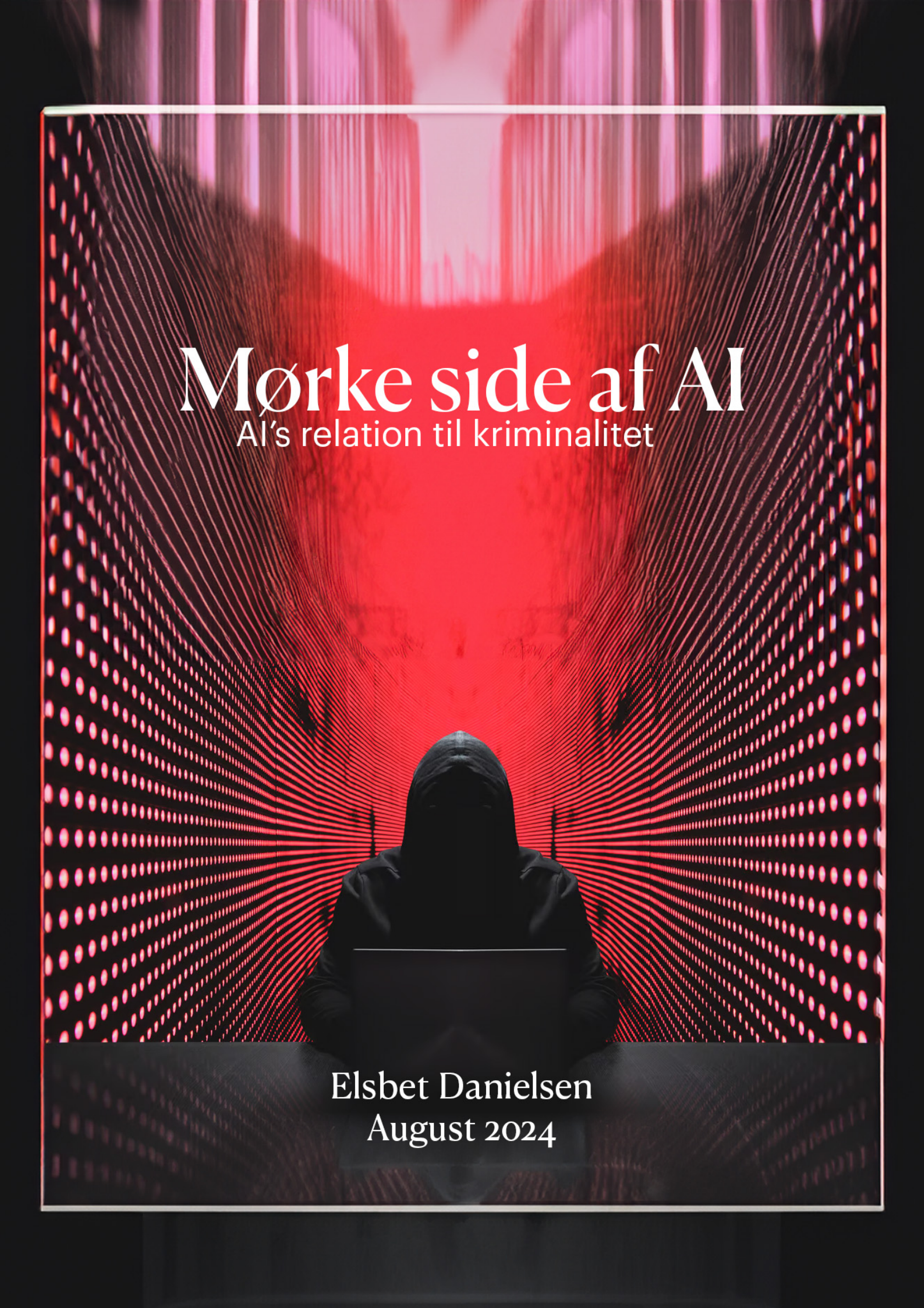


Mørke side af AI

AI's relation til kriminalitet

Elsbet Danielsen

August 2024

Cybermanipulation

Deepfake

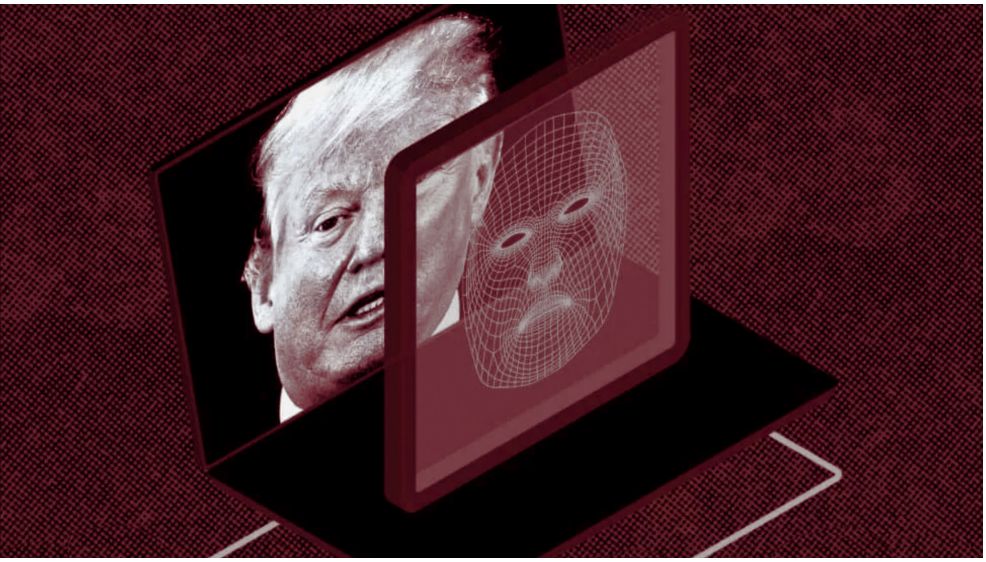
Du tænker nok; Hvad er deepfake? Jamen deepfake er den form for teknologi der, ved hjælp af kunstig intelligens, kan skabe billeder, videoer, stemmer, der forestiller virkeligheden, men aldrig har fundet sted. Du kan finde en persons ansigt på nettet, sagtens også finde deres stemmer et sted, og derfra snyde den persons nærmeste og få dem til at tro at det er den person der sender denne video og siger alle de ting. Kriminelle kan også deepfake en stemme og ringe til dig, og fremstå som et familiemedlem, din chef eller en tredje person som man har tillid til.

Deepfake startede ud med porno; man ser ofte at nogle kendisser får leaket en video af dem selv, men det viser sig så at det er en deepfake porno. De har fået sat sit ansigt på en pornoskuespillers krop, så det ser ægte ud og får alle til at tro det er den kendis. Det bliver også ofte brugt til hævnporno. Hvis man ikke er for glad for en person eller man bare er et usselt menneske og vil ødelægge nogens liv, så kan man bruge denne deepfake porno propaganda til at manipulere en hævnporno mod den valgte person.

AI-genereret og deepfaket seksuelt materiale, der bliver delt uden personens samtykke, er en skadelig og foruroligende form for misbrug. Dette kan være en metode for gerningsmanden til at ydmyge, nedgradere og umenneskeliggøre sine ofre. Disse seksuelt manipulerede deepfakes er primært rettede mod kvinder og piger, og det forstærker skadelige kønsstereotyper, hvilket fremmer kønsbaseret vold. Der er mange onlinefællesskaber på bland andet Discord, hvor der bliver oprettet private grupper hvor man så deler deepfakede nøgenbilleder af diverse kendte kvinder.

Alle disse pornovideoer og nøgenbilleder der bliver AI-genererede af helt uskyldige individer, er meget skamfuldt og særdeles grænseoverskridende for dem. Man bliver jo bange for om folk og deres nærmeste tror at det er ægte, og at det virkelig er dem der er nøgen på deres skærme.
Tilbage i 2021 blev 25 danske kvinder udsat for netop dette. Danske influencerer fik manipuleret deres instagrambilleder hvor de fjernede deres tøj og delte det så uden nogens samtykke. Det er selvfølgelig strafbart at lave og dele sådanne deepfakes, og det har det været i mange år. Man kan ende i fængsel i 6-7 år for det.

Men hvor kommer navnet Deepfake egentlig fra? Deepfake er en sammensætning af "deep learning" og "fake". Deep learning er en undergenre til "maskinlæring" hvor en maskine bliver programmeret til at huske ting, den har lært. Det der adskiller deep learning fra maskinlæring og AI er at den er skabt til at ligne en menneskehjerne; handle og tænke som en menneskehjerne. Så når deepfake bliver lavet, så er det denne hjerne der er ansvarlig. Det er dog ikke kun denne "AI-hjerne" der er alene ansvarlig, det er jo et menneske der finder på ideen og sætter det hele i gang.



Der er en meget kendt video der cirkulerer nettet. Det er videoen af Barack Obama der kalder Donald Trump en "total and complete dipshit". Det var komikeren Jordan Peele der manipulerede denne og den er selvfølgelig ikke ægte, men der er mange mennesker der troede den var. Det er faren ved deepfakes, der er desværre en masse mennesker der falder for det. Han lavede den for at demonstrere fremtidens "fake news" og hvor nemt det er at skabe og sprede falsk information.
Tv-værten Ditte Haue har også været ud for et deepfake bedrag. Der var blevet lavet en video om youtuberen MrBeasts nye app, hvor der siges at den app er lavet til at folk kan vinde store summe af penge. Ditte siger så at titusinde af mennesker har allerede hørt om den og hvor let det er at vinde penge med dens hjælp. Det

ligner et helt normalt TV2 indslag med Ditte Haue, men det er dog ikke virkelighed og kun lavet for at narre folk til at downloade denne app og snydes for alle deres penge.

AI Svindel

Sigende at være den første AI svindelsag, var da en gerningsmand lod som om han var en CEO af et energiselskab med base i Storbritanien. Han generede en deepfake-audio til at efterligne stemmen på den administrerende direktør for selvskabets Tysklands-baseret moderselvskab for at sætte en ulovlig pengeoverførsel i spil. De ringede til den CEO der havde base i Storbritanien og sagde det var meget vigtigt at de overførte $243,000 med det samme. Det skulle sendes til en leverandør i Ungarn, og at pengene nok skulle blive refunderet. Da pengene var blevet sendt så kørte de rundt til forskellige lokationer så gerningsmanden var sværere at identificere. De ringede en anden gang og bad om flere penge, men det sagde firmaet nej til, og blev mistænkelige. De ringede en trejde gang, men firmaet var allerede mistænkelige og sagde selvfølgelig nej. Heldigvis fik de erstatning for det da de var forsikrede imod AI-svindel.

Det nye indenfor cyberkriminalitet er deepfake-audio. Det er sværere at identificere og derfor kan kriminelle lave svindel der er sværere at opdage. Der er mange firmaer og kunder der bliver narret af dette, da det er så svært at skilne fra virkeligheden. En mor får et opkald fra sin datter som ringer fra et ukendt nummer og siger at hun har mistet sin telefon og kreditkort, og spørger om moren ikke kan overføre nogle penge til hendes venindes kort. Moren spørg om datteren er okay og om alt er vel, og datteren siger selvfølgelig ja, så moren overfører straks

pengene. Det viser sig så at det slet ikke var datteren, det var en deepfake-audio. Det er nemt at ringe til uskyldige individer og udgive sig for at være en af deres nærmeste, for hvorfor skulle de ikke stole på deres barn, partner eller bedste ven?

Der er blevet lavet en måling og det viser sig at seks ud af ti danskere er blevet udsat for svindel ved køb og salg på nettet. Ved hjælp af kunstig intelligens så er det blevet nemmere at lave falske e-mails, sms'er, hjemmesider, opslag, profiler mm., som ligner de ægte firmaer og salgspersoner, og det vil også gøre det meget sværere for menneskeheden i forhold til svindel og bedrag i fremtiden.

Digital Skurk

De forskellige svindelhandlinger har nogle begreber man kan sammensætte dem med. "Phishing" er når gerningsmanden prøver at franarre personlig information. "Spoofing" er når gerningsmanden udgiver sig for at være myndighederne, en virksomhed eller en privatperson, og sker over et falsk nummer eller e-mail. "Smishing" er over sms og "vishing" er over telefonopkald. Alt dette kan gå indenunder "social energieering" eller på dansk "social manipulation". Her er det overordnede at svindleren prøver på at manipulere personen så de kan få deres personlige informationer. Det er lighed med tricktyveri, i modsætning til traditionel hacking som kan sammenlignes med indbrud i computeren. I stedet for dirke låsen op, så manipulerer gerningsmanden personen til at lægge nøglen under måtten eller direkte åbne døren for dem. Svindleren spiller meget på personens følelser; vække nysgerrighed eller skabe tillid. Ofte vil der også være et komponent af forhastethed så der føles som om der ikke er tid til at dobbelttjekke noget og bare stole blindt på gerningsmanden.

Ransomware er en virus, som gerningsmanden planter, der krypterer ens data på computeren og låser alting, og derfra kræver en løsesum for du kan få det tilbage. Det er en af de største cybersikkerhedstrusler mod diverse firmaer. Det kan være store selskaber, hospitaler, såvel som privatpersoner. Truslen på sundhedsvæsenet er især alamerende. Det er en stor chikane for alle og enhver på hospitalet, hvis hele systemet lukker ned, og bliver holdt for løsesum. Dette kan være dødsensfarligt for de mennesker der er på hospitalet og har brug for hjælp. En sag tog sted i juli 2019 på Springhill Memorial Hospital i Alabama, USA. En mor var ved at føde sit barn, da et ransomware angreb begyndte. Alt udstyr slukkede og lægerne kunne ikke følge med i kvindens vitale tegn, eller fosterhjertemonitoren. Hospitalspersonalet var helt hjælpeløse imod dette skræmmende angreb. Lægerne lagde ikke mærke til at navlestrengen var viklet rundt om barnets hals og det resulterede i hjerneskade. Kvinden gav hospitalet skylden, og sagde at de skulle have advaret hende om hvad der skete. Barnet døde ni måneder efter. Mellem 2016 og 2021 var der 374 ransomware angreb på sundhedsvæsenet der afslørede 42 millioner menneskers personlige information. Der dør normalt 3 ud af hver 100 indlagte Medicare patienter. Under et ransomware angreb, så øger dette tal til 4 ud af hver 100. Mellem 2016 og 2021 forårsagede ransomware dødsfald på mellem 42 og 67 Medicare patienter.

AI bekæmper kriminalitet

Kunstig intelligens bliver i stigende grad brugt til at generere nøgenbilleder, lave online svindel mm. AI har ændret alt indenfor kriminalitet, men giver os også en mulighed for at bekæmpe disse trusler og kriminelle handlinger. I januar 2024 var der en sag i Indien, hvor politiet opklarede en blind sag vedrørende et

uidentificeret offer. De brugte AI til at genskabe offerets ansigt så de kunne identificere personen. Politiet satte plakater op og udgav et billede på deres officelle hjemmeside. Offeret viste sig at være en mand ved navn Hitendra. Hans bror så en plakat udenfor politistationen og genkendte ham med det samme. AI rekonstruerede Hitendras ansigt så hans øjne var åbne og han var mere genkendelig. Politiet fandt ud af at offeret havde været oppe at slås med to individer over en kvinde, to måneder før. Personerne, i en beruset tilstand, mødte Hitendra den 9. januar 2024, hvor de kvalte ham til døde. Det er blevet offentliggjort at fire personer, en kvinde og en taxichauffør er blevet anholdt for dette mord. AI kan bruges til at åbne blinde sager og identificere personer, som ikke kunne blive identificeret førhen.

Der er blevet lavet en AI-nødopkaldsoftware, designet til at hjælpe og løse vold i hjemmet. Systemet er blevet fodret flere års data omkring sådanne sager og er trænet til at genkende mønstre og finde effektive løsninger. Et andet fremskridt i AI er at den kan bruges til at identificere ukendte skyderier. Sensorer er installeret rundt omkring, og optager hvor og hvornår diverse skyderier tager sted. AI kan også lokalisere gerningsmanden. Data'en fra sensorer bliver herefter videresendt til politiet, hvor de så kan foretage en anholdelse.

http://localhost:5173/