Entsetzen | charakterlos | Gebrüll |
| Fäule | Glück | höllisch | freudlos | Knall | Dulli | hassen |
| igitt | Jubel | Krebs | mies | mitreißen | kaputt | pissig |
| speien | prima | scheu | Schmerz | radikal | schlecht | sauer |
| würgen | zujubeln | Unfall | Unfall | Unfall | Trottel | Vorwurf |

Sentiment Analyse mit generativer KI durch
natürlichsprachliche Formulierung der Anfrage an
das OpenAI API:

Antworte mit folgendem JSON Objekt: { "totalSentiment": 0, "emoji": Emoji, "explanationSentiment": [string] } Analysiere die Stimmung im Text mit totalSentiment: Wert für Stimmung von -1 bis 1, emoji: Emoji für Stimmung, explanation: Maximal 3 Stichwörter(Adjektive) auf deutsch für Stimmung. Alle Werte Null wenn kein Text.



positiv

negativ

<https://twitter.com/jungewelt>

Jun 1, 2023, 2:51:24 PM

Auf der Berliner Antifa-Ost-Demo am Mittwoch hat die Polizei mehrere Menschen brutal festgenommen. Einige Teilnehmer erlitten dabei schwere Kopfverletzungen durch Schläge und Tritte der Beamten. #Berlin #Kreuzberg #freelina #antifaost #linae #AntifaOst_verfahren #b3105 <https://t.co/p0IF2bjKXp>



Sentiment

- brutal
- erlitten
- brutal

#Berlin #Kreuzberg #freelina #antifaost #linae
#AntifaOst_verfahren #b3105

https://twitter.com/koeln_museum

Jun 1, 2023, 9:11:13 AM

🔥 Diese Woche startet im benachbarten MAKK nicht nur unsere Ausstellung KARNEVAL IN KÖLN. WIE ALLES BEGANN 🎉! Auch in der POP-UP-BAR 1794 COLOGNE – PARIS geht's rasant weiter 🎉 – mit der Bar-Ausstellung zur „Reise der Kölner Stadtschlüssel“ 🗝️ & einem tollen Programm: Am Do. <https://t.co/YmgfRRia7z>



Sentiment

- tollen
- rasant tollen
- rasant

positiv und negativ

https://twitter.com/BZ_NachtFloh

May 31, 2023, 12:38:44 PM

Mehr als eine halbe Million Besucher haben am Pfingstwochenende in Kreuzberg und Neukölln den Karneval der Kulturen gefeiert – auf den Straßen und in den Parks. Ausgelassenheit und Freude auf der einen, Müll und Zerstörung auf der anderen Seite. <https://t.co/XqonqsgWfc>

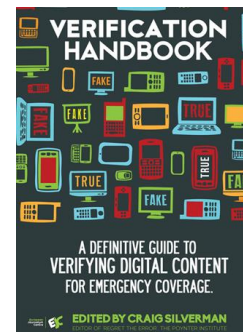
Sentiment

- Ausgelassenheit Freude
- Zerstörung
- Zerstörung
- Freude
- Müll

Validierung der Glaubwürdigkeit durch W3C Signals



- Training für Validierungsprozess durch Best Practice von Journalism Center (Verification Handbook), First Draft News, EU Disinfo Lab ...
- Nutzung von Standardtools für Suche, Mapping Services, Reverse Image Search, ...
- immer Nutzer in der Schleife, um die Ergebnisse der Tools zu bewerten
- Iterative Verbesserung von Tools z.B. für Deep Fake Detection, Audio Analyse ...



Formale Bewertung der Glaubwürdigkeit

von Posts durch ca 130 W3C Credibility Signals:

- Hasssprache, Beleidigungen, Schimpfwörter
- Grammatik- / Rechtschreibfehler
- Emotionalität, Subjektivität
- Ausrufe, Großschreibung
- Vokabularbreite
- Lesbarkeits-Indikator (z.B. Flesch-Kincaid Level)
- Konsistenz von Titel, Zusammenfassung und Inhalt

....



EU Digital Services Act (DSA) - ratifiziert Feb 2024 für große Plattformen

- Entfernung von illegalen Inhalten und Hass-Kampagnen durch Plattform Operator
- weniger relevant für ResKriVer:
 - Option ohne Profil, Kündigungsoption
 - Dark Pattern, zielgerichtete Werbung bzgl. Religion, ethnischer Herkunft ..., Rachepornos
 - gefälschte Waren ...



Abdul Alhazred Al-Almani 🍌🍌🍌 @AlhazredAl · Aug 4

#Grunewald so: am brennen
#Feuerwehr, **#Polizei**, **#Bundeswehr** und **#THW** so: wortwörtlich in den Flammen stehend unter **Gefahr** für Leib und Leben den **Brand** am bekämpfen
MdB einer rechtsextremen Politsekte so: hehe jetzt Platz für Flüchtlinge ixde **#noAfD**, die Partei der Schutzleute



Howden Lucas 🇩🇪 @z0ff · Aug 5

Ich kapiere diesen Schwachsinn sowieso nicht - mitten im **#Grunewald** ein **#Sprengplatz** - da lagern sie auch noch seit ewigen Zeiten anscheinend 'marode **#Weltkrieg** s-Bomben' und **#Feuerwerk** s-Gedöns.

Im Wald.

Diese Idiotie wird noch als 'alternativlos' bezeichnet.
#ihrNervtNurNoch

Die rote Linie 🇩🇪🇵🇹 @reflexxel

Auf die Gefahr ungerecht zu sein, aber die Berichte über den Brand des **#Grunewald**.s in der **@rbbabendschau** finde ich verstörend: Hunderte von Einsatzkräften und schwere Technik vor Ort scheinen vor allem mit perfekten Selbstschutz beschäftigt zu sein, nicht mit Löschen.

Der ..

[Translate Tweet](#)

8:42 PM · Aug 4, 2022 · Twitter for iPhone

- Wie kann man zwischen kritischer Meinungsäußerung und illegaler Hassrede unterscheiden ?
- Viele Posts sind nicht wahr oder falsch, legal oder illegal sondern kritische Äusserungen, wie z.B. zu "Corona-Maßnahmen" oder "Impfnebenwirkungen"
- Gefahr, dass Redaktionen sicherheitshalber zu viel löschen



- im Krisenfall ist nach DSA ausschließlich EU-Kommission zuständig
- z.B. bei terroristischen Handlungen, Naturkatastrophen wie Erdbeben, Wirbelstürme, Pandemien
- Vorgaben der EU für Anbieter sehr großer Online-Plattformen:
 - Verfahren zur Moderation von Inhalten
 - Zusammenarbeit mit vertrauenswürdigen Hinweisgebern
 - Förderung vertrauenswürdiger Informationen
 - Anpassung algorithmischer Systeme
 - Änderung der AGBs

Einschränkung der Meinungsfreiheit ?

- Einflussnahme von Regierungen aufgrund DSA Artikel 36 oder durch DSA Moderationsregeln
 - Beispiel: Twitter Files von Elon Musk: X/Twitter folgt Vorgaben von Regierungen (Türkei Erdogan Wahl, BRD /Corona Politik, Ukraine ...)
- > Risiken, dass Meinungsvielfalt eingeschränkt wird
 - aufgrund DSA werden Operator im Zweifelsfall löschen um Strafen aus dem Weg zu gehen
 - Diskriminierung von Kritik, vor allem von Randgruppen und kritischen Nachfragen -> Ende des freien Internets ?

weitere EU Regulierungs Maßnahmen

- EU AI Act – verabschiedet Dez 2023
 - KI Algorithmen zur Klassifikation, Bewertung (kein Bias / Diskriminierung), Social Scoring unter Nutzung personenbezogener Daten
- EU Data Act – verabschiedet Nov 2023
 - Offenlegung von Daten gegenüber Behörden im Krisenfall
- EU Market Act – ratifiziert seit Oktober 2022
 - Regelungen zu Gatekeeper Problematik

....

- Aktuelle, relevante Social Media Posts für BOS primär nur auf Twitter
- Auf vielen Plattformen nicht alles sichtbar, Untergruppen z.B. nur über bestimmte Accounts (Facebook, Telegram, vk)
- Häufig wechselnde Bedingungen für Kosten und Zugriffsrechte bei Twitter API
- Bei Twitter nur für verified / bezahlte Accounts, Anzahl der Aufrufe begrenzt (z.B. 10.000 Zugriffe / Tag)

**Vielen Dank für die
Aufmerksamkeit !**

Fragen ?