Tipps gegen Wahlmanipulation im Netz

(Stand 25.09.2024)

Die Europäische Kommission, die im Rahmen des Digital Services Act (DSA) für die Aufsicht über sehr große Online-Plattformen und sehr große Suchmaschinen wie Facebook, Instagram, X, TikTok und andere[1] im Rahmen einer Mitteilung Leitlinien für Anbieter sehr großer Online-Plattformen und sehr großer Suchmaschinen zur Minderung systemischer Risken in Wahlprozessen, wie die Nationalratswahlen, die demnächst in Österreich stattfinden, gemäß Art. 35 Abs. 3 der Verordnung (EU) 2022/2065, erlassen.

Warum? Die Art und Weise, in der diese Plattformen genutzt werden können, hat großen Einfluss auf die öffentliche Meinungsbildung und den öffentlichen Diskurs. Ihre Gestaltung ist im Allgemeinen auf eine Optimierung ihres in der Regel werbegestützten Geschäftsmodells ausgerichtet. Damit geht einher, dass diese Dienste aufgrund ihrer hohen Nutzerzahl, ihrer Gestaltung, Funktionsweise und Nutzung systemische Risken, nämlich Risken zur Erzeugung gesellschaftlicher Verwerfungen, auslösen können.

Die Europäische Kommission kann als für die Aufsicht über die genannten Dienste ausschließlich zuständiger Regulator von diesen verlangen, sogenannte Risikominderungsmaßnahmen zu ergreifen.

Welche konkreten systemischen Risken bestehen bei Wahlen auf diesen Plattformen? Böswillige Akteure können die genannten Risiken der Plattformen ausnutzen, um freie und faire Wahlen zu manipulieren oder zu unterlaufen. Dabei können etwa Konten von Politiker:innen nachgeahmt werden, falsche Angaben zum Wahlvorgang (die Wahlen wurden verschoben, täuschende Angaben zur Gültigkeit von Wahlzetteln) propagiert und falsche Informationen zum Ausgang der Wahlen äußerst rasch verbreitet werden, um den demokratischen freien Wahlprozess direkt anzugreifen.

Hierbei muss man berücksichtigen, dass globale Konflikte längst nicht mehr auf militärische Auseinandersetzungen beschränkt sind. Die Strategien der Beteiligten reichen nun von wirtschaftlichem Druck über Cyberangriffe bis hin zur Verbreitung von Propaganda in Medien und sozialen Netzwerken. Diesen vielschichtigen Ansatz bezeichnet man als "hybride Kriegsführung". Diese Bedrohung durch gezielte und koordinierte Manipulation von Informationen im digitalen Raum um Gesellschaften, etwa das westliche Gesellschaftsmodell, durch Konflikte zu destabilisieren und in Frage zu stellen ist aufgrund der relativen Neuheit der Allgegenwart digitaler Dienste ein junges, aber nicht weniger ernstzunehmendes Phänomen der letzten Jahre. Neu ist die Geschwindigkeit und die Präzision, mit der sie sich im digitalen Raum verbreiten kann und die Bedeutung, die die Online-Umgebung im Informationsökosystem, besonders in einer Wahlauseinandersetzung, gewonnen hat.

[1] https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses

WAS KANN ICH TUN, UM MICH VOR MANIPULATION ZU SCHÜTZEN?

Es ist nicht so, dass man „sowieso nichts machen kann". Ganz im Gegenteil, im Kampf gegen Desinformation kommt jedem einzelnen Verantwortung zu, denn jeder kann unfreiwillig zum Gehilfen werden, indem er Informationen weitergibt, ohne sie vorher zu überprüfen. Mit etwas Aufmerksamkeit und kritischer Lektüre ist es oftmals möglich, gezielte Desinformation zu enttarnen und durch Meldung an das betreffende soziale Netzwerk deren virale Verbreitung zu verhindern.

Dieser Leitfaden soll einen praktischen Überblick geben und die Abwehrfähigkeit jedes/r Einzelnen stärken.

WAS IST DESINFORMATION?

In einer grundlegenden Mitteilung der Europäischen Kommission zur Bekämpfung von Desinformation wird Desinformation wie folgt definiert:

„Desinformation sind nachweislich falsche oder irreführende Informationen, die mit dem Ziel des wirtschaftlichen Gewinns oder der vorsätzlichen Täuschung der Öffentlichkeit konzipiert, vorgelegt und verbreitet werden und öffentlichen Schaden anrichten können."

Die Techniken, die zum Einsatz kommen sind vielfältig: Manipulation von Bildern, Verwendung gefälschter Konten in sozialen Netzwerken, irreführende oder aus dem Zusammenhang gerissene Aussagen, gefälschte Webseiten usw.

WELCHE GEFAHREN GEHEN DESINFORMATION AUS?

Sie untergräbt das öffentliche Vertrauen. Desinformation untergräbt das Vertrauen der Öffentlichkeit in die Medien, Regierungsinstitutionen und die Demokratie selbst. Und oft bezweckt sie bzw. jene, die sie verbreiten genau das.

Sie untergräbt soziale Stabilität und nationalen Zusammenhalt. Desinformation kann soziale und ethnische Spannungen verschärfen, indem sie Stereotypen und Unwahrheiten verbreitet, die Misstrauen und Spaltung in der Gesellschaft schürt bzw. diese erheblich verstärkt.

Sie verursacht wirtschaftlichen Schaden. Die Verbreitung von Desinformation kann negative wirtschaftliche Auswirkungen haben, indem Verbraucher in die Irre geführt werden, Finanzmärkte gestört und der Ruf von Unternehmen gefährdet oder untergraben wird.

Sie schwächt die Landesverteidigung. Ausländische Mächte können Desinformation verbreiten, um für gesellschaftliche Unruhe zu sorgen vor allem zu besonderen Anlässen (Sportereignisse, Wahlen), um eigene Interessen durchzusetzen, die dem betroffenen Land schaden.

Online-Desinformation hat zudem das Potential, zu Gewalt in der analogen Welt zu führen. Sie bezweckt eine Verhaltensänderung von Menschen häufig, indem sie auf das Auslösen negativer Emotionen wie Wut und Empörung zielt.

Desinformation ist im Internet allgegenwärtig und nimmt viele Formen an.

Abbildung 1: Ein im Rahmen des Doppelgänger Netzwerk, einer pro-russischen Desinformationskampagne, verbreiteter Artikel, der das Design eines großen deutschen Medienhauses imitiert.

WIE KANN MAN DESINFORMATION ERKENNEN?

Seien Sie misstrauisch, wenn ein Inhalt ...

- zwar „Fakten“ enthält, aber keine Quellen und/oder Beweise angegeben oder genau wiedergegeben werden
- mit besonders starken Emotionen oder Gefühlen behaftet ist
- einen besonders reißerischen und emotionalen Titel trägt, um geteilt zu werden
- reißerisch ist und Ungereimtheiten enthält
- unbestätigte „Insiderinformationen“ enthält, die nirgends sonst zu finden sind
- keinen Aufschluss über den Urheber erlaubt

ÜBERPRÜFEN SIE DAHER, WER DER URHEBER DER INFORMATION IST!

Eine Quelle ist vermutlich bzw. möglicherweise zuverlässig, wenn:

- Sie allgemein anerkannt, schon lange bekannt ist und zitiert wird. Aber ACHTUNG! Auch dann ist eine Überprüfung erforderlich, denn Konten und Links sowie die Webseiten, zu denen sie führen, können gefälscht sein. Wenn man die Quellen nicht kennt, sollte man sie im Internet suchen, und zwar mehrere Einträge. Bezüglich der Inhalte sollte darauf Bedacht genommen werden, mehrere verschiedene Quellen einzusehen.
- Sie äußert sich über einen bekannten Kanal, wie etwa eine Zeitung, Fernsehsender, Radiosender oder Magazine. Auch hier muss man achtsam sein: Es gilt zu recherchieren, wer hinter diesen Veröffentlichungen steckt. Medien wenden in der Regel die journalistische Sorgfalt an, das heißt sie prüfen bestmöglich die Fakten und geben die verschiedenen Standpunkte wieder. Wenn das Medium glaubwürdig ist, muss auch geprüft werden, ob die Webseite bzw. der Link nicht gefälscht ist. Leider entspricht es auch einem Trend, dass Desinformation ihren Weg in traditionelle Medien findet.

ERKENNEN SIE GEFÄLSCHTE INTERNETSEITEN:

Eine gefälschte Webseite, ist eine Internetseite, die so gestaltet wurde, dass sie wie die Webseite einer echten Organisation oder einer bekannten Person aussieht, mit dem Ziel, Nutzer:innen in die Irre zu führen. Eine gefälschte Website kann durch die folgenden Hinweise erkannt werden:

- Das Fehlen der typischen Abbildung eines Vorhängeschlosses neben der URL (http statt https):

- Die Struktur der URL: Achten Sie darauf, dass die sogenannten Second Level- (google, bing) und Top-Level-Domain (.eu, .com, .at) immer mit der der Website übereinstimmt, die man besucht. Beispiele:

Echt	Nicht echt
www.paypal.com	www.paypa1.com
www.apa.at	www.apa.art

- Grammatik- oder Rechtschreibfehler
- Das Vorhandensein von "aggressiven" Werbeanzeigen
- „Schlampiges“ Design (pixelige oder unscharfe Bilder und Logos, Links, die ins Leere führen, uneinheitliches Design)
- Wortgleiche, generische Kommentare

WAS TUN, WENN ZWEIFEL AN EINER INFORMATION BESTEHEN?

- Es sollte eine andere zuverlässige Quelle im Internet aufgesucht werden, um die Informationen abzugleichen.
- Fact-Checking-Websites enthalten oft die gesuchte Information (siehe Links am Ende der Seite).
- Wenn es sich um ein Bild handelt, kann das Bild in die Google-Suchleiste kopiert werden, und so seine Herkunft festgestellt werden.

WIE ERKENNT MAN EIN VON KÜNSTLICHER INTELLIGENZ GENERIERTES BILD?

Künstliche Intelligenz (KI) wird hinsichtlich der visuellen Darstellungen immer besser, dennoch haben KI-generierte Bilder noch häufig erkennbare Makel. Manchmal ist KI bei der Darstellung bestimmter Gegenstände oder bestimmter Körperteile (besonders Hände und Haare) nicht stimmig und präzise genug. Accessoires wie Ohrschmuck, Halsketten, Gläser sind oft nicht korrekt dargestellt:

Abbildung 2 Bildquelle: Pocol et al. (2024). Seeing is No Longer Believing: A Survey on the State of Deepfakes, AI-Generated Humans, and Other Nonveridical Media. In: Advances in Computer Graphics. Springer. https://doi.org/10.1007/978-3-031-50072-5_34

Abbildung 3 Bildquelle: Kamali et al. (2024). How to Distinguish AI-Generated Images from Authentic Photographs. https://doi.org/10.48550/arXiv.2406.08651

Wenn eine Menschenmenge oder eine große Anzahl von Personen KI-generiert ist, können Gesichter an Konsistenz und Schärfe verlieren.

KIs habe manchmal noch Probleme komplexe Texturen darzustellen. Es ist auch schwierig für sie, Lichtverhältnisse konsistent zu halten, sodass unterschiedliche Lichtverhältnisse bzw. Schatten in verschiedenen Bereichen des Bildes aber auch unnatürliche Unschärfen einen Hinweis geben können:

Abbildung 4 Bildquelle: Bild generiert mit Dreamshaper XL Lighting

KI generierte Darstellungen wirken oft sehr symmetrisch:

Abbildung 5 Bildquelle: Bild generiert mit Dalle-3

Die KI hat Schwierigkeiten, Text in ein Bild zu integrieren (z. B. ein Markenname auf einem Kleidungsstück oder ein Name eines Lebensmittelprodukts):

Abbildung 6 Bildquelle: Kamali et al. (2024). How to Distinguish AI-Generated Images from Authentic Photographs. https://doi.org/10.48550/arXiv.2406.08651

Die Verbesserung künstlicher Intelligenz schreitet rasant voran. Das führt dazu, dass eine Identifizierung basierend auf den genannten Kriterien zunehmend schwieriger wird bzw. auch vielleicht nicht möglich ist.

WIE ERKENNT MAN EIN DEEPFAKE VIDEO?

Deepfakes sind mit KI erstellte, sehr realistisch wirkende Fälschungen. Diese Technologie kann verwendet werden, um Personen in Videos darzustellen, die Dinge sagen oder tun, die sie in Wirklichkeit nie gesagt oder getan haben.

Die Erkennung solcher Videos ist schwierig (siehe oben), aber es gibt einige Anhaltspunkte, die helfen können, Deepfakes zu identifizieren:

- ➔ Unnatürliche Gesichtsbewegungen und Mimik: Zu achten ist auf unnatürliche oder ruckartige Bewegungen des Gesichts, die oft nicht mit der restlichen Körperbewegung übereinstimmen. Besonders die Augen, Lippen und die Gesichtsproportionen können hier Hinweise geben.
- ➔ Unstimmigkeiten bei der Beleuchtung: es sollte geprüft werden, ob die Beleuchtung auf dem Gesicht mit der Beleuchtung der Umgebung übereinstimmt. Inkonsistente Schatten oder Lichtreflexe können ein Zeichen für eine Manipulation sein.
- ➔ Augenblinzeln: Deepfake-Algorithmen haben oft Schwierigkeiten, das Blinzeln der Augen korrekt darzustellen. Zu seltenes oder unnatürliches Blinzeln kann ein Hinweis auf ein Deepfake sein.
- ➔ Audio-Video-Synchronisation: Es muss darauf geachtet werden, ob die Lippenbewegungen exakt mit dem gesprochenen Wort übereinstimmen. Abweichungen können auf ein manipuliertes Video hindeuten.
- ➔ Fehlende oder unscharfe Details: Bei genauer Betrachtung können manche Details im Gesicht, wie Hautunreinheiten oder Falten, fehlen oder unscharf erscheinen.
- ➔ Hintergrundverzerrungen: Oftmals können im Hintergrund des Videos Verzerrungen oder Unschärfen auftreten, die nicht mit den Bewegungen im Vordergrund übereinstimmen.
- ➔ Kanten und Übergänge: Genau zu betrachten sind Kanten und Übergänge, besonders an den Rändern des Gesichts. Unnatürliche Kanten oder flimmernde Ränder sind häufige Anzeichen für Deepfakes.
- ➔ Haarbewegungen: Haare sind schwierig zu animieren. Es ist darauf zu achten, ob die Haarbewegungen realistisch sind oder sich vom restlichen Video abheben.
- ➔ Unnatürlicher Ausdruck: Gesichtsbewegungen können manchmal steif oder übertrieben wirken, besonders wenn Emotionen dargestellt werden.