



AI

Ethics, Law, and Medicine

Magnus Rosell

Rosell Enterprises

Content

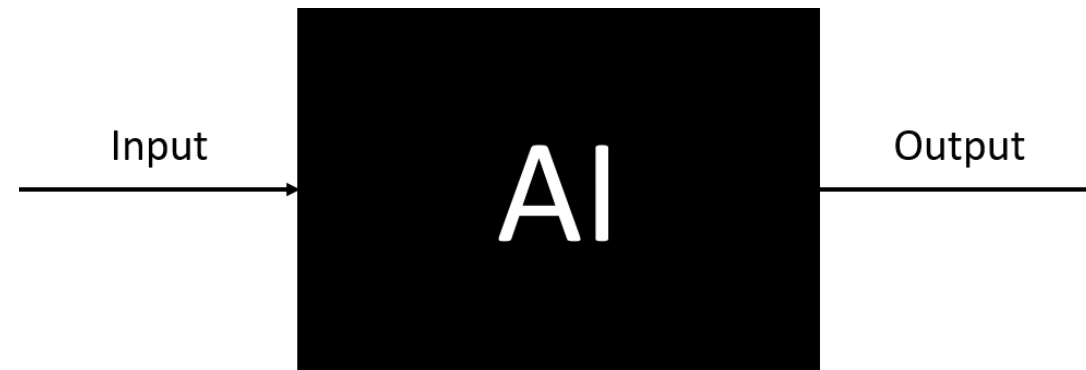
1. Introduction
2. AI and Ethics
3. AI Legislation
 - GDPR
 - EU AI Act

Introduction



Superhuman

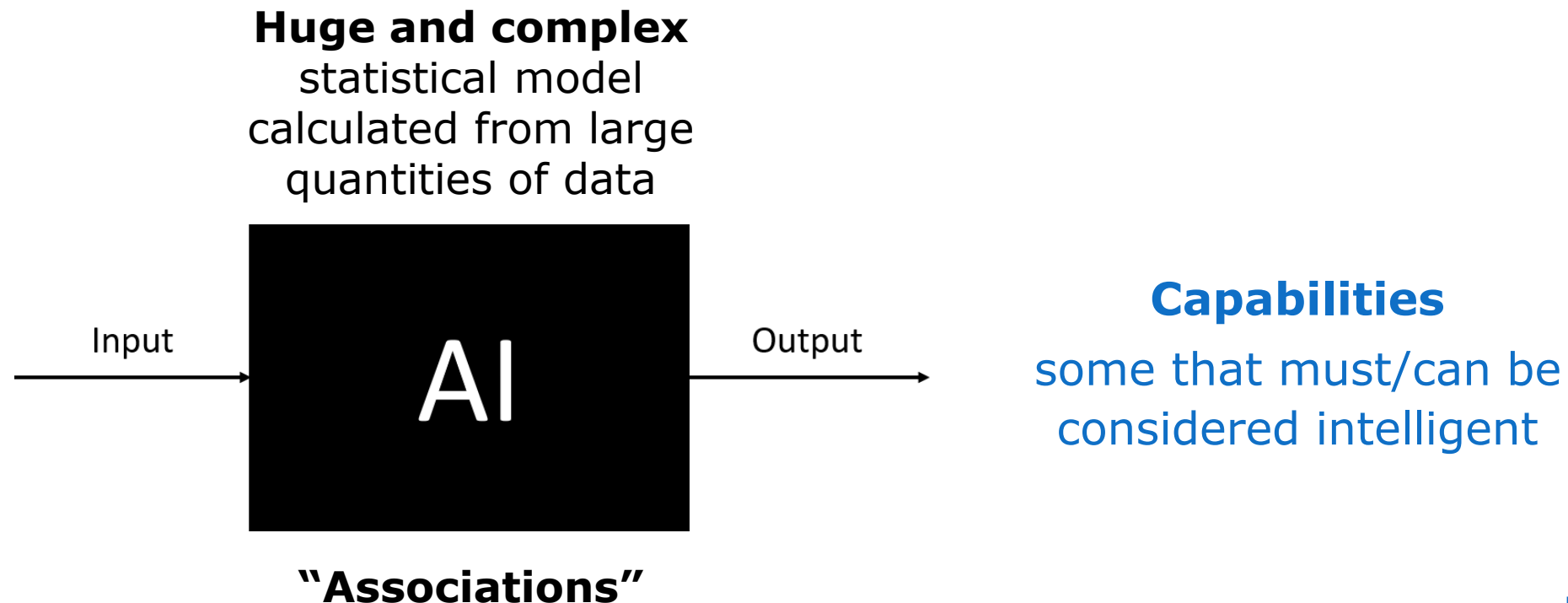
“Associations”

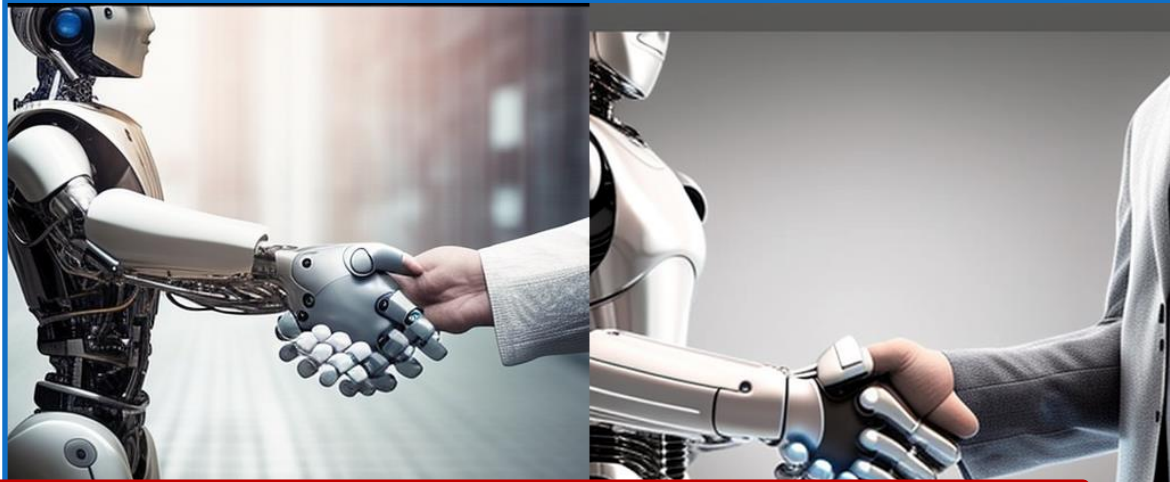
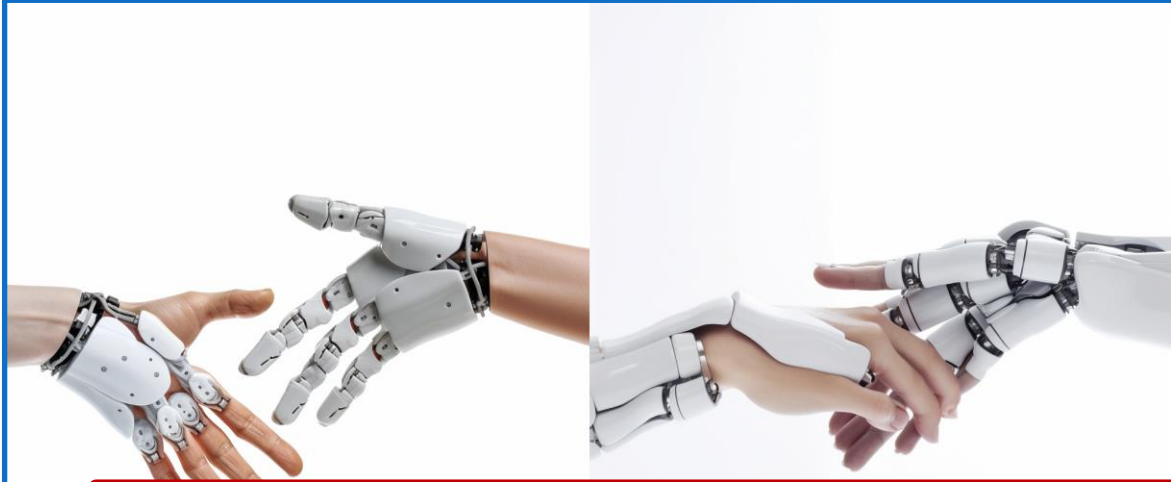


Huge and complex
statistical model calculated
from large quantities of data

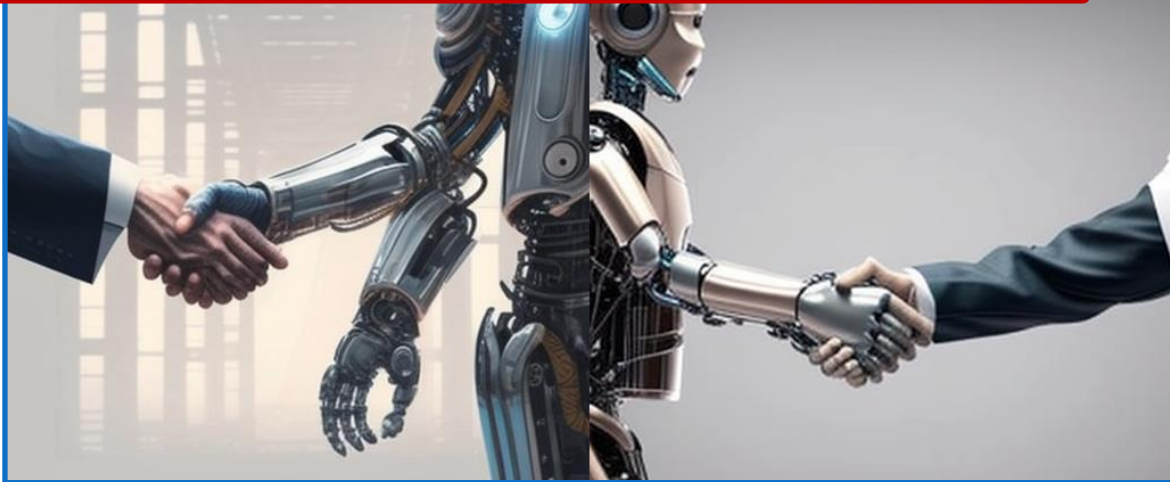
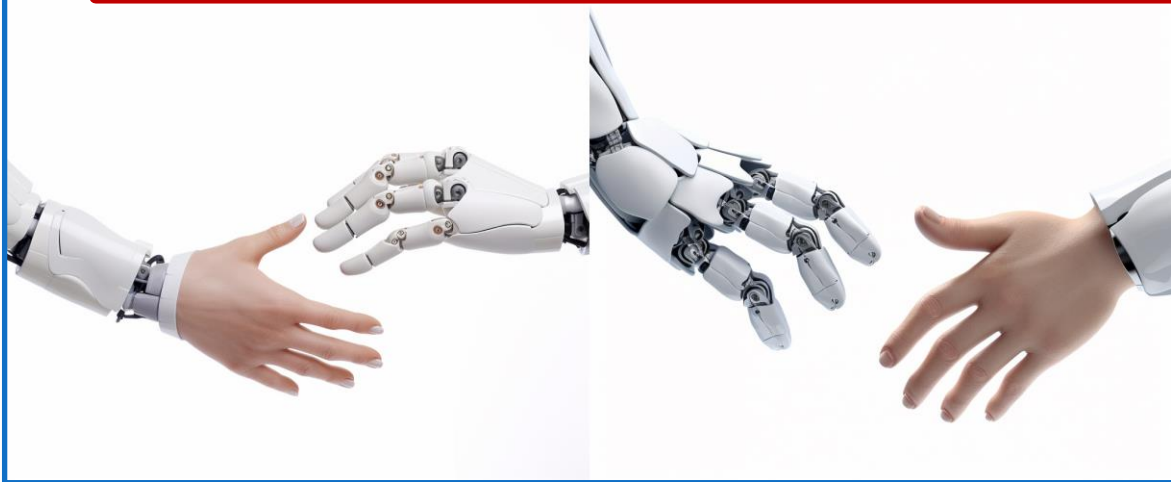
Artificial Intelligence

- Know what it **IS**
- See what it **DOES** and use it





Superhuman hallucinator?



- Inte: vad kan AI användas till?
- Utan: vad **bör** AI inte användas till?

- Inte: vad kan AI användas till?
- Utan: vad **bör** AI inte användas till?

Den AI vi använder idag är den sämsta vi någonsin kommer att ha...

Exempel



Övervakningssystem

- kontinuerliga och intermittent uppmätta vitalparametrar
- läkemedelsdata

Vid beräknat hög risk för vitalfunktions-svikt larmas en Mobil Intensivvårdsgrupp för bedömning och eventuell överflyttning till intensivvård.

Prompt: "Ett AI-system övervakar en inläggande patient baserat på kontinuerliga och intermittent uppmätta vitalparametrar, samt läkemedelsdata. När systemet beräknar hög risk för vitalfunktionssvikt ges ett alarm."

AI and Ethics

AI and Ethics

New AI systems exacerbate previously well-known problems



AI and Ethics

New AI systems exacerbate previously well-known problems

Three areas

- Value Alignment*
- Intentional Misuse
- More Concerns

* värdeanpassning / värdejämknig



The Value Alignment Problem 1(3)

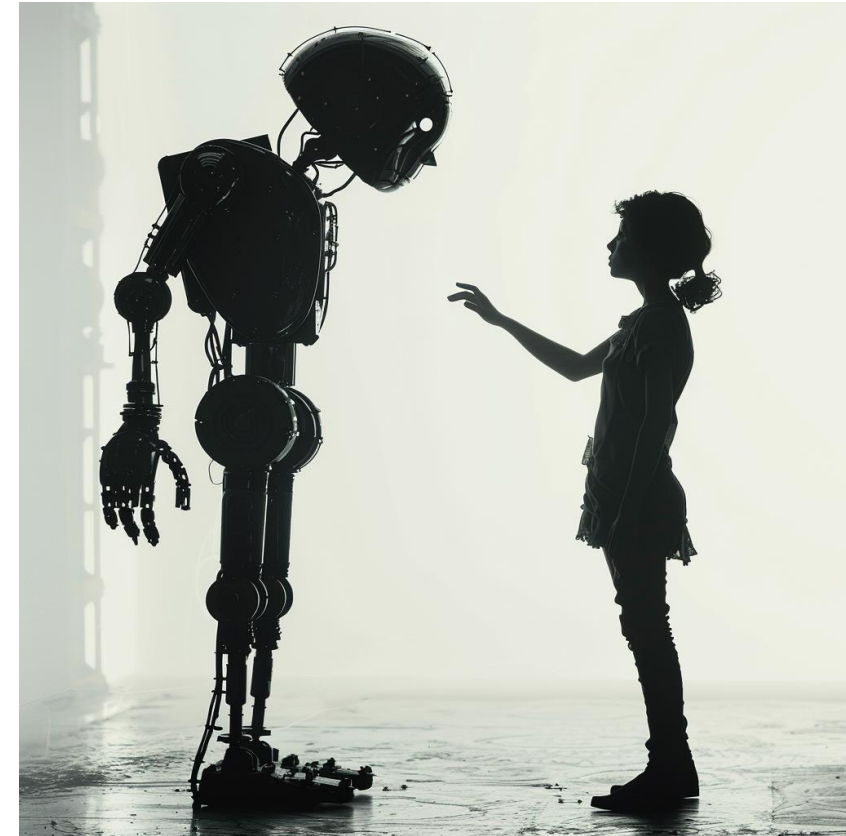
Delegation – interaction between

- human (principal) and
- AI (agent)

Two parts

- The explicit task
- (associated) values, conventions, etc.

Cf: principle-agent problem in economics

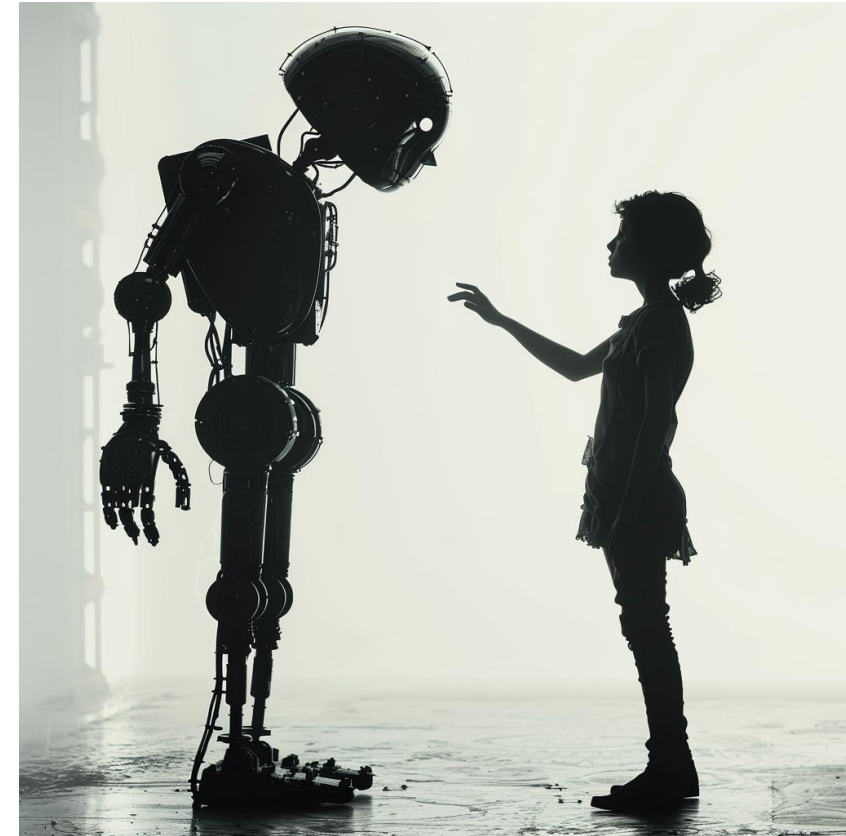


The Value Alignment Problem 2(3)

Two main problems:

1. Misaligned/mis-specified objectives (Felanpassade / feldefinierade "mål")

- Bias (\approx partiskhet)
- Artificial moral agency? (\approx moralisk handlingsförmåga)



The Value Alignment Problem 2(3)

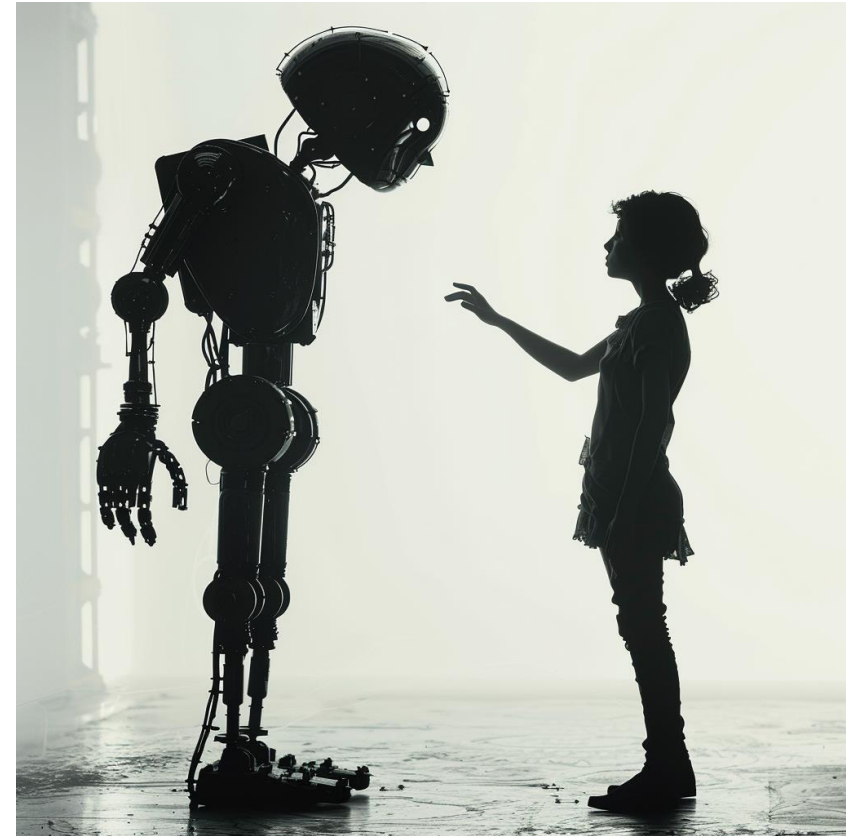
Two main problems:

1. Misaligned/mis-specified objectives (Felanpassade / feldefinerade "mål")

- Bias (\approx partiskhet)
- Artificial moral agency? (\approx moralisk handlingsförmåga)

2. Information asymmetries

- Transparency (\approx öppenhet) – details are (un)known
- Explainability (\approx förklarbarhet) – can humans understand decisions?



The Value Alignment Problem 3(3)

Normative component

- What are the correct values?

Technical components

- Can we, should we, and how do we encode values in AI?
- Implementation
 - **Outer** alignment problem: does the technical specification* correctly model the values?
 - **Inner** alignment problem: can we fulfil the technical specification?

* loss function in deep learning

Intentional Misuse

Examples:

- (Excessive?) Monitoring of citizens
- Military applications
- Fraud (bedrägeri), deception (vilselning), disinformation (and misinformation?), etc.
- Data privacy (dataskydd)
(extraction of information not meant to be available)

More Concerns

Some further issues:

- Intellectual properties
- Trust – not verifying things manually
- Deskilling – not doing things manually
- Environmental impact
- Employment
- Concentration of power

When to use AI?

Trust
certainty
understanding
ethics
laws ...

Time
productivity
competition





I want the nurse to be a human,
but I want the surgeon, one day,
to be a robot.

Eiso Kant,
CTO of poolside.ai

AI Legislation

I am not a lawyer; this is not legal advice.

AI Legislation

At the beginning of regulation.

Two major considerations

- **Data privacy**
 - Application of AI models on data.
 - EU: GDPR
 - Other countries
 - e.g. California Consumer Privacy Act (CCPA)
- **AI specific issues**
 - EU: AI Act
 - Other countries
 - laws taking form in, e.g., UK and USA.



GDPR



GDPR – Dataskyddsförordningen

”Dataskyddsreglerna grundar sig i de mänskliga rättigheterna. **Alla människor har rätt till respekt för privat- och familjeliv och till skydd av sina personuppgifter.**”

<https://www.imy.se/privatperson/dataskydd/>

Kompletterande svensk lagstiftning, bl.a. Dataskyddslagen (2018:218).

Tillsynsmyndighet: Integritetsskyddsmyndigheten (IMY)

GDPR



Application of AI models on data has to follow GDPR.

- "Konsekvensbedömning avseende dataskydd"
(DPIA – Data Protection Impact Assessment)

Hälso- och patientdata är extra skyddsvärda
(s.k. "särskilda kategorier" i artikel 9 GDPR).

- Exempel: Storskalig eller kontinuerlig monitorering klassas nästan alltid som högriskbehandling (av data).

EU AI Act



EUs AI-förordning

- "... AI systems may jeopardize fundamental rights..."
- Exclusions:
 - military, defence or national security, and scientific research and development.
 - prior to systems being put into service.

Timeline

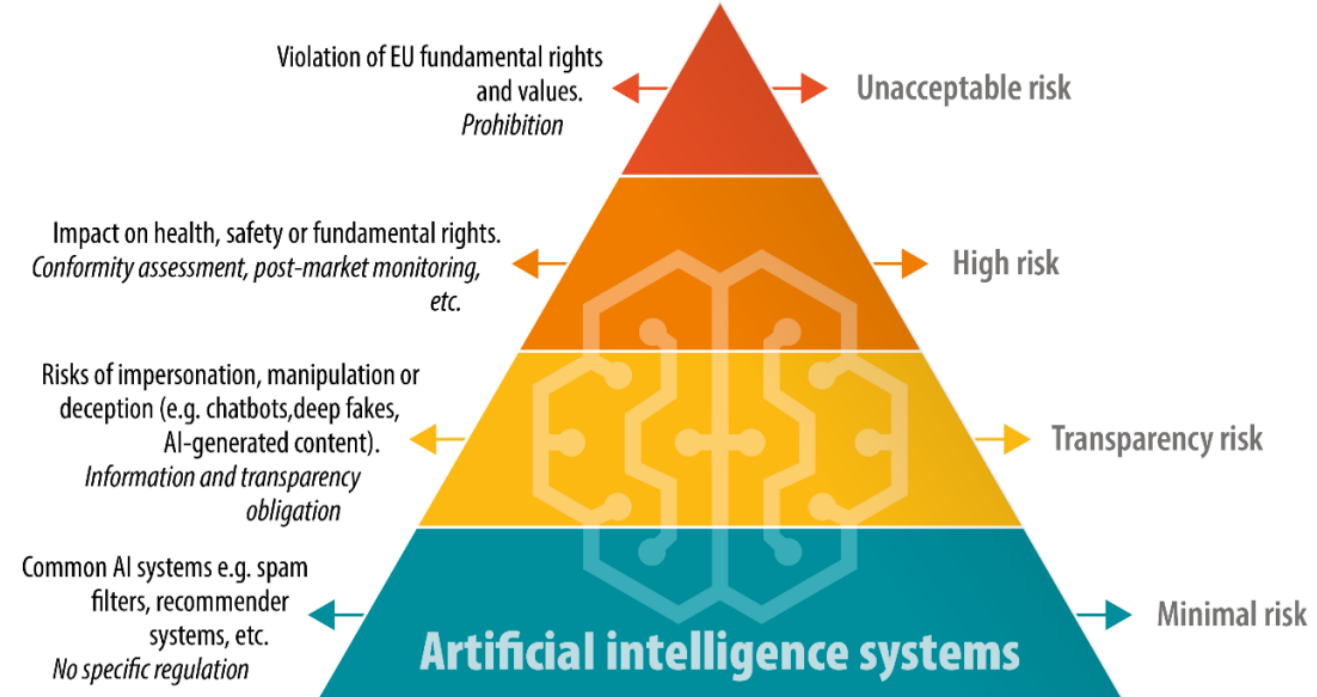
- Entered into force on August 1, 2024 (progressively 6-36 months).
- Member States designate authorities by August 2, 2025.
 - Integritetsskyddsmyndigheten (IMY) sköter GDPR???

EU AI Act

Four levels of risk

- **Unacceptable risk** -> Prohibition
- **High risk** -> Assessment and monitoring
- **Transparency** -> Information obligation (e.g. watermarking)
- **Minimal risk** -> already used applications

EU AI act risk-based approach



[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)

EU AI Act

Four levels of risk

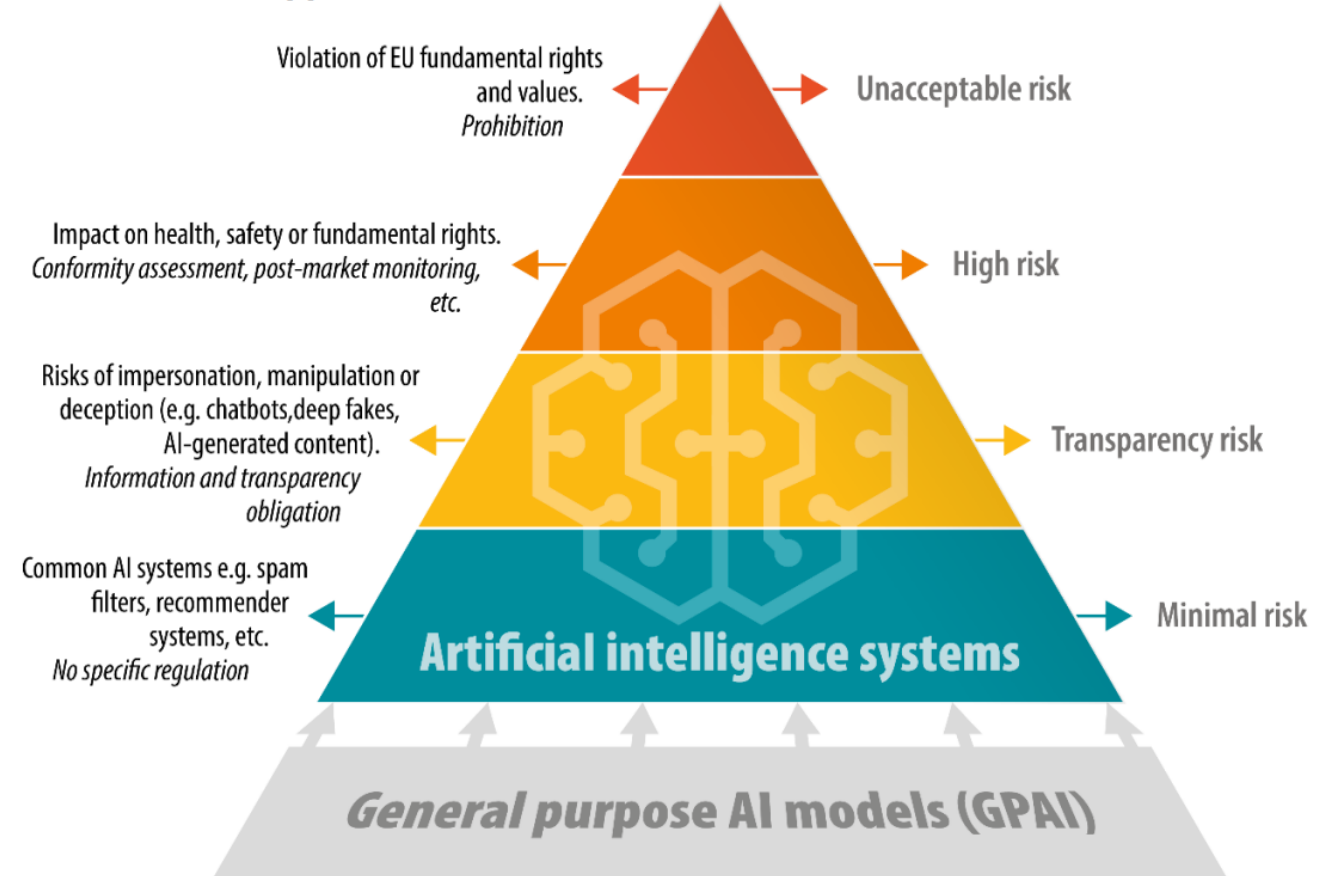
- **Unacceptable risk** -> Prohibition
- **High risk** -> Assessment and monitoring
- **Transparency** -> Information obligation (e.g. watermarking)
- **Minimal risk** -> already used applications

Requirements for General Purpose AI*

- Transparency
 - Documentation and "sufficiently detailed summary of the content used in training"
- Risk assessment and mitigation

* can be used in system developed by third party

EU AI act risk-based approach



GPAI models - Transparency requirements

GPAI with systemic risks - Transparency requirements, risk assessment and mitigation

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)

EU AI Act: Actor responsibilities

Simplified for medicine

- **Providers** (of software and devices)
 - responsible that systems comply with the Act's requirements: assessment and monitoring.

Notes:

- There are also actors that are responsible for verification of provider compliance.

EU AI Act: Actor responsibilities

Simplified for medicine

- **Providers** (of software and devices)
 - responsible that systems comply with the Act's requirements: assessment and monitoring.
- **Deployers** (hospitals, individual medical practitioners, etc.)
 - responsible for the professional use of the AI systems (and monitoring).

Notes:

- There are also actors that are responsible for verification of provider compliance.
- Deployers can be considered providers if they modify the AI system.

Deployer responsibilities

EU AI Act: explicitly outlines responsibilities for deployers using high-risk AI systems.

Four main points for medical practitioners:

- Use according to instructions.
- Inform patients that AI systems are used.

Deployer responsibilities

EU AI Act: explicitly outlines responsibilities for deployers using high-risk AI systems.

Four main points for medical practitioners:

- Use according to instructions.
- Inform patients that AI systems are used.
- Ensure practitioners have necessary competence.
 - “an adequate level of AI literacy, training and authority”.
- Practitioners are responsible for decisions.
 - “AI systems are tools that assist human decision-making.”

- Patientsäkerhetslagen
 - Användning av beslutsstödssystem

Analys av exempel

Exempel



Övervakningssystem

- kontinuerliga och intermittent uppmätta vitalparametrar
- läkemedelsdata

Vid beräknat hög risk för vitalfunktions-
svikt larmas en Mobil Intensivvårdsgrupp
för bedömning och eventuell överflyttning
till intensivvård.

“Ett AI-system övervakar en inläggande patient baserat på kontinuerliga och intermittent uppmätta vitalparametrar, samt läkemedelsdata. När systemet beräknar hög risk för vitalfunktionssvikt ges ett alarm.”

Etiska och legala problem

- 1. Datakvalitet & representativitet**
 - Kontinuerligt monitorerade patienter dominerar träningen → modellen kan undervärdera intermittenta mätserier.
- 2. Tröskelbalans (falska + missade larm)**
 - För många falska larm tränger undan resurser; för få larm riskerar patientskador.
- 3. Automation bias**
 - Personal kan överlita sig på AI-”grönt ljus” och ignorera klinisk intuition.
- 4. Transparens & förklarbarhet**
 - Svårt att förstå varför just **den** patienten flaggas/inte flaggas; gör det svårt att ifrågasätta beslut.
- 5. Integritet & GDPR**
 - Kontinuerliga vitaldata är känsliga hälsodata → kräver laglig grund, dataminimering & DPIA.
- 6. EU AI Act – hög-risk-krav**
 - AI-systemet klassas som “high-risk” medicinteknik och måste uppfylla krav på riskhantering, datastyrning, loggning och **human-in-the-loop**-övervakning. [ScienceDirect](#)
- 7. Rättvisa i resursfördelning**
 - Larmlogiken kan skapa undanträngningseffekter för patienter som inte flaggas men ändå är allvarligt sjuka.

Ansvarsfördelning ("shared but not diluted")

Aktör	Ansvar
Tillverkare/leverantör	Säker design, CE-märkning (MDR) + AI-Act-konformitet, prestandadata, post-market surveillance.
Sjukhus/vårdgivare	Lokal validering, val av larmtröskel, utbildning, löpande kvalitetsuppföljning och incidentrapportering.
Enskild kliniker	Integrera AI-stödet i egen professionell bedömning enligt Patientsäkerhetslagen; dokumentera skäl vid avvikelser.
Myndigheter (Läkemedelsverket / IVO)	Tillsyn över medicinteknik och vårdgivares hantering av AI-beslutsstöd.

Sammanfattning

Avändning av AI i vård och medicin.

GDPR: även **Dataskyddslagen**

- Personlig data

EU AI Act: jämför med **Patientsäkerhetslagen**

- *provider*
- *deployer*, inklusive enskilda kliniker



Prompt: "Thank you"

Questions
Feedback
Contacts

Magnus Rosell

magnus@rosell-enterprises.com

Några referenser

- Prince and LaCroix: Chapter 21 Deep learning and ethics, in Understanding Deep Learning, 2023.
- BRIEFING, EU Legislation in Progress: Artificial intelligence act
 - [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)
- Integritetsskyddsmyndigheten (IMY)
 - <https://www.imy.se/>

Extramaterial

Tidslinje EU AI Act

- It was adopted by the European Parliament on March 13, 2024,
- Approved by the EU Council on May 21, 2024.
- **Entered into force on August 1, 2024**, with its provisions becoming applicable progressively over the following 6 to 36 months.
- The AI Act mandates that all EU Member States designate these authorities by August 2, 2025, to oversee the implementation and enforcement of the regulation.
- Sverige har inte implementerat än.
 - Den svenska myndigheten som ansvarar för tillsyn av GDPR är **Integritetsskyddsmyndigheten (IMY)**. IMY är Sveriges nationella dataskyddsmyndighet och har till uppgift att säkerställa att personuppgifter behandlas korrekt och att individers integritet skyddas i enlighet med dataskyddsförordningen (GDPR) och kompletterande svensk lagstiftning, såsom dataskyddslagen (2018:218).

EU AI Act



General

- "... AI systems may jeopardize fundamental rights... "
- Exclusions:
 - military, defence or national security, and scientific research and development.
 - prior to systems being put into service.

Organizations

- **European AI Office** <https://digital-strategy.ec.europa.eu/en/policies/ai-office>
- **European Artificial Intelligence Board**, composed of representatives from member states and the Commission.
- Member states are required to designate **national authorities**, including a supervisory authority.
 - Encouraged to establish **regulatory sandboxes**.

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf)