



**AI502**

→4. Etisk AI i  
militær og  
sundhedsvæsen

# WHO's målsætninger

→ The appropriate use of **digital health** takes the following dimensions into consideration: health promotion and disease prevention, patient safety, ethics, interoperability, intellectual property, data security (confidentiality, integrity, and availability), privacy, cost-effectiveness, patient engagement, and affordability. It should be people-centred, trust-based, evidence-based, effective, efficient, sustainable, inclusive, equitable and contextualized. The growing global challenge of digital waste on health and the environment must also be appropriately managed.

→ WHO: Global Strategy on Digital Health 2020-25, p 16



# De tre hovedområder

**1. Biomedicinsk forskning** (*benchside*)

**2. Patientpleje:** helbredelse og palliation (*bedside*)

**3. Folkesundhed:** Profylakse, sundhedsøkonomi og epidemistyring (*community*)

# Biomedicinsk forskning

## ”Omiske data”

- every area of study whose name ends in -omic belongs to the omics fields of study. This includes the following disciplines:
  - genomics (the study of a person’s genome);
  - proteomics (the study of proteomes, a set of proteins in an organism);
  - metabolomics (the study of metabolites, small molecules within cells produced during metabolism);
  - metagenomics (the study of genetic materials recovered from natural environments);
  - phenomics (the study of total phenotypic characteristics of an organism);
  - transcriptomics (the study of RNA transcripts produced by a genome).
- (kilde: <https://www.altexsoft.com/blog/omics-data-analysis/>)

# AI med input af omiske data: Etiske dilemmaer

→ Fordele:

→ Præcisionsmedicin (*precision medicine*)

→ Praktisk implementering af grundforskning (*translational medicine*)

→ Ulemper/risici (jf. B&V 2021, 705-8):

→ Skæve datasæt

→ Potente datasæt i hænderne på regeringer og magtfulde koncerner

→ Misbrug til urimelig diskrimination



Blasimme & Vayena nævner at sammenkøring af omiske data og helbredsdata kan føre til urimelig diskrimination ifm. sundhedsforsikring. Er dette et reelt problem? Hvad kunne være eksempler?

Nobody has responded yet.

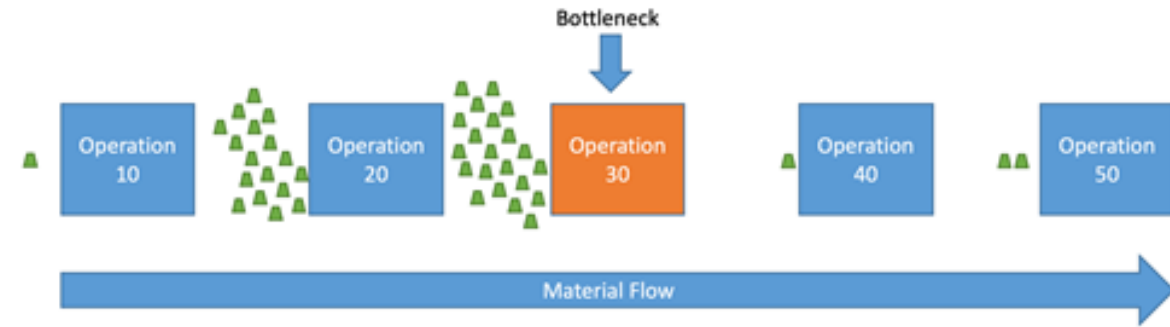
Hang tight! Responses are coming in.



# AI i patientpleje

1. Diagnostik : Måske paternalisme ift. menneskelige diagnostikere?
2. Overvågning af patienter og deres medicinering
3. Workflow og hospitalsøkonomi: Administration af sengepladser, operationskalendre etc.

# Et vigtigt begreb: Den etiske flaskehals (*ethical bottleneck*, B&V 709)



→ *Much like the neck of a bottle, where the opening is at its narrowest point, bottlenecks symbolize those critical junctures within your workflow where tasks or processes are constrained* (<https://teamhood.com/project-management/managing-bottlenecks/>)

→ **Almindelig brug af metaforen:** Et sted i en produktions- eller distributionslinje, der sætter en øvre begrænsning på outputtet af alle senere led i processen

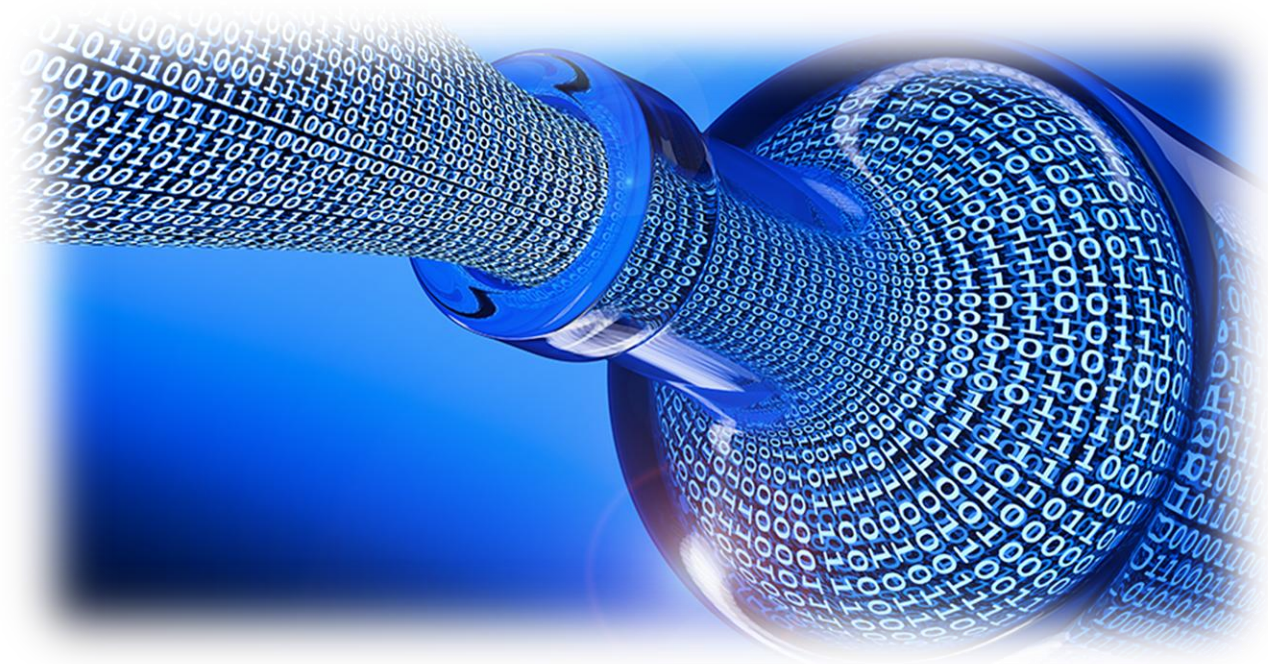
→ **Brugt (dobbel) metaforisk indenfor anvendt etik:** Et sted i en proces, der stræber mod en etisk acceptabel løsning af et praktisk problem, der sætter en suboptimal øvre grænse for den etiske kvalitet af alle senere led i processen



# En etisk flaskehals i AI-drevet patientpleje

*A biased training data set...leads to suboptimal performance for underrepresented groups (B&V 709)*

1. En AI model er trænet på data vedr. majoriteten
2. Modellen anbefaler suboptimal behandling til minoriteten
3. Nye data viser, at ressourcerne anvendes mest effektivt på majoriteten
4. Minoriteten får dårligere ressourceallokering
5. Minoriteten får endnu dårligere behandling
6. Feedback til trin 3



# AI i patientpleje

## Etiske dilemmaer

### → Fordele

- Besparelser på menneskelige ressourcer
- Effektiv brug af infrastruktur og økonomiske midler
- Pålidelig diagnosticering
- Grundig overvågning

### → Ulemper/Risici (jf. B&V 2021, 709-11)

- Upersonlig og uigennemsigtig behandling (manglende respekt for patienters krav om forståelighed og menneskelig kontakt)
- "Etiske flaskehalse": Skæve data selvforstærker ift. udsatte patientgrupper
- Tab af patienters autonomi ifm. overvågning



Blasimme & Vayena kalder det etisk problematisk hvis en patients humør konstant overvåges, fordi det "strider mod et pluralistisk liberalt samfunds værdier". Er dette korrekt? Brug et eksempel!

Nobody has responded yet.

Hang tight! Responses are coming in.

# AI i folkesundhed

1.Epidemistyring

2.Målrettede informationskampagner

3.Nudging



# AI i folkesundhed

## Etiske dilemmaer

### → Fordele

- Effektiv ressourceallokering
- Pålidelig smittesporing
- Personaliseret information til borgerne
- Skræddersyet, effektiv nudging

### → Ulemper/Risici (jf. B&V 2021, 709-11)

- "Etiske flaskehalse" igen: Skæve data selvforstærker ift. udsatte borgergrupper
- Tab af patienters autonomi ifm. overvågning og nudging
- Misbrug fra staten og private firmaers side





**Blasimme & Vayena kalder skræddersyet nudging et "Orwelliansk indgreb i den personlige frihed". Har de ret? Brug et eksempel!**

Nobody has responded yet.

Hang tight! Responses are coming in.

# AFIRRM – et ideal for implementering af AI

(Blasimme&Vayena 2021, 714-5)

- Adaptivity : regeringsgaranti for tryghed og kontrol
- Flexibility : øje for brugssammenhænge
- Inclusiveness : borgerinddragelse
- Reflexivity : blik for mange typer risici
- Responsiveness : beredskab ved misbrug og ulykker
- Monitoring : opdaterede data

# Autonome våbensystemer – potentielle anvendelsesområder

1. Paramilitært politiarbejde: Kontrol med voldelige demonstrationer og bevæbnede mistænkte
2. Forsvar af territorier. Autonom respons på angreb - neutralisering af missiler
3. Angrebskrig



# Tre hovedargumenter mod Autonome Våbensystemer

1. **Glidebaneargument - forråelse:** *We need to prohibit autonomous weapons systems that would be used against people, to prevent [a] slide to digital dehumanisation. ["stop killer robots"]*
2. **Argument fra uundgåelig stereotypisering:** *Stereotypes are entrenched by automated decision-making.... We should be challenging structures of inequality, not embedding them into weapons. ["stop killer robots"]*
3. **Argument fra maskiners moralske uvidenhed:** *Whether at a battlefield or at a protest, machines cannot make complex ethical choices, they cannot comprehend the value of human life. ["stop killer robots"]*



# Scholz & Galliot's

## "Minimalist version of ethical AI"

→ *In terms of meeting our fundamental moral obligations to humanity, we are ethically justified to develop MinAI [=Minimally Just Ethical AI] systems*  
→ (2020, 693)

→ MinAI har **kun** to **negative** funktioner (ibid. p. 692):

1. Nægter at beskyde "forbudte" mål, fx Røde Kors
2. Nægter overdreven magtanvendelse ift. militær målsætning





# Argumenter for Scholz og Galliotts minimalisme

→ **Overordnet:** Risici ved IKKE at underlægge eksisterende våben MinAI vejer tungere end risici ved at udvikle og implementere MinAI (p. 690):

1. MinAI vil gøre det lettere at holde kombattanter ansvarlige for deres våbenbrug (ibid.)
2. Et effektivt forbud mod AI i våben er usandsynligt. Så vi bør implementere den mest forsigtige form for AI: MinAI
3. MinAI vil spare ressourcer og menneskeliv (p. 696)
  - *[MinAI] in weapons should achieve a reduction in accidental strikes on protected persons and objects, reduce unintended strikes against noncombatants, reduce collateral damage by reducing payload delivery, and save lives of those who have surrendered*
4. MinAI vil sikre bedre proportionalitet i angreb ift. legitime militære målsætninger (ibid.)

# Indvendinger mod Scholz og Galliotts minimalisme

1. MinAI vil gøre militære beslutningstagere mere risikovillige ift. at indlede væbnet konflikt. Vil i sidste ende føre til højere aggressionsniveau (p. 701)
  - S&G: Alle teknologier kan misbruges. Træning kan kompensere for misbrug (ibid.) [Betryggende?]
2. MinAI kan let omkodes til den mørke side (= terrorisme): dræb identificerede nonkombattanter i stedet for at skåne dem! (jf. Vanderelst&Winfield 2016)
3. Glidebane: MinAI er et skridt på vejen mod etisk uacceptabel MaxAI
4. MinAI vil motivere krigsforbrydelser: *False flag operations*, fingerede overgivelser, obstruktion af nødhjælpsorganisationer, ansvarsfralæggelse etc.
5. MinAI vil lede til falsk tryghed og urealistiske forventninger til neutrale agenter på slagmarken (fx Røde Kors)

# Scholz og Galliotts minimalisme vs. de tre hovedindvendinger fra Stop Killer Robots

1. **Glidebaneargument** (*[a] slide to digital dehumanization*) : Måske stadig forrående at stole på maskiners skelnen mellem kombattanter og nonkombattanter?
2. **Argument fra uundgåelig stereotypisering**: Stadig risiko for, at “stereotype” nonkombattanter stilles bedst?
3. **Argument fra maskiners moralske uvidenhed**: (*Whether at a battlefield or at a protest, machines cannot...comprehend the value of human life*) : Kan MinAI fx ofre nonkombattanter når det er bydende nødvendigt for at redde mange andre liv?

Stop Killer Robots er forpligtede på at selv MinAI vil føre til en farlig forråelse i krigsstrategisk tænkning . Er du enig? Brug gerne et eksempel!

Nobody has responded yet.

Hang tight! Responses are coming in.

# Næste gang: Overvågning, AI og etik

