

# AI och hållbarhet

Vägledning för Region Stockholms verksamheter för  
hållbar användning av AI



Foto framsida: Ulf Huett Nilsson

Övriga foton i vägdelningen: Hans Berggren Johnér Bildbyrå, Dan Lepp, iStock, Yanan Li,  
Henrik Möller, Johnér Bildbyrå

## Innehållsförteckning

<b>1. Inledning</b> .....	<b>4</b>
Syfte och användning .....	5
<b>2. Risker med AI kopplat till social hållbarhet, diskriminering och mänskliga rättigheter</b> .....	<b>6</b>
Ökad risk för diskriminering av invånare .....	6
Ökad risk för att AI-lösningarna inte är representativa för vår befolkning och kontext .....	8
Ökad risk för att tilliten till våra verksamheter urholkas .....	9
Ökad risk för försämrad tillgänglighet och användarvänlighet.....	9
Lärande exempel: Beslutsstöd på myndighet diskriminerade kvinnor och utrikes födda .....	10
<b>3. Risker med AI kopplat till klimat och miljö</b> .....	<b>12</b>
Risk för ökad klimatpåverkan .....	12
Risk för ökad mängd elektroniskt avfall .....	13
Risk för ökad miljöpåverkan .....	14
Lärande exempel: När ökad AI-användning kan driva upp koldioxidutsläppen och varför .....	14
<b>4. Möjligheter med AI utifrån ett hållbarhetsperspektiv</b> .....	<b>16</b>
Digitala tvillingar i bygg- och anläggningsprocesser .....	16
Smart journalföring i vården .....	17
AI-baserade tolkverktyg i vården .....	19
Hållbarhetsrapportering och klimatberäkning .....	19
Realtidsövervakning av fastigheter och anläggningar .....	20
Datadriven leverantörsanalys .....	20
<b>5. Att använda AI på ett hållbart sätt</b> .....	<b>22</b>
Samhällsnyttan och värdeskapande i centrum .....	22
Kvalitetssäkra för att minska risken för diskriminering.....	22
Se till att Region Stockholms AI-användning går att förstå och förklara	23
Optimera genom medvetna val och resurseffektivitet .....	23
Arbeta aktivt med AI-leverantörer .....	24
<b>Bilaga 1: Hållbarhetsbedömning av AI-lösningar</b> .....	<b>26</b>

## 1. Inledning

Utvecklingen av artificiell intelligens (AI) skapar stora möjligheter för nya lösningar, effektivare processer och förbättrade beslut inom Region Stockholms kärnområden hälso- och sjukvård, kollektivtrafik, regional planering och kultur. Nya AI-verktyg kan vara till stor nytta genom att bidra till högre kvalitet i verksamheterna, stärka personalens förutsättningar att fatta välgrundade beslut samt frigöra resurser för kärnuppgifter med större samhällsnytta. Olika typer av AI-verktyg kan till exempel användas för att stödja medicinska bedömningar genom analys av stora mängder patientdata, effektivisera schemaläggning, automatisera administrativa flöden och dokumentation, förbättra planeringen av kollektivtrafik med hjälp av prediktiva modeller, analysera fritextsvar samt ge underlag för mer träffsäkra prognoser.

Detta kan vara till stor nytta för Region Stockholms verksamheter. Samtidigt medför AI nya risker och utmaningar, inte minst kopplade till hållbarhet. AI innebär många nya möjligheter, men riskerna bör alltid analyseras och hanteras. Ibland är ett AI-verktyg inte rätt lösning, utan enskilda bedömningar behöver alltid göras.

Offentlig sektor har ett unikt samhällsuppdrag: att skapa nytta för alla medborgare, skydda rättigheter och förvalta gemensamma resurser.

När AI nu blir en del av hur offentlig sektor tar fram beslutsunderlag, välfärdstjänster och planering måste tekniken användas på ett sätt som stärker dessa värden, inte utmanar dem. Beslut som fattas med stöd av AI eller processer som effektiviseras med hjälp av AI kan påverka människors liv, rättigheter och tillgång till offentliga tjänster.

Till skillnad från privata aktörer ska offentlig verksamhet fungera för alla invånare, och Region Stockholms samhällsservice är till för alla i Stockholmsregionen. Därför har vi ett extra stort ansvar att se till att all digitalisering är till nytta för vårt kärnuppdrag och inte motverkar våra mål. Att utveckla och använda AI inom Region Stockholm handlar därför inte bara om effektivisering och innovation. Det handlar också om samhällsnytta, förtroende, ansvar och långsiktig hållbarhet.

Region Stockholm har högt ställda mål när det gäller hållbarhet och ska kraftigt minska klimatpåverkan, verka för en giftfri miljö och säkerställa socialt hållbar utveckling i Region Stockholms verksamheter så att alla invånare i Stockholmsregionen kan ta del av samhällsservicen på ett jämlikt, jämställt och inkluderande sätt.

Samtidigt har Region Stockholm höga ambitioner kopplat till att använda AI för att förbättra sina verksamheter. Ansvarsfull och hållbar användning av AI kan stärka både kvalitet, effektivitet och samhällsnytta.

## Syfte och användning

Denna vägledning syftar till att ge verksamheter inom Region Stockholm stöd i hur AI kan planeras, utvecklas, upphandlas och användas på ett hållbart sätt.

Vägledningen syftar till att stötta medarbetare, chefer och politiker i att bedöma när och hur AI kan användas i Region Stockholms verksamheter till faktisk nytta för invånare, patienter, resenärer eller medarbetare, och hur hållbarhetsrisker kopplade till AI kan hanteras. Vägledningen ger exempel på hur hållbarhet kan integreras i hela kedjan, från behovsanalys och upphandling till användning, uppföljning och avveckling av AI-lösningar.

Vägledningen ger råd utifrån två perspektiv.

1. **Risker med AI utifrån ett hållbarhetsperspektiv**, det vill säga hur Region Stockholm kan använda mer AI i våra verksamheter utan att äventyra våra hållbarhetsmål om minskad klimat- och miljöpåverkan samt socialt hållbar utveckling med minskad risk för diskriminering.
2. **Möjligheter med AI utifrån ett hållbarhetsperspektiv**, det vill säga områden där hållbarhetspotentialen bedöms som stor om vi använder AI för att stärka vårt arbete.

Vägledningen avgränsas till verksamhetsnivå och inkluderar inte detaljerade råd om hur enskilda medarbetare ska förhålla sig till olika AI-verktyg. Denna typ av råd och vägledning finns tillgänglig på Region Stockholms intranät och i varje enskild verksamhet.



## **2. Risker med AI kopplat till social hållbarhet, diskriminering och mänskliga rättigheter**

AI bygger på data. Kvaliteten på dessa data avgör i hög grad hur rättvis, inkluderande och tillförlitlig tekniken blir. Olika AI-verktyg är tränade på olika datamängder, dataset från olika länder och utvecklade med skilda tekniska och funktionella utgångspunkter. Detta innebär att olika verktyg kan ge olika svar på samma fråga. Skillnaderna kan i vissa fall spegla varierande perspektiv, värderingar eller normer som finns representerade i de data som används. Det är därför viktigt att användare inte betraktar AI-verktygens svar som neutrala eller objektiva, utan som resultat av underliggande val i data, design och styrning.

Region Stockholms verksamheter omfattas av flera lagar och förordningar som reglerar hur vi får använda data och i förlängningen hur vi kan använda AI. För AI-lösningar kan det till exempel röra sig om AI-förordningen (AI-Act), dataskyddsförordningen (GDPR), sekretesslagar eller patientdatalagen.

AI-förordningen (AI Act) bygger på en så kallad riskpyramid. Ju högre risk en AI-lösning bedöms ha, desto striktare blir kraven. AI-förordningen delar in AI-system i fyra nivåer: oacceptabel risk, hög risk, begränsad risk och minimal risk. För Region Stockholms verksamheter är det viktigt att förstå vilken risknivå en specifik AI-lösning landar i, så att verksamheten kan följa reglerna på rätt sätt och samtidigt använda AI på ett tryggt sätt. Flertalet av de risker som presenteras nedan behöver hanteras mer noggrant om verksamhetens AI-verktyg klassas som hög risk i enlighet med AI-förordningen.

Nedan presenteras risker med ökad AI-användning kopplad till antidiskriminering och mänskliga rättigheter.

### **Ökad risk för diskriminering av invånare**

Det finns en risk att AI-verktyg tränas på historiska data som speglar tidigare orättvisor, till exempel skillnader kopplade till kön, ålder, etnicitet, funktionsnedsättning eller socioekonomi, vilket kan leda till att vissa grupper missgynnas i exempelvis beslut om vård, stöd eller service.

Detta kallas ofta för bias inom AI och betyder att ett AI-system kan ge snedvridna eller orättvisa resultat på grund av hur det har tränats eller byggts. Eftersom AI lär sig från data som människor har skapat, kan den ärva mänskliga fördomar, till exempel kopplade till kön, ålder eller bakgrund.

Exempelvis om en språkmodell har tränats på fler texter som beskriver chefer som män än kvinnor, kan den omedvetet börja koppla "chef" till "man". Eller om bilder på sjukvårdspersonal i träningsdata oftast visar vita personer, kan AI:n återge en skev bild av verkligheten.

Om träningsdatan inte är balanserad och kvalitetssäkrad för att rensa bort bias kan AI börja gynna vissa grupper och missgynna andra, ofta utan att det märks direkt. Därför är det viktigt att vara medveten om bias i AI för att kunna utveckla och använda tekniken på ett mer rättvist och ansvarsfullt sätt.

Det går redan nu se att grupper som historiskt har diskriminerats ofta löper särskilt stor risk att missgynnas, eftersom AI-modeller kan förstärka rådande marginalisering och diskriminering. Det beror dels på kvaliteten på träningsdatan och bristen på kontroll över den, dels på att generativa AI-tjänster kan vara konfigurerade för att vara användaren till lags och generera det som sannolikt uppfattas som önskvärt.

Bias i stora språkmodeller, som ChatGPT, Gemini, Copilot, Mistral, DeepSeek och Perplexity, uppstår genom hur de tränas. Modellerna lär sig språk genom att analysera enorma mängder text från internet, böcker och andra källor för att hitta mönster i hur ord och idéer används. Eftersom dessa texter är skrivna av människor speglar de också mänskliga värderingar, normer och fördomar. Mycket av det tillgängliga materialet kommer från engelskspråkiga och västerländska källor, vilket gör att perspektiv från västvärlden och majoritetskulturer dominerar, medan röster från andra delar av världen och minoritetsgrupper blir underrepresenterade. Modellen lär sig alltså inte vad som är rätt eller fel eller hur olika avvägningar bör göras, utan bara vad som oftast förekommer i datan, vilket gör att den genom sina instruktioner kan återskapa snedvridna eller begränsade perspektiv.

Det finns flera mekanismer som kan förklara hur bias uppstår i AI-system. Några konkreta exempel är:

- Snedvridna träningsdata. Om den data som används för att träna modellen innehåller mönster från verkligheten där vissa grupper redan övervakas mer, misstänkliggörs oftare eller anmäls i högre grad, kommer AI-modellen att lära sig dessa mönster och riskerar att förstärka dem i sina rekommendationer och beslut.
- Indirekta variabler, så kallade proxyvariabler. AI-modellen kanske inte använder kön, etnicitet eller bakgrund som direkta kriterier, men andra variabler, som exempelvis inkomstnivå, utbildningsnivå, bostadsområde eller yrkestyp, kan ha stark koppling till dessa faktorer. På så sätt kan modellen ändå särskilja

grupper indirekt, trots att känsliga variabler formellt sett inte används.

- Obalans i datamängden. Om vissa grupper är överrepresenterade i en viss kategori, av historiska eller strukturella skäl, kan modellen börja tolka detta som ett starkare samband än vad som faktiskt finns.
- Överdrivet fokus på prediktiv precision, alltså att maximera träffsäkerheten i att hitta sannolika avvikelser, utan att väga in rättvise-, jämlikhets- eller diskrimineringsaspekter. AI kan inte föra semantiska resonemang utan dessa parametrar måste läggas in i modellen och testas.
- Monokulturbias uppstår när en organisations eller samhällssektors många individuella perspektiv och erfarenheter ersätts av en enskild teknologileverantörs synsätt, antaganden och värderingar. I praktiken innebär detta att den mångfald av bias som annars kan identifieras, balanseras eller justeras inom en verksamhet ersätts av en homogen, ofta implicit världsbild som byggts in i den tekniska lösningen.

Om AI används för att effektivisera exempelvis beslutsfattande, kontroller eller handläggning kan det innebära att ett AI-system ersätter många människors olika fördomar och bias. Ökad AI-användning, utan tillräcklig kvalitetssäkring, ökar därför risken för att befintliga ojämlikheter och bias byggs in i system och därmed drabbar fler invånare.

## **Ökad risk för att AI-lösningarna inte är representativa för vår befolkning och kontext**

AI-system kan tränas på olika typer av data, och det är viktigt att fråga sig hur representativ träningsdatan är i förhållande till det specifika sammanhang där AI-lösningen ska användas. Det kan exempelvis handla om vilket land träningsdatan kommer från, hur representativ den grupp människor som inkluderats är och hur gammal datan är.

Utgångspunkten för oss behöver vara att träningsdatan ska vara så representativ och kontextnära som möjligt. Annars riskerar de sociala normer och värderingar som reproduceras genom träningsdatan att bli ett problem.

Exempelvis är det viktigt att AI-lösningar som ska användas i en viss kontext speglar befolkningen som lever där. Många länder har en helt annan syn på hbtqi-personer eller jämställdhet än Sverige. I många andra länder förväntas till exempel inte personer med funktionsnedsättning ha ett arbete. Om data har importerats är det inte säkert att de är representativa för vår kontext. Äldre data kan också bli problematiska

eftersom normer och förhållanden i samhället utvecklas över tid, vilket gör att historiska data inte alltid är representativa för dagens samhälle. Om datan endast inkluderar personer av ett visst kön eller en viss ålder är den kanske inte överförbar till andra grupper.

Var data kommer från, vilka som inkluderats och hur gammal den är påverkar alltså hur AI-modellen reproducerar och analyserar.

## **Ökad risk för att tilliten till våra verksamheter urholkas**

Invånarnas data behöver hanteras med stort ansvar och stor försiktighet. Om det finns tillgång till stora mängder data behövs riskerna med att AI-modellen hanterar personuppgifter och känsliga uppgifter analyseras. Det innebär att den data som ska användas behöver riskbedömas och att det behöver finnas tydliga rutiner för hur data samlas in, lagras och används.

Bristande insyn i hur modellerna fungerar gör det dessutom svårt för enskilda att förstå eller ifrågasätta beslut, vilket riskerar att urholka rättssäkerheten och förtroendet för det offentliga.

Som offentlig aktör behöver vi kunna granskas och ha tydliga mekanismer som gör det möjligt för invånare att utkräva ansvar vid eventuella fel eller oönskade konsekvenser. Detta inkluderar behov av rutiner för granskning, återkoppling och möjligheten att snabbt åtgärda problem som kan uppstå.

Dessutom kan vissa AI-system vara svåra för invånare att förstå och fungera som "svarta lådor" som fattar beslut och gör bedömningar utan att allmänheten vet exakt vad de grundar sig på. Detta kan minska förtroendet mellan Region Stockholms verksamheter och patienter, resenärer och invånare.

Förklarbarhet, transparens och hur risker hanteras vid ökad AI-användning blir därmed centralt för att bibehålla förtroendet för Region Stockholms verksamheter.

## **Ökad risk för försämrade tillgänglighet och användarvänlighet**

När nya AI-verktyg introduceras finns en risk att vissa grupper i samhället eller inom arbetsgrupper inte har, eller ges, samma möjligheter eller förutsättningar att ta till sig den nya tekniken eller de nya arbetssätten. Liknande mönster kan kännas igen vid andra stora teknikskiften, och vid introduktion av nya digitala verktyg finns det en risk att olika segment skapas i samhället, där vissa använder ny teknik och andra inte.

För AI-lösningar som ska användas av många olika grupper, till exempel både invånare och medarbetare, är det viktigt att tidigt analysera hur tillgänglig och användarvänlig lösningen faktiskt är i praktiken. Frågan bör inte enbart handla om teknisk funktion, utan om lösningen underlättar eller försvårar för både personal och invånare med olika förutsättningar att ta del av tjänsten.

När AI-användningen ökar blir risken större att nya gränssnitt, automatiserade beslut eller otydliga återkopplingar skapar fler hinder än de tar bort, särskilt om lösningen optimeras för "genomsnittsanvändaren" och inte för variationen i behov, språk och digital vana.

För Region Stockholm är det viktigt att hantera denna risk eftersom lagen om tillgänglighet till digital offentlig service (DOS-lagen, lag 2018:1937) kräver att offentliga aktörers webbplatser och appar är tillgängliga, vilket innebär att de ska kunna uppfattas, hanteras och förstås av alla användare, inklusive personer med funktionsnedsättning.

Bristande tillgänglighet kan direkt försämra jämlik tillgång till samhällsservice, öka trycket på alternativa kontaktvägar (till exempel telefon och fysisk service) och i förlängningen påverka kvaliteten i Region Stockholms tjänster.

### **Lärande exempel: Beslutsstöd på myndighet diskriminerade kvinnor och utrikes födda**

Ett exempel på risker med AI i offentlig verksamhet var när en stor, svensk myndighet utvecklade en AI-modell för att identifiera personer som potentiellt fuskat med bidrag, med syftet att spara resurser och rikta kontroller mer träffsäkert.

Vid en extern granskning framkom dock att modellen oftare flaggade vissa grupper än andra. Det rörde sig inte om en medveten diskriminering, utan om en oavsiktligt inbyggd bias i modellen. Granskningen visade att kvinnor och personer med utländsk bakgrund flaggades i högre utsträckning än andra grupper. Samtidigt visade verkliga utfall att kvinnor och män stod för ungefär lika stor andel felaktigheter. Trots detta pekade AI-modellen ut kvinnor som sannolika fuskare mer än dubbelt så ofta som män, vilket skapade en systematisk snedfördelning i hur kontroller riktades.

Förklaringen låg i hur AI-modellen hade tränats och i den data som låg till grund för modellen, som troligtvis inte hade kvalitetssäkrats i tillräckligt hög utsträckning. AI-modellen hade lärt sig mönster från historiska data om vilka som tidigare misstänkts och kontrollerats. Eftersom dessa

processer redan påverkats av samhällsliga strukturer och ojämlikheter reproducerade och förstärkte modellen dessa mönster.

Exemplet visar att AI inte är neutralt, utan kan spegla och förstärka tidigare beslut och strukturer om inte data, metod och resultat granskas ur ett hållbarhets- och rättighetsperspektiv.

I bästa fall upptäcks detta redan i träningsfasen och kan korrigeras innan en AI-modell tas i bruk. Diskriminering sker även utan AI-modeller, men då kanske det handlar om en enskild handläggare och relativt få drabbade. Med en AI-modell som inte är tillräckligt kvalitetsgranskad riskerar de drabbade att bli många.



### 3. Risker med AI kopplat till klimat och miljö

AI-utvecklingens klimat- och miljöpåverkan är en aspekt som ibland förbises i organisationers strategiska bedömningar. Vid en genomlysning av stora techbolags principer och policyer för "ansvarsfull AI" under 2025 var det endast några få som inkluderade klimat- och miljöpåverkan. Ansvarsfull AI handlade i större utsträckning om andra viktiga frågor, som etik, transparens, informationssäkerhet och förklarbarhet, men inkluderade inte klimat- och miljöpåverkan i samma utsträckning.

Samtidigt går det att utläsa i de stora techbolagens hållbarhetsrapporter att deras klimat- och miljöpåverkan har ökat kraftigt under de senaste åren, och flera av dem hänvisar till ökad användning av AI och utbyggnad av datacenter som en förklaring. Särskilt utvecklingen av stora språkmodeller är mycket resurskrävande och ställer höga krav på superdatorer med kraftfulla grafikprocessorer och omfattande minneskapacitet. Driften av dessa system kräver stora mängder energi och användning av sällsynta jordartsmetaller, vilket sammantaget bidrar till ökad klimat- och miljöbelastning.

För en offentlig aktör som Region Stockholm, med ett uttalat ansvar att bidra till Sveriges klimatmål, är det därför viktigt att noggrant hantera de risker som en ökad användning av AI kan medföra. Nedan beskrivs centrala klimat- och miljörelaterade risker kopplade till AI-teknikens utveckling och användning.

#### **Risk för ökad klimatpåverkan**

AI-modeller, särskilt stora språk- och transkriptionsmodeller, kräver mycket beräkningskraft både vid träning och drift. Det kräver i sin tur mycket energi, och efterfrågan på datacenter för att utveckla AI har ökat snabbt.

De stora techbolagen kapprustar för att möta den växande AI-efterfrågan, och allt fler stora datacenter byggs runt om i världen. Exempelvis förbrukade de stora datacentren sammantaget 415 terawattimmar under 2025, vilket är mer än hela Storbritanniens förbrukning. Detta väntas mer än fördubblas till 2030, enligt Internationella energiorganet IEA.

En utmaning är att expansionen av antalet datacenter sker på flera olika platser runt om i världen där tillgången till fossilfri el är begränsad. Många datacenter drivs därför med fossil energi, såsom olja, naturgas och kol.

För Region Stockholm, som har ambitiösa mål om klimatneutralitet, innebär den ökade klimatpåverkan från AI en dubbel utmaning. Om Region

Stockholms verksamheter använder allt fler AI-baserade lösningar som i sin tur bygger på datacenter med hög klimatpåverkan riskerar de att öka de konsumtionsbaserade utsläppen, så kallade scope 3-utsläpp, vilket försvårar möjligheten att nå klimatmålen. Om Region Stockholm använder egenutvecklade lösningar on premise (lokal lösning) eller molnlösningar som drivs av fossilfri energi blir påverkan inte lika stor. Val av lösning, hur energieffektiv den är och krav på fossilfri drift är därför avgörande för hur stor risken för ökade utsläpp blir.

AI-verktyg kan samtidigt vara ett verktyg för hållbarhetsomställning, exempelvis genom energieffektivisering och bättre beslutsunderlag. Men om själva driftförutsättningarna för AI medför omfattande utsläpp finns en risk att digitaliseringsarbetet i praktiken motverkar hållbarhetsmålen.

Bristande transparens kring var och hur AI-tjänster körs begränsar Region Stockholms möjlighet att bedöma klimatpåverkan och ställa tydliga krav i upphandlingar, vilket försvårar kontrollen över de indirekta effekterna av digitaliseringen.

Sammantaget kan ökad AI-användning därmed bli en förbisedd men betydande riskfaktor i Region Stockholms hållbarhetsarbete, om inte klimatpåverkan vägs in redan vid planering, upphandling och drift av digitala tjänster.

## **Risk för ökad mängd elektroniskt avfall**

Implementering och drift av nya AI-lösningar kräver ofta kraftfull och uppdaterad utrustning och hårdvara, vilket leder till ett kontinuerligt behov av att byta ut servrar, processorer, datorer och annan utrustning. När dessa komponenter blir föråldrade uppstår ett växande inflöde av elektroniskt avfall (e-avfall). Elektroniskt avfall är komplext att återvinna. Många it-produkter innehåller både värdefulla metaller och farliga ämnen som kan orsaka skador på miljö och hälsa om de inte hanteras korrekt.

För Region Stockholm är detta en risk eftersom regionen har nästan 50 000 medarbetare och är en stor användare av avancerad medicinteknisk utrustning och it-utrustning. Om hårdvara och utrustning ersätts i snabb takt utan tydliga rutiner för återanvändning, materialåtervinning och säker avveckling riskerar Region Stockholm att bidra till ökade utsläpp, resursförluster och negativa miljöeffekter. Det skulle stå i strid med målet om en klimatneutral och kostnadseffektiv verksamhet.

Genom att integrera krav på återbruk, längre livscyklar och cirkulära leverantörsled i både upphandling och intern planering kan Region

Stockholm hålla nere sin klimatpåverkan, öka de cirkulära flödena av material och produkter och samtidigt hålla ner kostnaderna.

## **Risk för ökad miljöpåverkan**

Utveckling och drift av avancerad AI kräver omfattande datorkraft och specialiserad hårdvara, framför allt grafikkort (GPU:er) och servrar. Produktionen av sådan utrustning är starkt resurskrävande och bygger på metaller som litium, kobolt och andra sällsynta jordartsmetaller. Dessa material är avgörande för det digitala samhället, men också för andra teknikområden som elbilar och batterilagring, vilket skapar en global konkurrens om begränsade resurser för hållbar omställning.

Gruvdriften av dessa metaller medför både klimat- och miljöbelastning samt sociala risker, som otrygga arbetsvillkor och påverkan på lokalsamhällen i utvinningsländer. Den globala AI-infrastrukturen kräver också stora mängder vatten, bland annat genom de kylsystem som krävs i datacenter. Enligt organisationen EESI kan stora datacenter konsumera uppemot 5 miljoner gallon per dag, vilket motsvarar den dagliga vattenförbrukningen i en stad med 10 000 till 50 000 invånare.

För Region Stockholm är denna risk viktig att hantera eftersom ökad AI-användning påverkar miljö- och resursavtryck och kan leda till försämrade villkor i leverantörsled. En medveten styrning av upphandling, livscykelhantering och återanvändning av hårdvara kan minska risken.

## **Lärande exempel: När ökad AI-användning kan driva upp koldioxidutsläppen och varför**

Ett globalt techbolag har rapporterat att dess koldioxidutsläpp ökat kraftigt sedan 2019, vilket i huvudsak kopplas till högre energiförbrukning och användning av fossil energi i bolagets datacenter. Som bolaget själva beskriver beror detta på att AI-tekniken kräver mycket beräkningskraft och därmed mer el. Bolaget låg länge i framkant med ambition om nettonollutsläpp till 2030, men har nu reviderat sina egna hållbarhetsmål och sänkt ambitionerna.

Samtidigt lyfter bolaget fram att AI kan vara en viktig möjliggörare för att minska utsläpp i framtiden, exempelvis genom effektivare processer och smartare styrning. I den egna omställningen synliggörs däremot en tydlig målkonflikt. När AI integreras brett i produkter och tjänster kan det bli svårt att minska utsläppen i samma takt som energibehovet ökar.

För att göra klimatpåverkan mer begriplig har bolaget även försökt synliggöra utsläpp på användarnivå, till exempel genom uppskattningar

per prompt. Det är pedagogiskt, men helhetsbilden kvarstår: bolaget har ökat sina koldioxidutsläpp kraftigt, och bolagets egen förklaring är den ökade användningen av AI. När infrastrukturen, volymerna, modellstorlekarna och datacenterkapaciteten skalar upp blir den totala energianvändningen och vilken typ av energi som används en strategisk hållbarhetsfråga.

### **Tre bakomliggande anledningar till ökade koldioxidutsläpp vid ökad AI-användning**

För det första börjar stora datacenter bli alltmer energieffektiva, men AI ökar efterfrågan på el snabbare än effektiviseringar hinner kompensera. Även när ett datacenter är väl optimerat med exempelvis lågt PUE<sup>1</sup> kan den totala energiförbrukningen stiga kraftigt om mängden beräkningar och kapaciteten växer (till exempel AI-träning/inferens).

För det andra är elbehovet lokalt svårt att matcha med ny förnybar produktion och nätkapacitet när efterfrågan växer snabbt. I sådana lägen blir marginaelen ofta fossilbaserad (naturgas, kol eller olja), vilket gör att ökade datacenterlaster ger tydlig klimatpåverkan även om företaget har övergripande mål om fossilfri drift. För techbolaget i exemplet var användningen av fossilbaserad energi avgörande för de kraftigt ökade utsläppen.

För det tredje är en del resursanvändning "onödig" ur ett verksamhetsperspektiv. Tung kod, överdimensionerade applikationer, onödig datatrafik och stora modeller för små uppgifter driver upp elanvändningen. Här finns många energieffektiviseringsvinster att göra framgent.



---

<sup>1</sup> Power Usage Effectiveness är ett mått på hur energieffektivt ett datacenter är.

## 4. Möjligheter med AI utifrån ett hållbarhetsperspektiv

Ökad användning av AI i Region Stockholm rymmer stora möjligheter ur ett hållbarhetsperspektiv. Rätt utformad AI kan göra vården, kollektivtrafiken och den övergripande förvaltningen mer jämlik, träffsäker, hållbar och resurseffektiv.

I detta kapitel presenteras väl utvalda exempel där ökad AI-användning har stor hållbarhetspotential.

Samtidigt är tekniken under utveckling och, som med många nya initiativ, behöver en rad avvägningar göras och riskanalyser beaktas när lösningarna ska införas. Nya AI-verktyg kan komma med lösningar på vissa problem, men andra problem kan också uppstå och behöva bedömas och hanteras. Gedigen riskanalys och riskhantering ska ses som en förutsättning för att de exempel som presenteras nedan verkligen ska resultera i hållbarhetsnytta.

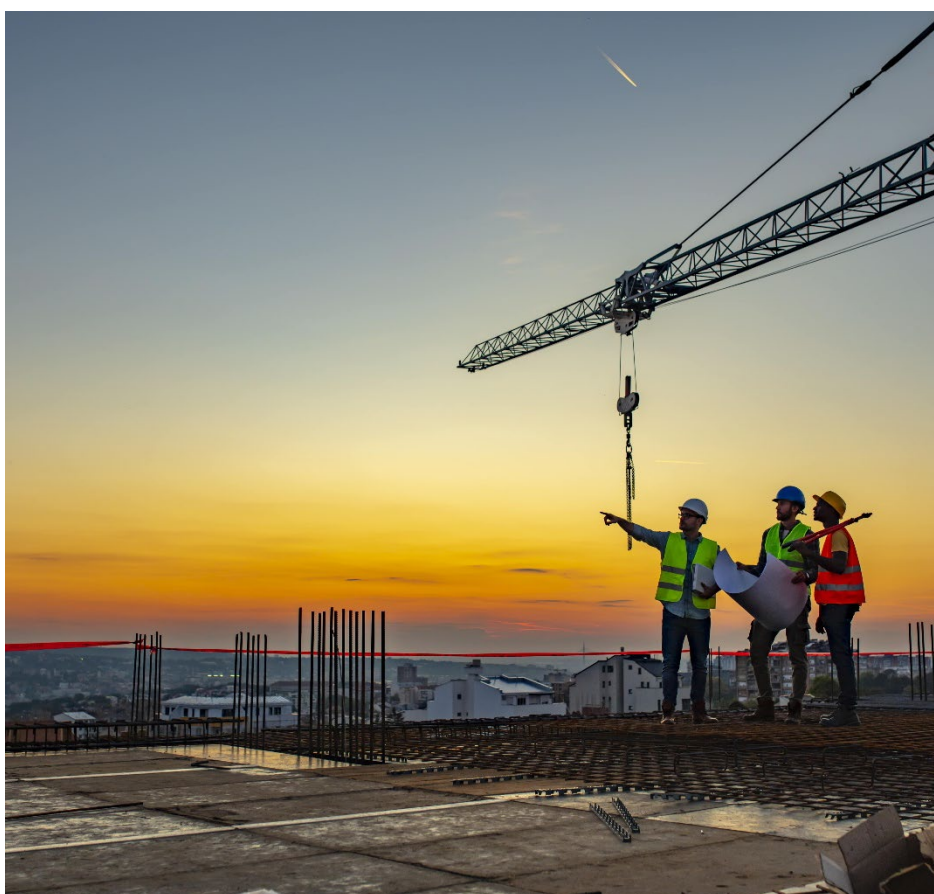
### Digitala tvillingar i bygg- och anläggningsprocesser

Digitala tvillingar i bygg- och anläggningsprocesser kan ge Region Stockholm en tydlig hållbarhetspotential genom minskade utsläpp, effektivare resursanvändning och bättre livscykelstyrning av stora investeringar i kollektivtrafik och fastigheter.

En digital tvilling är en digital representation av en fysisk anläggning eller byggnad, ofta en kombination av 3D-modell och uppdaterade data från projektering, drift och sensorer, som gör det möjligt att följa, analysera och simulera hur något fungerar över tid. Genom digitala tvillingar kan Region Stockholms verksamheter simulera olika utformningar, materialval och byggmetoder innan byggstart, så att koldioxidutsläpp, energiåtgång och resursflöden kan optimeras redan i planeringsskedet. I stora projekt kan tvillingen användas för att planera logistik, schakt, transporter och maskinanvändning på ett sätt som minimerar störningar, spill och klimatpåverkan. Det kan även möjliggöra klimatanpassning utifrån risk och sårbarhetsanalyser kring hur den kan påverkas vid ett ändrat klimat.

I praktiken kan digitala tvillingar stödja livscykelanalyser av stationer, anläggningar, fastigheter, tunnlar och installationer och därmed hjälpa Region Stockholm att styra mot lägre livscykelutsläpp. Erfarenheter från regional utveckling i exempelvis Aachen visar också hur digitala tvillingar och modeller i större skala kan koppla sensordata till beslutsunderlag för mer hållbara stads- och infrastrukturbeslut, vilket Region Stockholm kan bygga vidare på i sina egna investeringar.

Detta är särskilt viktigt eftersom Region Stockholm står inför mycket stora klimat- och investeringsåtaganden inom kollektivtrafik och fastigheter, där en stor del av klimatpåverkan och kostnaderna uppstår över hela livscykeln. Med digitala tvillingar kan Region Stockholm simulera olika alternativ för att få bättre beslutsunderlag och minska omprojektering och driftstörningar, vilket sparar pengar och reducerar onödiga utsläpp och materialförbrukning. Tekniken kan även stärka transparens och uppföljning mot hållbarhetsmål genom att synliggöra energiprestanda, materialflöden och utsläpp på ett sätt som kan följas upp systematiskt.



### **Smart journalföring i vården**

AI-baserad smart journalföring kan erbjuda möjligheter att effektivisera och förbättra vårdens arbete inom Region Stockholm. Exempelvis testas tal till text-verktyg som kan sammanfatta samtal mellan vårdpersonal och patienter automatiskt till strukturerade journalanteckningar. Detta minskar den administrativa bördan, men framför allt frigörs tid och utrymme till vårdpersonalens kärnuppdrag, mötet med patienten.

Det utvecklas i dag olika verktyg som kan skapa korta sammanfattningar av patientens tidigare sjukdomshistorik inför ett vårdbesök, vilket ger personalen bättre förutsättningar att förbereda sig. Det bidrar till ökad trygghet, minskad risk för förbisedda uppgifter och ett gott och professionellt bemötande. Patienten kommer då troligen att uppleva att vårdpersonalen är mer närvarande och insatt, vilket stärker tilliten och kvaliteten i vårdkontakten och i slutändan även vårdutfallet.

Exempel på konkreta nyttor är kortare tid för journalföring efter patientmöten, bättre sammanhållning av patientdata och snabbare introduktion av ny personal som lättare får en överblick över patienters sjukdomshistorik. Dessutom kan minskad pappershantering och mer resurseffektiv informationshantering bidra till ökad miljömässig hållbarhet.

För Region Stockholm skulle detta kunna innebära att vårdpersonalen kan ägna mer tid åt patientmöten i stället för administration, vilket ökar både arbetsglädjen och vårdens värde för invånarna. På sikt kan smart journalföring bidra till en mer socialt hållbar vård där teknik används för att förstärka och förenkla, inte ersätta, vårdpersonalens insatser.



## **AI-baserade tolkverktyg i vården**

AI-baserade tolkverktyg testas på flertalet platser inom svensk och nordisk hälso- och sjukvård. AI-tolkning översätter talat språk eller skriftliga inspel i realtid och kan vara ett snabbt och skalbart komplement till mänsklig tolkning. AI-baserade tolkverktyg har omedelbar tillgång till många språk, oavsett plats eller tid på dygnet. Det stöder även sällsynta och inhemska språk som är svåra att täcka med mänskliga tolkar, och uppdateras kontinuerligt med nya språk.

AI-baserade tolkverktyg passar kanske extra bra för kortare patientmöten vid enklare dialoger, till exempel vid provtagning eller omvårdnad, och ger personalen möjlighet att säkerställa att information förstås korrekt. AI-baserade tolkverktyg går att använda på mobiltelefoner, surfplattor och datorer, och det finns lösningar där informationen inte sparas på enheterna och där samtalen som översätts inte används för att träna AI.

Det finns stor potential i att AI-baserade tolkverktyg kan främja jämlik tillgång till vård, höja patientsäkerheten och korta väntetider. AI-baserade tolkverktyg kan därmed bli ett viktigt komplement till mänsklig tolk, som ändå bör prioriteras för djupgående medicinska samtal, särskilt när precision, erfarenhet och empati behövs.

## **Hållbarhetsrapportering och klimatberäkning**

AI har stor potential att förenkla, möjliggöra och automatisera insamling, kvalitetssäkring och analys av data från många verksamheter och system. Det kan minska manuellt arbete, korta ledtider och höja datakvaliteten genom att identifiera luckor, avvikelser och dubbelräkning.

AI kan stödja mer träffsäkra klimatberäkningar genom att matcha inköps- och aktivitetsdata mot relevanta emissionsfaktorer. Genom att automatisera insamling och strukturering av data från inköp, fakturor, avtal, logistik och resor, även när informationen ligger i fritext eller är otydligt kategoriserad, kan särskilt den konsumtionsbaserade klimatpåverkan (scope 3) beräknas på ett mer effektivt sätt. AI kan också kvalitetssäkra beräkningarna genom att upptäcka avvikelser, dubbelräkning och dataluckor samt peka ut vilka leverantörer eller produktgrupper som bör prioriteras eftersom de står för störst utsläpp.

Sammantaget ger det bättre beslutsunderlag, mer transparent uppföljning mot klimatmål och en mer resurseffektiv rapporteringsprocess.

## **Realtidsövervakning av fastigheter och anläggningar**

Realtidsövervakning innebär att fastigheter och anläggningar följs och styrs löpande med hjälp av sensorer, uppkopplade system (Internet of Things) och analys i stället för att man främst utgår från manuella kontroller, fasta tidsscheman eller historiska genomsnitt.

Det kan handla om att kontinuerligt mäta och förstå energiförbrukning, temperatur, luftkvalitet, ventilation, fukt, tryck, beläggning i lokaler, driftstatus i tekniska installationer samt avvikelser i maskiner och system. När dessa data samlas i en gemensam plattform kan AI användas för att upptäcka mönster, förutse behov och föreslå eller automatiskt genomföra justeringar inom säkra och fördefinierade ramar.

För Region Stockholm kan realtidsövervakning bidra till minskad klimatpåverkan genom att energi, lokaler och tekniska system används mer effektivt. I sjukhus, vårdbyggnader, anläggningar och administrativa fastigheter kan AI exempelvis optimera och energieffektivisera ventilation och värme utifrån faktisk beläggning, aktuellt väder och önskade komfortnivåer. Genom att analysera historiska data tillsammans med realtidsdata kan systemen också förutsäga framtida energibehov och anpassa driftstrategier i förväg, vilket ger lägre energianvändning, stabilare inomhusklimat och ofta också lägre kostnader.

Realtidsövervakning stödjer även prediktivt underhåll, där AI hjälper till att förutse när komponenter eller system är på väg att falla eller tappa prestanda. I stället för att åtgärda problem först när de orsakar driftstopp kan underhåll planeras mer träffsäkert och i rätt tid. Exempelvis hissar eller rulltrappor kan utrustas med sensorer som mäter drift, temperatur och vibrationer. Datan skickas till en säker molntjänst som analyserar mönster och kan förutse fel, visa felkoder och automatiskt skapa en serviceorder innan de går sönder. Det kan förlänga livslängden på installationer, minska oplanerade avbrott och bidra till en mer robust infrastruktur med lägre total kostnad över tid.

## **Datadriven leverantörsanalys**

AI-modeller kan analysera inköpsdata tillsammans med öppna källor för att matcha leverantörers energianvändning, koldioxidutsläpp, transportkedjor och materialval, och därmed ge bättre beslutsunderlag i upphandlings- och inköpsprocesser. Genom att väga in historiska data, marknadsfaktorer och säsongsvariationer kan AI också bidra till mer tillförlitlig planering och prediktiva modeller för resursoptimering, vilket kan minska både kostnader och klimatpåverkan. Utöver detta kan AI generera konkreta förbättringsförslag för varje leverantör, exempelvis

materialbyten eller effektivare logistikupplägg som minskar utsläpp av växthusgaser.

AI kan användas för att granska leverantörer genom att automatiskt samla in och analysera information från flera källor och omsätta den till en uppdaterad riskbild. Den kan läsa och jämföra leverantörers dokument, som uppförandekoder, certifikat, revisionsrapporter och kontraktsbilagor, mot Region Stockholms hållbarhetskrav och internationella standarder, och flagga saknade delar eller avvikelser. AI kan också bevaka öppna källor som nyheter och rapporter på flera språk för att upptäcka tecken på risker kopplade till exempelvis arbetsvillkor, barnarbete eller miljöskandaler. På så sätt kan Region Stockholm få tidiga varningar, prioritera vilka leverantörer som behöver fördjupad granskning och följa upp mer kontinuerligt i stället för enbart genom punktvis revisioner.

På så sätt kan Region Stockholm gå från punktvisa revisioner till mer kontinuerlig uppföljning, där AI hjälper till att identifiera risker och sårbara platser, ge tidiga varningar när nya risker uppstår och föreslå åtgärder som både stärker mänskliga rättigheter och minskar klimatpåverkan.



## 5. Att använda AI på ett hållbart sätt

Att använda AI på ett hållbart sätt handlar om vad en verksamhet använder AI till och på vilket sätt den använder AI. Genom att göra medvetna val och bedöma hållbarhetsaspekterna av AI kan AI bli ett viktigt verktyg som stärker, snarare än utmanar, Region Stockholms hållbarhetsarbete.

I detta avslutande kapitel presenteras ett antal områden där Region Stockholms verksamheter kan göra medvetna val och bedöma hållbarhetsaspekterna av olika AI-initiativ och verktyg.

Region Stockholms verksamheter kan sedan använda det verktyg för hållbarhetsbedömning som tagits fram (se bilaga 1).

Hållbarhetsbedömningen möjliggör fördjupad reflektion och identifiering av hållbarhetsrisker kopplade till ett specifikt AI-verktyg eller en specifik AI-lösning.

### Samhällsnyttan och värdeskapande i centrum

Region Stockholms användande av AI ska gynna invånarna i Stockholmsregionen och bidra till hållbar utveckling. Samhällsnyttan ska vägleda när och hur Region Stockholm använder AI.

- Använd AI som främst bidrar till bättre vård, kollektivtrafik, service, kultur eller regional utveckling för invånarna.
- Använd AI som påskyndar hållbarhetsarbetet och gör att Region Stockholms verksamheter når hållbarhetsmålen om att minska klimatpåverkan, skapa en giftfri miljö och främja socialt hållbar utveckling.
- Använd AI som gör vår samhällsservice ännu mer effektiv och effektfull.

### Kvalitetssäkra för att minska risken för diskriminering

Region Stockholms användning av AI ska säkerställa att risken för diskriminering inte ökar. AI-modeller eller verktyg bör ha testats för att minimera inbyggda bias i exempelvis algoritmer, språkhantering, visuell hantering, utfall och resultat.

- Gör stickprov och testa riskhypoteser. Grupper i samhället som historiskt missgynnats eller misstänkliggjorts riskerar att påverkas negativt av bias. Genom att testa för oönskade bias kopplat till exempelvis kön, etnicitet, sexualitet, ålder, socioekonomi och funktionsnedsättningar kan risken minskas.

- Säkerställ att den data som används för att träna AI är representativ för befolkningen eller användarna, exempelvis utifrån vilka länder datan kommer från, hur gammal den är och vilken typ av data som inkluderas.
- Bedöm risken för om vissa samhällsgrupper kan få sämre samhällsservice eller kvalitet, och vidta åtgärder för att förebygga detta.

## **Se till att Region Stockholms AI-användning går att förstå och förklara**

I synnerhet för offentliga aktörer behöver det vara tydligt hur AI-verktyg skapas och används för att tilliten till Region Stockholms verksamheter ska upprätthållas. AI-användningen ska vara möjlig att förstå, gå att granska och gå att förklara. Utgångspunkten ska vara att AI-användningen stärker allmänhetens tillit till Region Stockholms verksamheter.

- Säkerställ att det finns tydlig information till användare, personal, patienter, resenärer och invånare om när, hur och i vilket syfte AI används.
- AI-lösningen bör testas med relevanta användargrupper och, när det är lämpligt, genom samråd med civilsamhället.

## **Optimera genom medvetna val och resurseffektivitet**

Genom att välja modell klokt, skriva effektivare kod, välja rätt algoritmer och lagra data utifrån behov kan beräkningskraften och därmed energiförbrukningen minskas avsevärt. En AI-lösning ska vara skalbar och flexibel för att möta framtida behov, vilket minskar risken för teknologisk inlåsning eller föråldring. Om det krävs helt ny IT-utrustning och löpande uppgraderingar är klimat- och miljöbelastningen självklart högre än om befintlig IT-utrustning kan användas.

- Gör medvetna val kring vilken typ av AI som ska användas och vilken kapacitet som behövs. Bedöm om AI-lösningen är tillräckligt resurseffektiv eller om en enklare lösning med lägre energianvändning och klimatpåverkan bör väljas.
- Bedöm om lösningen kräver tung modellträning eller om en befintlig modell som blir mer energisnål kan användas i stället.
- Se över möjligheterna att ställa in att minsta möjliga energianvändning är förvalt läge (default) vid användning av exempelvis språkmodeller.
- Säkerställ att det finns en plan för ansvarsfulla inköp, reparation, återbruk, uppgraderingar och avveckling som minimerar it-avfall.

## Arbeta aktivt med AI-leverantörer

Region Stockholm kommer troligen att köpa in flera AI-tjänster och verktyg samt vidareutveckla egna AI-verktyg på befintliga och färdigtränade AI-modeller.

Därför bör Region Stockholms verksamheter ha löpande dialog och ställa krav på leverantörer om hur hållbarhetsperspektivet är integrerat i olika AI-lösningar som ska användas.

I vissa fall kommer befintliga system och it-leverantörer att vara de primära leverantörerna även för nya AI-verktyg. I andra fall kan det vara helt nya AI-verktyg som löser ett specifikt problem inom Region Stockholms samhällsservice. Oavsett vilket bör Region Stockholms verksamheter ha löpande dialog och ställa krav på leverantörer om hur hållbarhetsperspektivet är integrerat i olika AI-verktyg.

Nedan följer några områden som kan vara relevanta att inkludera vid kravställning och dialog med leverantörer. Bedöm om leverantören kan:

- Redovisa att den data som används är kvalitetssäkrad och representativ för målgruppen.
- Redovisa datas ursprung, kvalitet och aktualitet tydligt.
- Redovisa att bias-tester genomförts och dokumenteras kopplat till exempelvis kön, etnicitet, ålder, socioekonomi eller funktionsnedsättningar.
- Tillhandahålla tydliga mekanismer för mänsklig överprövning, loggning och spårbarhet.
- Tillhandahålla möjlighet att ingripa före kritiska beslut.
- Säkerställa att datacenter drivs med förnybar eller fossilfri energi.
- Redovisa koldioxidutsläpp, minst scope 1 och 2 enligt Green House Gas-protokollet.
- Är certifierad eller arbetar i enlighet med relevanta ISO-standarder, som exempelvis ISO 14001 (miljöledning), ISO 50001 (energieffektivitet) eller ISO 42005 (ansvarsfull AI).
- Löpande uppdatera om var/när beräkningar körs och hur energiförbrukning kan minskas under både utveckling och drift.
- Arbeta i enlighet med cirkulära principer och ha möjlighet att uppdatera hårdvara samt mjukvara löpande för att minska risken för onödigt it-avfall.

Avslutningsvis bör Region Stockholms användning av AI utgå från att teknikutveckling och hållbar utveckling ska gå hand i hand. AI ska användas där den skapar verkligt värde för invånare, patienter, resenärer

och medarbetare, samtidigt som risker för diskriminering, bristande transparens, hög resursförbrukning och negativ miljöpåverkan förebyggs.

Genom medvetna val, tydliga krav, löpande uppföljning och aktiv dialog med leverantörer kan Region Stockholm använda AI på ett sätt som stärker kvaliteten i samhällsservicen och bidrar till en socialt, miljömässigt och ekonomiskt hållbar utveckling. På så sätt kan AI bli ett verktyg som stödjer, snarare än motverkar, Region Stockholms långsiktiga hållbarhetsarbete.



## Bilaga 1: Hållbarhetsbedömning av AI-lösningar

### Varför göra en hållbarhetsbedömning av AI-lösningar och verktyg?

Det är ett konkret verktyg för att analysera hur en AI-lösning eller ett AI-verktyg kan användas på ett hållbart sätt. Genom bedömningen kan ni identifiera de hållbarhetsområden som behöver stärkas innan ni arbetar vidare med er AI-lösning eller ert AI-verktyg.

### När kan ni göra en hållbarhetsbedömning av AI-lösningar och verktyg?

Region Stockholms verksamheter kan använda hållbarhetsbedömningen vid idéstadiet, inför upphandling, vid utveckling, piloter eller drift av AI-lösningar och verktyg. Det är alltså ett verktyg som kan användas löpande och gärna i samband med andra bedömningar, exempelvis klassning i enlighet med AI-förordningen.

### Hur kan ni göra en hållbarhetsbedömning av AI-lösningar och verktyg?

Hållbarhetsbedömningen identifierar ett antal områden som innehåller bedömningsfrågor. Dessa frågor ska hjälpa verksamheten att bedöma hållbarheten i en specifik AI-lösning eller ett specifikt AI-verktyg. Ju fler frågor som verksamheten kan bedöma som uppfyllda med hög mognadsgrad, desto starkare blir hållbarhetsbedömningen totalt.

### Vad gör ni efter att hållbarhetsbedömningen är klar?

Arbeta vidare med att hantera de risker ni har identifierat eller de områden som ni behöver stärka. Förslagsvis sammanställer ni er hållbarhetsbedömning så att den finns tillgänglig att dela och förankra med andra i organisationen, exempelvis ledning och upphandling. Den utgör en viktig del av riskanalysen och riskhanteringen för hur ni ser på användningen av AI-verktyget på ett hållbart sätt och hur ni förhåller er till riskerna.

## Så gör ni

- Identifiera och avgränsa till vilken AI-lösning eller verktyg som ni ska bedöma.
- Samla de personer eller den arbetsgrupp hos er som arbetar med den kommande AI-lösningen eller det kommande AI-verktyget för att göra bedömningen. Förslagsvis är det en grupp där både medarbetare från it/utveckling och hållbarhet gör bedömningen tillsammans. Det kan även vara värdefullt att samverka med medarbetare från upphandling/inköp och juridik.
- Avsätt 1–2 timmar för att diskutera igenom frågorna och göra bedömningen tillsammans.
- Använd reflektionsfrågorna för att bedöma varje område. Ju fler tillfredsställande svar och former av riskhantering per reflektionsfråga ni har, desto högre hållbarhetsbedömning får

området. Det är gruppen som tillsammans bedömer om ni har låg, medel eller hög hållbarhetsbedömning för varje område.

- Fyll i era svar för varje område med en kort motivering.
- Förslagsvis använder ni AI-transkribering under mötet för att fånga diskussionen och sammanställa er bedömning.
- När hållbarhetsbedömningen är färdig beslutar ni hur ni går vidare för att stärka upp inom de områden där ni hade en otillräcklig hållbarhetsbedömning. Dessa områden bör ni arbeta vidare med och åtgärda snarast.

## Hållbarhetsbedömning

### Område 1: Samhällsnyttan och värdeskapande i centrum

Region Stockholms användande av AI ska gynna invånarna i Stockholmsregionen och bidra till hållbar utveckling. Samhällsnyttan ska vägleda när och hur Region Stockholm använder AI.

#### Reflektionsfrågor:

- Ser ni att AI-lösningen kommer att bidra till bättre vård, kollektivtrafik, service, kultur eller regional utveckling för invånarna?
- Kommer användningen av AI-lösningen att stärka hållbarhetsarbetet och göra att Region Stockholms verksamheter når hållbarhetsmålen om att minska klimatpåverkan, skapa en giftfri miljö och/eller främja socialt hållbar utveckling?
- Kommer användningen av AI-lösningen att göra vår samhällsservice ännu mer effektiv och effektiv?

#### Hållbarhetsbedömning (välj en):

- Tillräcklig nivå och dokumenterad
- Otillräcklig nivå och kräver vidare arbete

#### Förklaring till bedömning:

### Område 2: Kvalitetssäkra för att minska risken för diskriminering

Region Stockholms användning av AI ska säkerställa att risken för diskriminering inte ökar. AI-modeller eller verktyg bör ha testats för att minimera inbyggda bias i exempelvis algoritmer, språkhantering, visuell hantering, utfall och resultat.

Reflektionsfrågor:

- Har det gjorts stickprov/testning av riskhypoteser för att testa eventuella oönskade bias kopplade till exempelvis kön, etnicitet, sexualitet, ålder, socioekonomi och funktionsnedsättningar?
- Har det säkerställts att den data som används för att träna AI är representativ för målgruppen (exempelvis befolkningen eller användarna)?
- Har det gjorts någon bedömning av, eller åtgärder mot, risken för att vissa samhällsgrupper skulle kunna få sämre samhällsservice eller kvalitet?

Hållbarhetsbedömning (välj en):

- Tillräcklig nivå och dokumenterad
- Otillräcklig nivå och kräver vidare arbete

Förklaring till bedömning:

**Område 3: Region Stockholms AI-användning går att förstå och förklara**

I synnerhet för offentliga aktörer behöver det vara tydligt hur AI-verktyg skapas och används för att tilliten till Region Stockholms verksamheter ska upprätthållas. AI-användningen ska vara möjlig att förstå, gå att granska och gå att förklara. Utgångspunkten ska vara att AI-användningen stärker allmänhetens tillit till Region Stockholms verksamheter.

- Finns det en plan för hur AI-verktyget ska förankras hos personal som kommer att använda AI-verktyget?
- Finns det generell information till patienter, resenärer och invånare om när, hur och i vilket syfte AI-verktyget används?
- Har AI-lösningen testats med relevanta användargrupper och, när det är lämpligt, genom samråd med civilsamhället?

Hållbarhetsbedömning (välj en):

- Tillräcklig nivå och dokumenterad
- Otillräcklig nivå och kräver vidare arbete

Förklaring till bedömning:

**Område 4: Optimera användningen på ett resurssmart sätt**

Genom att välja modell klokt, skriva effektivare kod, välja rätt algoritmer och lagra data utifrån behov kan beräkningskraften och därmed energiförbrukningen minskas avsevärt. En AI-lösning ska vara skalbar och flexibel för att möta framtida behov, vilket minskar risken för teknologisk inlåsnings eller föråldring. Om det krävs helt ny it-utrustning och löpande uppgraderingar är klimat- och miljöbelastningen självklart högre än om befintlig it-utrustning kan användas.

- Har vi bedömt vilken typ av AI som ska användas och vilken kraft som behövs? Är den påtänkta AI-lösningen tillräckligt resurseffektiv eller finns en enklare lösning med lägre energianvändning och klimatpåverkan?
- Kan vi använda en AI-lösning utan tung modellträning genom att använda en befintlig modell som blir mer energisnål?
- Har vi sett över möjligheterna att ställa in att minsta möjliga energianvändning är förvalt läge (default) vid användning av exempelvis språkmodeller?
- Finns det en plan för ansvarsfulla inköp, reparation, återbruk, uppgraderingar och avveckling som minimerar it-avfall?

Hållbarhetsbedömning (välj en):

- Tillräcklig nivå och dokumenterad
- Otillräcklig nivå och kräver vidare arbete

Förklaring till bedömning:

### **Område 5: Hållbar upphandling och inköp av AI-tjänster och verktyg**

Region Stockholm kommer troligen att köpa in flera AI-tjänster och verktyg eller vidareutveckla egna AI-verktyg på befintliga och färdigtränade AI-modeller.

Därför bör Region Stockholms verksamheter ha löpande dialog och ställa krav på leverantörer om hur hållbarhetsperspektivet är integrerat i olika AI-lösningar som ska användas.

Nedan finns reflektionsfrågor kopplade till hållbar upphandling och inköp av AI-tjänster och verktyg. Bestäm först vilka frågor som är relevanta för er att bedöma och vilka som faller utanför ert uppdrag.

Har vi ställt krav kring eller har dialog med leverantören om att:

- Redovisa att den data som används är kvalitetssäkrad och representativ för målgruppen?
- Redovisa datas ursprung, kvalitet och aktualitet tydligt?
- Redovisa att och hur bias-tester genomförts och dokumenteras kopplat till exempelvis kön, etnicitet, ålder, socioekonomi eller funktionsnedsättningar?
- Tillhandahålla tydliga mekanismer för mänsklig överprövning, loggning och spårbarhet?
- Tillhandahålla möjlighet att ingripa före kritiska beslut?
- Säkerställa att datacenter drivs med förnybar eller fossilfri energi?
- Redovisa koldioxidutsläpp, minst scope 1 och 2 enligt Green House Gas-protokollet?
- Certifiering eller arbete i enlighet med relevanta ISO-standarder, som exempelvis ISO 14001 (miljöledning), ISO 50001 (energieffektivitet) eller ISO 42005 (ansvarsfull AI)?
- Löpande uppdatera om var/när beräkningar körs och hur energiförbrukning kan minskas under både utveckling och drift?
- Arbeta i enlighet med cirkulära principer och ha möjlighet att uppdatera hårdvara samt mjukvara löpande för att minska risken för onödigt it-avfall?

Hållbarhetsbedömning (välj en):

- Tillräcklig nivå och dokumenterad
- Otillräcklig nivå och kräver vidare arbete

Förklaring till bedömning:

**Sammanfattning:**

- Dessa områden bedömer vi som tillräckligt analyserade och dokumenterade:

- Dessa områden bedömer vi som otillräckliga i dagsläget och behöver vidareutveckla och arbeta vidare med:

**Regionledningskontoret**  
Hållbarhetsavdelningen  
Besöksadress: Lindhagensgatan 98  
Stockholm  
08-123 100 00  
[www.regionstockholm.se](http://www.regionstockholm.se)