

*Komet kommenterar 2020:04 publicerad 2020-03-03*

Utvecklingen inom artificiell intelligens (AI) går fort, vilket lyfter fram vikten av att beakta etik. Enligt en systematisk kartläggning av nästan hundra internationella riktlinjer för etik inom AI saknas samsyn i vad som krävs för att AI ska betraktas som etisk. Kartläggningen, som här kommenteras, ger en överblick och gör det lätt att hitta riktlinjer för vidare läsning.

## Kommenterad rapport

Jobin, A., M. Ienca, et al. (2019). *The global landscape of AI ethics guidelines*. Nature Machine Intelligence 1(September 2019): 389-399.

## Komets kommentarer

- Kartläggningen visar att AI och etik är ett område under utveckling, där det ännu inte hunnit växa fram någon gemensam internationell samsyn. Detta i kombination med mängden riktlinjer inom området gör att den som läser en riktlinje bör vara medveten om att det krävs egen eftertanke och ställningstagande innan man går vidare med implementering i den egna verksamheten.
- Kartläggningen har identifierat flest underlag från Nordamerika och Europa, vilket kan ha påverkat analysen av innehållet i riktlinjerna (underlag från till exempel Asien och Afrika saknas i stort sett).
- Sverige gör flera satsningar inom AI. Ett exempel är *AI Innovation of Sweden*<sup>1</sup>, som invigdes i Göteborg början av 2019 och sedan öppnat noder på flera orter. Ett syfte är att snabba upp implementering av AI i Sverige. Etik, transparens och säkerhet lyfts fram som viktiga utgångspunkter. Ett annat exempel är *WASP-HS*<sup>2</sup>, som startades i början av 2020 som uppföljning av tidigare program om autonoma system, programvara och AI. Syftet är bland annat att öka förståelsen och bidra till utveckling av teori och praktik för mänskliga och samhällsliga aspekter av AI.
- Etiska frågeställningar kopplade till AI har uppmärksammats allt oftare i Sverige. Ett exempel är *Statens medicinsk-etiska råd* (Smer) som inriktade sin årliga etikkonferens 2018 på temat Artificiell intelligens – löftesrik teknik med etiska utmaningar<sup>3</sup>. Ett annat exempel är *Institutet för framtidsstudier* som hade artificiell intelligens, beslutsfattande och arbetsliv som tema för 2019 års stora framtidsdag<sup>4</sup>. Flera av föreläsningarna tog upp etiska aspekter av AI.
- En styrka med den kommentarerade rapporten är att den ger en överblick av de många riktlinjer som finns inom AI och etik. Den visar tydligt att det inte finns ett enda, utan i stället många olika, förhållsätt till etiska frågor inom detta snabbt framväxande område. Rapporten kan användas som ett uppslagsverk för att hitta vidare till de riktlinjer läsaren är intresserad av att fördjupa sig i.
- Författarna tar själva upp flera möjliga svagheter, bland annat metodologiska begränsningar i litteratursökning av så kallad grå litteratur (den typ av material som inte riktigt går att placera in i de gängse publikationstyper som ges ut av vanliga förlag). Eftersom sådana publikationer kan vara svåra att fånga upp vid litteratursökningar finns det risk att man inte hittat alla relevanta riktlinjer. Det kan också finnas riktlinjer på andra språk än de som författarna valt att arbeta med.

## Sammanfattning av originalrapporten

Under de senaste fem åren har det publicerats en stor mängd principer och riktlinjer som tar upp etik inom artificiell intelligens (AI). Som avsändare finns såväl universitet och forskningsinstitut som privata företag och organisationer inom offentlig sektor, se faktaruta.

Forskare vid Tekniska högskolan i Zürich har gjort en systematisk kartläggning av den internationella litteraturen inom området. Litteratursökningen inriktades på dokument som innehåller principer och riktlinjer för etik inom AI i bred mening. Forskarna hade noga specificerade och förutbestämda kriterier för vilka dokument som skulle accepteras.

De 84 inkluderade riktlinjerna är inte begränsade till något specifikt område, utan spänner över flera olika tillämpningsområden. Flera är generella och tar till exempel upp ramverk för etik inom AI, etiska principer inom AI eller ansvarfull AI. Några inriktas mot olika delområden inom AI, såsom maskininlärning, algoritmer eller robotik. Därtill innehåller kartläggningen riktlinjer som är inriktade på specifika områden, såsom hälsa eller juridik.

Majoriteten av de inkluderade riktlinjerna var från USA (21), Storbritannien (13) respektive andra länder inom EU (19). Endast enstaka riktlinjer kom från Asien och Australien och ingen från varken Afrika eller Sydamerika. Ingen svensk riktlinje finns med, men tre från Finland och en från vardera Norge och Island. Översikten innehåller en förteckning över samtliga inkluderade riktlinjer.

Författarna som har utarbetat riktlinjerna är överens om utgångspunkten att AI ska vara "etisk". De menar dock att det saknas samsyn i vilka krav som måste uppfyllas för att AI ska kunna betraktas som etisk, och hur standard och praxis bör utformas för att uppnå etisk AI.

Författarnas analys lyfter fram fem etiska principer (öppenhet, rättvisa, att inte göra skada, ansvar samt integritet) där det finns en betydande olikhet i tolkning, förståelse och implementering. För att hantera detta menar författarna att utveckling av riktlinjer kräver en omfattande etisk analys och att den som utvecklar riktlinjer också bör ha adekvata strategier för hur de ska implementeras.

1. <https://www.ai.se/en>
2. <https://wasp-hs.org/sv/>
3. [http://www.smer.se/wp-content/uploads/2019/06/Smer-konferensrapport\\_2\\_webb.NY-REV.pdf](http://www.smer.se/wp-content/uploads/2019/06/Smer-konferensrapport_2_webb.NY-REV.pdf)
4. <http://storaframtidssdagen.nu/>
5. <http://www.codex.vr.se/regler.shtml>

### *Faktaruta – exempel på aktörer*

*Bland de som har utfärdat de riktlinjer kartläggningen bygger på finns*

- **Företag**, såsom Google, IBM, Microsoft och Tieto
- **Akademi**, såsom universitetet i Montreal och The Royal Society
- **Offentlig förvaltning**, såsom Ministry of Economic Affairs (Finland) and Employment och Department of Industry Innovation and Science (Australien)
- **Organisationer**, såsom World Economic Forum, Internet Society och Association for Computing Machinery

### Om Komet Kommenterar

Komet kommenterar aktuella internationella rapporter som rör regelverk, teknikutveckling och innovation. Syftet är att ge ett svenskt perspektiv, sätta information i ett sammanhang och göra underlaget lätt tillgängligt.