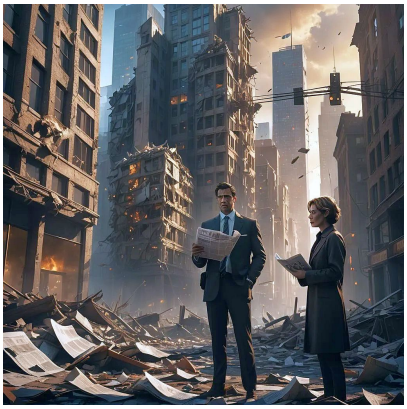


Desinformatie 1/â?¦

Beschrijving

Nog maar eens wat over de plaag van onze tijd: desinformatie of â??*fake news*â?• zoals dat wordt gegenereerd door de diverse AI mogelijkheden (voor tekst, beeld en geluid).

Zelfs het onderzoek naar die desinformatie is de afgelopen jaren onder vuur komen te liggen. Politici van rechts hebben ongefundeerde beweringen gedaan over hoe onderzoek naar desinformatie conservatieve standpunten het zwijgen oplegt, terwijl anderen hebben betoogd dat desinformatie moeilijk te identificeren is en dat fact-checking-initiatieven te politiek zijn om effectief te zijn.

Een analyse

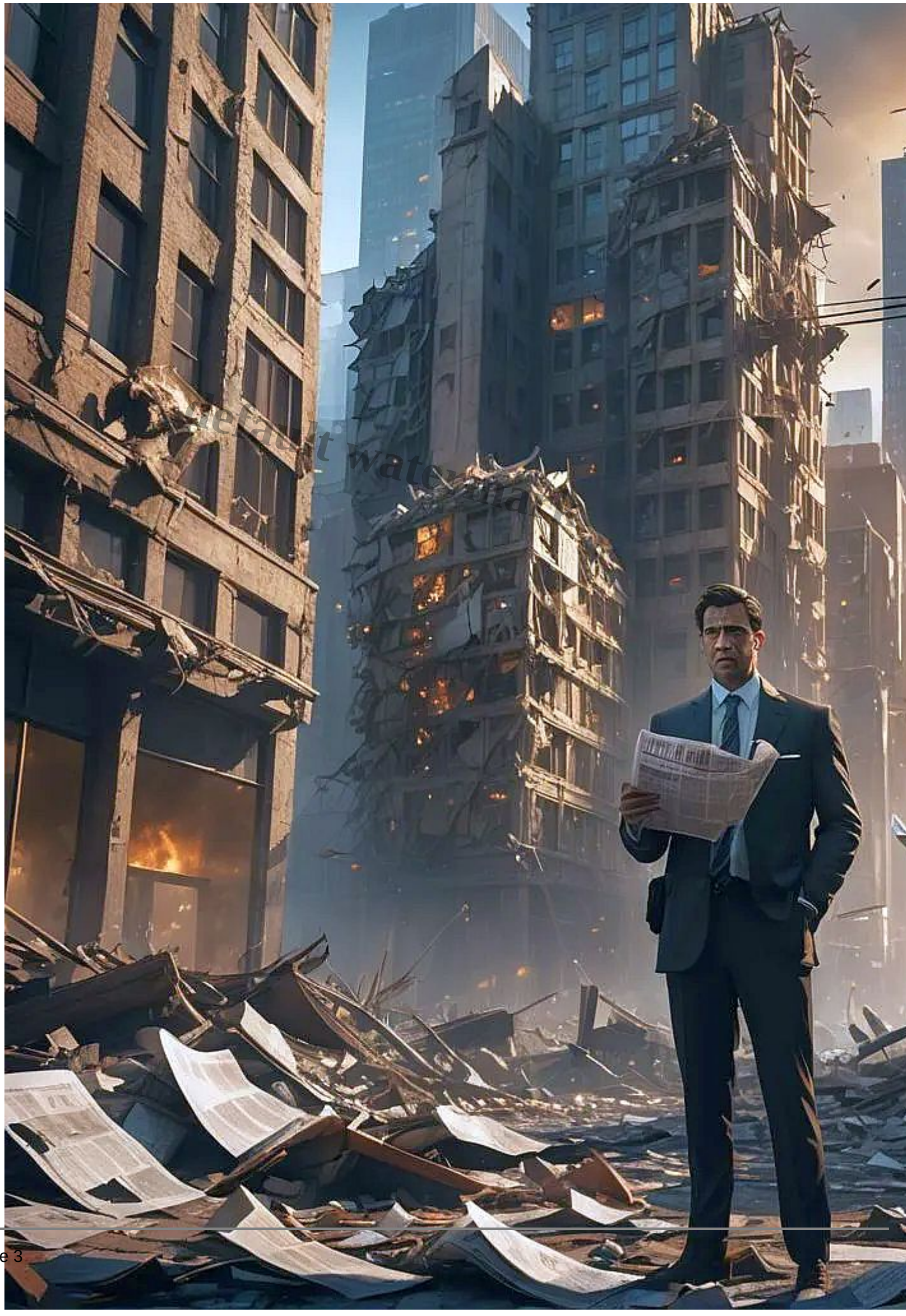
Een artikel gepubliceerd in het tijdschrift *Nature Humanities and Social Sciences Communications* met de titel â??*Liars Know They Are Lying: Differentiating Disinformation from Disagreement*â?• weerlegt deze beweringen en biedt een diepgaand analyse van de vaak gepolitiseerde wereld van onderzoek naar desinformatie.

De auteurs, *Stephan Lewandowsky, Ullrich K. H. Ecker, John Cook, Sander van der Linden, Jon Roozenbeek, Naomi Oreskes* en *Lee C. McIntyre*, betogen dat opzettelijke desinformatie (d.w.z. valse beweringen met de intentie om te misleiden) â??*aantoonbaar schadelijk is voor de volksgezondheid, op bewijs gebaseerde beleidsvorming en democratische processen.*â?• De auteurs stellen ook manieren voor waarop maatschappelijke organisaties en beleidsmakers opzettelijke desinformatietactieken kunnen identificeren en erop kunnen reageren zonder hun toevlucht te nemen tot censuur.

De auteurs leveren ruim empirisch bewijs dat de bewering ontkracht dat desinformatieonderzoek erop gericht is conservatieven het zwijgen op te leggen. Sterker nog, zo schrijven ze, recent onderzoek dat ***208 miljoen Amerikaanse Facebook-gebruikers*** bestudeerde, ontdekte dat een â??*substantieel segment van het nieuws-ecosysteem uitsluitend door conservatieven wordt geconsumeerd en dat de meeste desinformatie zich binnen deze ideologische bubbel bevindt.*â?•

Critici het zwijgen opgelegd

Dat heeft rechtse politici er echter niet van weerhouden om angst voor de vrijheid van meningsuiting onder hun aanhangers aan te wakkeren. Politici hebben ook onderzoekers op de korrel genomen en in diskrediet gebracht: van openbare veroordelingen tot wetgevende maatregelen om de academische vrijheid in te perken. Hierdoor vinden onderzoekers het steeds moeilijker om onpartijdige onderzoeken uit te voeren. Zo werd *Kate Starbird*, een onderzoeker naar desinformatie aan de *Universiteit van Washington*, opgeroepen om te getuigen voor de *Select Subcommittee on the Weaponization of the Federal Government,* waar ze te maken kreeg met valse beschuldigingen dat ze samenspande met de regering-Biden om conservatieve stemmen het zwijgen op te leggen.
Ze vertelde de *New York Times*: â??*De mensen die profiteren van de verspreiding van desinformatie hebben effectief veel van de mensen het zwijgen opgelegd die hen juist zouden proberen aan te spreken.*â?•



Is objectiviteit mogelijk?

Het artikel behandelt ook wat de auteurs de â??postmoderneâ?• kritiek op desinformatie-onderzoek noemen. In plaats van beschuldigingen van desinformatie te weerleggen met feitelijk bewijs, vallen ze terug op een aanval op â??*het idee dat objectieve kennis zelfs mogelijk is*.â?•
In het artikel worden verschillende voorbeelden van deze tactiek aangehaald, met name van voormalig president *Donald Trump* en zijn adviseurs, die vaak verwezen naar het idee van alternatieve feiten â?? of de beroemde opmerking van voormalig burgemeester van New York City en bondgenoot van Trump, *Rudy Giuliani*: â??*Waarheid is geen waarheid.*â?•
Deze strategieÃ«n dienen als manieren om desinformatie te verspreiden om het publieke vertrouwen te ondermijnen, vaak met weinig tot geen publieke gevolgen, vooral in het licht van beslissingen die door sociale media-bedrijven zijn genomen om:

- hun vertrouwen en veiligheidsinspanningen te verminderen en
- tegelijkertijd werknemers te ontslaan die werken aan het detecteren van haatzaaierij en verkiezingsintegriteit.

Belangrijkste strategieÃ«n

Naast het leveren van bewijs voor waarom desinformatie nog steeds een urgente beleidskwestie is, bieden de auteurs nuttige inzichten in hoe desinformatie en de onderliggende intentie om te misleiden te identificeren. Ze bieden drie belangrijke strategieÃ«n.

- **Statistische en linguÃ¯stische analyse:**Ã©Ã©n effectieve methode is gebaseerd op statistische en linguÃ¯stische analyse van tekst. Hoewel mensen notoir onbetrouwbaar zijn in het detecteren van leugens â?? ze presteren slechts iets beter dan toeval â?? hebben ontwikkelingen in natuurlijke taalverwerking (NLP) dit vermogen aanzienlijk verbeterd. Machine-learningmodellen kunnen nu linguÃ¯stische signalen analyseren om teksten als misleidend of eerlijk te classificeren. De auteurs citeren bijvoorbeeld onderzoek naar een model dat de distributie van verschillende soorten woorden onderzocht en een nauwkeurigheidspercentage van 67% behaalde, waarmee het beter presteerde dan menselijke beoordelaars die slechts 52% scoorden. Een soortgelijk model werd gebruikt om tweets van Donald Trump te classificeren als waar of onwaar op basis van onafhankelijke feitencontroles. Opmerkelijk genoeg behaalde dit model een nauwkeurigheid van meer dan 90%.
- **Analyse van interne documenten:** Een andere methode om opzettelijke misleiding te detecteren, is het analyseren van de interne documenten van instellingen, zoals overheden of bedrijven. Door de interne kennis van deze entiteiten te vergelijken met hun openbare verklaringen, kunnen onderzoekers actieve misleiding ontdekken, vooral wanneer deze op grote schaal plaatsvindt. Deze aanpak is met name effectief gebleken bij het blootleggen van wanpraktijken van bedrijven, waarbij discrepanties tussen interne communicatie en openbare standpunten opzettelijke misleiding onthullen. Hoewel deze techniek veel geld kan kosten, kan het ook leiden tot belangrijke resultaten, zoals de â??*veroordeling van Phillip Morris op grond van de federale afpersingswet (RICO)*.â?•
- **Vergelijking van verklaringen met officiÃ«le getuigenissen:** De derde benadering richt zich op het identificeren van discrepanties tussen publieke verklaringen en die welke in een rechtbank zijn afgelegd. Een opvallend voorbeeld is Donald Trumps â??*grote leugen*â?• over de

presidentsverkiezingen van 2020. Terwijl Trump publiekelijk en herhaaldelijk beweerde dat er sprake was van wijdverbreide verkiezingsfraude, ondersteunden zijn advocaten, die meer dan 60 rechtszaken aanspanden met betrekking tot de verkiezingen, deze beweringen ***niet*** in de rechtbank. Sterker nog, Trumps advocaten ***ontkenden*** vaak elke vermelding van fraude toen ze door rechters werden ondervraagd.

Het genoemde artikel ondersteunt het idee dat onderzoek naar desinformatie essentieel is voor een democratisch discours. Belangrijk is dat ook dat er onderscheid gemaakt wordt tussen een gezond democratisch debat op basis van *betwiste feiten* en *regelrechte leugens*. Zoals de auteurs opmerken, geeft het oneens zijn over feiten â??*geen toestemming voor het gebruik van regelrechte leugens en propaganda om het publiek opzettelijk te misleiden*â??, en is het mogelijk om â??*onwaarheden, desinformatie en leugens te identificeren en ze te onderscheiden van te goeder trouw gegeven politieke en beleidsmatige argumentatie.*â??

Kwantificeren van de impact van desinformatie

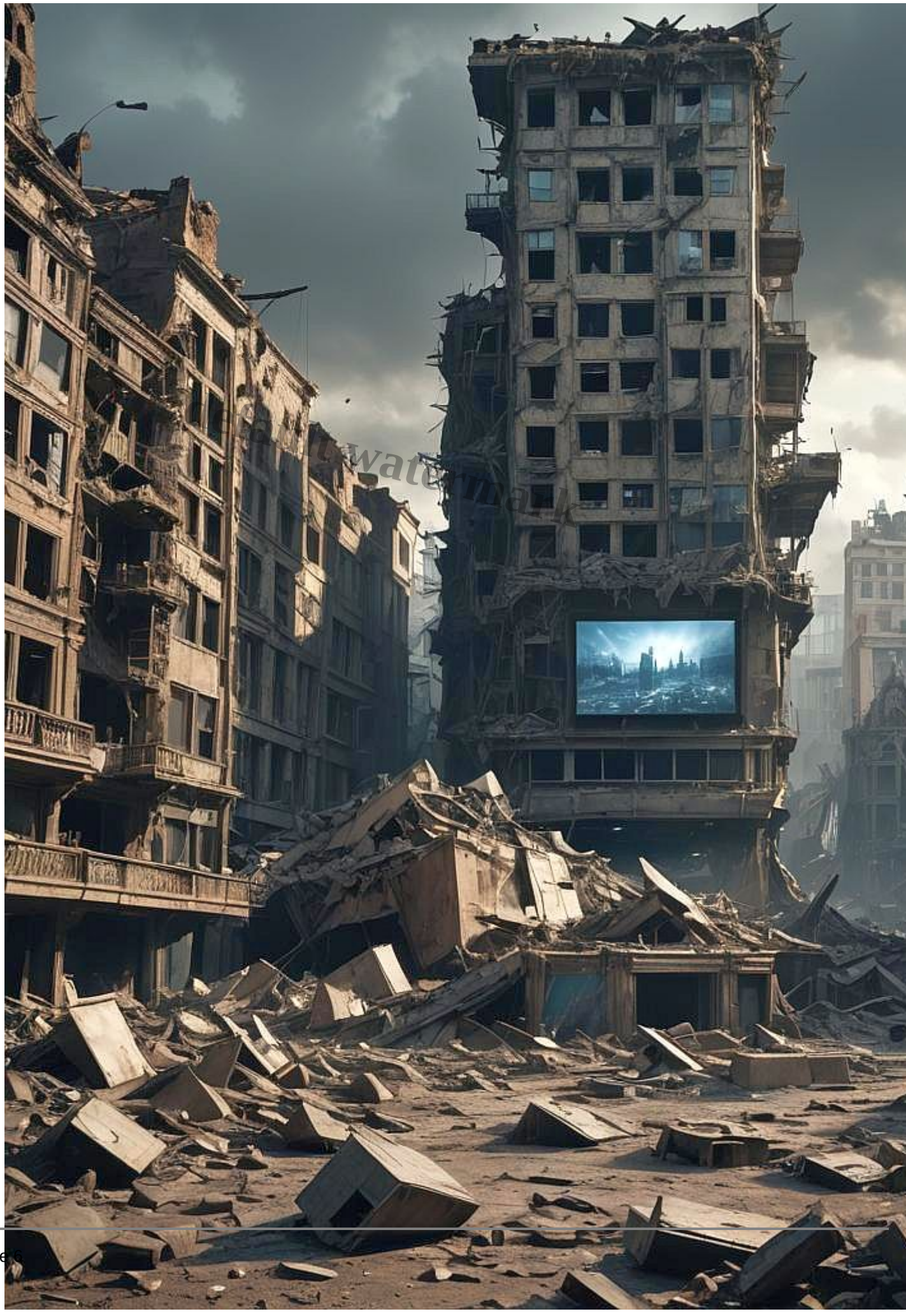
Om een â??â??kwantitatief inzicht te krijgen in desinformatie en de impact ervan op het informatie-ecosysteem, voerde een aparte groep onderzoekers van het *Observatory on Social Media* van de *Universiteit van Indiana* een onderzoek uit om de groeiende kwetsbaarheid van sociale media-bedrijven voor manipulatie door kwaadwillenden te onderzoeken. Hun onderzoek werd aangestuurd door de vraag: *hoe beÃ¯nvloeden manipulatietactieken de kwaliteit van informatie op socialemedianetwerken?*
In het resulterende artikel, getiteld â??*Quantifying the vulnerabilities of the online public square to adversarial manipulation tactics*â??, betogen ze dat â??*gebruikers van sociale media kwetsbaar zijn voor adversarial manipulation tactics, waarmee kwaadwillenden de blootstelling aan content die bijvoorbeeld democratische verkiezingen en de volksgezondheid bedreigt, kunnen vergroten.*â?•

Met behulp van een simulatiemodel genaamd *SimSoM* onderzocht het onderzoek tactieken die door kwaadwillenden worden gebruikt om de kwaliteit van informatie te verslechteren. De onderzoekers definiÃ«ren â??kwaadwillendenâ?• als die â??*accounts die worden aangestuurd door kwaadwillende (adversarial) actoren om content van lage kwaliteit te verspreiden onder authentieke agenten.*â?•
Dergelijke accounts kunnen worden aangestuurd door *mensen, bots* of *cyborgs*.
Aan de andere kant zijn er gebruikers die hoogwaardige content willen delen en consumeren. Het model beschouwt drie mogelijke manipulatietactieken:

- infiltratie,
- misleiding en
- overvloed (â??floodingâ?•).

Infiltratie verwijst naar hoe â??*slechte actoren de blootstelling aan hun berichten vergroten door authentieke accounts hen te laten volgen*â??, terwijl *flooding* ontstaat als slechte actoren gebruikers *spammen* met content van lage kwaliteit. Misleiding, tenslotte, verwijst naar de aantrekkingskracht van de content.

De onderzoekers beschouwen gevallen waarin informatie van lage kwaliteit een hoge misleidende aantrekkingskracht heeft op basis van hoe deze wordt gecommuniceerd.



Belangrijkste bevindingen

De studie concludeert dat *infiltratie de meest effectieve* manipulatietactiek is. Als de waarschijnlijkheid om het account van een slechte actor te volgen 10% is, â??*wordt de gemiddelde kwaliteit in het systeem teruggebracht tot minder dan de helft.*â?• Wanneer slechte actoren uitsluitend content genereren met een maximale misleidende aantrekkingskracht, wordt de kwaliteit teruggebracht tot ongeveer 70%. Als slechte actoren infiltratie combineren met misleiding of flooding, wordt de â?? *gemiddelde informatiekwaliteit teruggebracht tot 40%.*â?•
De auteurs ontdekten ook dat niet-authentieke accounts niet gericht hoeven te zijn op â??*invloedrijke* â?• accounts (accounts met veel volgers en veel activiteit). In plaats daarvan â??*kunnen ze meer schade aanrichten door verbinding te maken met willekeurige accounts.*â?• In tegenstelling tot wat algemeen wordt aangenomen, kunnen desinformatie-campagnes dus meer polarisatie creÃ«ren als de doelen willekeurig zijn. Zoals de auteurs opmerken, â??*is de kwaliteitsverdeling ongelijk, zodat de beoogde bevolking er slechter aan toe is, maar andere delen van de gemeenschap worden gespaard. Targetingtactieken hebben daarom de neiging averechts te werken als we ervan uitgaan dat kwaadwillenden van plan zijn om de verspreiding van hun content over de hele gemeenschap te maximaliseren*.â?•

Op dezelfde manier ontdekt het onderzoek ook dat het targeten van accounts waarvan bekend is dat ze misinformatie verspreiden, niet bijzonder effectief is. In dit geval ontdekten ze dat dergelijke gebruikers zich in een â??*echo chamber*â?• bevinden waarin â??*berichten van lage kwaliteit worden gedeeld en snel verouderd raken binnen Ã©Ã©n dicht verbonden partijdige cluster, waardoor de rest van het netwerk wordt gespaard*.â?•

Conclusie

Deze twee recente artikelen pakken de tegenwerking van desinformatieonderzoek op verschillende manieren aan en bieden belangrijke inzichten.
Het artikel van *Stephan Lewandowsky et al.* werpt licht op de bredere sociale en politieke dynamiek die onderzoek compliceert, en benadrukt de uitdagingen waarmee onderzoekers worden geconfronteerd en het belang van intentie bij het identificeren van desinformatie.
De kwantitatieve studie door onderzoekers aan de *Indiana University* biedt numerieke gegevens over de impact van desinformatie op het vertrouwen en gedrag van het publiek, en onthult in hoeverre desinformatiecampagnes democratische processen en de kwaliteit van online gedeelde informatie kunnen ondermijnen.
Samen bevestigen deze studies het belang van het bestuderen en beperken van desinformatie, vooral in een verkiezingsjaar voor de Verenigde Staten.

Datum aangemaakt
2024/08/17