

An abstract geometric pattern in the top left corner, consisting of a network of thin yellow lines connecting various sized yellow and grey circular nodes, creating a mesh-like structure.

Riktlinjer för ansvarsfull AI

En verktygslåda för civilsamhället

Juni 2025

Introduktion



OM RIKTLINJERNA

Dessa riktlinjer riktar sig till ideella organisationer som vill förstå vad artificiell intelligens (AI) innebär i praktiken.

Vi tar upp hur AI påverkas av lagar som GDPR och dokumentet går in på de speciella etiska utmaningar som civilsamhället står inför.

Riktlinjerna avslutas med gemensamma principer och konkreta råd. Dessa anser vi ska vägleda civilsamhällets organisationer till en ansvarsfull och etisk AI-användning. Vi vill att tekniken ska stödja organisationernas viktiga uppdrag, inte undergräva dem.

Syftet med riktlinjerna är att ge civilsamhällets organisationer praktisk kunskap och råd. Målet är att hjälpa dem att navigera i AI-landskapet på ett säkert sätt. Vi tittar på olika metoder, verktyg och tips för att använda AI ansvarsfullt.

BAKGRUND

Artificiell intelligens är smarta digitala tekniker som kan lära sig och bli bättre. AI förändrar samhället och skapar nya möjligheter, men också utmaningar. Civilsamhällets organisationer i Sverige möter AI allt mer. De ser fördelar med att AI kan göra deras arbete mer effektivt och nå fler. AI-verktyg kan hjälpa till med mycket. De kan automatisera återkommande uppgifter, analysera stora mängder data för att hitta viktiga insikter, och anpassa kommunikation med olika grupper. Om AI används smart så kan det hjälpa organisationer att få ut mer av sina resurser och frigöra tid för andra viktiga arbetsuppgifter.

Men att använda AI i civilsamhället har också risker. Viktiga frågor handlar om hur rättvis AI är, om tekniken är transparent, vem som ansvarar för den och hur personuppgifter hanteras. Detta är extra viktigt för organisationer som jobbar för social rättvisa. Om AI-system tränas på orättvis data kan de förstärka orättvisor i samhället. Det kan leda till diskriminering, vilket går emot många organisationers grundvärden. Dessutom kan vissa AI-system vara svåra att förstå och fungerar som "svarta lådor" som fattar beslut och gör bedömningar utan att vi vet exakt vad det grundar sig i. Detta kan minska förtroendet mellan organisationer och deras medlemmar, givare eller allmänhet. Många organisationer har också ont om pengar och teknisk kunskap, vilket gör det svårare att investera tid och pengar.



VAD ÄR ANSVARFULL AI?

I detta dokument innebär begreppet "Ansvarsfull AI" utvecklingen och användningen av AI-system på ett sätt som följer lagar (som GDPR och AI-förordningen) och är etiskt korrekt. Det handlar också om att systemen ska vara tekniskt stabila och respektera mänskliga rättigheter. Viktiga principer är rättvisa, öppenhet, ansvar, integritet och säkerhet.

För civilsamhället är det avgörande att ansvarsfull AI hjälper organisationen att nå sitt mål. Det ska också skydda de marginaliserade grupper de jobbar med. Ofta handlar det om att minska skador, men för organisationer som jobbar för social rättvisa, innebär ansvarsfull AI att aktivt främja positiv förändring. Tekniken ska inte oavsiktligt förstärka de orättvisor de vill bekämpa.

“

För organisationer som jobbar för social rättvisa, innebär ansvarsfull AI att aktivt främja positiv förändring.



CIVILSAMHÄLLET'S FORUM FÖR ANSVARSFULL AI

Civilsamhällets Forum för Ansvarsfull AI startade hösten 2024. Det är en arbetsgrupp med tolv utvalda personer från olika organisationer. De representerar olika roller och synsätt, till exempel inom funktionsrätt, barnrätt, folkbildning, idrott och mänskliga rättigheter. Forumet vill vara en plats för dialog, kunskapsutbyte och lärande. Deltagarnas olika bakgrunder är viktiga för att kunna ta fram relevanta riktlinjer för hela civilsamhället.

Forumets huvuduppgift har varit att tillsammans ta fram riktlinjer och goda exempel för hur AI kan användas. Målet med bredden i forumet är att hantera de utmaningar och möjligheter som civilsamhället möter med AI.

AI Sweden tog initiativ till Civilsamhällets Forum för Ansvarsfull AI. Det skedde inom projektet AI for Impact, med stöd från Google.org. Carl Norling Markai, Irina Glass Balog och Francisca Lundborg från AI Sweden har samordnat och lett arbetet.

Deltagare i forumet:

- Mia Ahlgren, Funktionsrätt Sverige
- Elisabet Borg, Riksidrottsförbundet och SISU Idrottsutbildarna, RF-SISU
- Martin Tägtström, Svenska Röda Korset
- Max Horttanainen, Konsult, tidigare Sverok
- Rebecca Vinthagen, Ideell Arena
- Jeremija Isakovic, Nod Sverige
- Samuel Somo, Nykterhetsrörelsens bildningsverksamhet
- Moa Marklund, Bris
- Maria Jacobson, Antidiskrimineringsbyrån Väst
- Anna Eklöf, Barncancerfonden
- Rodolfo Zúñiga, Rädda Barnen (tidigare Alireza Malakuti & Jacob Hellman)
- My Salama Meiton, Fremia (tidigare Klara Sahleström)

Metod

Arbetet i Civilsamhällets Forum för Ansvarsfull AI skedde under sex workshops hösten 2024 och våren 2025. AI Sweden ledde mötena. Workshoparna fokuserade på att utforska och diskutera viktiga AI-utmaningar, särskilt om personuppgifter, säkerhet, skevhet (bias) och rättvisa.

Målet var att hitta utmaningar och möjligheter, utveckla riktlinjer för hållbarhet och samarbete, dela erfarenheter och hitta bästa metoder. Resultatet ska fungera som en vägledning för civilsamhället och kanske även andra sektorer.

Genom samarbete, dialog och diskussioner ville forumet klargöra vad ansvarsfull AI betyder för civilsamhället.

De ville också se hur sektorn kan bidra till en ansvarsfull AI-utveckling.





Ansvarsfördelning

Forumets deltagare har delat med sig av insikter, erfarenheter, problem och idéer för lösningar.

AI Sweden har valt ut deltagare, tagit fram material och arrangerat workshops. AI Sweden har också samlat ihop diskussionerna från workshoparna i denna slutrapport.

Om urvalet av deltagare

Vi är medvetna om att alla synsätt och röster inom civilsamhället inte är representerade i urvalet. Att skapa en grupp som representerar hela det svenska civilsamhället och samtidigt fungerar bra som arbetsgrupp är svårt, för att inte säga omöjligt. Vårt mål har varit att inkludera representanter med kunskap om AI eller som jobbar med övergripande policyfrågor.

Deltagarna har valts ut för att spegla forumets mål: att samla breda perspektiv för att hantera frågor om ansvarsfull AI.

Vi välkomnar att fler organisationer och grupper deltar i framtida diskussioner om ansvarsfull AI i civilsamhället och tar egna initiativ.



Grundläggande principer för ansvarsfull AI i civilsamhället

Att skapa och använda AI på ett bra sätt kräver en tydlig etisk grund. Civilsamhällets organisationer måste följa internationella principer. Samtidigt måste de hantera unika etiska frågor som hör till deras specifika uppdrag och arbete. Det finns en global samsyn om etiska principer för AI. Dessa principer finns i riktlinjer från EU^{1,2}, UNESCO³ och andra källor. De har använts i forumets diskussioner och kompletterats med perspektiv från civilsamhället. En klar värdegrund med uttalade principer underlättar förståelsen och användningen av AI för organisationer.



Människan i centrum

AI ska stärka människor och deras beslut, inte ta över dem. Människor måste kunna övervaka AI och ingripa när det behövs, under hela AI:s livscykel. Människan har alltid det yttersta ansvaret.



Stabila och säkra system

AI-system måste vara pålitliga och säkra mot attacker. De ska vara korrekta och fungera som de ska. De ska också orsaka så lite oavsiktlig skada som möjligt.



Skydd av data och privatliv

Att respektera personlig integritet och skydda data är avgörande. Detta innebär att följa GDPR:s regler. Det handlar också om att ha bra metoder för att hantera data under hela AI-processen.

1. European Commission, "AI Act", hämtad 17 juni 2025. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
2. European Commission, "Guidelines for Trustworthy Artificial Intelligence, Shaping Europe's Digital Future", published 8 April 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
3. UNESCO, "Ethics of Artificial Intelligence", hämtad 17 juni 2025. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

**Mångfald, rättvisa, och ingen diskriminering**

AI-system måste behandla alla människor och grupper rättvist. Vi måste aktivt hitta och minska orättvis snedvridning i data och algoritmer. Vi ska säkerställa att AI inte förstärker diskriminering. Inkludering ska vara ett viktigt mål när AI designas.

**Öppenhet och förståelse**

Det ska vara tydligt hur AI-system skapas och används. Beslut som AI fattar ska vara möjliga att förstå och förklara. Detta gör det möjligt att granska och ifrågasätta dem.

**Nytta för samhälle och miljö**

AI ska gynna människor och samhället. Den ska också bidra till hållbar utveckling och ett fungerande ekosystem. Negativa effekter ska förebyggas eller minimeras enligt principen *Gör ingen skada*.

**Tydligt ansvar**

Det måste finnas system för att kunna ställa någon till svars för AI-system och deras resultat. Detta betyder att man ska kunna kontrollera och spåra vad AI gör. Det ska också finnas möjlighet till upprättelse om skada uppstår.



Specifika utmaningar för civilsamhället

Civilsamhällets organisationer står inför särskilda etiska utmaningar. Dessa kommer från deras unika roll, de grupper de finns till för och hur de arbetar.

Snedvridning (bias) och marginaliserade grupper

Detta är kanske den största etiska utmaningen för civilsamhället. AI-system lär sig av data som kan innehålla fördomar. Dessa fördomar kan handla om kön, etnicitet, funktionsnedsättning eller ekonomi. Då kan AI-systemet förstärka befintliga fördomar. Detta hotar organisationer som jobbar med marginaliserade grupper. Orättvis AI kan leda till orättvis fördelning av resurser, att personer nekas tjänster, eller att skadliga stereotyper förstärks.

Ett exempel är Försäkringskassans AI-system från 2024.⁴ Det användes för att hitta bedrägerier men det visade sig att systemet i oproportionerligt hög grad pekade ut folk baserat på utländsk bakgrund och inkomst. Detta är en tydlig varning. Sådana system kan bryta mot rätten till jämlikhet och social trygghet. Civilsamhällets organisationer måste därför vara mycket vaksamma. De måste aktivt arbeta för att hitta och minska snedvridning i data och algoritmer. Detta ska anpassas till deras specifika arbete och de grupper de tjänar.

Principen om rättvisa för civilsamhället handlar om mer än att bara undvika snedvridning. Den ska också aktivt främja jämlikhet och rättvisa. AI-användningen ska kopplas direkt till organisationens mål om social förändring. När man bedömer AI-rättvisa, bör civilsamhället titta på om AI hjälper till att minska orättvisor eller om den aktivt främjar rättigheter för utsatta grupper. Det handlar om att se bortom teknikens siffror och bedöma den större samhällspåverkan.



4. Lighthouse Reports, "Sweden's Suspicion Machine", publicerad 27 november 2024.
<https://www.lighthousereports.com/investigation/swedens-suspicion-machine/>

Att inte tappa bort uppdraget

Det finns en risk att civilsamhällets organisationer använder AI främst för att spara pengar eller bli mer effektiva. De kanske inte kritiskt granskar om AI verkligen passar deras uppdrag och värderingar. Detta kan leda till att man prioriterar att vara effektiv framför huvuduppdraget. Civilsamhällets organisationer behöver ha rutiner för att se till att AI-beslut drivs av uppdraget och är etiskt korrekta, inte bara tekniskt smarta.

Brist på resurser och kunskap

De flesta civilsamhällesorganisationer, särskilt de små, har små budgetar. De saknar ofta egen expertis inom AI, dataanalys eller etik. Kostnaden för AI-verktyg, teknik och utbildad personal kan vara för hög. Detta är inte bara ett praktiskt problem, utan också en viktig etisk fråga.

Att arbeta ansvarsfullt med AI är ett pågående arbete. Man måste göra riskbedömningar, ha bra datahantering, övervaka snedvridning, prata med intressenter och utbilda personal. Att använda AI-system utan dessa resurser kan leda till dåliga resultat. Det kan skada just de samhällen som organisationen vill hjälpa. Det är därför etiskt viktigt att organisationer realistiskt bedömer sin förmåga innan de använder AI. Att fortsätta utan tillräcklig kontroll riskerar att bryta mot principen *gör ingen skada*.

Förtroende

Civilsamhällets organisationer bygger ofta på förtroende från finansiärer, medlemmar, volontärer och allmänheten. Att använda AI-system som är otydliga, komplexa eller svåra att förstå kan lätt skada detta förtroende. Särskilt om beslut som påverkar människor känns orättvisa eller ansvarslösa. Att behålla förtroendet kräver mycket öppenhet. Det kräver tydlig kommunikation om hur AI används. Det kräver också en meningsfull dialog med alla intressenter.

Känslig data

Civilsamhällets organisationer samlar ofta in mycket känslig personlig information. Det kan handla om hälsa, ekonomi, personliga förhållanden, åsikter eller erfarenheter av diskriminering och trauman. Att använda sådan data i AI-system ger ett stort etiskt ansvar. Det handlar om privatliv, säkerhet, samtycke och risken för skada om data läcker ut eller missbrukas. Att följa GDPR och ha bra rutiner för dataskydd är helt avgörande.

Ansvarsfull AI i praktiken

Att använda AI på ett ansvarsfullt sätt kräver ett tydligt tillvägagångssätt. Vi kan lära oss av hur andra har gjort och använda de resurser som finns. Civilsamhällets organisationer kan använda olika ramverk, verktyg och fallstudier som stöd. Detta hjälper dem att använda AI på ett bra sätt.

I detta avsnitt tittar vi först på bredare ramverk. Sedan ger vi exempel på en mall för etiska frågor. Organisationer kan använda denna mall tidigt när de funderar på att införa ett AI-system. Vi går också igenom några specifika studier och etiska frågor.

Använd befintliga ramverk och verktyg

Många internationella och branschspecifika resurser ger råd om hur man använder AI-etik i praktiken. Dessa ramverk och principer är en viktig grund att bygga på. Civilsamhällets organisationer behöver dock praktiska tips. De behöver veta hur de kan omsätta dessa idéer till handling. Detta gäller för deras specifika arbete, begränsade resurser och hur de styrs.

Konkreta resurser som verktygslådor och guider är viktiga här. De innehåller checklistor, steg-för-steg-metoder (som bedömningar av mänskliga rättigheter) och relevanta exempel. Dessa hjälper till att koppla ihop abstrakta principer med den dagliga användningen i ideella organisationer. Att fler sådana praktiska resurser dyker upp visar på ett stort behov. Organisationer behöver hjälp för att omvandla teori om ansvarsfull AI till konkreta handlingar.

INTERNATIONELLA RAMVERK

UNESCO:s rekommendation om AI-etik: Detta är ett globalt godkänt ramverk. Det bygger på mänskliga rättigheter. Ramverket beskriver viktiga värden (mänskliga rättigheter, fred, mångfald, miljö) och tio principer (till exempel säkerhet, privatliv, öppenhet, rättvisa, ansvar, mänsklig tillsyn). Det föreslår också politiska åtgärder inom områden som etik, data, jämställdhet, miljö och hälsa. Det uppmanar särskilt till etiska konsekvens-bedömningar (EIA) under hela AI-processen..

EU:s AI High-level Ethics Groups etiska riktlinjer för pålitlig AI: Dessa riktlinjer lyfter fram sju nyckelkrav för pålitlig AI. Dessa är: mänsklig kontroll, stabilitet/säkerhet, privatliv/datahantering, öppenhet, mångfald/icke-diskriminering/rättvisa, samhälls-/miljönytta och ansvar. De betonar att pålitlig AI måste följa etiska principer. Den ska också vara laglig, etisk och robust.

BSR:s guider baserade på mänskliga rättigheter (HRBA): Dessa guider är gjorda för dem som implementerar ansvarsfull AI i det dagliga arbetet. De är användbara även för civilsamhället. De ger praktiska råd om hur man integrerar internationella mänskliga rättigheter i AI-utveckling och -användning. De fokuserar på exempel från generativ AI, men är bredare användbara.





Europarådets ramkonvention om AI: Detta är det första lagligt bindande internationella avtalet om AI. Det är i linje med EU:s AI-förordning. Konventionen förstärker vikten av att AI-system följer mänskliga rättigheter, demokrati och rättsstatens principer.

Diggs riktlinjer för generativ AI inom offentlig sektor: Riktlinjerna från Myndigheten för digital förvaltning innehåller relevanta etiska riktlinjer. Dessa passar väl ihop med den offentliga service som många civilsamhällesorganisationer utför.

OECD:s AI-principer: Principerna är det första internationella ramverket som främjar utveckling och användning av artificiell intelligens som är innovativ, pålitlig och respekterar mänskliga rättigheter samt demokratiska värderingar.

EU:s AI Act: Världens första heltäckande lagstiftning för artificiell intelligens. Den trädde i kraft den 1 augusti 2024 och syftar till att främja tillförlitlig AI inom EU genom att skydda hälsa, säkerhet och grundläggande rättigheter, samtidigt som den ger rättslig klarhet för företag och organisationer.

RESURSER ANPASSADE FÖR CIVILSAMHÄLLET

Digital Security Lab Ukraine (DSLU) Toolkit for Civil Society: Denna verktygslåda är direkt riktad till civilsamhällets organisationer. Den hjälper till att förebygga och minska risker från oansvarig AI-användning. Den säkerställer också att lagar och mänskliga rättigheter följs. Den täcker områden som privatliv/säkerhet (personuppgifter, interaktion med offentliga AI-verktyg, systemträning), jämlikhet (snedvridning i träning/användning, diskriminering), yttrandefrihet (desinformation, innehållsmoderering), immateriella rättigheter och mänsklig tillsyn (processkrav, personalutbildning). Den har en praktisk checklista för att välja AI-system på ett ansvarsfullt sätt.

UNICEF:s policy för AI och barn: Ger nio krav för AI som är anpassat för barn. Den fokuserar på att stödja barns utveckling, inkludering, icke-diskriminering och rättvisa. Den prioriterar också dataskydd, säkerhet, öppenhet och ansvar.

TechSoups AI-policyriktlinjer: Ger en mall och råd för civilsamhällets organisationer som vill skapa egna AI-policyer. Den täcker hur man redovisar AI-användning, skyddar egen information och användardata, använder ett etiskt ramverk, begränsar AI till arbetsuppgifter och delar kunskap internt.

Building AI playgrounds – 8 principles for designing AI for children: Detta är ett bra exempel på hur man kan tänka kring AI utifrån en specifik målgrupp. Det visar hur verktyg kan utvecklas med barns perspektiv i fokus.

Accessibility Standards Canada:s tekniska vägledning om tillgängliga och rättvisa AI-system: Många civilsamhällesorganisationer arbetar mot diskriminering och för marginaliserade gruppers rättigheter. Vi måste leva som vi lär. Denna vägledning är gjord av forskare som vill att ingen ska lämnas utanför. Vägledningen vill bidra till lösningar som motverkar statistisk diskriminering. Den vill skapa förutsättningar för att inkludera människor som är annorlunda och skiljer sig från normen.

Lärdomar från användningsfall

Det är bra att bygga sitt arbete på tydliga värderingar men att använda AI i praktiken är ofta svårt. Det handlar om att hitta en balans mellan olika intressen. Att se hur andra civilsamhällesorganisationer utforskar AI ger viktiga insikter. Det kan också ge inspiration och visa på möjliga etiska dilemman.

Att tänka på känsliga frågor tidigt i ett projekt är viktigt. Genom att formulera etiska överväganden tidigt är det lättare att hitta och hantera problem innan de blir stora. Det sparar både tid och pengar. Det gör också att projektet följer organisationens värderingar och minskar risken för dåliga följder.

Exempel på användning av AI

- AI-modererat forum för barn
- AI-baserad plattform för bättre diagnos
- AI-baserad transkribering för mer fokus på samtalet
- Datadriven folkbildningsutveckling
- AI-stöd för att sammanställa material
- AI-baserade beslutsunderlag för styrelser
- Transkribering av samtal
- AI-baserad prestationsanalys inom barnidrott
- Analys av patientdata för bättre behandlingar

AI-MODERERAT FORUM FÖR BARN

En barnrättsorganisation startar ett forum för barn på sin webbplats där de kan stötta varandra. Tanken är att skapa en trygg plats för barn som övervakas av vuxna kuratorer.

Problemet är att det tar lång tid innan inläggen hinner granskas och publiceras. Barnen använder därför inte forumet.

Med ett nytt AI-verktyg kan organisationen låta AI sköta en del av övervakningen. Då kan inlägg och svar publiceras snabbare. Forumet blir mer aktivt och engagerande.

Etiska överväganden

- **Beslut:** AI-systemet riskerar att publicera inlägg som bryter mot reglerna. Detta är extra riskfyllt med nya ämnen som AI:n inte tränats på.
- **Förtroende:** Barn ska lita på att organisationens kanaler har vuxna kuratorer. De ska få hjälp när de behöver det. Att AI delvis sköter forumet får inte förstöra detta förtroende.
- **Anonymitet:** Forumet är anonymt och ska förbli så. Om AI-systemet gör fel, kan inlägg med personlig information publiceras.



AI-BASERAD PLATTFORM FÖR BÄTTRE DIAGNOS

En patientförening startar en plattform för att dela data. Den har AI-stöd för att analysera och sammanfatta information. Den kan också ge rekommendationer. Plattformen ska förbättra precisionsmedicin och personlig vård för patienter med en viss sjukdom.

Målet är att förbättra diagnos, behandling och överlevnad. Detta ska ske genom att koppla ihop biologiska data, AI och öppen datadelning.

Etiska överväganden

- **Datahantering:** Risk att patientdata används fel eller läcks.
- **Samtycke:** Svårt att se till att patienter/vårdnadshavare förstår hur data används.
- **Global orättvisa:** Avancerad medicin kan öka klyftan mellan rika och fattiga länder.
- **AI-beslut:** Oklart vem som ansvarar om AI-systemet "hallucinerar" (hittar på) och fattar fel beslut.
- **Vem äger data?** Konflikter kan uppstå om vem som äger insamlad data och forskningsresultat. Hur blir det tydligt vilken data som delas och hur partner hanterar den?

STATISTISK DISKRIMINERING I SJÄLVKÖRANDE FORDON

En forskare hjälper till med en rapport om självkörande fordon för en myndighet. I testerna visar det sig att en person i rullstol, som korsar gatan baklänges, skulle bli överkörd. Forskaren får veta att det inte fanns tillräckligt med data. Men när mer data används, visar det sig att fordonet med större säkerhet skulle köra över personen. En förening med fokus på funktionsrätt har tillgång till värdefull men känslig data som skulle kunna förbättra systemet.

Etiska överväganden

- **För lite träningsdata:** Om data inte visar normbrytande beteenden, riskerar det diskriminering. Det kan även leda till dödsfall i kritiska system.
- **Dela känslig data?** Ska organisationer för funktionsrätt dela icke-stereotyp data? Detta med tanke på privatliv och begränsade resurser. Hur säkerställer man att sådan data används etiskt och ansvarsfullt?



AI-BASERAD TRANSKRIBERING FÖR MER FOKUS PÅ SAMTALET

En ideell organisation leder många möten mellan olika sektorer. Idag lägger de mycket tid på att sammanfatta samtalen. Detta material används sedan av deltagarna. Med AI kan de frigöra tid från detta arbete. Då kan organisationen fokusera mer på att leda samtalen. Detta ska öka kvaliteten på dialogen och förbättra materialet.

Etiska överväganden

- **Precision och återgivning:** AI-text kan baseras på muntliga utsagor och minnesanteckningar. Det finns risk att informationen inte återges korrekt. Detta kan leda till felaktiga ord eller begrepp, vilket påverkar tillförlitligheten. Hur säkerställs att AI-transkriberingen är tillräckligt exakt? Och att den inte förvränger det sagda?
- **Förutfattade meningar:** AI-modellen eller den som skrivit minnesanteckningarna kan ha fördomar. Detta kan påverka hur samtalen sammanställs. Det kan leda till fel slutsatser. Organisationen bör vara medveten om denna risk. De bör införa strategier för att minska snedvridning. Till exempel genom att använda flera källor och låta fler granska AI-materialet.
- **Öppenhet och förtroende:** Deltagarna ska få veta att AI-transkribering används. De ska också veta hur informationen används. Brist på öppenhet kan leda till misstro. Det är viktigt att deltagarna känner sig bekväma med hur deras samtal hanteras.
- **Privatliv och känslig information:** Om samtalen innehåller känslig information (t.ex. personuppgifter) måste GDPR följas. Informationen måste hanteras säkert och ansvarsfullt.
- **Ansvar och kvalitet:** Vem ansvarar för att AI-materialet är korrekt och pålitligt? Organisationen ska ha tydliga rutiner för att granska och rätta fel.
- **Minskad mänsklig kontakt:** AI kan frigöra tid, men det finns en risk att det minskar mänsklig interaktion. Det kan påverka hur väl man förstår samtalet. Organisationen bör balansera effektivitet med behovet av mänsklig närvaro.

DATADRIVEN FOLKBILDNINGsutveckling

Ideella organisationer inom folkbildning har mycket data. De har information om deltagare, verksamhet, geografi, demografi och aktiviteter. Denna data används mest för att följa regler och intern kontroll. Frågan är: Om denna data används smart, skulle den då kunna visa framtida trender och behov? Skulle den kunna vägleda strategiska beslut om vilken verksamhet som ska prioriteras? Vilka ledare som är relevanta? Och hur kommunikationen kan bli bäst? Vilka framtida behov inom folkbildningen skulle kunna förutses med dataanalys?

Etiska överväganden

- **Medlemskap och privatliv:** En persons medlemskap i olika föreningar ger insikter om intressen och nätverk. Detta kan användas för att förstå målgrupper bättre. Men det väcker frågor om privatliv. Är sådan spårning etiskt försvarbar?
- **Demografi och tillväxt:** Demografisk data (ålder, kön, utbildning, plats) kan visa på marknader. Frågan är hur vi ansvarsfullt kan använda denna data för att växa. Detta utan att diskriminera eller utnyttja grupper. Vilken tillväxt är önskvärd och hållbar?
- **Känslig data (etnicitet, kön, funktionshinder):** Att samla in sådan data är mycket känsligt på grund av tidigare och nuvarande diskriminering. Robusta säkerhetsåtgärder och tydliga regler är avgörande. Detta för att skydda privatlivet och förhindra missbruk. Hur säkerställer vi öppenhet och samtycke när sådan data hanteras?
- **Våra egna värderingar:** Hur påverkar våra egna värderingar synen på känslig datahantering? Skulle vi acceptera att andra gör samma sak, även om de säger sig ha goda avsikter? Vi måste vara konsekventa i våra etiska principer.
- **Normalisering av datahantering:** När vi samlar in och analyserar data systematiskt, bidrar vi till att normalisera vissa beteenden. Vi måste tänka på konsekvenserna. Särskilt när data delas med andra som inte delar våra etiska värderingar. Vilka långsiktiga effekter kan vår datahantering ha på samhället?

AI-STÖD FÖR ATT SAMMANSTÄLLA MATERIAL

Ideella organisationer använder AI allt mer för att bli effektivare. Ett exempel är hur AI kan sammanfatta möten, workshops och forskningsrapporter. Detta sparar tid för personalen. De kan då fokusera på sitt huvuduppdrag.

Med AI-verktyg kan organisationer snabbt få överblick över diskussioner. AI kan transkribera ljud och video. Den kan hitta viktiga teman och skapa sammanfattningar. Detta skulle annars kräva manuellt arbete. AI kan också analysera stora forskningsmaterial och hitta relevant information. Detta underlättar kunskapsinhämtning och beslut.

Denna AI-användning hjälper till att skapa en mer datadriven verksamhet. Men det är viktigt att vara medveten om etiska utmaningar. Vi måste se till att tekniken används ansvarsfullt. Den ska följa organisationens värderingar och lagar.

Etiska överväganden

- **Risk för förenklingar:** Hur kan komplexiteten i problem och sanningar gå förlorad när informationen sammanfattas? Vilka nyanser och detaljer kan missas? Vilka blir följderna av sådan informationsförlust?
- **AI:s roll i urval:** Vilka etiska tankar ska göras när AI får en stor roll i att välja vad som är viktigt? Särskilt för föreningens uppdrag och värderingar? Hur kan vi säkerställa att AI:s urval inte förstör organisationens huvuduppdrag? Vilka system behövs för att behålla mänsklig kontroll?



AI-BASERADE BESLUTSUNDERLAG FÖR STYRELSE

AI förändrar hur beslutsunderlag skapas för styrelser. AI kan analysera intern data. Det kan inkludera transkriberingar från styrelsemöten, medlemsenkäter och intervjuer. AI kan sedan skapa omfattande rapporter och exakta rekommendationer. Denna förmåga att snabbt bearbeta stora datamängder leder till snabbare beslut. Det säkerställer också att underlagen är mer välgrundade.

AI:s analysförmåga gör det möjligt att hitta trender, avvikelser och möjligheter. Sådant skulle annars kunna missas. Detta ger en effektivare arbetsmiljö för styrelse och kansli. Tid frigörs för strategisk planering. AI stärker organisationens förmåga att agera smart och fatta informerade beslut. Detta främjar långsiktiga mål.

Etiska överväganden

- **Datatäckning och slutsatser:** Har all viktig data tagits med? Kan man dra säkra slutsatser, även om AI har tillgång till mer data än en person? Finns det risk att algoritmen förenklar beslutsfattandet för mycket?
- **Informationsförlust och sammanhang:** När information bearbetas, riskerar viktiga nyanser att försvinna? Kan viktig kontext filtreras bort? Hur säkerställer vi att underlagen behåller den detaljrikedom som fanns i originalmaterialet?



TRANSKRIBERING AV SAMTAL



En medlemsorganisation har undersökt AI för att transkribera och sammanfatta rådgivningssamtal. Målet är att rådgivarna ska kunna fokusera mer på samtalet. Samtidigt vill de ha bra dokumentation. AI-genererad dokumentation kan användas som anteckningar. Detta ska leda till effektivare och bättre rådgivning.

Etiska överväganden

- **Samtalets påverkan:** Hur påverkas samtalet när medlemmen och rådgivaren vet att det spelas in? Kan det leda till att känsliga ämnen undviks? Eller att språket blir mer formellt?
- **Sekretess:** Hantering av konfidentiell information är centralt. Vem inom organisationen får se transkriptionerna? Hur skyddas informationen mot obehörig åtkomst eller läckor?
- **Bias i AI-modeller:** Risk att AI-modeller har fördomar som påverkar hur samtal tolkas? Kan vissa dialekter missförstås?

AI-BASERAD PRESTATIONSANALYS INOM BARNIDROTT

I barnidrotten används nu avancerad sportteknik, som kom från elitidrott. AI-kameror filmar matcher och analyserar spelarnas prestationer. Inspelningarna visas ofta online. Kamerornas analys fokuserar på elitidrottens parametrar. Dessa skiljer sig från barnidrottens värderingar. Ändå använder tränare, ledare och föräldrar dessa analyser som grund för att ge feedback till barnen.

Etiska överväganden

- **Träning och speltid efter elitperspektiv:** Hur påverkar AI-tjänsterna träning och speltid? Baseras det på elitidrottarens perspektiv istället för barnens utveckling? Vilka beteenden uppmuntras hos ledare som styrs av tekniken?
- **Feedback och fysisk utveckling:** Hur påverkas barns fysiska utveckling när feedback baseras på elitidrottarens fysik? Istället för barnens egna förutsättningar?
- **Privatliv och data:** Hur skyddas barnens privatliv när matcher filmas? Och när tekniken samlar in detaljerad data om deras prestationer?



ANALYS AV PATIENTDATA FÖR BÄTTRE BEHANDLINGAR

En ideell organisation behandlar personer med trauman, inklusive krigs- och tortyröverlevare. De undersöker AI för att förbättra sina behandlingsmetoder. Patienterna har ofta komplexa och utsatta bakgrunder. Organisationen ser en möjlighet att använda AI för att analysera behandlingssessioner. Målet är att hitta mönster som leder till framgångsrik rehabilitering. Detta för att utveckla metoder och optimera behandlingsstrategier.

Ett annat område är att tidigt upptäcka tecken på ökad risk för självskada. Genom att analysera behandlingsdata kan AI hitta små förändringar. Dessa kan missas av människor. AI kan då ge behandlare viktigt stöd i riskbedömning och förebyggande arbete.

Etiska överväganden

- **Dataskydd och sekretess:** Hantering av patientdata är mycket känsligt. Risk för dataintrång och obehörig åtkomst måste hanteras säkert. Det är avgörande att all datahantering följer lagar och etiska riktlinjer.
- **Bias och tillförlitlighet i AI-modeller:** Att utveckla AI-modeller för att hitta riskfaktorer för självskada kan vara svårt. Det kan vara svårt att hitta tillräckligt med träningsdata. Risk finns att snedvridning i datan leder till felaktiga resultat. Det är viktigt att modellerna är pålitliga innan de används kliniskt.
- **Professionell självständighet och tillit:** Hur påverkas behandlarnas självständighet av AI-stöd? Risk finns att de blir för beroende av AI:s rekommendationer. Detta kan göra att viktig klinisk bedömning glöms bort.
- **Informerat samtycke och kommunikation:** Att etiskt informera patienter om hur deras data används i AI-analyser är avgörande. Att få deras samtycke är svårt med marginaliserade patienter. Kommunikationen måste vara tydlig, enkel och respektfull.
- **Värdet av mänsklig kontakt:** Mänsklig interaktion är kärnan i behandling av traumatiserade. AI kan ge insikter, men får inte underminera denna viktiga del av vården.

PROJEKT FRÅN AI FOR IMPACT:S TALANGPROGRAM 2024

AI Sweden har tidigare stöttat civilsamhällesorganisationer att utforska AI. Ett initiativ var ett talangprogram där tidig etisk reflektion var ett krav. Nedan beskrivs projektet och de etiska övervägandena.



Förebygg motorhaverier med stöd av AI

Sjöräddningssällskapet (SSRS) räddar liv till sjöss. De deltog i AI Swedens talangprogram. Målet var att förbättra pålitligheten hos deras räddningsbåtar. Huvudproblemet var att förhindra motorhaverier under räddningsuppdrag. Ett studentteam arbetade med SSRS. De modellerade motorerna för att förutsäga servicebehov och förebygga haverier. Målet var att båtarna skulle kunna underhållas innan problem uppstod. Detta är mycket viktigt då en båt som inte fungerar inte kan rädda liv. De undersökte också möjligheten att förutsäga GPS-störningar i skärgården.

Teamet använde data från sensorer på båtmotorerna. Det inkluderade varvtal, oljetryck, temperatur, motortimmar, motorlast, hastighet, väder, vind, vindriktning, position och kompassriktning. De byggde en modell för att hitta mönster och avvikelser i data. Eftersom de saknade data om faktiska motorhaverier, använde de oövervakad maskininlärning. De såg ett spretigare mönster i data innan en båt togs in för service. Detta indikerade möjliga problem. Men det var svårt att bekräfta dessa mönster som faktiska föregångare till fel utan specifik haveridata. Trots detta visade projektet att insamlad data är värdefull. Den kan användas för att hitta mönster. Modellen var bra på att härma en båts normala beteende. Avvikelser från detta visade att något var fel.

Projektet visar på ansvarsfull AI genom fokus på samhällsnytta och säkerhet. Ansvarsfull AI handlar om att utveckla och använda AI säkert och fördelaktigt. Det ska vara i linje med mänskliga värderingar. Här används AI för att stödja SSRS livräddande arbete. Det ökar tillförlitligheten hos räddningsutrustningen. Att förhindra att båtar går sönder bidrar direkt till säkerheten för volontärer och de som behöver räddas.



AI för att förhindra att givare slutar ge

Projektet skulle minska antalet givare som slutar ge pengar. De analyserade data för att hitta mönster för givarlojalitet. Målet var att göra insamlingen smartare. Det är mer ekonomiskt att behålla befintliga givare än att värva nya. Projektet använde data om månadsgivare, inklusive demografi och donationshistorik. En webbtjänst skapades för att ladda upp data och förutsäga givarbortfall. Detta på en säker server hos UNICEF.

En viktig etisk fråga var hanteringen av givarnas personliga information (PII). Data (ålder, postnummer, kön) ansågs tillräckligt anonymiserade. Men det fanns en liten risk för indirekt identifiering. Givarna hade inte informerats specifikt om att deras data användes för AI-träning. Projektet diskuterade hur fokus på givarbehållning med AI kan påverka rekryteringen. Det konstaterades att projektets mål var att minska bortfall.



AI-stöd för att göra forskning mer tillgänglig

Mind är en ideell organisation för mental hälsa och självmordsprevention. De driver chatt- och telefonjourer. Med tre masterstudenter gjorde Mind ett projekt. Målet var att göra ny forskning om mental hälsa mer tillgänglig för volontärer. Inte alla volontärer är experter på forskning. Projektet använde AI för att sammanfatta publikationer. De använde ett RAG-verktyg (liknar en chattjänst men för ett urval av dokument). Verktöget matades först med relevanta publikationer som Mind granskat.

Projektet bidrog till positiv samhällspåverkan genom att göra forskning mer tillgänglig. Det främjade informativt agerande. Etiska frågor handlade främst om att begränsa användare till anställda. Experter inom psykisk hälsa skulle granska applikationen. Detta säkerställer att informationen används av personer med relevant bakgrund. Det minskar risken för feltolkningar eller skadlig användning. Särskilt inom självmordsprevention. Experter från Minds nätverk utvärderade också AI:ns förmåga att återge forskningsartiklar korrekt. Denna kvalitetskontroll minskar risken för spridning av felaktig information.



Underlätta integration och kommunikation med AI

Projektet med Svenska Röda Korsets Umeåkrets skapade en prototyp. Den skulle underlätta tillgången till information för flyktingar och nyanlända i Sverige. Den skulle också stödja Röda Korsets volontärer. Många nyanlända har svårt att förstå information om Sverige och myndighetsprocesser. Volontärer behöver bättre resurser för att ge relevant och enkel vägledning.

För att se om AI kunde hjälpa, utvecklades en webbaserad prototyp. Det var en AI-assistent eller chatbot. Prototypen använde AI-modeller via API:er för funktioner som transkribering (OpenAI Whisper), översättning och svar (GPT-3.5-turbo), text-till-tal (Microsofts API) och vektorisering av samtal för statistik (OpenAI Embeddings). Prototypen hade också en översättningsfunktion mellan ukrainska och svenska. Och en "Community Guide" som svarade på frågor om Sverige baserat på Röda Korsets information. En unik funktion var att den sparade konversationshistorik. Den analyserade data statistiskt för att hitta vanliga frågor och behov. Gränssnittet var användarvänligt för volontärer med olika teknisk bakgrund.

De etiska övervägandena i detta projekt ska ses som en prototyp. Huvudsyftet var att utvärdera AI:s potential. Ett centralt etiskt fokus var informerat samtycke. Användare informerades tydligt om att samtal kunde sparas för statistik. De kunde aktivt välja att samtycka eller inte vid varje ny konversation. Prototypen fungerade även utan samtycke för lagring. Detta minskade risken för påtryckningar. Röda Korsets volontärer förklarade syftet med datalagringen. De försäkrade användarna om stöd oavsett val.

Dataflödet i prototypen använde Azure-tjänster för serverdrift och datalagring. Även om prototypen inte direkt lagrade personlig information som kunde kopplas till individer, var medvetenheten om känsligt innehåll viktigt. Öppenheten kring dataanalysen var ett steg mot etisk AI i en sårbar kontext. Men en fullskalig implementering skulle ha krävt djupare etiska överväganden och policyutveckling. Prototypen var ett första steg för att förstå möjligheterna, inte en färdig lösning med utredda etiska följder för bred användning.



AI-stöd för att förenkla socialt entreprenörskap

Reach for Change vill stärka sociala entreprenörer. Detta genom att förbättra deras tillgång till och användning av forskning och data. Projektet blev en webbapplikation med AI-drivna verktyg. Dessa verktyg byggde på Reach for Changes egen verktygslåda. De inkluderade analys av grundorsaker, uppföljning av effekter, problemanalys, skapande av intervjuguider och enkätverktyg.

Projektets mål var att göra forskning mer tillgänglig för sociala entreprenörer. Så att de kan använda pålitlig forskning som underlag. De skulle snabba på sin forskning och tydligt presentera sina samhällsproblem. Genom att enklare hitta källor och förbättra tillgången till data, förväntas verktyget ha stor påverkan.

Projektet hanterade etiska frågor genom att tidigt vara medvetna om sin låga AI-mognad. De föreslog experiment med befintliga, kommersiella språkmodeller. Detta för att förstå användarbehov och potentialen med AI-lösningar. Denna försiktiga metod visar etisk medvetenhet. De ville inte införa AI utan att förstå sina möjligheter och begränsningar. Särskilt inom socialt entreprenörskap, där utsatta grupper påverkas.

Vidare betonades vikten av pålitliga källor. Åtgärder vidtogs för att minimera "hallucinationer" från språkmodellen. Genom att bygga ett RAG-system med trovärdiga källor och anpassa språkmodellen att bara använda hämtad information. De inkluderade också källhänvisningar och en friskrivning om möjliga hallucinationer. Detta visar en etisk inställning till öppenhet och ansvarsfull användning av AI.

Börja utvecklingen med etiken i fokus

Det är viktigt att tänka etiskt tidigt i ett AI-projekt. Detta leder till mer genomtänkta och hållbara lösningar. Det skapar också öppenhet och tydligt ansvar från början. Du blir medveten om hur teknik kan påverka människor och samhället. Det hjälper dig att hitta problem och snedvridningar i tid. Detta är avgörande för att skapa rättvisa lösningar. En strukturerad mall ser till att alla viktiga saker tas med. Det gör att dina etiska funderingar blir systematiska och effektiva.

För att underlätta detta har vi tagit fram en etisk mall som vi uppmuntrar din organisation att använda. Mallen är ett praktiskt stöd för dig som jobbar med AI inom civilsamhället. Den hjälper dig att ställa rätt frågor när en idé har formulerats, men innan utveckling eller implementering. Målet är att säkerställa att AI används på ett sätt som är rättvist, öppet och följer era värderingar.



OM MALLEN

Mallen består av reflektionsfrågor som hjälper er att fundera över viktiga etiska frågor. Du behöver inte ha svar på allt. Det viktiga är att börja tänka i dessa banor och dokumentera era resonemang. Frågor som inte är relevanta kan ni hoppa över, eller lägga till egna.

Diskutera varje del i mindre grupper eller i hela gruppen. Använd gärna post-its eller digitala verktyg för att samla era tankar.

Mallen tar cirka två timmar att genomföra.

När och hur använder du mallen?

Använd mallen i idéstadiet av ett projekt. Gärna under en workshop eller gruppdiskussion. Den passar både för enklare och djupare etiska bedömningar. Detta beror på hur långt ni kommit och hur stort projektet är. Mallen kan användas både fysiskt (med whiteboard och post-its) och digitalt (som Miro eller Mural).

Vem bör vara med i diskussionen?

Inkludera personer med olika synsätt för att få en bredare analys:

- Projektledare eller den som har idén
- Teknik- eller AI-experten
- De som ansvarar för etik eller policys
- Kanske externa rådgivare

Ju fler olika perspektiv ni har med, desto bättre blir analysen.

“

Mallen är ett praktiskt stöd för dig som jobbar med AI inom civilsamhället. Den hjälper dig att ställa rätt frågor när en idé har formulerats, men innan utveckling eller implementering.



Etisk mall

INLEDNING OCH SYFTE (15 MIN)

Varför gör vi detta?

Här presenterar ni AI-projektet eller idén. Det är viktigt att förstå syftet med granskningen. Se till att ni har resurser att hantera etiska problem.

Koppling till organisationens mål:

Hur passar AI-idén ihop med er organisations mål och värderingar? Presentera kort hur AI-projektet bidrar till organisationens mål

REFLEKTIONSFRÅGOR I GRUPP (90 MIN)

Spendera cirka 8 till 10 minuter på varje fråga.



Värdegrund och syfte: Hur bidrar AI-lösningen till vår verksamhet? Finns det risk att tekniken går emot våra värderingar?



Risk för diskriminering: Kan någon grupp påverkas negativt av AI? Hur? Vad kan vi göra för att minska den risken?



Ansvar: Hur ser ansvarsfördelningen ut vid fel i AI-systemets beslut eller agerande? Har vi tydliga roller och rutiner?



Målgruppens delaktighet: Behöver målgruppen involveras? Om ja, hur och varför? Hur har deras synpunkter tagits tillvara?



Data och träning: Vilken data har använts för AI:n? Kan den innehålla snedvridning (bias)?



Transparens: Hur tydliga är vi med hur AI:n fungerar? Förstår användare varför beslut fattas?



Uppföljning och utvärdering: Har vi en plan för att följa upp hur AI:n fungerar? Hur vet vi om den gör nytta?



Risker med att inte agera: Finns det risker med att inte använda AI? Hur väger vi dessa mot riskerna med att gå vidare med den?



Mänsklig tillsyn: Hur ser beslutsprocessen ut? Hur mycket tar AI beslut på egen hand? Finns mänsklig tillsyn, och när/hur sker den?



Lagar och riktlinjer: Vilka lagar (som AI Act, GDPR) eller interna policyer behöver vi följa? Bifoga gärna relevanta dokument eller länkar till era diskussioner.

RESULTAT OCH NÄSTA STEG (15 MINUTER)

Efter workshopen bör ni ha:

- en tydligare bild av de etiska aspekterna i ert AI-projekt.
- identifierade risker och förslag på åtgärder.
- en plan för uppföljning och eventuell vidare granskning.

Avsluta workshopen med att sammanfatta vad ni kommit fram till. Vem ansvarar för att ta arbetet vidare? Skriv namn på alla som deltagit. Bestäm ett datum för uppföljning. Då kan ni gå igenom resultatet och era slutsatser. Ändra om ni missat något, komplettera och framförallt – lär er av processen!

Checklista

- ☐ Vi har sammanfattat våra insikter
- ☐ Vi har utsett en ansvarig för att ta arbetet vidare
- ☐ Vi har satt ett datum för uppföljning av resultat och slutsatser



Rekommendationer för ansvarsfull AI i civilsamhället

För att använda AI på ett bra och etiskt sätt behöver civilsamhällets organisationer agera smart. De bör kombinera olika delar: tydliga styrmodeller, praktiska bedömningar, utbildning och samarbete.



Ha klara policies: Det är viktigt att tidigt ta fram interna regler för hur ni använder AI. Ramverk från TechSoup eller DSLU Toolkit kan vara bra att utgå ifrån. Vänta inte tills problem uppstår. Dessa regler ska tydligt beskriva:

- Vad AI får användas till inom organisationen. Detta ska följa ert uppdrag och era värderingar.
- Hur ni ska hantera data i AI-system, särskilt känsliga uppgifter. Detta ska följa GDPR.
- Hur ni granskar och bedömer risker med AI-projekt.
- Krav på öppenhet mot anställda och andra intressenter.
- Vem som ansvarar för AI-systemets utveckling, användning och resultat.



Starta en etikgrupp: Beroende på hur stor och komplex er organisation är, fundera på en särskild AI-etikkommitté. Ni kan också lägga till AI-etik i befintliga grupper, till exempel dem för riskhantering. Gruppen bör ha olika synsätt. Den ska inkludera teknik-, juridik-, etik- och programansvariga. Kanske även representanter från dem ni hjälper. Detta säkerställer en helhetssyn. Gruppen måste ha ett tydligt uppdrag och klara befogenheter.



Säkra ledningens stöd: Att ledningen engagerar sig är avgörande. Ledare måste visa att ansvarsfull AI är viktig. De måste ge tillräckligt med pengar och personal för etiskt arbete och övervakning. De ska också skapa en kultur där etiska frågor prioriteras.

PRAKTISKA STEG FÖR ATT INFÖRA AI ETISKT

1

Börja i liten skala

Undvik stora, riskfyllda AI-projekt direkt. Starta med pilotprojekt inom säkrare områden. Då kan ni lära er mer internt, testa om det fungerar och upptäcka utmaningar. Gör detta innan ni satsar stora resurser.

2

Gör konsekvensbedömningar

Bedöm risker och effekter systematiskt innan AI-system används. Detta bör inkludera:

- Bedömningar av dataskydd (DPIA) enligt GDPR för behandling av känsliga personuppgifter.
- Bedömningar av mänskliga rättigheter enligt EU:s AI-förordning för vissa högrisksystem.
- Bredare etiska bedömningar (som UNESCO rekommenderar). Dessa ska titta på större etiska och samhällsliga följder, bortom dataskydd och mänskliga rättigheter.
- Använd strukturerade metoder som checklistor och bedömningar baserade på mänskliga rättigheter.

3

Prioritera datahantering

Ha robusta metoder för att hantera data i AI-system. Följ GDPR strikt och fokusera på:

- God datakvalitet, relevans och representativitet.
- Att bara samla in nödvändig information.
- Säkra åtgärder för att skydda data och sekretess.
- Tydlig dokumentation av datakällor och hantering.
- Att aktivt hitta och minska fördomar i data.

4

Var transparent

Kommunicera öppet med alla intressenter (personal, de ni hjälper, givare, allmänheten) om hur och varför AI används. Förklara systemets funktioner, begränsningar och risker på ett enkelt språk. Se till att användare vet när de pratar med ett AI-system. Märk tydligt AI-genererat innehåll, särskilt "deep fakes" eller material som är viktigt för allmänheten.

5

Behåll mänsklig tillsyn

Utforma AI-system så att människor kan påverka och kontrollera. Speciellt när AI-beslut har stora följder för individer. Använd modeller där en människa måste ingripa ("human-in-the-loop") eller övervakar och kan ändra ("human-on-the-loop"). AI ska stärka människors förmågor och omdöme, inte helt ersätta dem. Särskilt i situationer som kräver empati eller etiska överväganden.

6

Granska leverantörer

När ni överväger AI-verktyg eller tjänster från andra företag, gör en noggrann kontroll. Bedöm leverantörens etiska engagemang, deras datasäkerhet, hur öppna deras algoritmer är och om de följer GDPR. Se till att ni har bra avtal som skyddar er.

7

Övervaka och anpassa

Att använda AI ansvarsfullt är en pågående process. Det är ingen engångsuppgift. Övervaka AI-systemets prestanda löpande. Utvärdera deras verkliga effekter (inklusive fördomar eller oväntade följder). Samla in feedback från användare och berörda. Anpassa rutiner, regler och system baserat på det ni lär er.

Checklistor för organisationer som ska köpa, utveckla och använda AI

Funderar er organisation eller förening på att börja med AI-verktyg? Eller har ni redan kommit igång? Oavsett vilket, tänk kritiskt och etiskt. Detta gäller både när ni köper in eller utvecklar AI-tjänster. Det finns nu lagar och etiska riktlinjer för AI. Det är viktigt att er organisation har kunskap om dessa. Lagstiftningen skyddar mänskliga rättigheter och här finns AI-förordningen, EU:s grundläggande rättigheter, den svenska grundlagen och diskrimineringslagen.

Civilsamhällets forum för ansvarsfull AI har tagit fram dessa checklistor. De är tänkta att stödja svenska civilsamhällesorganisationer. De hjälper er att införa AI-system på ett ansvarsfullt sätt. Dessa checklistor täcker olika delar av AI-användning. Från strategi och styrning till datahantering, etik, öppenhet, säkerhet och hur ni hanterar leverantörer. Genom att använda dem kan ni se till att viktiga saker tas med. Potentiella risker minimeras under hela AI-projektet.

Gå igenom varje checklista noggrant. Markera de punkter som är viktiga för er. Om en punkt inte passar, skriv varför. Om en punkt inte är uppfylld, skriv vilka åtgärder ni behöver göra. Se dessa checklistor som ett levande dokument. Uppdatera och anpassa dem löpande.



CHECKLISTA: STRATEGI OCH STYRNING

Vet ni varför ni vill använda AI? Att klargöra syftet är bland det första som krävs. Både ledning, anställda och volontärer mår bra om syftet är klart och tydligt innan arbetet dras igång.

- ☐ Värdet av att ha en AI-tjänst är tydligt för organisationen.
- ☐ AI-tjänsten är i linje med organisationens strategi.
- ☐ Den tänkta AI-användningen utgår från organisationens uppdrag och värderingar.
- ☐ Det finns ett formellt beslut om att AI-tjänsten ska användas.
- ☐ Organisationens ledning har avsatt tillräckligt med resurser för att genomföra ansvarsfull AI och tillsyn. Resurser kan till exempel vara kompetens, personal och tid.
- ☐ Organisationen har antagit en intern AI-policy som täcker tillåtna användningsområden, datahantering, etik och ansvarsskyldighet.
- ☐ Det finns en grupp eller en person som har ansvar för att övervaka AI-etik i organisationen, till exempel en etikkommitté, granskare eller etiskt ansvarig.
- ☐ Organisationen har gjort en konsekvensanalys av vilka effekter användningen av AI får för personalens arbetsuppgifter.
- ☐ Organisationens ledning har kommunicerat att enskilda personers tjänster och/eller arbetsuppgifter eventuellt kan förändras.

CHECKLISTA: MÄNSKLIGA RÄTTIGHETER, JÄMLIKHET OCH HÅLLBARHET

Civilsamhällets organisationer bärs av värderingar. Vårt syfte är att förmedla budskap om det vi tror på och arbetar för. För oss är det viktigt att inte förstärka ojämlikhet och marginalisering samt att respektera mänskliga rättigheter.

- ☐ En konsekvensanalys har genomförts inför inköp av en AI-tjänst. Konsekvenserna gällde personal, användare och allmänheten.
- ☐ Tjänsten har testats i sandbox eller på annat sätt som är anpassat till organisationens användning.
- ☐ Tjänsten har testats för etisk riskhantering och rättighetspåverkan enligt EU:s förordning om AI.
- ☐ Tjänsten har testats för att minimera inbyggda bias (fördomar, diskriminering) i exempelvis algoritmer, språkhantering, visuell hantering och resultat.
- ☐ Tjänsten har testats för att minimera inbyggd strukturell och individuell diskriminering.
- ☐ Följande konventioner och andra internationella överenskommelser om mänskliga rättigheter är grundläggande som bakgrundsdocumentation för användning av AI i organisationen.
- ☐ Följande svensk och europeisk lagstiftning är väsentlig bakgrundsinformation, utöver de som rör digitalisering och AI.
- ☐ Följande maktanalyser ligger till grund för de bedömningar som görs och de beslut som tas när det gäller AI-tjänsten. En maktanalys kan exempelvis vara feministisk, normkritisk och antirasistisk.

**FORTS. CHECKLISTA: MÄNSKLIGA
RÄTTIGHETER, JÄMLIKHET OCH HÅLLBARHET**

- ☐ Grupper som diskriminerats historiskt och som kan vara särskilt utsatta för diskriminering genom AI-teknik är identifierade.
- ☐ Följande specifika behov och sårbarheter bland användare av tjänsten behöver organisationen ta hänsyn till.
- ☐ AI-tjänsten är designad och konstruerad på ett sätt som inte förstärker rådande marginalisering och diskriminering.
- ☐ Det framgår tydligt till vem enskilda ska vända sig till vid frågor om organisationens användning av AI.
- ☐ Tjänsten följer organisationens hållbarhetsstrategi.
- ☐ En analys av AI-tjänstens klimatpåverkan har gjorts, enligt EU:s klimatlag⁵ och energilagstiftning⁶.

5. Europeiska unionen, Europaparlamentets och rådets förordning (EU) 2021/1119 av den 30 juni 2021 om inrättande av en ram för att uppnå klimatneutralitet och om ändring av förordningarna (EG) nr 401/2009 och (EU) 2018/1999 (europeisk klimatlag), Europeiska unionens officiella tidning L 243 (9 juli 2021): 1–17.

6. Europeiska unionen, Europaparlamentets och rådets förordning (EU) 2018/1999 av den 11 december 2018 om styrningen av energunionen och av klimatåtgärder samt om ändring av förordningar och direktiv och om upphävande av förordning (EU) nr 525/2013, Europeiska unionens officiella tidning L 328 (21 december 2018): 1–77.

CHECKLISTA: DATAHANTERING

Data är hårdvaluta i AI-världen. Något att vara rädd om och försiktig med. Sedan 2015 finns GDPR, EU:s dataskyddsförordning. Följ den. Även AI-förordningen betonar god dataförvaltning.

- ☐ GDPR respekteras för att behandla personuppgifter för det syfte organisationen vill använda AI.
- ☐ Behandlingsregistret har uppdaterats och behandlingsansvarig har utsetts. Data minimeras till vad som är strikt nödvändigt för syftet med användningen av AI.
- ☐ Datans kvalitet, noggrannhet och representativitet har bedömts och säkerställts.
- ☐ Det finns tekniska och organisatoriska åtgärder för att skydda data. En konsekvensbedömning av dataskydd (DPIA) har genomförts.
- ☐ Säkerhetsklassificering och mandat inom organisationen har gjorts tydliga för dem som arbetar med AI-tjänsten.
- ☐ Lagringsperioder för data är definierade.
- ☐ Det finns en flaggning inbyggd i systemet när lagringsperioden närmar sig slutet.

CHECKLISTA: TRANSPARENS OCH KOMMUNIKATION

EU:s AI-förordning säger att AI-tjänster och system ska vara begripliga och tillgängliga. Hur tjänsten används ska också vara tydligt. Varje algoritm ska kunna granskas.

- ☐ Det framgår tydligt för olika intressenter och aktörer vilka AI-verktyg som används, exempelvis för ny personal, förmånstagare, givare och allmänheten.
- ☐ AI-tjänstens beslut eller resultat är begripliga för de som berörs av dem.
- ☐ Användaren är informerad om att den interagerar med ett AI-system eller en AI-tjänst, exempelvis en virtuell agent.
- ☐ AI-genererat innehåll är tydligt märkt.

CHECKLISTA: MÄNSKLIG TILLSYN OCH ANSVAR

AI-teknik ska inte styra sig själv. En människa ska vara högst ansvarig för teknik, etik och korrigering av fel och brister.

- ☐ Roller och ansvar för AI-tillsyn och ansvarsskyldighet är tydligt tilldelade.
- ☐ Det finns mänsklig granskning och möjligheter till ingripande vid upptäckt av brister i AI-tjänsten.
- ☐ Medarbetare som använder eller övervakar AI-tjänsten har fått tillräcklig utbildning om dess kapacitet, begränsningar och etiska användning.
- ☐ Det finns möjlighet för personer att bestrida AI-drivna beslut eller söka upprättelse för skada.

CHECKLISTA: SÄKERHET

AI-tjänster ska vara säkra att använda. Men de är också marginaliserade för cyberhot, exempelvis hackning, datastöld, identitetsstöld och försök till utpressning.

- ☐ AI-systemets risknivå har fastställts enligt EU:s AI-förordning – som låg, mellan eller hög.
- ☐ Om högrisk-AI används, ska alla tillämpliga krav i AI-förordningen uppfyllas, bland annat riskhantering, datastyrning, dokumentation, loggning, mänsklig tillsyn, noggrannhet/robusthet/säkerhet, bedömning av överensstämmelse, registrering, enligt Artikel 27 i AI-förordningen, Fundamental Rights Impact Assessment for high-risk systems.
- ☐ Det finns personal, tekniska åtgärder och rutiner för att skydda AI-systemet mot cyberattacker.

CHECKLISTA: ÖVERVAKNING OCH ANPASSNING

Alla tekniska system, hårdvara och mjukvara uppdateras då och då. Så är det med AI också. Det kommer att finnas buggar, oavsiktliga men inbyggda fördomar och diskriminering eller andra konstruktionsfel. Därför krävs regelbunden övervakning och träning av AI-tjänsten.

- ☐ Det finns en plan för att övervaka och utvärdera AI-systemets prestanda efter att det införts.
- ☐ Resultaten granskas regelbundet med avseende på rättvisa, noggrannhet och felaktigheter.
- ☐ Det finns en mekanism för att samla in feedback från användare och andra som främst berörs av AI-tjänsten och dess resultat.
- ☐ Organisationen och leverantören är beredda att anpassa eller avbryta användningen av AI-systemet om det upptäcks att systemet eller tjänsten orsakar skada.

CHECKLISTA: LEVERANTÖRSHANTERING

I regel saknar föreningar och organisationer tillräcklig teknisk kompetens för att själva utveckla en AI-tjänst. Den köps in. Inköpare och beställare måste därför veta vilka krav som ska ställas, bland annat utifrån checklistorna här och all relevant lagstiftning. Det är organisationens ansvar att göra en korrekt beställning.

- ☐ Organisationen har tillräcklig kunskap om upphandling av tekniska tjänster för att kunna hantera de krav som ställs på AI-tjänsten.
- ☐ AI-leverantören har granskats utifrån gällande lagstiftning och specificerade krav som är unika för organisationen.
- ☐ AI-leverantören är tydlig och transparent om sitt etiska arbete. AI-leverantören är tydlig med hur tjänsten tränas och underhålls.
- ☐ Det finns ett tydligt avtal mellan organisationen och leverantören eller ett personuppgiftsbiträdesavtal som beskriver ansvar och datahanteringsprocedurer.
- ☐ Det finns lämpliga överföringsmekanismer om data överförs till tredje part inom EU eller utanför unionen.

Bygg kapacitet, samarbeta, och påverka

Enskilda civilsamhällesorganisationer kan inte hantera alla AI-frågor ensamma. Därför är det viktigt att bygga gemensam kunskap och samarbeta med andra.



Satsa på utbildning och kunskap: Ge grundläggande AI-kunskap och etikträning till all personal och styrelsemedlemmar. Detta gör tekniken lättare att förstå. Det ökar också den kritiska medvetenheten. Ge mer specialiserad utbildning till dem som jobbar direkt med AI-system. Hantera personalens oro för hur AI påverkar deras jobb. Gör det genom öppen kommunikation. Fokusera på hur AI förstärker, inte ersätter, deras roller.



Använd nätverk och samarbeta: Delta aktivt i olika samarbeten och nätverk. Dela erfarenheter, utmaningar och lärdomar med andra organisationer. Detta gemensamma lärande är mycket viktigt. Särskilt eftersom det finns begränsad offentlig information om etiska processer. Samarbete är inte bara bra, det är troligen nödvändigt för att AI ska användas ansvarsfullt i hela sektorn. Detta gäller särskilt för att lösa problem som brist på resurser och den komplexa AI-frågan. Annars kan ansvarsfull AI bli splittrad och bara finnas hos de största organisationerna.



Utforska gemensamma resurser: Se över möjligheterna att samarbeta och utveckla gemensamma resurser. Till exempel etiska riktlinjer för nischade frågor, bedömningsmallar eller utbildningsmaterial. Dela teknisk kunskap och verktyg om möjligt för att minska dubbelarbete. Det kommer att göra ansvarsfull AI mer tillgänglig för mindre organisationer.



Delta i debatten: Civilsamhället har en viktig uppgift i att vara med och forma AI-policy och reglering. Detta gäller både i Sverige och EU. Organisationer kan till exempel bidra med riktlinjer och förespråka civilsamhällets specifika behov och synpunkter. Se till att marginaliserade grupperas rättigheter skyddas. Förespråka också särskild finansiering och stöd. Detta för att möjliggöra ansvarsfull AI-användning i den ideella sektorn med begränsade resurser.



Rättighetsorganisationernas roll: Rättighetspolitiska organisationer ska kritiskt granska makt och orättvisor. För att göra detta inom AI krävs kunskap för att kunna utkräva ansvar. Hur använder din kommun AI? Vad används AI till? Kanske i socialtjänsten? Eller som roliga chattbotar? Hur använder polisen ansiktsgenkänning? Eller annan biometrisk data? Vem beställer och köper in tjänster och data? Det ska gå att granska och ifrågasätta AI-användning etiskt. Är det högrisk-AI? Kan vi få ut data om hur klassificering sker hos myndigheter som samlar in personuppgifter? Kan domstolar hantera bevis som samlats in med AI? Kom ihåg att nästan alla appar i din telefon spårar och sparar data. Detta för kommersiella syften eller AI-träning. Samma gäller för sociala plattformar.



Viktigt intressentengagemang: Gå längre än bara interna processer. Involvera aktivt de som gynnas, berörda samhällen och andra relevanta externa intressenter. Detta ska ske i design-, test-, implementerings- och utvärderingsfaserna av AI-projekt. Det främjar förtroende och ger värdefulla insikter. Det ser också till att AI-system verkligen är till nytta och passar kulturellt.

Slutsats

AI-utvecklingen innebär spännande möjligheter för civilsamhället. Effekter av smart användning är bättre effektivitet, djupare insikter från data och ökat engagemang. Det kan också ge större genomslagskraft för organisationens uppdrag. Men dessa möjligheter är kopplade till etiska utmaningar, som risk för att marginaliserade grupper diskrimineras. Privatliv och dataskydd kan hotas. Organisationernas drifta bort från sitt syfte. Förtroendet kan minska. Dessutom finns praktiska hinder som brist på resurser och kunskap.

För att hantera detta behövs ett proaktivt och etiskt tillvägagångssätt. Civilsamhällets organisationer måste se AI som mer än bara ett tekniskt verktyg. De måste agera strategiskt och se till att AI används i linje med deras värderingar. AI ska också tjäna deras grundläggande syfte att bidra till ett rättvist samhälle. Detta kräver att man följer nya lagar som EU:s AI-förordning och GDPR. Men också att man engagerar sig djupare i etiska principer. Till exempel rättvisa, öppenhet och ansvar.

**Omfamna ansvarsfull
AI för ett starkare
civilsamhälle**



Nästa steg

Vägen framåt innebär att införa starka interna styrmodeller. Dessa ska inkludera tydliga AI-policys och etiska granskningar. Det kräver också att man systematiskt gör praktiska saker. Som noggranna konsekvensbedömningar, bra datahantering samt att synliggöra och identifiera fördomar. Man ska också behålla mänsklig kontroll och ha öppenhet i AI-arbetet.

Ensam är inte stark. Med tanke på utmaningarna är samarbete, kunskapsdelning och gemensam kapacitetsutveckling nödvändigt. Detta gäller inom hela civilsamhället. Att använda nätverk, delta i debatter, svara på remisser och investera i AI-kunskap är avgörande. Detta är delar av en gemensam strategi för ansvarsfull AI-användning.

Arbetet med ansvarsfull AI är en pågående process. Den handlar om att lära, anpassa sig och reflektera etiskt. Genom att hantera denna resa klokt och i samarbete, kan civilsamhället nyttja AI:s kraft. Inte bara för att förbättra sitt arbete. Utan också för att stärka sin positiva påverkan och sina uppdrag. De kan bidra effektivare till samhället och främja mänskliga rättigheter i en digital värld. Vaksamhet, kritiskt tänkande och en stark etisk hållning kommer att vara avgörande. Detta för att AI verkligen ska tjäna mänskligheten.



Civilsamhällets forum för ansvarsfull AI

Referenser

Avdieieva, Tetiana and Oleksandr Batkhin. "Human-Rights-Compliant Use of Artificial Intelligence Systems: Toolkit for Civil Society". Digital Security Lab. 2024.

https://dslua.org/wp-content/uploads/2024/10/Toolkit_AI-for-CSOs_DSLU.pdf

BSR – Sustainable Business Network. "Overview of the Responsible AI Practitioner Guides for Taking a Human Rights-Based Approach to Generative AI." Februari 2025.

<https://www.bsr.org/files/BSR-Overview-Responsible-AI-Practitioner-Guides-Human-Rights-Based-Approach-to-Generative-AI.pdf>

Council of Europe. "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." 9 maj 2024.

<https://rm.coe.int/1680afae3c>

Davies, Sally. "Building AI playgrounds." 27 februari, 2024.

<https://scrapbook.yarn.family/building-ai-playgrounds-6367259b3824>

Europeiska kommissionen. "AI Act". Hämtad 17 juni, 2025.

<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

Europeiska kommissionen. "Ethics guidelines for trustworthy AI." 8 april, 2019.

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Europeiska unionen. Europaparlamentets och rådets förordning (EU) 2021/1119 av den 30 juni 2021 om inrättande av en ram för att uppnå klimatneutralitet och om ändring av förordningarna (EG) nr 401/2009 och (EU) 2018/1999 (europeisk klimatlag), Europeiska unionens officiella tidning L 243 (9 juli 2021): 1–17. <https://eur-lex.europa.eu/legal-content/SV/TXT/HTML/?uri=CELEX:32021R1119>

Europeiska unionen. Europaparlamentets och rådets förordning (EU) 2018/1999 av den 11 december 2018 om styrningen av energunionen och av klimatåtgärder samt om ändring av Europaparlamentets och rådets förordningar (EG) nr 663/2009 och (EG) nr 715/2009, Europaparlamentets och rådets direktiv 94/22/EG, 98/70/EG, 2009/31/EG, 2009/73/EG, 2010/31/EU, 2012/27/EU och 2013/30/EU samt rådets direktiv 2009/119/EG och (EU) 2015/652 och om upphävande av Europaparlamentets och rådets förordning (EU) nr 525/2013. Europeiska unionens officiella tidning L 328 (21 december 2018): 1–77.

<https://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:32018R1999>

Government of Canada. “Accessible and Equitable Artificial Intelligence Systems - Technical Guide.” Senast ändrad 3 juni 2025.
<https://accessible.canada.ca/creating-accessibility-standards/accessible-and-equitable-artificial-intelligence-systems-technical-guide>

Lighthouse Reports. “Sweden’s Suspicion Machine.” 27 november, 2024. Hämtad 17 juni, 2025.
<https://www.lighthousereports.com/investigation/swedens-suspicion-machine/>

Myndigheten för digital förvaltning (DIGG). “Riktlinjer för generativ AI inom offentlig förvaltning.” Hämtad 17 juni 2025.
<https://www.digg.se/ai-for-offentlig-forvaltning/riktlinjer-for-generativ-ai>

OECD. “AI principles”. Hämtad 17 juni, 2025.
<https://www.oecd.org/en/topics/sub-issues/ai-principles.html>

TechSoup. “Generative AI Guidelines for TechSoup Team.” Hämtad 17 juni, 2025.
https://pages.techsoup.org/hubfs/Downloads/2023.08.16_GenerativeAIGuidelines_TechSoup.pdf

UNESCO. “Ethics of Artificial Intelligence.” Hämtad 17 juni, 2025.
<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

UNICEF. *Policy guidance on AI for children.* 2nd ed., november 2021.
<https://www.unicef.org/innocenti/media/1341/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf>