
DESINFORMATIONEN UND MESSENGERDIENSTE: HERAUSFORDERUNG UND LÖSUNGSANSÄTZE

Martin Steinebach



MOTIVATION

- Desinformationen in den heute bekannten Ausprägungen sind unter anderem auch durch die technischen Möglichkeiten bedingt, in denen sie verbreitet werden
 - User Generated Content erlaubt Erzeugen von Inhalten durch Bürger
 - Soziale Medien erzeugen Verbreitungsnetze
 - Bewertende Algorithmen erzeugen Filterblasen
 - NLP ermöglicht Bots, die Nachrichten verstärken
- Messenger sind Kanäle, die an Bedeutung gewonnen haben
 - Abwanderung der Nutzer von Sozialen Medien
- Welche neuen Aspekte bringen Messenger in Desinformationen ein
 - Überschneidung von öffentlicher und privater Kommunikation

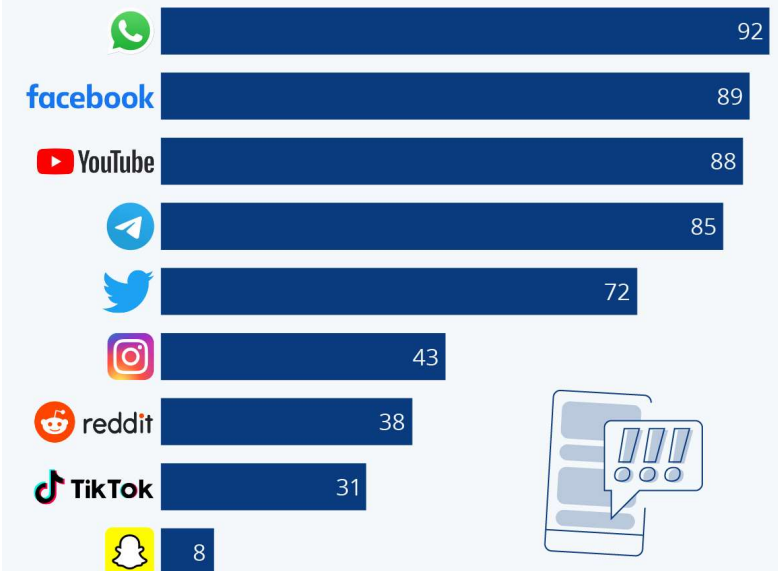


RÄUME UND UNTERSCHIEDUNGSMERKMALE

- Sichtbarkeit von Mitteilungen
 - Gruppen vs. Kanäle, öffentlich vs. Geschlossen
- Typische Kommunikationsarten
 - P2P vs. Broadcast
- Kontrolle
 - Inhaltsbezogen oder technisch
- Altersgruppen
- Diese Eigenschaften können auch auf einer Plattform unterschiedliche vorhanden sein

Facebook führt bei Fake News

Anteil der Befragten, die folgende Plattformen als relevant für Verbreitung von Desinformation sehen (in %)



Basis: Qualitative Befragung von 63 Expert:innen (in Deutschland); 08.05.-05.07.2020
Quelle: Vodafone Stiftung



statista

TELEGRAM

- Populäre, zugängliche Kanäle in Telegram als Ausgangspunkt für Desinformationen



Mit Fake News gegen die Bundestagswahl: Telegram schießt gegen die Briefwahl



Am 26. September findet die diesjährige Bundestagswahl statt. Bereits seit einigen Wochen kann auch per Post abgestimmt werden (Symbolbild).
© Sven Hoppe / DPA


Darknet für die Hosentasche



VERBREITUNG

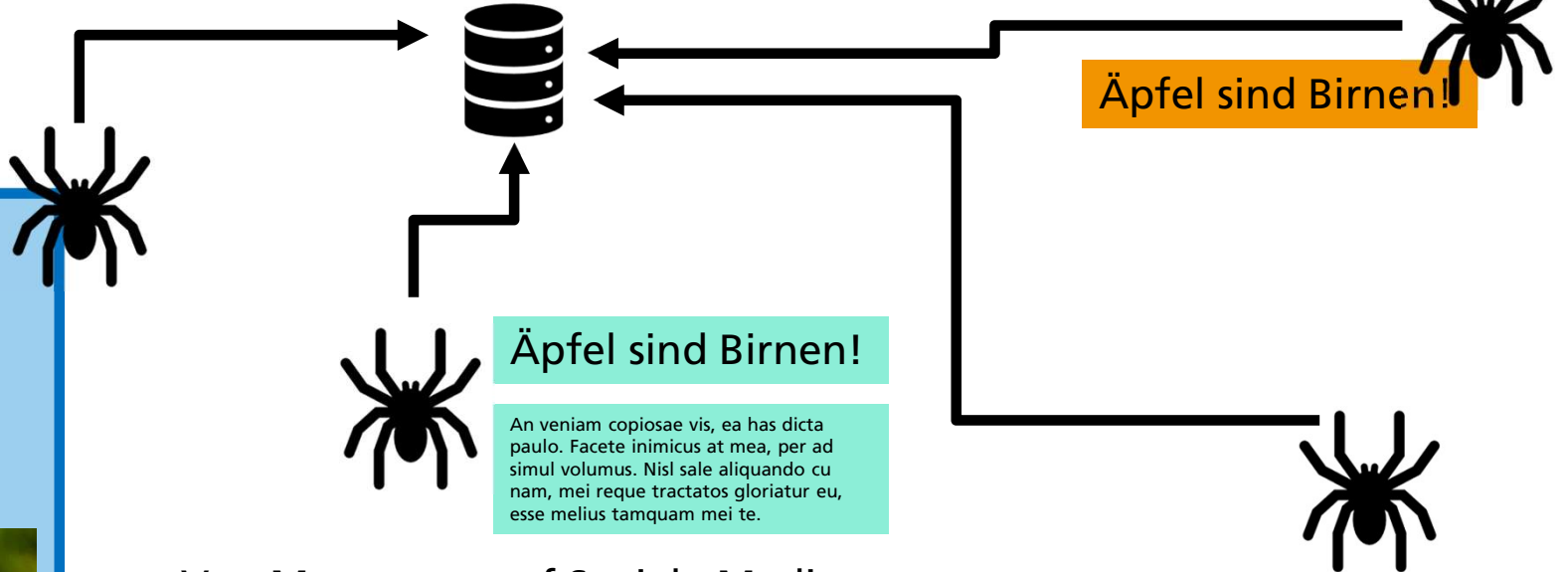
Äpfel sind Birnen!

Lorem ipsum dolor sit amet, omnis accusata conclusionemque ea pro. Ad cum nostro voluptatum, quis viris ex vel. Ne case copiosae sea, mea no meis oporteat consecetur. An veniam copiosae vis, ea has dicta paulo. Facete inimicus at mea, per ad simul volumus. Nisl sale aliquando cu nam, mei reque tractatos gloriatur eu, esse melius tamquam mei te. Falli constituam in nec.



▶

● ● ●



- Von Messenger auf Soziale Medien
 - Text
 - Medien
- Verbreitungsformen
 - Zitat
 - Verkürzung
 - Abwandlung

Äpfel? Birnen!



▶

VIDEOS IN DESINFORMATIONEN

- Deepfakes in Desinformationen spielen noch keine wichtige Rolle
 - Eher: Videos aus dem Kontext gerissen
- Möglichkeiten der Manipulationserkennung
 - Wiedererkennen von Videos, auch nach Schnitt, Einblendungen,...
 - Robuste Hashverfahren, Feature Matching
 - Allgemeine Manipulationserkennung
 - Mehrfache Kompression, Homogene Sensorfingerprints
 - Deepfake-Erkennung
- Perspektivisch: Änderungen von Aussagen in Videos (audiovisuelle Deepfakes)
 - Erkennen von Synchronisation Bild/Ton
 - Wiedererkennen Ausgangsmaterial
- Sprachnachrichten gewinnen an Bedeutung

ANALYSE VON INVERSEN BILDERSUCHEN: BEISPIEL



Deutsch sein ist kein Verbrechen ...
Ayer a las 13:49 · 🌐

Paris aktuell! 🇫🇷🇫🇷🇫🇷
Frankreich wehrt sich!!! 🇫🇷🇫🇷🇫🇷
Wann fällt Macron endlich??? 🙄

👍 771 78 comentarios 346 veces compartida

👍 Me gusta 💬 Comentar ➦ Compartir

Más relevantes ▾

Mathias Bachmann
So müsste es in Deutschland auch sein 🇩🇪 77

Me gusta · Responder · Ver traducción · 23 h

Se ha seleccionado la opción "Más relevantes", por lo que es posible que algunas respuestas se hayan filtrado.

Autor
Deutsch sein ist kein Verbrechen
In Deutschland wartet man lieber auf die Wahlen und verlässt sich sowieso lieber auf andere anstatt seinen fetten feigen Arsch selbst zu erheben!!!!!! 🇩🇪 🙄 17

Me gusta · Responder · Ver traducción · 19 h

➦ 1 respuesta 🗨️

ANALYSE VON INVERSEN BILDERSUCHEN: ANALYSE DER PERFORMANZ

■ Bewertung der inversen Bildersuchen basierend auf 25 Bildern

1. Google Bilder	(92% / 4% / 4%)
2. Microsoft Bing	(80% / 16% / 4%)
3. TinEye	(36% / 8% / 55%)
4. Yandex	(52% / 36% / 12%)
5. Baidu	(32% / 0% / 68%)
6. Sogou	(19% / 8% / 73%)

Grün = Rel. Häufigkeit, dass man die Falschmeldung widerlegen könnte

Rot = Rel. Häufigkeit, dass die inv. Bildersuche nur Beiträge auflistet, bei denen die Falschmeldung weiterverbreitet wird (negativer Effekt)

Schwarz = Rel. Häufigkeit, dass das Eingabebild oder das Original nicht gefunden wurde.

GEPLANTER ABLAUF



VIELEN DANK.

- Projekt DYNAMO
- <https://www.forschung-it-sicherheit-kommunikationssysteme.de/projekte/dynamo>
- Partner
 - Fraunhofer SIT, Darmstadt
 - Universität Duisburg-Essen, Essen
 - Universität Kassel
 - Hochschule der Medien Stuttgart
- Laufzeit: 11/2021 - 08/2024
- Kontakt: martin.steinebach@sit.fraunhofer.de