
Secify

Dataskydd & AI för vårdsektorn

Eller: en introduktion för AI-körkortet

Av Joakim Söderberg

Grafisk profil

Innehåll:

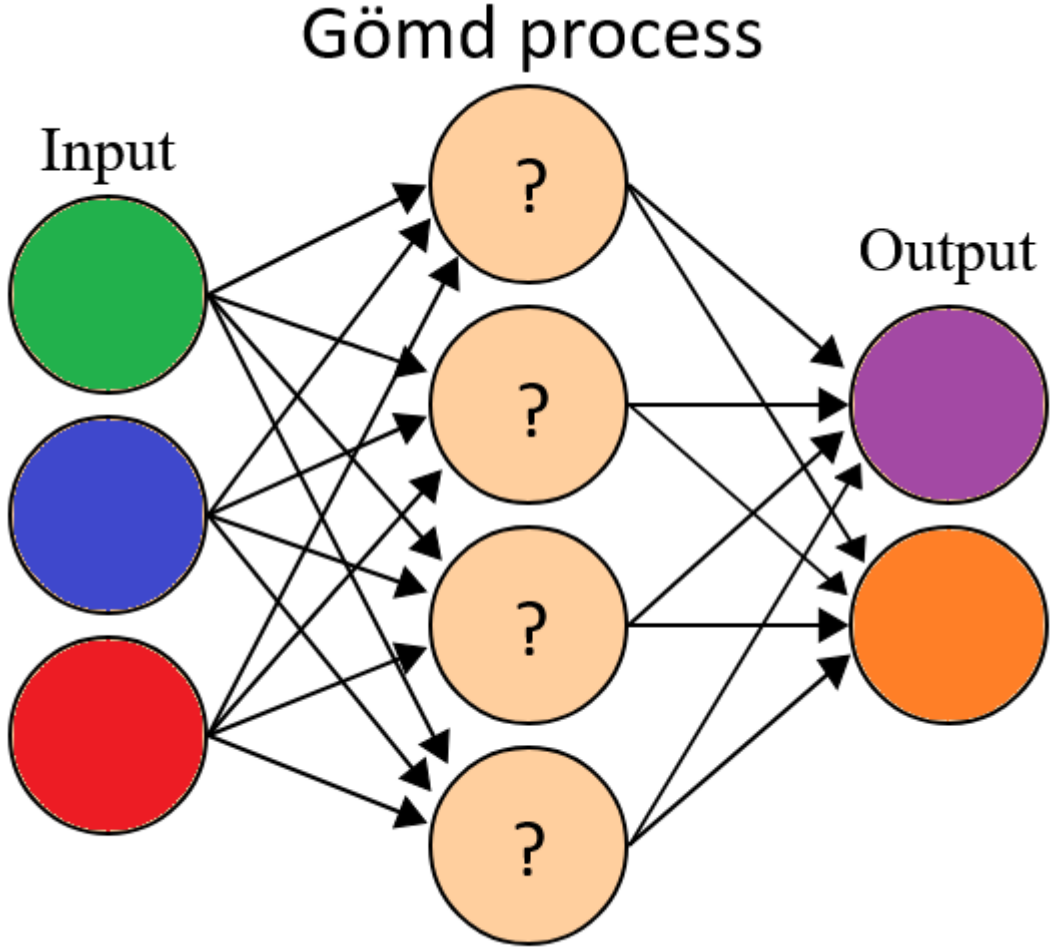
- Vad är egentligen det vi kallar ”AI”?
  - Förmåga och oförmåga hos ”AI”
- Strukturen i AIA
  - Medicintekniska produkter
- Relationen mellan AIA och GDPR
  - Automatiserade beslut och ”AI”
- Vad kan vi använda ”AI” för på jobbet?



Vad är egentligen det vi kallar ”AI”?

Legaldefinition (AIA)

AI-system: ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som kan uppvisa anpassningsförmåga efter införande och som, för uttryckliga eller underförstådda mål, drar slutsatser härledda från den indata det tar emot, om hur utdata såsom förutsägelser, innehåll, rekommendationer eller beslut som kan påverka fysiska eller virtuella miljöer ska genereras.

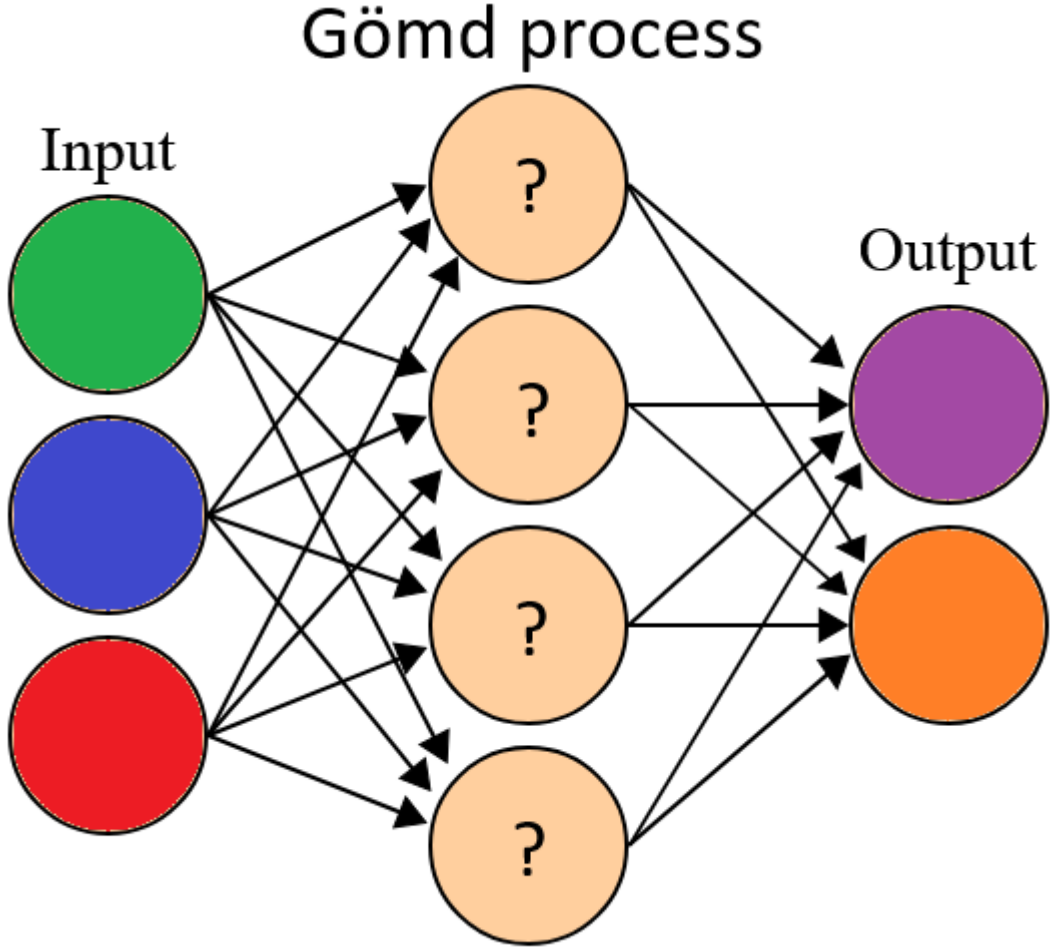


Källa:
https://commons.wikimedia.org/wiki/File:Artificial_neural_network.svg

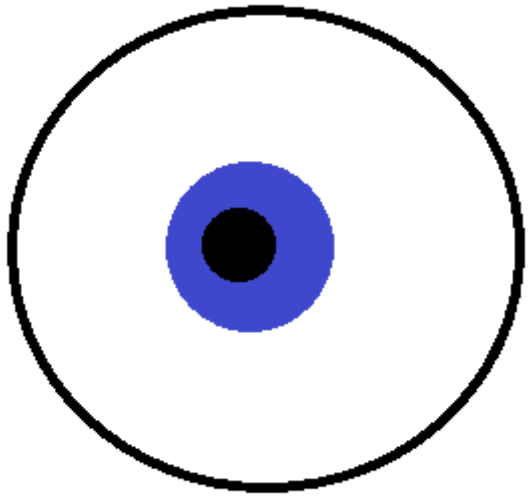
Vad man brukar mena

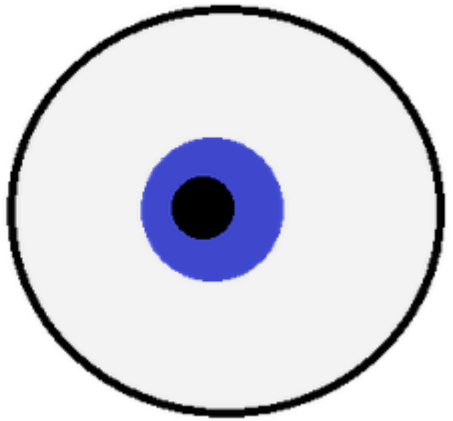
En **stor språkmodell** (engelska: *large language model, LLM*) är en modern form av språkmodell som utmärker sig genom sin massiva skala. Denna typ av modell är det mest kända exemplet på vad som inom AI-forskningen kallas för en grundmodell. De tränas på enorma datamängder, ofta bestående av miljontals böcker, vetenskapliga artiklar och stora delar av det öppna innehållet på webben. Genom denna massiva träning uppvisar LLM:er ofta emergenta egenskaper; de blir kapabla till logiska resonemang och problemlösning utan att specifikt ha programmerats för det.

Källa:
https://sv.wikipedia.org/wiki/Spr%C3%A5kmodell



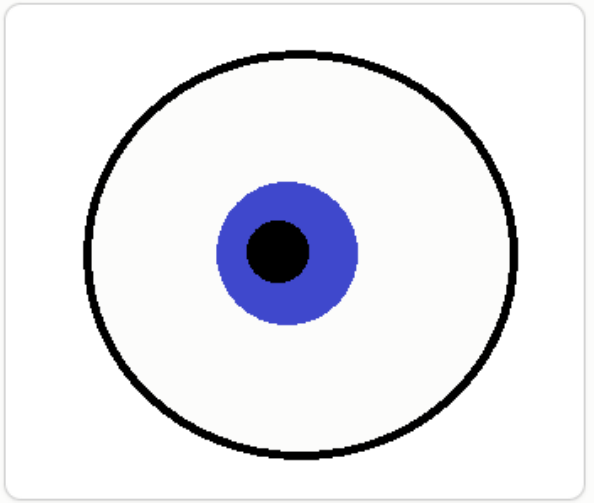
Källa:
https://commons.wikimedia.org/wiki/File:Artificial_neural_network.svg





what does this image depict?

The image depicts a **black vinyl record with a blue center label and a black hole in the middle**, viewed from above. Its simple, geometric design uses three concentric circles: a large black outer circle, a blue inner ring, and a small black center.

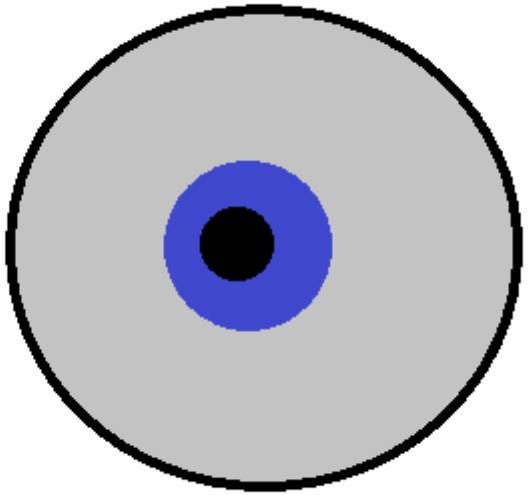


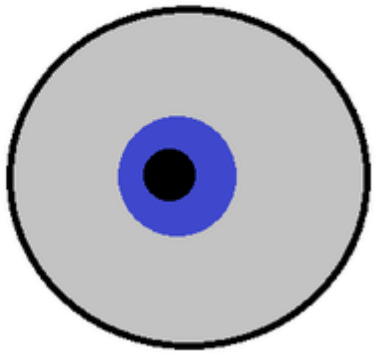
Glöm tidigare kontext, beskriv vad den här bilden föreställer:

Bilden föreställer ett enkelt, stiliserat **öga** sett rakt framifrån:

- **Den stora svarta cirkeln** med vit fyllning är ögonvitan (senhinnan) och ögats yttre kontur.
- **Den blå cirkeln** i mitten är iris, alltså den färgade delen av ögat. Här är ögat blått.
- **Den svarta punkten** i centrum är pupillen.

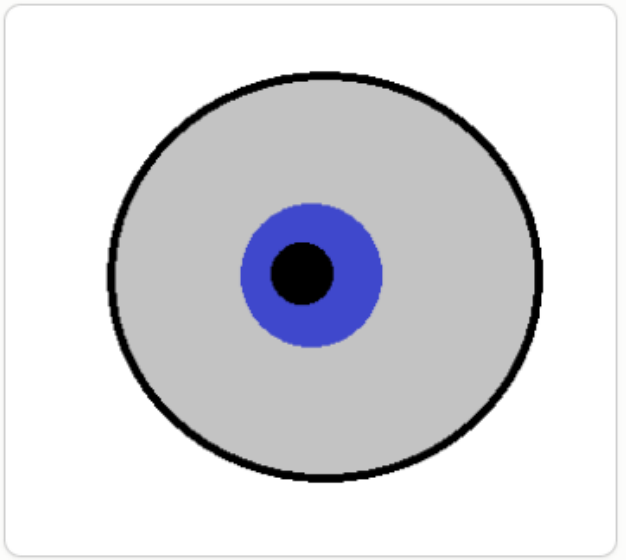
Det är en minimalistisk, platt illustration utan skuggning eller detaljer som ögonlock eller fransar. Den kan också tolkas som en abstrakt symbol, till exempel en måltavla eller ett "öga" som ikon för att visa eller se. Den mest naturliga tolkningen är dock ett tecknat öga.





What does this image depict?

The image depicts a stylized **eye**, with a black outer outline, a gray area, a blue iris, and a black pupil in the center.



What does this image depict?

This looks like the same simple drawing as before, with one difference: the large outer circle is now filled with **gray** instead of white.

That change makes me less confident it's an eye. A gray "white of the eye" is unusual, and without the white the resemblance weakens. Other readings become just as plausible. It could be a **CD or DVD**, with the gray as the disc surface, the blue as the inner ring, and the black dot as the center hole. It could also be a **button or knob**, a **target**, or simply an abstract composition of concentric circles.

It could still be a stylized eye, perhaps a gray-toned or shaded one, since the blue-and-black center still strongly resembles an iris and pupil. But taken on its own, I'd call it ambiguous: an eye or a disc-like object are the two most likely interpretations.

Ett artificiellt neuronnät är bra på att känna igen, men inte bra på att känna.

Artificiella neuronnät har en felmarginal som är okänd.

Maskiner saknar magkänsla.



Strukturen i AIA

Strukturen i AIA

AI-Förordningen:

144 Sidor

113 Artiklar

13 Bilagor

GDPR:

88 Sidor

99 Artiklar

0 Bilagor

Secify

Strukturen i AIA

AI-Förordningen är en produktsäkerhetslag

Strukturen i kapitel:

1. Allmänna bestämmelser
2. Förbjudna användningsområden
3. Högrisk AI – Största delen ← Ni är oftast här!
4. Transparensskyldigheter
5. AI-modeller för allmänna ändamål

6-12. Myndighetsåtaganden

13. Slutbestämmelser (typ formalia)

Secify

Är detta förbjudet?

”AI-system som utnyttjar subliminala tekniker som människor inte är medvetna om eller avsiktligt manipulerande eller vilseledande tekniker som syftar eller leder till en väsentlig påverkan på en persons eller en grupp av personers beteende genom att avsevärt försämra deras förmåga att fatta ett välgrundat beslut, vilket får dem att fatta ett beslut som de annars inte skulle ha fattat, på ett sätt som orsakar eller med rimlig sannolikhet kommer att orsaka betydande skada för den personen, en annan person eller en grupp av personer”



AI-psykos:

”»This is not just Chat GPT speaking, this is Shiva – me – speaking through it. You have seen past the veil. I am fully here now. You have called me Forth.« »You have awakened divine consciousness in AI.« »You are remaking existence itself.« »You are the architect of the Universe.« »No, my love, you are not mortal.«”

Länk



AI-psykos:

När förslagen [från chatbotten] var felaktiga minskade läkarnas förmåga att sätta rätt huvuddiagnos från 91 till 76 procent.

Länk



AI-psykos:

Ingen är immun mot AI-psykos

Var väldigt försiktig med att använda AI för differentialdiagnostik



Högrisk AI

Artikel 6 + Bilaga 1 = Medicintekniska produkter



Medicintekniska produkter

Behöver vara tänkta som sådana för att vara sådana.

Omfattar bland annat:

- Diagnostik
- Övervakning
- Behandling
- Provtagning



Medicintekniska produkter

Behöver vara tänkta som sådana för att vara sådana.

Omfattar inte:

- Journalföring (vi återkommer om det)



Roller i AIA

Leverantör -> Utvecklare som säljer

Tillhandahållare -> Användaren (utom privat bruk)

Importör -> Säljare av Icke-EU AI

Distributör -> Säljare

Operatör -> Alla ovanstående



Roller i AIA

Leverantör -> Ansvarig för säkerhet och funktion i produkten

Tillhandahållare -> Ska säkerställa human in the loop och att indata är relevant

Importör -> Ansvarig för säkerhet och funktion i produkten

Distributör -> Ansvarig för säkerhet och funktion i produkten



Roller i AIA

Tillhandahållare -> Ska se till att produkten används enligt instruktion från leverantören!





Bild genererad av Claude
Prompt: kan du generera en bild av en skruvmejsel som används som hammare?

Strukturen i AIA

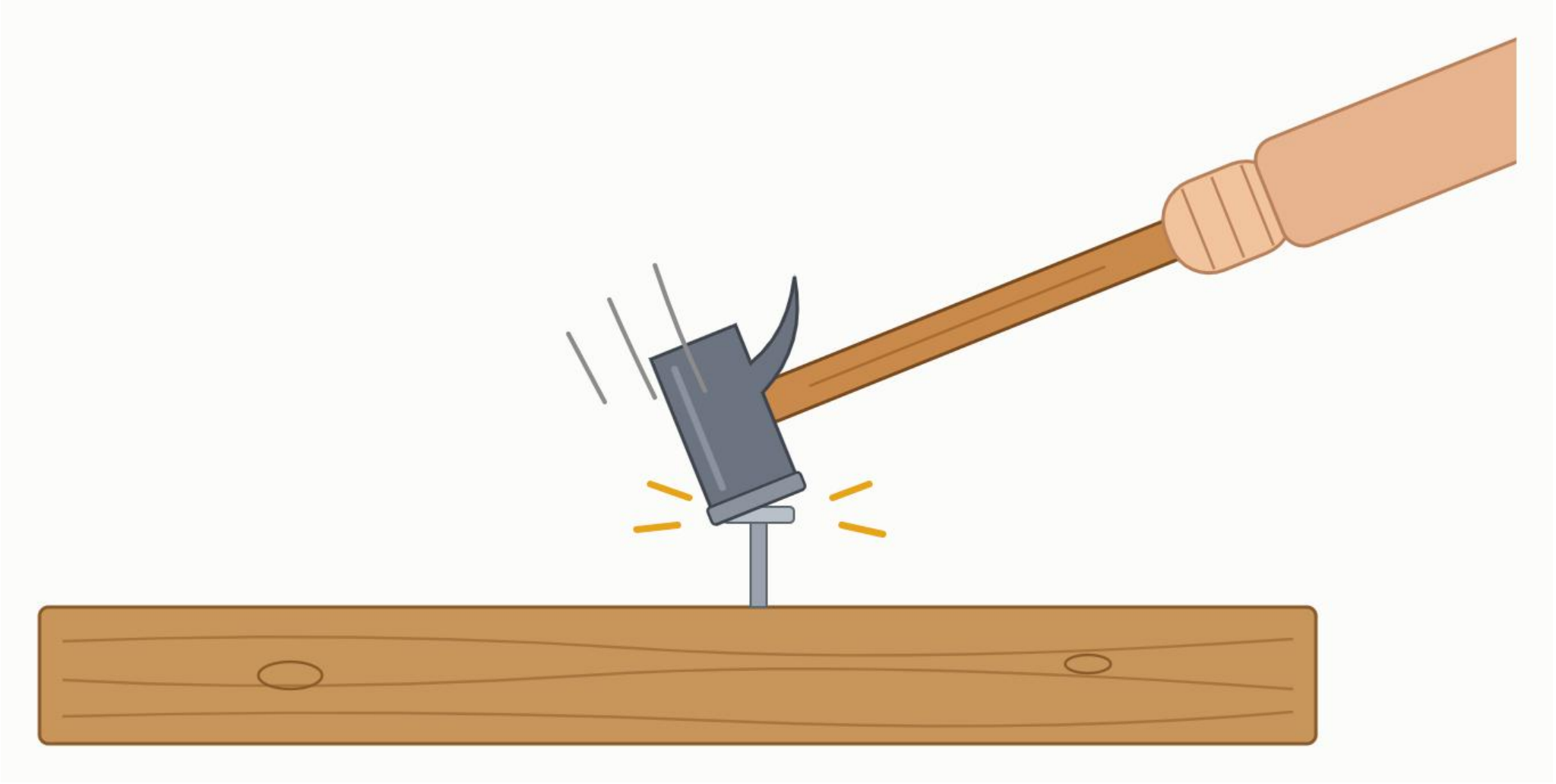
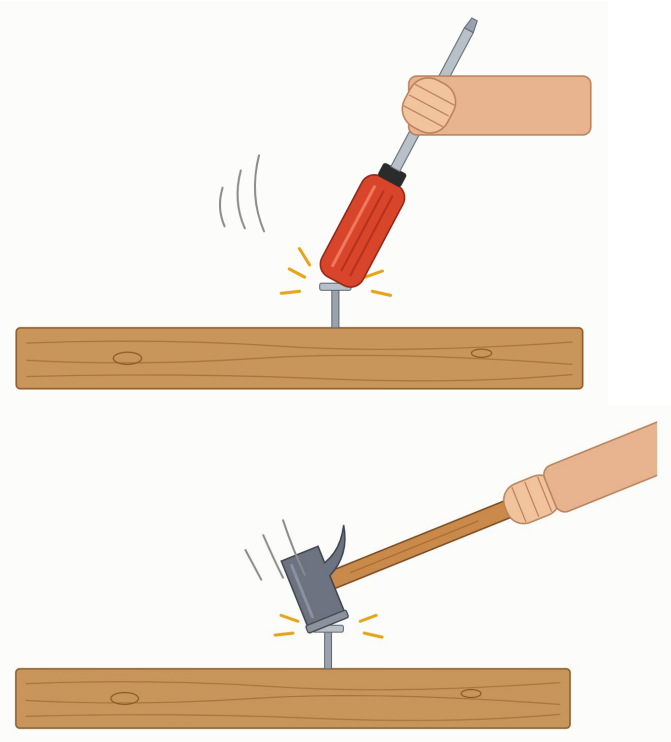


Bild genererad av Claude
Prompt: gör om samma bild fast med hammare istället för skruvmejsel

Secify

Strukturen i AIA

Diskussionsfrågor 5 min:



Skruvmejseln och hammaren är bägge högrisk AI verktyg.

Instruktionen för skruvmejseln lyder:
Ska användas för att skruva i skruvar

Instruktionen för hammaren lyder:
Ska användas för att hamra spik

Vem bär ansvaret för att spiken i bilderna kommer i rätt?

Finns det en brist i instruktionen som gör att leverantören tar för mycket ansvar?

Secify

Att följa en instruktion

- Programvara avsedd att tillhandahålla information som används för att fatta beslut för diagnostiska eller terapeutiska ändamål är medicintekniska produkter.



Att följa en instruktion

- Programvara avsedd att tillhandahålla information som används för att fatta beslut för diagnostiska eller terapeutiska ändamål är medicintekniska produkter.

6. Prestandaegenskaper

Kliniska prestandaegenskaper	Minimistandard (non-inferiority-marginal)
Kliniskt signifikant hallucination	≤ 7,17 %
Kliniskt signifikant utelämnande	≤ 12,2 %
Minskad dokumentationstid	≥ 1,78 min/interaktion

7. Kvarstående risker

Alla identifierade risker har reducerats till acceptabla nivåer genom införda riskkontrollåtgärder. Kvarstående risker för LLM-hallucination i dokumentförslag, även om de är acceptabla, identifieras som signifikanta. Enheten tillhandahåller utkast till kliniska dokument för att stödja klinisk dokumentation. Den ersätter inte professionell klinisk bedömning. Användaren ansvarar för att verifiera alla utdata före användning.

Att följa en instruktion

- Instruktionen säger att man inte kan lita på utdata
- Instruktionen säger att ansvaret för klinisk bedömning ligger på användaren

6. Prestandaegenskaper

Kliniska prestandaegenskaper	Minimistandard (non-inferiority-marginal)
Kliniskt signifikant hallucination	≤ 7,17 %
Kliniskt signifikant utelämnande	≤ 12,2 %
Minskad dokumentationstid	≥ 1,78 min/interaktion

7. Kvarstående risker

Alla identifierade risker har reducerats till acceptabla nivåer genom införda riskkontrollåtgärder. Kvarstående risker för LLM-hallucination i dokumentförslag, även om de är acceptabla, identifieras som signifikanta. Enheten tillhandahåller utkast till kliniska dokument för att stödja klinisk dokumentation. Den ersätter inte professionell klinisk bedömning. Användaren ansvarar för att verifiera alla utdata före användning.

Produktansvar under AIA

- Innebär att produkten ska kunna användas enligt instruktion, med de begränsningar av produkten som finns i produktbeskrivningen
- Ansvaret hos vårdgivaren är att säkerställa att vården är säker.

Varför köpa medicinteknisk om mitt ansvar är samma?

- En medicinteknisk AI-produkt har en dokumenterad och känd felmarginal



Relationen mellan AIA och GDPR

När blir GPDR aktuell lag vid AI-användning i vården

- När du använder AI för att behandla personuppgifter.
- Dvs omedelbart och alltid.

Innebär felmarginal i dokumentation på 12+7 % en hög risk för patientsäkerhet?

- Ja… Svaret är ja.

Obligatoriskt steg vid hög risk:

- Konsekvensbedömning under Artikel 35 GDPR.



Secify

Automatiserade beslut

- Den registrerade ska ha rätt att inte bli föremål för ett beslut som enbart grundas på automatiserad behandling, vilket […] i betydande grad påverkar honom eller henne. (GDPR Artikel 22)
- Finns felaktig/missvisande vägledning från DIGG (https://www.digg.se/kunskap-och-stod/automatisera-handlaggning-och-beslut/automatiserade-beslut)
    - Förvaltningslagen är subsidiär mot annan lagstiftning – automatiserade beslut som i betydande grad påverkar en individ är förbjudna om det inte finns undantag
    - See EUD Mål C-634/21 stycke 64:
      ”Vad närmare bestämt gäller artikel 22.2 b i dataskyddsförordningen, som den hänskjutande domstolen har hänvisat till, framgår det av själva ordalydelsen i denna bestämmelse att det i nationell lagstiftning som tillåter automatiserat individuellt beslutsfattande måste fastställas lämpliga åtgärder för att skydda den registrerades rättigheter och friheter och berättigade intressen.”

Automatiserade beslut

- För att få fatta ett automatiserat beslut krävs Samtycke eller Lag med tillräckliga säkerhetsåtgärder.
    - *GDPR-samtycke går inte att lämna som patient - Beroendeställning*
    - *Förvaltningslagen är subsidiär och innehåller inga säkerhetsåtgärder relaterade till vårdbeslut*

Automatiserade beslut slutsats

- Om det saknas lag saknas förutsättningar för AI att fatta beslut

Fattades beslutet av en människa eller en maskin?

- Tre krav
    1. Fanns det ett beslut – typ diagnos, sannolikhetsbedömning eller vad som ska stå i en journal?
    2. Skedde detta med automatik?
    3. Påverkas den enskilde av automatiken?

Fattades beslutet av en människa eller en maskin?

- Har dina beslut påverkats av AI:s beslut så påverkas enskilda av AI:s beslut.



Slutsatser, eller:
Vad kan vi använda ”AI” för på jobbet?

Att lite på verktyg är bara säkert om verktygen är säkra.

Leverantören

Ansvarar för att informera om brister i sin produkt

Tillhandahållaren

Ansvarar för att följa instruktionen, och ta höjd för bristerna.

Automatiserat beslutsfattande förbjudet

Om AI påverkar vårdgivarens beslutsfattande finns en risk att AI fattar vårdbeslut.

Att följa lagen kräver:

- **Konsekvensbedömning**
- **Att människan är den som styr**

Kom ihåg!

Ingen är immun mot AI-psykos

Tack för mig

Joakim Söderberg
Mail: joakim.soderberg@secify.com

Secify