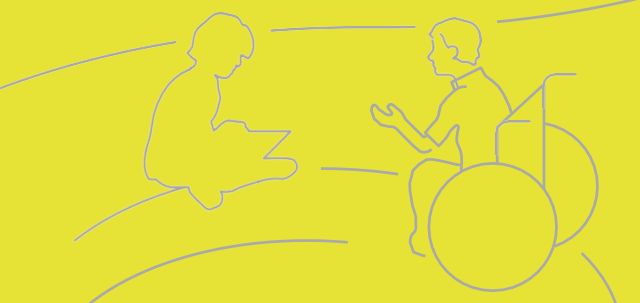


Unescos
rekommendation om

Etik och AI: Att forma framtiden för våra samhällen



unesco

Svenska Unescorådet

Unescos
rekommendation om
etik och AI: Att forma
framtiden för våra
samhällen

Kort om rekommendationen

→ Unescos rekommendation om etik och artificiell intelligens (AI) är en globalt antagen juridisk text som ska säkerställa att AI utvecklas och implementeras på ett etiskt sätt och i linje med mänskliga rättigheter av alla organisationer, företag och enskilda personer.

→ Etik inom AI innebär mänsklig interaktion med AI-system som inte är skadlig, utan bidrar till fred, mänsklig värdighet, hållbarhet och säkerhet.

→ En förutsättning för det är att alla barn och vuxna över hela världen ges möjlighet att uppnå en grundläggande nivå av AI-kompetens.

→ Rekommendationen innehåller riktlinjer som regeringar kan använda i lagstiftning, strategi- och policyutveckling av AI.

→ Unesco tar regelbundet fram rekommendationer. Unescos samtliga medlemsstater har gemensamt kommit överens om innehållet i dessa normativa instrument.

→ När en rekommendation har antagits måste medlemsstaterna rapportera tillbaka till Unesco om vad de har gjort för att implementera rekommendationen.

Varför behövs en global överenskommelse om en människo-orienterad AI?

→ Artificiell intelligens har blivit en del av vårt dagliga liv. Här finns en enorm potential till samhällsnytta och AI kan bidra till att uppnå de globala målen för hållbar utveckling (Agenda 2030). Till exempel kan AI användas för att förutsäga skördar, vilket hjälper till att säkerställa livsmedelsförsörjning. Men användningen av AI innebär också risker. I flera fall har AI-system väckt etiska frågeställningar, eftersom de kan bidra till ökad ojämställdhet mellan könen, kränka människors integritet och underlätta manipulation av medieinnehåll. Eftersom de flesta AI-system används globalt, och världens länder står inför samma utmaningar, är det viktigt med ett gränsöverskridande perspektiv på den etiska utvecklingen och användningen av AI. I november 2021 antog Unescos medlemsstater det första globala standardiseringsinstrument om etik för AI.

Vilka är rekommendationens huvudsakliga mål?

→ Rekommendationen vill främja att AI-system tjänar mänskligheten, bidrar till en god miljö och är oskadliga. Rekommendationen är en globalt accepterad juridisk text som uttrycker värderingar och principer, samt hur de kan omsättas i praktiken. Rekommendationen föreslår konkreta åtgärder för Unescos medlemsstater inom elva policyområden, inklusive hälsa, miljö och jämställdhet. Rekommendationen syftar till att skydda och främja mänskliga rättigheter och grundläggande friheter, värdighet och jämlikhet. Etiska frågor kring AI-system är relevanta i alla stadier av ett AI-systems livscykel: från forskning, design och utveckling till implementering och användning. Här ingår också underhåll, hantering, kontroll, utvärdering och utveckling.

Vad är unikt med Unescos rekommendation?

→ Rekommendationen har antagits i konsensus av Unescos medlemsstater. Särskild vikt läggs vid behoven i låg- och medelinkomstländer (LMIC) vad gäller etisk AI-utveckling och användning. Men utvecklingen av en människoorienterad AI är en angelägen uppgift för hela samhället, vilket går långt bortom enbart reglering av teknologin. Respekt för och främjande av mänskliga rättigheter måste vara grunden för all AI-reglering.

→ Rekommendationen beskriver utförligt alla områden av mänskligt liv som påverkas av AI. Särskilt betonas områden som hittills har blivit försummade i diskussionerna om AI, såsom ...



... länders olika utvecklingsnivåer

... behoven
hos det
globala syd

... utsatta
gruppers behov

... könsfördomar
och
diskriminering

... och
hållbarhet



Vad är AI?

→ Artificiella intelligenssystem kan bearbeta data och information på ett sätt som liknar eller verkar imitera mänskliga tankeprocesser, inklusive aspekter som resonemang, lärande, perception, förutsägelse, planering och styrning.

De fungerar delvis genom vad som kallas maskininlärning och maskinresonemang.

Vilka är rekommendationens viktigaste krav?

För att säkerställa en etisk utveckling och användning av AI fastställer rekommendationen krav inom elva policyområden. Följande exempel representerar ett urval av dessa bestämmelser och visar på rekommendationens omfattning:

- 1 Lagar och strategier för datahantering måste skydda rätten till integritet, hantera frågor om AI-baserad övervakning och säkerställa att individer behåller rättigheterna över sin data.
- 2 Internationellt samarbete och samverkan kring AI måste intensifieras för att hantera akuta utmaningar och minska den digitala klyftan mellan det globala syd och det globala nord.
- 3 Etiska konsekvensbedömningar av AI-system måste genomföras under systemens hela livscykel för att säkerställa att mänskliga rättigheter främjas och förhindra att människor skadas genom AI.
- 4 Styrningsmekanismer, som exempelvis tillsättning av ansvariga för arbetet med etik och AI, är nödvändigt för att förebygga och minimera potentiella skador orsakade av AI-system
- 5 Incitament måste införas för att utveckla etiskt AI-assisterade applikationer för övervakning, skydd samt återställning av miljön och ekosystemen.
- 6 AI måste utvecklas och användas på ett sätt som främjar och säkerställer mångfald och inkludering (inklusive kulturell mångfald), säkerställer icke-diskriminering, främjar yttrandefrihet och bidrar till jämställdhet mellan könen.

- 7 Inför handlingsplaner med syfte att stärka individer och minska den digitala klyftan genom att främja allmänhetens grundläggande AI-kompetens.
- 8 AI-system måste förbättra tillgången till information och kunskap.
- 9 Regeringar och privata företag måste säkerställa en rättvis övergång för anställda vars jobb är hotade på grund av nya AI-teknologier och system.
- 10 Riktlinjer för människa-robot-interaktioner och deras påverkan på människa-människa-relationer måste utvecklas, särskilt med hänsyn till människors mentala och fysiska hälsa.



Call for action inom fem centrala policyområden

Rekommendationen ger en tydlig etisk ram som regeringar kan använda för att utveckla konkreta handlingsplaner för implementering. Rekommendationen betonar hur viktigt det är med sektorsövergripande samarbete mellan stat, civilsamhälle och den privata sektorn.

Följande paragrafer sammanfattar fem av rekommendationens elva policyområden med ett övergripande samhällsperspektiv. Konkreta policyförslag presenteras inom respektive område för att uppnå en etisk utveckling och användning av AI.



Hälsa

Inom hälso- och sjukvården är de potentiella riskerna och möjligheterna med AI mer påtagliga än inom andra områden. Tack vare AI kan exempelvis sjukdomar upptäckas tidigare och nya behandlingsalternativ har utvecklats. AI kan också utgöra en resurs för att hantera den växande personalbristen inom vårdsektorn.

Regeringar och aktörer inom hälso- och sjukvården uppmanas att:

- Utveckla riktlinjer för människa-robot-interaktioner och deras påverkan på människa-människa-relationer, med särskilt fokus på människors mentala och fysiska hälsa. Människa-robot-interaktioner måste vara i enlighet med mänskliga rättigheter och grundläggande friheter, främja mångfald och skydda sårbara grupper, särskilt vid användning av AI-assisterade system för neuroteknologi och nya tekniker för att styra datorn direkt med hjärnan.
- Tillämpa AI-system för att främja människors hälsa genom att övervaka sjukdomsutbrott och dela data globalt.
- Främja forskningssamverkan om effekterna av långvarig interaktion mellan människor och AI-system, med särskilt fokus på de psykologiska och kognitiva effekter som dessa system kan ha på barn och unga.

Kultur

AI har betydande påverkan på kulturlivet, till exempel på konsumenternas betalningsvilja för kultur och kulturutövares möjligheter att försörja sig på sitt arbete. Algoritmer påverkar redan vår kulturkonsumtion genom att exempelvis styra vilken musik vi lyssnar på.

Unescos medlemsstater uppmanas att:

- Samarbeta med teknikföretag och andra aktörer för att främja ett mångsidigt utbud av och tillgång till kulturella uttryck, samtidigt som synlighet och tillgänglighet för lokalt innehåll säkerställs.
- Främja ny forskning i skärningspunkten mellan AI och immateriella rättigheter (IP, intellectual property) för att fastställa om och hur AI-assisterade verk ska skyddas av immateriella rättigheter. En viktig fråga är också hur AI-teknologier påverkar upphovspersonernas rättigheter och intressen, inte minst när deras verk används för att utveckla, träna och implementera AI-applikationer.
- Uppmärksamma AI-systemens påverkan – särskilt applikationer för språkbehandling som automatisk översättning och röstassistenter – på mänskligt språk och kulturellt uttryck.





● Kön

Användningen av AI-system kan introducera, reproducera och förstärka fördomar. När organisationer använder översättningsprogram kan ord och termer som är könsneutrala på ett språk, exempelvis "läkaren", könskodas ("el doctor" på spanska). Ett annat exempel är när företag inom könsdominerade branscher rekryterar nya kandidater baserat på datamängder med befintliga CV. I en mansdominerad bransch kommer det troligtvis att gynna manliga kandidater. En viktig princip i rekommendationen är att minimera och undvika att diskriminerande och förstärkande strukturer förstärks och upprätthålls under hela AI-systemets livscykel.

Beslutfattare, arbetsgivare inom AI-forskning och -utveckling samt andra AI-aktörer bör arbeta för:

- Att en handlingsplan för jämställdhet alltid ingår i nationella digitala riktlinjer.
- Att öka investeringarna i viktiga arbetsmarknadspolitiska områden för att bredda rekryteringen, till exempel satsningar som främjar kvinnors karriärutveckling inom STEM- och IKT-områdena.
- Att säkerställa att könsstereotyper och diskriminerande fördomar inte överförs till AI-system (till exempel genom mångsidig, öppen och tillförlitlig data).

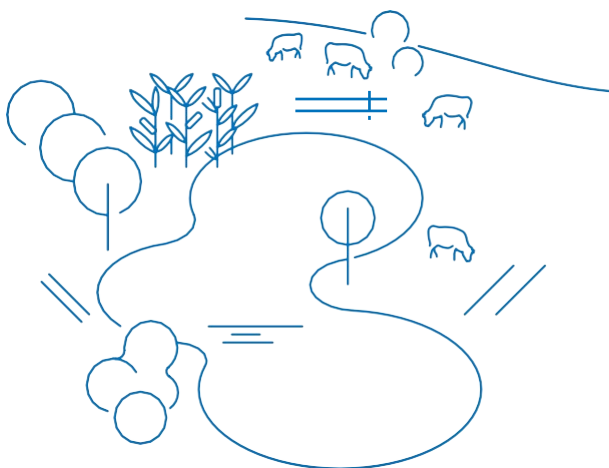


● Utbildning och forskning

Skolor och lärmiljöer är avgörande för att främja medvetenhet om AI-etik och minska digitala klyftor mellan individer och grupper av individer. Utbildning är ett viktigt verktyg för att säkerställa att både barn och vuxna över hela världen utvecklar de AI-kompetenser som krävs för de etiska och tekniska aspekterna av AI. Samtidigt omgärdas digitalt lärande i skolan av etiska dilemman. Till exempel bör elevers personuppgifter skyddas i enlighet med tillämpliga standarder för dataskydd.

Medlemsstaterna, arbetsgivare inom utbildningssektorn, lärare och ideella organisationer uppmanas att:

- Främja en allmän kunskapshöjning om AI, dess utveckling och påverkan på barns rättigheter.
- För in AI-etik i tekniska utbildningar, samt främja tvärvetenskaplig AI-forskning där discipliner utanför STEM-området (naturvetenskap, teknik, ingenjörsvetenskap och matematik) ingår.
- Arbeta för en vetenskapligt grundad policyutveckling inom AI, för att främja medvetenhet om etik och AI.



● Miljö och ekosystem

AI bidrar redan idag till att skydda miljön. Till exempel används artificiell intelligens för att övervaka och förbättra biologisk mångfald genom analys av drönar- och satellitbilder. Dock kan AI-system också ha negativ påverkan på miljön, till exempel på grund av det stora energibehov som krävs för att hantera stora datamängder inom generativ AI.

Rekommendationen efterlyser åtgärder, bland annat inom följande områden:

- Utvärdera den direkta och indirekta miljöpåverkan under hela AI-systemets livscykel och säkerställ att miljöregler följs.
- Säkerställa en utveckling och användning av rättighetsbaserade och etiska AI-drivna lösningar för katastrofberedskap, samt för övervakning skydd och bevarande av miljön, ekosystemen, och planeten.
- Förbud av AI-system som har en hög sannolikhet att orsaka oproportionerligt negativa effekter på miljön.

Det är dags att stödja en etisk och människoorienterad AI-utveckling och tillämpning

Unescos rekommendation syftar till att säkerställa att AI utvecklas och tillämpas för samhällets bästa. För att säkerställa en människoorienterad AI-utveckling behövs insatser från alla aktörer; inte bara beslutsfattare, utan även civilsamhällsorganisationer, privata företag och allmänheten. Trots uppmaningar om konkreta åtgärder understryker rekommendationen att en kollektiv insats behövs. Alla AI-aktörer, inte bara de som jobbar inom hälso- och sjukvård, kultursektorn eller utbildning, bör ta denna möjlighet att tillsammans skapa system som säkerställer en etisk och människoorienterad utveckling och användning av AI. Låt arbetet börja!

Publicerad av

Tyska Unescokommissionen

Slovenska Unescokommissionen

Nederländska Unescokommissionen

Översatt av

Svenska Unescorådet

Redaktörer för originalet

Anne Diessner, Jeannine Hausmann
(responsible), Maximilian Müngersdorff,
Davor Orlic, Jon Verriet

Status

april 2023

ISBN

978-3-947675-35-7

Copyright

Texterna i denna publikation är licensierade
under Creative Commons License
Attribution ShareAlike 4.0 International
(CC BY-SA 4.0).  www.creativecommons.org/licenses/by-sa/4.0/

Design och layout

Panatom

stöttat av



Federal Foreign Office

Med stöd av:

