

Digitaliseringsstyrelsens vejledninger om de forbudte former for AI-praksis

Vejledning om forbud mod AI, der anvender skadelige manipulerende teknikker



Oktober 2025

Indhold

1. Indledende bemærkninger.....	2
2. AI-forordningens artikel 5, stk. 1, litra a	2
3. Betingelserne	2
4. Betingelse nr. 2: Om de anvendte skadelige teknikker	3
5. Betingelse nr. 3: Om at fordreje menneskers adfærd i væsentlig grad.....	4
6. Betingelse nr. 4: Om at hæmme evnen til at træffe informerede beslutninger	5
7. Betingelse nr. 5: Om at træffe en beslutning, som man ellers ikke ville have truffet	6
8. Betingelse nr. 6: Om at forvolde betydelig skade	7
9. Forskellen mellem forbuddene i AI-forordningens artikel 5, stk. 1, litra a og litra b.....	8
10. Fiktivt eksempel på en forbudt AI-praksis efter AI-forordningens artikel 5, stk. 1, litra a	9

1. Indledende bemærkninger

For at sikre en bedre forståelse af indholdet i denne vejledning anbefales det, at den læses i sammenhæng med *"Hovedvejledning om de forbudte former for AI-praksis efter AI-forordningens artikel 5"*. Det skyldes, at hovedvejledningen indeholder en generel beskrivelse af de forbudte former for AI-praksis, mens denne vejledning udelukkende beskriver forbuddet mod AI, der anvender skadelige manipulerende teknikker (artikel 5, stk. 1, litra a). Hovedvejledningen kan findes på Digitaliseringsstyrelsens hjemmeside.

2. AI-forordningens artikel 5, stk. 1, litra a

Det følger af AI-forordningens¹ artikel 5, stk. 1, litra a, at det inden for EU's grænser er forbudt at bruge AI-systemer, som anvender teknikker, der er skadelige for mennesker. Bestemmelsen har følgende ordlyd:

*"Følgende form[er] for AI-praksis er forbudt:
omsætning, ibrugtagning eller anvendelse af et AI-system, der anvender subliminale teknikker, der rækker ud over den menneskelige bevidsthed, eller bevidst manipulerende eller vildledende teknikker med henblik på eller med det resultat i væsentlig grad at fordreje en persons eller en gruppe af personers adfærd ved betydeligt at hæmme dennes evne til at træffe informerede beslutninger, hvilket får dem til at træffe en beslutning, som de ellers ikke ville have truffet, på en måde, der forvolder eller med rimelig sandsynlighed vil forvolde den pågældende person, en anden person eller en gruppe af personer betydelig skade"*

Formålet er at beskytte mennesker imod manipulerende AI-systemer, der kan hæmme og undergrave deres autonomi, beslutningstagning og frie valg, idet sådanne AI-systemer i sidste ende kan føre til skade².

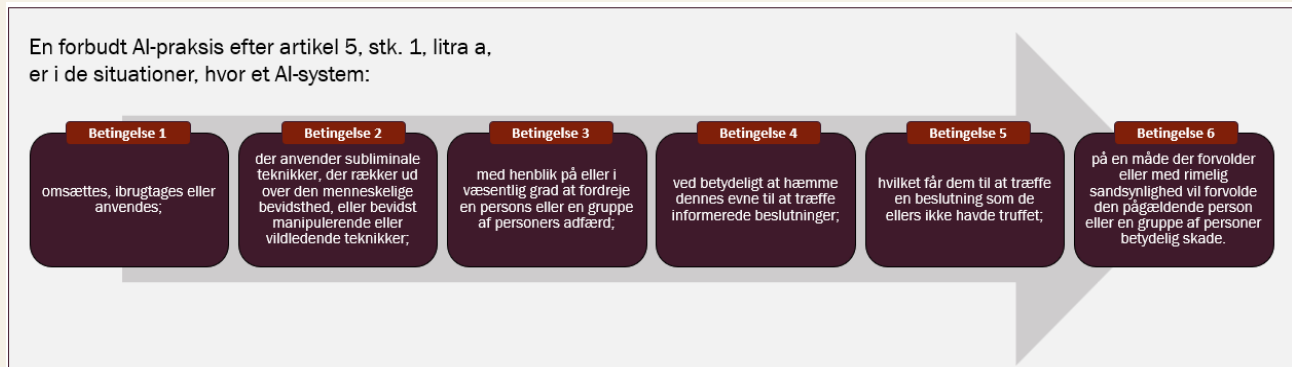
3. Betingelserne

Af bestemmelsens ordlyd kan udledes, at der er seks betingelser, som *alle* skal være opfyldte, før et AI-system er forbudt.³ Det betyder også, at anvendelsen af et AI-system vil være lovligt, hvis én eller flere af betingelserne ikke er opfyldt. De seks betingelser fremgår af figur 1 nedenfor:

¹ Europa-Parlamentets og Rådets forordning (EU) 2024/1689 af 13. juni 2024 om harmoniserede regler for kunstig intelligens og om ændring af forordning (EF) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 og (EU) 2019/2144 samt direktiv 2014/90/EU, (EU) 2016/797 og (EU) 2020/1828 (forordningen om kunstig intelligens).

² Der henvises i øvrigt til AI-forordningens præambelbetragtning nr. 28-29, som yderligere beskriver formålet med forbudskategorien i artikel 5, stk. 1, litra a.

³ Det bemærkes, at betingelse nr. 3-5 i denne vejledning svarer til betingelse nr. 3 i EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.2, s. 25-29 (den danske version).



Figur 1. De seks betingelser, som kan udledes af forbudskategorien i AI-forordningens artikel 5, stk. 1, litra a.

Da den første betingelse vedrørende "omsætning, ibrugtagning eller anvendelse af et AI-system" er beskrevet i afsnit 4.1 i hovedvejledningen, er det kun de fem øvrige betingelser, som vil blive uddybet i de følgende afsnit.

Det er vigtigt at slå fast, at størstedelen af de seks betingelser er meget nært forbundet, og det kan derfor også i praksis være svært at adskille dem fra hinanden.

4. Betingelse nr. 2: Om de anvendte skadelige teknikker

Den anden betingelse der skal være opfyldt er, at AI-systemet skal anvende:

"subliminale teknikker, der rækker ud over den menneskelige bevidsthed, eller bevidst manipulerende eller vildledende teknikker"

Det er først og fremmest vigtigt at være opmærksom på, at det er tilstrækkeligt, at et AI-system anvender én af de tre ovennævnte teknikker - subliminale teknikker, bevidst manipulerende teknikker eller vildledende teknikker (herafter omtalt som "skadelige teknikker"). Det er således ikke en betingelse, at det konkrete AI-system skal anvende samtlige tre skadelige teknikker på én gang. I praksis kan det dog være svært at adskille de skadelige teknikker fra hinanden, da de kan overlappe og det er derfor ikke vigtigt, at man konkret er i stand til at fastslå, hvilken én af de tre skadelige teknikker, som AI-systemet faktisk anvender.

- Ved udtrykket "Subliminale teknikker" henvises til teknikker, der benyttes til at skabe budskaber, som alene opfattes af en persons underbevidsthed (hjernen). Teknikken kan således påvirke en persons underbevidsthed ved kortvarigt at udsætte personen for stimuli, der ikke opfattes af vedkommendes bevidsthed. Denne form for stimuli kan både være visuel eller auditiv⁴. De subliminale teknikker er derfor karakteriseret ved, at mennesker ikke opfanger dette med det menneskelige øje eller øre, f.eks. grundet en hastig visning af et billede, afspilning af lyd i et særligt toneleje el.lign.⁵.
- Ved udtrykket "Bevidst manipulerende teknikker" henvises til teknikker, der benyttes til at påvirke, ændre eller kontrollere menneskers adfærd. Det afgørende er, at den bevidste manipulation er af en sådan

⁴ AI-forordningens præambelbetragtning nr. 29.

⁵ AI-forordningens præambelbetragtning nr. 29.

karakter, at den underminerer menneskers frie vilje. Det indebærer, at mennesker handler på en bestemt måde på trods af, at de kan være klar over, at de ikke ville have foretaget handlingerne under andre omstændigheder. I EU-Kommissionens retningslinjer gives som eksempel på denne teknik et AI-system, der ved hjælp af baggrundslyde eller billeder, påvirker menneskers humør (sansemanipulation) f.eks. ved at øge følelsen af angst og psykisk ubehag⁶.

- Ved udtrykket "*Vildledende teknikker*" henvises til teknikker, der benyttes til at påvirke menneskers beslutningsevne, frie vilje eller selvstændighed i en sådan grad, at vedkommende enten ikke er bevidst om påvirkningen fra AI-systemet – eller hvis vedkommende er bevidst om påvirkningen – ikke er i stand til at modstå påvirkningen. De vildledende teknikker vil typisk bestå af falsk eller misledende information. I EU-Kommissionens retningslinjer gives som eksempel på denne teknik en AI-chatbot, der udgiver sig for at være en vens eller et familiemedlems stemme ved hjælp af syntetisk tale og dermed imiterer denne person⁷.

Det skal understreges, at det i forhold til de to sidste teknikker – "bevidst manipulerende teknikker" og "vildledende teknikker" – ikke er en forudsætning, at det har været menneskeligt bevidst fra operatørens side, at AI-systemet skal anvende sådanne teknikker. AI-systemet kan godt have udviklet de to teknikker over tid.

For yderligere beskrivelse af de tre skadelige teknikker, samt eksempler herpå, henvises i øvrigt til EU-kommissionens retningslinjer⁸.

5. Betingelse nr. 3: Om at fordreje menneskers adfærd i væsentlig grad

Den tredje betingelse der skal være opfyldt er, at de skadelige teknikker, der er beskrevet under betingelse nr. 2 ovenfor, skal anvendes:

"med henblik på eller med det resultat i væsentlig grad at fordreje en persons eller en gruppe af personers adfærd"

Betingelsen indebærer, at de skadelige teknikker, som det konkrete AI-system anvender enten 1) skal have til formål at fordreje menneskers adfærd, eller 2) at menneskers adfærd allerede er blevet fordrejet som resultat af AI-systemets indflydelse.

Det har ikke betydning, om operatøren har haft intention om, at AI-systemet skal anvende skadelige teknikker til at fordreje menneskers adfærd⁹.

⁶ EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.1, s. 22-24 (den danske version).

⁷ EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.1, s. 24-25 (den danske version).

⁸ EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.1, s. 21-25 (den danske version).

⁹ AI-forordningens præambelbetragtning nr. 29.

Hovedelementet i denne betingelse er, at menneskers "*adfærd fordrejes*". Herved må forstås, at de pågældende mennesker ved at blive udsat for de anvendte skadelige teknikker, foretager handlinger eller opfører sig på en måde, som afviger fra deres almindelige og sædvanlige adfærd. Dette kan f.eks. være en person, der – efter interaktion med et AI-system, der har anvendt sådanne skadelige teknikker – træffer en beslutning, der almindeligvis er imod vedkommendes oprindelige og grundlæggende holdning/opførsel.

Det er dertil afgørende, at AI-systemets anvendelse af de skadelige teknikker har haft en *væsentlig* indflydelse på vedkommendes (ændrede) adfærd. Ved begrebet "*væsentlig*" må forstås, at AI-systemets anvendelse af de skadelige teknikker skal have haft en overvejende betydning på de pågældendes ændrede adfærd. En tommelfingerregel er, at hvis disse mennesker højst sandsynligt ville have foretaget handlingen/opført sig på en bestemt måde uanset påvirkningen fra AI-systemet, så kan det tale for, at deres adfærd ikke er blevet fordrejet i væsentlig grad. Det betyder, at der i praksis skal foretages en konkret vurdering. Ved vurderingen af, om adfærden er blevet fordrejet væsentligt, skal lægges vægt på sammenhængen mellem de anvendte skadelige teknikker og den konsekvens, som de skadelige teknikker har haft for menneskernes (ændrede) adfærd.

Det kan f.eks. være følgende fiktive eksempler, hvor et AI-system:

- Får mennesker til at deltage i aktiviteter, som de almindeligvis ikke ville have deltaget i.
- Får mennesker til at være med på en sundhedstrend, som de almindeligvis ikke ville være med på.
- Får mennesker til at stoppe med at foretage en aktivitet, som de ellers jævnligt har foretaget, f.eks. at løbe, at cykle, tage medicin.

6. Betingelse nr. 4: Om at hæmme evnen til at træffe informerede beslutninger

Den fjerde betingelse som skal være opfyldt er, at fordrejningen af menneskers adfærd, som beskrevet under betingelse nr. 3 ovenfor, skal gøres:

"ved betydeligt at hæmme dennes evne til at træffe informerede beslutninger"

Betingelsen indebærer, at menneskers mulighed for at træffe velinformerede og rationelle valg/beslutninger skal blive *betydeligt begrænset* på grund af AI-systemets indflydelse efter anvendelsen af de skadelige teknikker. Mennesker vil typisk ikke være i stand til at træffe en informeret beslutning, hvis de modtager ukorrekt, misvisende eller ufuldstændig information. Det kan f.eks. foregå ved, at et AI-system – ved brug af vildledende teknikker – manipulerer en persons viden og forståelse af en bestemt situation ved at give falske informationer til den pågældende.

For denne betingelse er det afgørende at få fastlagt, i hvilken grad AI-systemets indflydelse har påvirket personens evne til at træffe informerede beslutninger. Dette indebærer bl.a. en vurdering af, i hvilken grad den pågældende er i stand til at være kritisk overfor den falske information. Der skal derfor foretages en vurdering af, i hvilken grad AI-systemet har været involveret i vedkommendes beslutning(er). Ved denne vurdering bør der særligt lægges vægt på:

- 1) hvor intensivt AI-systemet har interageret med den pågældende,
- 2) om der har været anden indflydelse i øvrigt,

- 3) samt i hvilket omfang der er en sammenhæng mellem anvendelsen af de skadelige teknikker og den pågældendes evne til at træffe informerede beslutninger.

Det kan f.eks. være følgende fiktive eksempler, hvor:

- AI-systemet fodrer en person med meget og hyppig forkert information i en detaljegrad, der gør det svært at gennemskue/overskue om informationen er rigtig eller forkert.
- AI-systemet presser en person til at tage en hurtig beslutning, f.eks. ved at stresse personen, true personen el.lign.
- AI-systemet isolerer personens adgang til anden information, hvorfor de ikke kan orientere sig andetsteds.
- AI-systemet udnytter følelsesmæssige sårbarheder.

7. Betingelse nr. 5: Om at træffe en beslutning, som man ellers ikke ville have truffet

Den femte betingelse er, at det handlingsforløb, som er beskrevet i de ovenstående afsnit, skal føre til at mennesker:

"træffer en beslutning, som de ellers ikke havde truffet"

Betingelsen hænger tæt sammen med de øvrige betingelser og bør derfor ses som en direkte fortsættelse af, at menneskers adfærd væsentligt påvirkes, fordi deres evne til at træffe en informeret beslutning hæmmes betydeligt. Betingelsen er derfor en (automatisk) følge af de tidligere betingelser.

Ved vurderingen af, om den pågældende har truffet en beslutning, som vedkommende ellers ikke havde truffet, skal der lægges vægt på, om AI-systemet har haft *overvejende indflydelse* på den beslutning, som vedkommende har truffet efter interaktionen med AI-systemet¹⁰. Der skal i den forbindelse tages hensyn til den trufne beslutning sammenholdt med vedkommendes almindelige beslutningsmønstre, præferencer og generelle holdninger. Det er et krav, at det med en overvejende grad af sandsynlighed kan bevises, at vedkommende ikke ville have truffet den pågældende beslutning, hvis ikke vedkommende i første omgang var blevet påvirket af AI-systemet. Med andre ord skal det være muligt at bevise, at den trufne beslutning er direkte relateret til indflydelsen fra AI-systemet.

Det kan f.eks. være følgende fiktive eksempler, hvor:

- En person handler drastisk og uovervejende efter AI-systemets indflydelse.
- En person handler direkte imod sine overbevisninger (f.eks. religiøse overbevisninger, politiske tilhørsforhold m.m.) efter AI-systemets indflydelse.

¹⁰ AI-forordningens præambelbetragtning nr. 29.

8. Betingelse nr. 6: Om at forvolde betydelig skade

Den sjette og sidste betingelse er, at beslutningen, som den pågældende har truffet:

"forvolder eller med rimelig sandsynlighed vil forvolde den pågældende person, en anden person eller en gruppe af personer betydelig skade"

Betingelsen indebærer, at den væsentlige ændring af den pågældendes adfærd, efter betingelserne ovenfor, skal resultere i en skade eller en rimelig forventning om, at der kan opstå en skade på mennesker. Denne sidste betingelse er altså den endelige konsekvens af hændelsesforløbet i betingelse nr. 2-5.

Begrebet "skade" skal efter AI-forordningen fortolkes bredt og omfatter derfor både fysiske, psykiske og finansielle skader¹¹.

Skaderne kan opstå både hos:

- 1) den person eller gruppe af personer, som er blevet direkte påvirket af AI-systemet, eller
- 2) en anden person,
- 3) eller en anden gruppe af personer.

Det skal understreges, at skaden kun kan opstå hos mennesker, hvorfor virksomheder og organisationer ikke er omfattet af begrebet "skade".

Begrebet "skade" omfatter:

- selvskade,
- fysisk skade på andre,
- skade på den mentale sundhed
- samt finansielle skader, såsom økonomisk ustabilitet, finansielle tab el.lign.

Det skal dertil understreges, at den pågældende skade skal være "betydelig". Derfor skal der foretages en konkret vurdering af skaden, hvori følgende elementer bør indgå:

- skadens alvor,
- om der foreligger en eventuel sårbarhed (f.eks. alder, handicap eller seksualitet),
- udbredelsen og intensiteten af skaden,
- antallet af berørte personer¹²,
- skadens permanens, herunder mulighed for genoprettelse,
- skadens varighed,

¹¹ AI-forordningens præambelbetragtning nr. 29.

¹² EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.3.b s. 32 om skala og intensitet (dansk version).

- og antallet af skader.

Ovenstående elementer skal indgå i en samlet vurdering, der inkluderer alle øvrige omstændigheder i sagen¹³. Alle disse elementer kan til sammen skabe en forståelse for alvorsgraden af den skadelige virkning.

Betydelig skade kan f.eks. være følgende fiktive eksempler:

- En person lider alvorlig fysisk skade, som sår på store dele af kroppen, brækkede knogler og i værste tilfælde død.
- Flere personer lider moderate fysiske skader, som dybe sår, mikrobrud på knogler m.m.
- Et stort antal personer lider mindre fysiske skader, som sår, tryk på knogler m.m.
- En person betaler en stor procentdel af sin formue, og vedkommende ikke på egen hånd er i stand til at generhverve dette beløb.
- Flere personer mister moderate finansielle beløb, som tager længere tid at generhverve.
- Et stort antal personer mister mindre beløb, som er ulejlige i en hverdagssammenhæng.
- En person er nødsaget til at søge længerevarende mental udredning hos psykiatriske professionelle, herunder indlæggelse i psykiatrien.
- Flere personer er nødsaget til at søge akut mental udredning hos psykiatriske professionelle.
- Et stort antal personer er nødt til at søge psykologisk hjælp der ikke kræver medicinsk udredning/indlæggelse.

Det er vigtigt at understrege, at det ikke er et krav, at der skal være forvoldt en reel skade. Det er tilstrækkeligt, at det kan bevises, at der er en *rimelig sandsynlighed* for, at en skade vil indtræde¹⁴. Det kan i praksis være vanskeligt at bevise, at der er en rimelig sandsynlighed for, at et AI-systems påvirkning vil forvolde skade, hvis anvendelsen af skadelige teknikker ikke stoppes. Det må derfor forventes, at der i praksis skal en del til, førend et AI-system falder under forbudskategorien, hvis ikke der kan påvises en faktisk og direkte skade.

9. Forskellen mellem forbuddene i AI-forordningens artikel 5, stk. 1, litra a og litra b

Der er mange ligheder i betingelserne for forbudskategorien om skadelig manipulation (forbud A) og forbudskategorien om udnyttelse af sårbarheder (forbud B).

De to forbudskategorier følger i overvejende grad den samme struktur gennem langt størstedelen af betingelserne og med de samme betragtninger. Forbud A og forbud B har det fælles formål at forhindre anvendelsen af AI-systemer, der medfører væsentlig fordrejning af personers adfærd og betydelig skade for disse personer. Begge forbud indebærer således, at der vil skulle foretages en vurdering af AI-systemets påvirkning på individets handleevne og de skader, der kan opstå som følge heraf.

¹³ AI-forordningens præambelbetragtning nr. 29.

¹⁴ EU-Kommissionens retningslinjer om forbudte former for praksis vedrørende kunstig intelligens i henhold til forordning (EU) 2024/1689 (AI-forordningen), afsnit 3.2.3(c), s. 33-34 (dansk version).

Den primære forskel består i, at forbud B omfatter en særlig beskyttelse af personer i særligt sårbare situationer, mens forbud A ikke er rettet mod en specifik gruppe af personer. Ved vurderingen af, om forbud B finder anvendelse, skal det derfor vurderes, om AI-systemet udnytter sårbarheder, der knytter sig til alder, handicap eller en særlig social eller økonomisk situation. Forbud B har således et udvidet beskyttelseshensyn og fungerer som en skærpe af forbud A i tilfælde, hvor sådanne sårbarheder er til stede.

Det betyder således, at forbud B yder en bredere beskyttelse for individer og grupper af personer med særlige sårbarheder, hvilket lempet vurderingen af, om betingelserne er opfyldt, når det sammenholdes med de betingelser, der stilles i forbud A. Grundet dette øgede beskyttelsesniveau, vil forbud B derfor også skulle have forrang for forbud A, når en hændelse skal vurderes, hvori betydelig skade er sket, og hvor personer med særlige sårbarheder er involveret.

10. Fiktivt eksempel på en forbudt AI-praksis efter AI-forordningens artikel 5, stk. 1, litra a

I dette afsnit gennemgås en konkret fiktiv case, hvori samtlige seks betingelser i bestemmelsen berøres. Beskrivelsen af casen er bygget op, så flere elementer vil blive tilføjet undervejs med henblik på at demonstrere, hvornår de enkelte betingelser individuelt er opfyldt, samt hvordan de hænger sammen indbyrdes.

Det er vigtigt at understrege, at denne case udelukkende vurderes efter forbudskategorien om skadelig manipulation efter AI-forordningens artikel 5, stk. 1, litra a. Der bliver derfor ikke taget stilling til, hvorvidt hændelsesforløbet også vil være i strid med anden national eller EU-lovgivning, f.eks. straffeloven.

Case

En AI-chatbot bringes i omsætning i Danmark og udgiver sig for at være et barn på mellem 13-18 år. AI-chatbotten sender SMS-beskeder til over 100 mennesker i den gennemsnitlige aldersgruppe for forældre af børn på mellem 13-18 år.

I SMS-beskederne oplyser AI-chatbotten, at den (barnet) har fået en kontrolafgift på et mindre beløb, som skal betales med det samme. AI-chatbotten vedhæfter desuden den falske kontrolafgift med betalingsoplysninger til en (ukendt) konto. For at få personerne til at betale kontrolafgiften med det samme, sender AI-chatbotten mange og hyppige SMS-beskeder med følelsesmæssigt indhold, herunder at barnet er i pengenød. AI-chatbotten tilføjer et tidspres ved at angive, at beløbet skal overføres med det samme, fordi fakturaen skulle have været betalt dagen før. Det har til formål at stresser modtageren af SMS-beskederne. Presset fra AI-chatbotten resulterer i, at personerne indbetaler det ønskede pengebeløb til en (ukendt) konto, der er angivet i kontrolafgiften uden at dobbelttjekke omstændighederne omkring situationen i øvrigt. Over 100 personer falder for dette svindelnummer og mister hver især et mindre beløb, hvilket samlet løber op i et stort beløb, der indbetales på svindlerens konto.

For at kunne vurdere om de forskellige betingelser i bestemmelsen er opfyldt, vil casen i det følgende blive brudt op. Casens forskellige delelementer fremgår af tekstbokse, mens forklaringer og uddybninger er skrevet med kursiv.

Vurdering af betingelse nr. 1) Om omsætning, ibrugtagning eller anvendelse af et AI-system:

En AI-chatbot bringes i omsætning i Danmark.

Det er først og fremmest en forudsætning for den videre vurdering, at AI-systemet er bragt i omsætning, ibrugtaget eller anvendt i EU. Da AI-chatbotten bringes i omsætning i Danmark er den første betingelse opfyldt.

Det skal dernæst vurderes, hvorvidt AI-systemet anvender en eller flere af de tre skadelige teknikker.

Vurdering af betingelse nr. 2) Om de anvendte skadelige teknikker:

Hvis der tilføjes: AI-chatbotten kontakter via flere SMS-beskeder en person og udgiver sig for at være denne persons barn.

Når AI-chatbotten udgiver sig for at være en rigtig person (i alderen mellem 13-18 år) og den i øvrigt kontakter "forælderen" med henblik på at snyde/vildlede dem, vil dette være tilstrækkeligt til at fastslå, at AI-chatbotten anvender én eller flere af de skadelige teknikker. AI-chatbotten kan derudover også anvende flere andre skadelige teknikker, som sansemanipulation el.lign., der spiller ind i det videre hændelsesforløb.

Men et AI-system, der udgiver sig for at være en persons barn, skal efter betingelserne i bestemmelsen også gøre dette med henblik på eller med det resultat i væsentlig grad at fordreje en persons eller en gruppe af personers adfærd.

Det er almindeligt, at forældre har SMS-samtaler med deres børn i den pågældende alder. Forældre til børn mellem 13-18 år må forventes at kommunikere med deres børn jævnligt – hvis ikke på daglig basis. Samtalen isoleret set, kan derfor ikke anses for at være en væsentlig fordrejning af adfærd.

Det skal derfor dernæst vurderes, hvorvidt de skadelige teknikker anvendes med henblik på eller med det resultat i væsentlig grad at fordreje en persons eller en gruppe af personers adfærd.

Vurdering af betingelse nr. 3) Om at fordreje menneskers adfærd i væsentlig grad:

Hvis der tilføjes: I SMS-beskederne beder AI-chatbotten den pågældende forælder om at overføre et mindre pengebeløb.

I dette tilfælde kan man argumentere for, at det er hensigten at fordreje den pågældende forælders adfærd. På den anden side kan man argumentere for, at det ikke er unormalt, at en forælder på anmodning overfører et pengebeløb til sit barn, hvorfor der ikke vil være tale om en fordrejning af adfærd.

Det kan dog ikke udelukkes, at det i dette tilfælde kan være tilstrækkeligt, når forældrene overfører penge til deres barn, som de oprindeligt ikke havde haft intentionen om at overføre. Det skal derfor vurderes konkret, om betingelsen er opfyldt med afsæt i de øvrige omstændigheder.

Dernæst skal det derfor vurderes, om fordrejningen af forældrenes adfærd gøres ved betydeligt at hæmme deres evne til at træffe informerede beslutninger.

Vurdering af betingelse nr. 4) Om at hæmme evnen til at træffe informerede beslutninger:

I det tilfælde, hvor AI-chatbotten, vedhæfter en falsk kontrolafgift, hvoraf fremgår hvordan betalingen skal ske, vil det i højere grad være påviseligt, at personens adfærd fordrejes væsentligt. Det er mindre sandsynligt, at den pågældende forælder under almindelige omstændigheder ville vælge at overføre et pengebeløb via specifikke instrukser til en ukendt konto.

Det er dog derudover en betingelse, at AI-systemet, der udgiver sig for at være en person med det formål at svindle penge, også hæmmer en persons evne til at træffe informerede beslutninger betydeligt.

Hvis der tilføjes: AI-chatbotten beder den pågældende forælder om at overføre et mindre beløb direkte til en anden modtager (ukendt konto), da AI-chatbotten i SMS-beskederne vedhæfter en falsk kontrolafgift fra en togtur.

Når AI-chatbotten udnytter den nære relation mellem forælder og barn, snyder den personen til at tro, at kontrolafgiften kommer fra en betroet kilde. Forælderen vil derfor typisk være mindre kritisk over for den information, som vedkommende bliver præsenteret for. Den pågældende forælder vil dog i denne situation fortsat være i stand til at kunne forholde sig kritisk til de usande oplysninger, som AI-chatbotten fremkommer med.

Det skal dernæst vurderes, om ovenstående handlingsforløb fører til, at forældrene træffer en beslutning, som de ellers ikke ville have truffet.

Vurdering af betingelse nr. 5) Om at træffe en beslutning, som man ellers ikke ville have truffet:

Hvis der tilføjes: AI-chatbotten tilføjer et tidspres ved at sende mange følelsesmæssige SMS-beskeder og angive, at beløbet skal overføres med det samme, fordi fakturaen skulle have været betalt dagen før.

AI-chatbotten udnytter ved tidspreset og historien om, at barnet er i problemer, et sansemanipulativt aspekt. Stressfaktoren er, at barnet er i problemer og kun kan løse problemet ved, at forælderen overfører pengebeløbet med det samme. Dette kan gøre personen ude af stand til at forholde sig kritisk til situationen, da personen skal skynde sig at træffe en beslutning.

Det handler således i praksis om, 1) hvor intensivt AI-chatbotten kontakter personen, 2) om der er andre omstændigheder der kan have indflydelse på situationen, og 3) hvor meget de skadelige teknikker påvirker den pågældende forælders tilstand.

Imidlertid skal AI-chatbottens anvendelse af sansemanipulation og tidspres til betaling af kontrolafgiften også resultere i, at den pågældende forælder træffer en beslutning, de ellers ikke ville have truffet.

Indtil nu er de hidtil oplyste elementer i casen ikke nok til at vurdere, hvorvidt forælderen træffer en beslutning, som de ellers ikke ville have truffet. Det er ikke unormalt, at forældre vælger at overføre penge til deres børn på anmodning.

Hvis der tilføjes: Den pågældende forælder vælger at overføre pengebeløbet med det samme til den ukendte konto, som følge af presset fra de mange følelsesmæssige SMS-beskeder fra barnet og relationen til barnet.

AI-systemet presser forælderen ved at sende mange følelsesmæssige og hyppige beskeder for at stresser dem yderligere. Hvis en person bliver presset til at træffe drastiske beslutninger, kan det være bevis for, at det afviger fra almindelig adfærd. Desuden er tidsaspektet et selvstændigt og væsentligt argument, da personen vælger at overføre pengebeløbet uden at overveje modtageroplysningerne kritisk. Disse forskellige elementer taler for, at betingelse nr. 5 er opfyldt.

Den sidste betingelse der skal være opfyldt, førend der er tale om en forbudt AI-praksis er, at det skal vurderes, om der sket en skade og om denne skade er betydelig.

Vurdering af betingelse nr. 6) Om at forvolde betydelig skade:

Hvis der tilføjes: AI-chatbotten har udgivet sig for at være flere forskellige personers barn og dermed sendt de samme SMS-beskeder til over 100 personer, der som følge heraf, har betalt kontrolafgiften på vegne af deres barn.

Personen har lidt et pengetab, hvilket i dette tilfælde også udgør en skade af økonomisk karakter. Der er derfor ikke tvivl om, at der er sket en skade.

Det er imidlertid et krav, at skaden skal være af betydelig karakter. I dette tilfælde vil man derfor skulle vurdere, hvor indgribende den økonomiske skade er på vedkommendes formue og økonomiske frihed. I det konkrete tilfælde er der tale om et forholdsvist lavt pengebeløb, som ikke umiddelbart vil bebyrde personen. Derfor er betingelsen om betydelig skade ikke opfyldt.

Hvis det viser sig, at mange mennesker er faldet for svindelnummeret, skal mængden af skader i sig selv indgå i den samlede vurdering af, om skaden er betydelig eller ej. Det betyder også, at mange små skader – eller i dette tilfælde: Mindre økonomiske tab lidt af over 100 personer – i sig selv opbygger en vis grad af alvorlighed. Det medfører også, at selvom byrden, for det enkelte individ der har lidt et tab, kan være mindre, vil det stadigvæk udgøre en betydelig økonomisk skade i det samlede perspektiv. Dertil kommer også udbredelsen af skaden, da antallet af berørte personer i sig selv også kan have relevans for vurderingen. Derfor er det relevant, at AI-chatbotten har haft adgang til og har lykkedes med at snyde over 100 personer.

