Erasmus School of Social and Behavioural Sciences

Autoriteit online Terroristisch en Kinderpornografisch Materiaal

Van blootstelling aan online terroristische content tot daadwerkelijke mobilisatie: inzicht in vatbaarheid, risicobeoordeling en het onderbouwen van interventies

samenvatting

Stavroula Spernovasili, DominikaGrudzień, Neelav Novo Parashar
Erasmus Universiteit Rotterdam

Managementsamenvatting

De Autoriteit online Terroristisch en Kinderpornografisch Materiaal (ATKM) beoordeelt of onlinemateriaal voldoet aan de wettelijke criteria voor terroristische content op grond van Verordening (EU) 2021/784 en daarom mogelijk verwijderd moet worden. Binnen dit mandaat wordt in dit rapport onderzocht hoe gedragswetenschappelijk en psychologisch bewijs context kan bieden bij het identificeren van kenmerken van terroristische content die mogelijk een groter risico met zich meebrengt om radicalisering of de beweging naar daadwerkelijke mobilisatie te faciliteren. Het rapport heeft niet tot doel de individuen die dergelijk materiaal delen te beoordelen of te profileren. In plaats daarvan brengt het rapport in kaart hoe verschillende vormen en functies van content, waaronder ideologische uitingen, gewelddadige beelden, wervingsmateriaal, gegamificeerde deelname en operationele of tactische informatie, kunnen worden geïnterpreteerd en geprioriteerd.

Het rapport beschrijft allereerst hoe de hedendaagse onlineomgeving verschilt van de gecentraliseerde propagandasystemen die van oudsher in radicaliseringsonderzoek worden bestudeerd. Extremistische content wordt tegenwoordig vaak verspreid via gedecentraliseerde en participatieve ruimtes waarin humor, memes, verwijzingen naar games, competitie, sociale erkenning en interactie met anderen kunnen bijdragen aan betrokkenheid. Deze dynamiek is ook zichtbaar in losjes georganiseerde netwerken die escalerende daden van geweld of zelfbeschadiging belonen. Kinderen en jongeren kunnen in het bijzonder aan deze omgeving worden blootgesteld omdat ze in aanraking kunnen komen met gewelddadig materiaal via sociale media, videogames, streamingplatforms en algoritmische aanbevelingssystemen, ook zonder dat daar actief naar wordt gezocht.

Vervolgens beoordeelt het rapport het psychologische, gedragsmatige en neurologische bewijs met betrekking tot de blootstelling aan gewelddadige digitale content. Herhaaldelijke blootstelling kan bijdragen aan ongevoeligheid, verminderd empathisch reageren, morele ontkoppeling, agressieve priming en een meer tolerante houding ten aanzien van geweld. Sommige onderzoeken brengen blootstelling aan gewelddadige media eveneens in verband met toegenomen agressie en andere nadelige uitkomsten. Deze effecten zijn over het algemeen echter bescheiden, de bevindingen zijn niet geheel eenduidig en een groot deel van het bewijs is correlatief van aard of gebaseerd op laboratoriummetingen die geen weerspiegeling vormen van ernstig daadwerkelijk geweld. Het bewijs onderschrijft derhalve niet het beschouwen van blootstelling aan, of het delen van, gewelddadige content als een directe voorspellende factor van mobilisatie. Blootstelling dient veeleer te worden opgevat als één probabilistische invloed waarvan de effecten afhankelijk zijn van de wisselwerking tussen individuele kwetsbaarheden, sociale omstandigheden en beschermende factoren.

Het rapport onderzoekt vervolgens factoren die de vatbaarheid voor gegamificeerde extremistische omgevingen kunnen vergroten. Hiertoe behoren psychologische spanning, sociaal isolement, onvervulde behoeften aan verbondenheid en erkenning, identiteitsonzekerheid, sensatiezucht en beperkte steun vanuit het gezin of de gemeenschap. Gegamificeerde structuren kunnen betrokkenheid versterken door deelname, status, competitie en gedragsescalatie te belonen. Omgekeerd kunnen ouderlijke betrokkenheid, ondersteunende relaties, digitale geletterdheid, verbondenheid met de gemeenschap en toegang tot alternatieve bronnen van identiteit en zingeving de vatbaarheid verminderen. Deze factoren worden gepresenteerd als contextuele verklaringen en niet als een checklist of een basis voor de ATKM om uit de content psychologische kenmerken van een individu af te leiden.

Het operationele deel bestudeert hoe content en, waar wettelijk beschikbaar, bijbehorende gedragssignalen systematischer kunnen worden geïnterpreteerd. Daarin worden vier heuristische typen van het delen van content besproken: ideologisch-operationele content, statusgedreven netwerkcontent, performatieve shockcontent en incidenteel verder verspreide content.

Voortbouwend op dit onderscheid stelt het rapport een voorlopig raamwerk voor risicoclassificatie voor dat is opgebouwd rond vier dimensies: escalatie, netwerk, ideologie en capaciteit en operationeel gedrag. Operationele indicatoren, zoals tactisch leren, logistieke planning, wapenverwerving, doelwitselectie, verkenning of aanvalsvoorbereiding, wegen het zwaarst omdat zij de beweging richting mobilisatie mogelijk betrouwbaarder onderscheiden dan louter ideologische uitingen. Het raamwerk deelt content op in niveaus van minimaal, lager, gemiddeld of hoog risico ter ondersteuning van de prioritering, maar het blijft conceptueel in plaats van diagnostisch en vereist proefprojecten, toetsing van de betrouwbaarheid van interbeoordelaars, juridische waarborgen en kalibratie aan de hand van relevante casusresultaten.

Tot slot doet het rapport aanbevelingen op drie preventieniveaus. Primaire preventie richt zich op het bewustzijn van ouders, mediageletterdheid, digitale weerbaarheid en educatie voor jongeren. Secundaire preventie omvat op zoekopdrachten gebaseerde omleiding, interventies door geloofwaardige boodschappers, contranarratieven, actieve-inoculatiegames en doorverwijzing naar psychosociale hulpverlening door belanghebbenden met het juiste mandaat en toegang tot informatie op gebruikersniveau. Tertiaire en operationele maatregelen omvatten het prioriteren van content die operationele of tactische indicatoren bevat, het versterken van getrainde menselijke beoordeling van gecodeerd of ambigu materiaal, het inbouwen van technische frictie in het platformontwerp en het ondersteunen van platformoverschrijdende hash-uitwisseling voor bekende extremistische content. Met al deze aanbevelingen wordt beoogd de preventie en de prioritering van content te versterken, zonder ideologische of provocerende uitingen op zichzelf te beschouwen als bewijs dat een individu zich ontwikkelt in de richting van geweld.