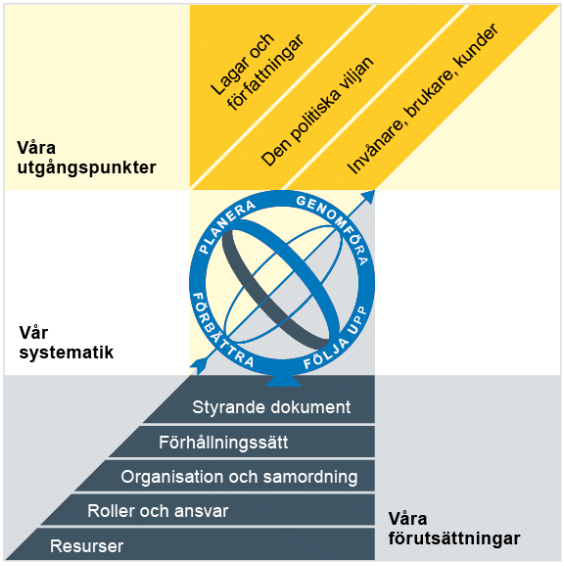
Göteborgs Stad

Förvaltningen för funktionsstöds anvisning för AI

Reglerande styrande dokument

- Policy
- Riktlinje
- Regel
- ► **Anvisning**
- Rutin
- Instruktion

Göteborgs Stads styrsystem



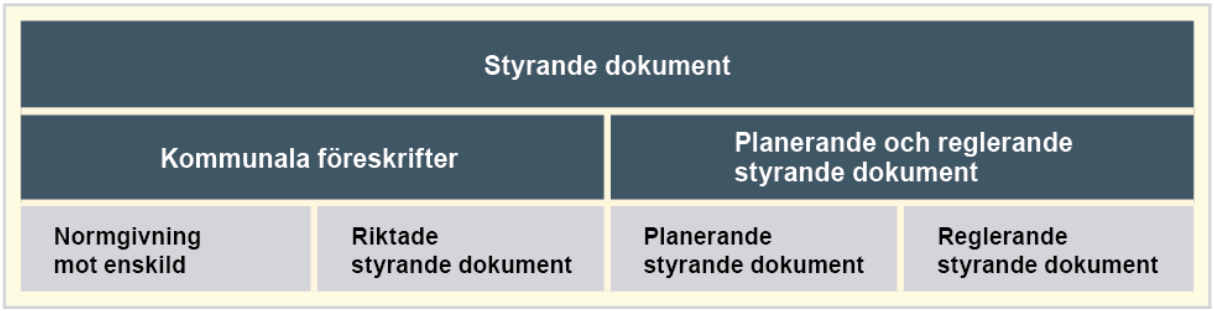
Utgångspunkterna för styrningen av Göteborgs Stad är lagar och författningar, den politiska viljan och stadens invånare, brukare och kunder. För att förverkliga utgångspunkterna behövs förutsättningar av olika slag. Stadens politiker har möjlighet att genom styrande dokument beskriva hur de vill realisera den politiska viljan. Inom Göteborgs Stad gäller de styrande dokument som antas av kommunfullmäktige och kommunstyrelsen. Därutöver fastställer nämnder och bolagsstyrelser egna styrande dokument för sin egen verksamhet. Kommunfullmäktiges budget är det övergripande och överordnade styrande dokumentet för Göteborgs Stads nämnder och bolagsstyrelser.

Om Göteborgs Stads styrande dokument

Göteborgs Stads styrande dokument är våra förutsättningar för att vi ska göra rätt saker på rätt sätt. De anger vad nämnder/styrelser och förvaltningar/bolag ska göra, vem som ska göra det och hur det ska göras. Styrande dokument är samlingsbegreppet för dessa dokument.

Stadens grundläggande principer såsom demokratisk grundsyn, principer om mänskliga rättigheter och icke-diskriminering omsätts i praktisk verksamhet genom att de integreras i stadens ordinarie beslutsprocesser. Beredning av och beslut om styrande dokument har en stor betydelse för förverkligandet av dessa principer i stadens verksamheter.

De styrande dokumenten ska göra det tydligt både för organisationen och för invånare, brukare, kunder, leverantörer, samarbetspartners och andra intressenter vad som förväntas av förvaltningar och bolag. De styrande dokumenten ligger till grund för att utkräva ansvar när vi inte arbetar i enlighet med vad som är beslutat.



Beslutad av:	**Gäller för:**	**Diarienummer:**	**Datum och paragraf för beslutet:**
Avdelningschef Stab och kommunikation	Förvaltningen för funktionsstöd	[Nummer]	[Text]
Dokumentsort:	**Giltighetstid:**	**Senast reviderad:**	**Dokumentansvarig:**
Anvisning	[Giltighetstid]	2024-12-05	Ronny Gustafsson

Bilagor:
[Bilagor]

Innehåll

Inledning

Syftet med denna anvisning

Anvisningen syftar till att göra dig som medarbetare trygg i ditt ansvar när du använder AI i ditt arbete och vid utveckling av nya AI-system och modeller. Målet är att på ett ansvarsfullt sätt främja innovation. Anvisningen grundar sig till stor del i de lagar, förordningar och regler som förvaltningen, dess chefer och medarbetare redan lyder under.

Viktiga perspektiv som alla medarbetare behöver ha med sig i arbetet med AI är integritet, förtroende, transparens, etik och informationssäkerhet.

Allt eftersom nya lagar, förordningar och regler kopplat till AI tas fram av andra myndigheter kommer denna anvisning att justeras. Säkerställ därför att du använder senaste versionen av anvisningen när du använder dig av AI och/eller utvecklar system med hjälp av AI.

Vem omfattas av anvisningen

Denna anvisning gäller för alla medarbetare i förvaltningen för funktionsstöd.

Bakgrund

Artificiell intelligens (AI) är ett område som väcker allt större intresse och påverkar vårt samhälle på många sätt. Exempel är AI-system eller AI-modeller för allmänna ändamål som ChatGPT eller andra chatbottar. Utvecklingen inom AI-området går extremt fort. AI kan bidra med nytta i vardagen genom att till exempel förbättra sjukvården, underlätta tillgången till information, utbildning och praktik samt göra arbetsplatser/organisationer säkrare. Samtidigt finns det också några centrala utmaningar och risker med AI. Bland annat kan AI-system inkräkta på människors integritet, vara ett hot mot organisationers säkerhet, förstärka och reproducera diskriminering samt sprida falsk information.

I samband med att det europiska rådet antog sin ståndpunkt om AI betonades å ena sidan betydande samhälleliga och ekonomiska fördelar med AI inom en mängd områden och å andra sidan vikten av att skydda integriteten och garantera säkerheten. Rådet ville ha en säker, laglig och tillförlitlig AI med respekt för de grundläggande rättigheterna. Det vill säga att AI-system som används inom EU är **människocentrerade** och **tillförlitliga** och att en **hög skyddsnivå** säkerställs för **individers hälsa, säkerhet och grundläggande rättigheter**. Denna ståndpunkt blev grunden till Förordningen om artificiell intelligens eller EU:s AI-akt, i Sverige kallad AI-förordningen.

Målet med förordningen är att skapa en trygg och etisk hållbar miljö för AI-innovation.

AI-förordningen

AI-förordningen trädde i kraft 21 maj 2024 men börjar tillämpas successivt därför är det viktigt att redan nu vara medveten om vad som kommer. Det gäller i stor utsträckning för AI-system med hög risk som kommer att omfattas av särskilda krav. Kommuner och regioner träffas i stor utsträckning av de områden som är relevanta för högrisk AI, exempel medicinsk utrustning och lösningar avsedda att användas för att bedöma om fysiska personer har rätt till bidrag. AI-förordning börjar tillämpas:

- 2 februari 2025 - förbuden, definitionerna och bestämmelserna om AI-kompetens börjar tillämpas.
- 2 augusti 2025 – reglerna om styrning och skyldigheter för AI för allmänna ändamål börjar tillämpas.
- 2 augusti 2027 – skyldigheter för AI-system med hög risk som klassificerad som system med hög risk (eftersom de är inbyggda i reglerade produkter) börjar tillämpas.

Grundtanken med ett harmoniserat regelverk för AI är att säkerställa att AI-system på EU-s marknad är säkra och respekterar mänskliga rättigheter. Det är en förutsättning för att underlätta innovation och investeringar i AI. Förordningens rättsliga ram omfattar hela AI-systemets livscykel (utveckling, saluföring och användning).

Kommuner kan enligt AI-förordningens kontext ses som både spridare (deployer) och leverantör (provider). En spridare är de som använder ett AI-system under eget överinseende, medan en leverantör är de som utvecklar AI-system eller modeller för allmänna ändamål.

Utgångspunkten i AI-förordningen är densamma som i Göteborgs Stads riktlinje för informationssäkerhet – ett riskbaserat förhållningssätt. AI-förordningen har bestämmelser om vilka typer av AI-system som inte ska vara tillåtna, har hög risk, begränsad risk eller minimal risk och vad som krävs (fyra risknivåer).

Oacceptabel risk	AI-systemet utgör ett direkt hot mot individer och är förbjudet.
Hög risk	AI-system som har en negativ påverkan på säkerhet eller grundläggande exempel AI-system som omfattas av EU:s produktlagstiftning för säkerhet som leksaker eller medicintekniska produkter eller andra högrisk AI-system som kräver strikta regler (exempel användning inom utbildning, sysselsättning, brottsbekämpning och migration).
Begränsad risk	AI-system som kan behöva särskilda transparensåtaganden för att informera användare om att de integrerar med ett AI-system, exempel chattbot
Minimal risk	De allra flesta AI-system faller inom denna kategori och behöver då minimala eller inga regleringar.

I samband med att bestämmelserna börjar gälla kommer EU kommissionen att ta fram riktlinjer för klassificering av hög risk AI.

Regeringen har gett två myndigheter (DIGG och IMY) i uppdrag att ta fram riktlinjer för användning av generativ AI inom den offentliga förvaltningen vilka kommer att redovisas den 20 januari 2025.

AI och annan lagstiftning och reglering

AI kräver stora datamängder, som ofta innehåller personuppgifter. Då behöver personuppgiftsbehandlingen utgå från dataskyddsförordningen, lag (2001:454) om behandling av personuppgifter inom socialtjänsten, förkortas SoLPuL samt förordning (2001:637) om behandling av personuppgifter inom socialtjänsten, förkortas SoLPuF.

Vid utveckling av AI uppstår också nya informationsmängder i form av exempelvis källkod, tränings- och valideringsdata. De nya informationsmängderna ställer krav på analys utifrån om nya allmänna handlingar skapats vilket utifrån arkivlag och tryckfrihetsförordningen (1949:105) samt offentlighets- och sekretesslagen behöver hanteras på ett korrekt sätt (till exempel registrering, gallring/krav på bevarande, sekretess osv.). Även sammanställningar vid användning av generativ AI behöver analyseras utifrån ovan aspekter.

Förvaltningslagen (2017:900) ställer också krav på legalitet, objektivitet och proportionalitet samt att vi kunna motivera hur vi har kommit fram till en slutsats i ett beslut.

Möjligheterna till effektivisering med hjälp av AI-system bedöms vara stora inom Staden och det är viktigt att ett införande sker i enlighet med gällande lagstiftning.

Koppling till andra styrande dokument

Styrande dokument	Koppling till denna anvisning
Göteborgs Stads plan för digitalisering 2023–2026	Göteborgs Stads styrdokument kring digitalisering och informationssäkerhet tar inte uttryckligen upp AI, men är lika giltiga för utveckling av AI som av andra digitala system.
Göteborgs Stads riktlinje för styrning, samordning och finansiering av digital utveckling och förvaltning	
Göteborgs Stads riktlinje för informationssäkerhet	

Anvisning

Perspektiv

Förtroende och transparens

I Göteborgs Stad uppmanas vi att vara nytänkande och modiga, vilket exempelvis kan innebära att testa ny teknik. Samtidigt måste vi, som med alla digitala system, förhålla oss kritiska och ta ansvar för eventuella konsekvenser.

För att kunna bygga och bibehålla användarnas förtroende för AI-system eller modeller måste vi kunna **förklara** hur AI-systemet fungerar. Det betyder att processerna måste vara transparenta och att det finns en öppen kommunikation om AI-systemens kapacitet och syfte. Beslut måste – i så stor utsträckning som möjligt – gå att förklara för dem som påverkas direkt eller indirekt för utan den typen av information går det inte att bestrida ett beslut.

Det är inte alltid möjligt att förklara varför ett AI-system har gett ett visst resultat eller beslut (och vilken kombination av inputfaktorer som bidrog till det). Dessa så kallade ”svarta lådan-algoritmer” behöver uppmärksammas särskilt. I sådana fall kan det behövas andra åtgärder för att skapa förklarbarhet (t.ex. spårbarhet, möjlighet till granskning och transparent kommunikation om systemets kapacitet), förutsatt att systemet som helhet uppfyller grundläggande rättigheter. Den nödvändiga nivån av förklarbarhet beror i stor utsträckning på sammanhanget och på hur svåra konsekvenserna blir om resultatet är felaktigt eller missvisande på annat sätt.

Detta ansvar har du som medarbetare:

- Du är ansvarig för den information du lägger in i ett AI-system-modell. Informationen du lägger in kan komma att lagras och återanvändas. Du kan inte ta bort eller ångra den data du har lagt till.
- Du ska vara medveten om, och kunna förklara hur ett AI-system/modell fungerar och varför AI-systemet/-modellen har gett ett visst resultat.
- Du ska tänka kritiskt, utvärdera och faktagranska resultatet.
- Du är ansvarig för hur du använder resultatet från ett AI-system/-modell.

Detta ansvar har du som utvecklar AI-system/-modell:

- Säkerställ datans integritet och att den data som systemet tränas på inte innehåller socialt konstruerade snedvridningar, fel och misstag.
- Processer och data som används ska testas och dokumenteras i varje steg.
- Ge tydlig och proaktiv information till berörda parter om AI-systemets/-modellens kapacitet och begränsningar, så att det går att skapa rimliga förväntningar, och kommunicera om hur kraven genomförs. Var öppen med att de har med ett AI-system/-modell att göra.

Etik

AI-system ska användas på ett sätt som respekterar de etiska principerna om respekt för människans autonomi, skadeförebyggande och förklarbarhet. Vid användande ska vi särskilt uppmärksamma situationer som rör sårbara grupper som barn, personer med funktionsnedsättning och andra som historiskt sett har varit missgynnade eller riskerar utanförskap. Det gäller även situationer som utmärks av bristande balans i fråga om inflytande och information.

Varje människa har ett ”värde i sig”, som aldrig får förminskas, skadas eller förtryckas av andra – eller av ny teknik, som AI-system.

AI-system bör förbättra det individuella och kollektiva välbefinnandet och uppfylla fyra etiska principer; respekt för människans autonomi (självbestämmande - vilket går hand i hand med socialtjänstlagens intentioner), förebyggande av skada, rättvisa (går att överklaga beslut fattat av AI) och förklarbarhet.

Detta ansvar har du som medarbetare:

- Du ska följa gällande lagstiftning och reglering. Är du osäker behöver du ta kontakt med stödfunktioner inom området för att säkerställa att det du tänkt är lagligt eller lämpligt.
- Du behöver överväga om det du gör innebär att nya allmänna handlingar skapas. Ta kontakt med stödfunktioner inom området och arkivredogörare.
- Du ansvarar för att göra en bedömning av tillförlitligheten hos AI-system/-modell under utveckling, spridning och användning och anpassa bedömningen efter det specifika område där systemet används.
- Du ska vara uppmärksam på att du kan få felaktiga uppgifter och slutsatser. De svar som AI-system/-modellen ger kan låta rimliga, men kan vara direkt felaktiga eller påhittade.
- Du behöver vara medveten och uppmärksam på att resultaten som AI-system/-modeller ger kan förstärka bias mot olika grupper.

Detta ansvar har du som utvecklar AI-system/-modeller:

- Du ansvarar för att AI-systemet är lagligt och säkert samt att etiska principer säkerställs under hela systemets livscykel.
- Du ansvarar för att AI-systemet tar särskild hänsyn till våra målgruppers situationer.
- Du ansvarar för att säkerställa mänsklig tillsyn och kontroll över arbetsprocesserna i AI-systemet.
- Säkerställa att det går att identifiera den enhet som ansvarar för beslut som görs i ett AI-system och att beslutsprocesserna går att förklara.

Informationssäkerhet

Informationssäkerhet handlar om att information är en värdefull tillgång, ibland till och med livsviktig. Följderna av att vi inte har tillgång till informationen när vi behöver den

(tillgänglighet), att den är felaktig (riktighet) eller att obehöriga kan ta del av den (konfidentialitet) kan bli mycket allvarliga. Därför behöver vi skydda den.

Informationssäkerhet är i grunden teknikneutralt och omfattar såväl skydd av muntlig, pappersbunden som digital information. Informationssäkerhet är en förutsättning för cybersäkerhet vilken syftar till att skydda digitaliserade system mot antagonistiska hot. Cybersäkerhet och IT-säkerhet samt dataskydd (omfattar skydd av enskildas personuppgifter) är en del av informationssäkerhetsarbetet.

AI är ett effektivt verktyg inom cybersäkerhet. Även angripare kan använda AI för att skapa automatiserade, anpassningsbara och effektivt AI-drivna cyberangrepp.

AI-system måste garantera integritets- och uppgiftsskyddet under systemets hela livscykel. Detta omfattar de uppgifter som användaren har lämnat inledningsvis samt information som användaren genererar under sitt samspel med systemet (till exempel output som AI-systemet har genererat för specifika användare eller hur användare har reagerat på vissa rekommendationer). AI-system behöver vara skyddat mot sårbarheter som gör att systemet kan utnyttjas av motståndare, till exempel genom hackning. Angrepp kan vara inriktade mot datan (dataförgiftning), modellen (läckage) eller den underliggande infrastrukturen.

Ju mindre tillsyn en människa kan utöva över ett AI-system, desto mer omfattande tester och strängare styrning krävs. Om AI-system blir angripet, till exempel genom fientliga angrepp, kan både data och systemets beteende förändras, vilket innebär att systemet fattar annorlunda beslut eller stängs av helt och hållet. System och data kan också korrumperas genom avsiktligt fientliga åtgärder. Bristfälliga säkerhetsprocesser kan också resultera i felaktiga beslut eller till och med fysiska skador.

Förvaltningens information ska vara informationsklassad och hanteras i system som har den säkerhetsnivå som informationen kräver. Det gäller även AI-system/-modeller.

Detta ansvar har du som medarbetare:

- AI-tjänster som du använder på Göteborgs Stads datorer, mobiler och surfplattor ska endast användas utifrån din roll och dina arbetsuppgifter.
- Du ansvarar för att ha säkerställt att AI-systemet/-modellen och dess leverantör endast kan komma åt den information som du avsett.
- Du måste veta att den information som du lägger in i AI-systemet/-modellen inte kommer att användas för att träna systemet och om det finns risk för att den information du lagt in kommer att återanvändas och spridas.

Detta ansvar har du som utvecklar AI-system/-modeller:

- Du ansvarar för att säkerställa att AI-system tas fram, utvecklas och tas i bruk i enlighet med lagkrav och säkerhetskrav.

Personuppgifter och sekretess

Att mata in personuppgifter i ett AI-system är att behandla personuppgifter oavsett om de lagras eller inte. All behandling av personuppgifter ska följa de grundläggande principerna i dataskyddsförordningen och annan lagstiftning för behandling av personuppgifter inom socialtjänst.

Om och hur bestämmelserna uppfylls måste analyseras *innan* AI-systemet används för den tänkta behandlingen. Det innebär exempelvis att genomföra en konsekvensbedömning (DPIA), ta fram informationstexter, ta reda på om personuppgifter överförs utanför EU/EES (i huvudregel förbjudet men kan vara tillåtet). De flesta AI-tjänster skickar data till USA. Om du inte är säker på att datan hanteras på ett säkert sätt bör du inte lämna ut den.

Utmaningar utifrån dataskyddsförordningen är exempelvis inom områdena; profilering och automatiserat beslutsfattande, korrekthet, transparenskrav, registrerades rättigheter, uppgiftsminimering, ändamålsbegränsning och konfidentialitet.

Det är inte tillåtet att mata in personuppgifter i en AI-modell för allmänna ändamål. För AI-modell för allmänna ändamål såsom Copilot och Chat GTP saknas vidare metoder för att kontrollera respektive leverantörs behandling eller säkerhet. Personuppgifter ska således inte hanteras i sådana system. Detta gäller självfallet även uppgifter som omfattas av sekretess.

Offentlighets- och sekretesslagen (2009:400) anger att det finns bestämmelser om tystnadsplikt i det allmännas verksamhet och ett förbud att lämna ut allmänna handlingar. Bestämmelserna avser förbud att röja uppgift, vare sig om detta sker muntligen eller genom informationen lämnas ut på något annat sätt. Om information delas med en IT-leverantör, till exempel genom att laddas upp i en molntjänst, utan att det finns en sekretessbrytande bestämmelse, är den som lagt upp information skyldig till röjande av sekretessuppgift. Ett sätt att säkert ladda upp information så att leverantören inte kan ta del av den är att kryptera. Att göra sekretessbelagd information tillgänglig för obehörig är bara tillåtet om det finns en sekretessbrytande bestämmelse.

Röjande av sekretessbelagda uppgifter genom att mata in dem i ett AI-system/-modell kan vara straffbart enligt brottsbalken som brott mot tystnadsplikten.

Detta ansvar har du som medarbetare:

- Du har ansvar för att säkerställa att du inte matar in personuppgifter eller gör sekretessbelagda uppgifter och personuppgifter tillgängliga för obehöriga.
- Du behöver följa de rutiner som finns avseende behandling av personuppgifter.

Detta ansvar har du som utvecklar AI-system:

- Om AI-systemet ska hantera personuppgifter och/eller sekretessbelagda uppgifter ansvarar du för att det görs riskanalys och konsekvensbedömning för att avgöra om riskerna kan åtgärdas.
- Du ansvarar för att AI-systemet kravställts/byggs rätt utifrån gällande lagstiftning.

Känslig eller skyddsvärd information

Känslig och skyddsvärd information kan till exempel vara ekonomisk information, mötesanteckningar, planeringar och analyser. Gemensamt är att informationen, om den kommer i orätta händer, kan riskera att skada organisationen, anställda eller enskilda. Att lägga in känslig eller skyddsvärd information i AI-modeller för allmänna ändamål innebär att informationen kan bli tillgänglig för andra. Det är inte tillåtet att lägga in skyddsvärd eller känslig information i AI-modeller för allmänna ändamål.

Detta ansvar har du som medarbetare:

- Ta reda på om AI-systemet är säkert för inmatning av känslig eller skyddsvärd information.
- Lägg aldrig in känslig eller skyddsvärd information i ett AI-system/AI-modell som inte är säkert.

Detta ansvar har du som utvecklar AI-system:

- Om AI-systemet ska hantera känslig eller skyddsvärd information så behöver du säkerställa att systemet har adekvat skydd och skyddsåtgärder.

Upphovsrättsskyddad information

Upphovsrätt uppstår i samma stund som ett litterärt eller konstnärligt verk skapas. Därmed har upphovspersonen en ensamrätt att bestämma vem som får framställa exemplar av verket och hur/om verket ska göras tillgängligt för allmänheten. När det finns upphovsrätt behövs tillstånd. Det gäller oavsett om verket publiceras på nätet eller sprids på annat sätt.

I upphovsrättslagen finns dock ett undantag som innebär att det är tillåtet att utvinna material från internet – en teknik som kallas text och datautvinning (TDM). TDM är en automatiserad teknik som används för att analysera text och data i digital form i syfte att generera information. Det är sannolikt att träningen av AI-verktyget kan ses som text- och datautvinning. Men detta har inte prövats av domstol ännu. Det är med andra ord inte tydligt om det är tillåtet eller otillåtet att mata in upphovsrättsskyddat material i ett AI-system.

I användarvillkor för exempelvis ChatGPT anges att det är den som använder tjänsten som ansvarar för att inte bryta mot någon annans upphovsrätt. Det innebär att det är du själv, och inte leverantören av systemet som gör sig skyldig till upphovsrättsbrott om det skulle visa sig att det är olagligt. Var därför försiktig och lägg inte in upphovsrättsskyddat material i AI-tjänsten.

Användningen av AI-genererat material styrs av användarvillkoren. Det är därför viktigt att läsa villkoren noggrant för att förstå exakt vad du får och inte får göra med AI-genererat material. Villkoren kan också innehålla regler om betalning för användning av material utanför plattformen.

Detta ansvar har du som medarbetare:

- Du behöver tillstånd från den som har upphovsrätten till ett verk för att få använda verket.
- Om du använder en AI-genererad bild ska det tydligt framgå att det är en AI-genererad bild. Bildspråket i Göteborgs Stads grafiska profil gäller även för AI-bilder.
- Du är ansvarig att känna till och agera utifrån användarvillkoren.

Detta ansvar har du som utvecklar AI-system:

- Du ansvarar för att upphovsrättsskyddat material används på ett korrekt sätt.

AI-modeller för allmänna ändamål

AI-modeller för allmänna ändamål (general purpose AI-model) används för att exempelvis generera text (ChatGPT), mjukvarukod (Copilot), bilder (DALL-E), video (SORA) eller ljud (Voice Engine). AI-modeller uppvisar betydande generalitet och kan på ett kompetent sätt utföra ett brett spektrum av distinkta uppgifter och integreras i en rad system eller tillämpningar i efterföljande led. AI-modeller är väsentliga komponenter i AI-system men utgör inte i sig själva ett AI-system. AI-modeller kräver tillägg av ytterligare komponenter, till exempel ett användargränssnitt, för att bli AI-system. AI-förordningen ställer olika krav på AI-modeller för allmänna ändamål och AI-system.

Stora generativa AI-modeller är ett typiskt exempel på en AI-modell för allmänna ändamål, eftersom de möjliggör flexibel generering av innehåll som lätt kan rymma ett brett spektrum av särskiljande uppgifter.

När du använder en AI-modell är det du personligen som tar ansvar för den information som du lägger in i AI-modellen och leverantören använder informationen på det sätt som framgår av användarvillkoren. Du har därmed ett personligt ansvar och du behöver vara medveten om att du använder AI-modell för allmänna ändamål på egen risk. Risken som du tar ökar ju mindre kunskap du har om modellen, hur informationen hanteras, bearbetas och lagras samt delas.

Detta ansvar har du som medarbetare:

- Du ansvarar för att känna till och följa de användarvillkor som gäller för AI-tjänsten.
- Lägg aldrig in personuppgifter, sekretessuppgifter, känslig eller skyddsvärd information i en AI-modell för allmänna ändamål.
- Säkerställ att du har en god kunskap om AI-modellen och hur den fungerar. Använd aldrig AI-modeller som är helt nya eller kommer från en okänd avsändare.
- Säkerställ att du aldrig ger AI-modellen åtkomst till den information som finns på din dator.

Copilot

Copilot är Microsofts AI-modell och en del av Microsofts 365-miljö (M365) som vi har i Staden. Det innebär att den information du matar in hanteras i Microsofts molntjänst på samma sätt som informationen i OneDrive eller Outlook.

När du är inloggad med ditt Staden-konto och använder Copilot (tidigare kallat för Bing Chat Enterprise) får du en ikon och texten ”Skyddad” i grönt. I detta skyddade läge räknas Copilot inte som ett allmänt tillgängligt AI-system. Om du däremot inte är inloggad är Copilot en AI-modell för allmänna ändamål och reglerna ovan gäller.

Copilot är avsett för att generera, strukturera, sammanfatta och förbättra text om olika ämnen. Det finns sällan anledning att i samband med det behandla personuppgifter.

Detta ansvar har du som medarbetare:

- Du ansvarar för att säkerställa att system med sekretessuppgifter är stängda i samband med att Copilot används. Detta för att minimera risken för överföring.
- Lägg aldrig in personuppgifter, sekretessuppgifter, känslig eller skyddsvärd information i en AI-modell för allmänna ändamål.

Använda resultatet

AI-genererade texter kan innehålla felaktigheter. Texterna kan vara tydliga och övertygande men ändå innehålla faktamässiga fel. Du behöver alltid ha ett kritiskt förhållningssätt till information, oavsett om den kommer från en chattbott eller från ett AI-system. Det är du som ansvarar för resultatet och hur du använder det. Därför är det viktigt att du behandlar AI-genererad text som ett utkast och att du kontrollerar efter faktafel, olämpliga budskap, fördomar och så vidare.

AI-generade texter tas fram med en språkmodell som baseras på sannolikhet (beräknar vilket kommande ord som är mest troligt) och baseras på källor som du inte skulle lite på i andra sammanhang. Resultatet du får kan också bli olika om du ställer samma fråga mer än en gång. AI-genererade texter och media kan innehålla fördomar, vinklade eller olämpliga budskap trots de filter som de använder.

Detta ansvar har du som medarbetare

- Du ansvarar för att granska text och säkerställa att den är korrekt, har rätt tonalitet och är fri från fördomar, vinklade eller olämpliga budskap samt är i linje med förvaltningens uppdrag och förhållningssätt.
- Du kan inte överlåta din professionella bedömning till resultatet från en AI-tjänst.

Anskaffa, utveckla och införa AI-system

När ett AI-system eller modell ska köpas, utvecklas och införas behöver det ske utifrån samma principer som en digital tjänst/system. Klassning av information och riskanalys är underlag till kravställning tillsammans med de sedvanliga kraven utifrån behov, arkiv, teknik och lagstiftning. Förutom riskanalys och konsekvensbedömning så ska det för ett AI-system även göras en analys utifrån AI-förordningen för att klassificera typ av AI-system. Typ av AI-system och risknivå styr vilka krav som ställs på dels leverantören, dels användaren.

Den som ansvarar för att anskaffa eller skapa system ska säkerställa både systemets och leverantörens förmåga att skydda informationen, att AI-systemet uppfyller alla krav som ska ställas utifrån tillämplig lagstiftning och funktionalitet.

Skapa AI-system

Det finns flera möjligheter att skapa eller utveckla egna AI-system eller modeller. Ett exempel är Power Platform som är en molnbaserad ”low-code” plattform från Microsoft och en del i Stadens M365- miljö och ger möjligheter att skapa egna AI-system. Plattformen gör det lätt att komma igång med att utforska AI i olika sammanhang samt att på ett enkelt sätt ta en idé till något som kan testas och utvärderas.

Plattformen innehåller många olika byggblock och fler AI-tekniker tillkommer i takt med teknikens utveckling. När du ska skapa ett AI-system med hjälp av Power Platform så behöver du göra det i enlighet med lagstiftning och styrande dokument.

Tänk på att det AI-system som du har skapat kanske inte är lämpligt att skalas upp och tas i bruk i Power Platform om den inte är förenlig med det som anges i denna anvisning. Din idé kan dock ligga till grund för utveckling i en annan plattform eller så måste den anpassas innan den kan skalas upp och tas i drift.

Detta ansvar har du som utvecklar AI-system-/modell:

- Du ansvarar för att aldrig lägga in verkliga personuppgifter, sekretessuppgifter, känslig eller skyddsvärd information när du laborerar med och testar AI-teknik i Power plattformen.
- Du ansvarar för att ha säkerställt alla säkerhets- och lagmässiga krav och aspekter innan systemet sätts i drift.

Definitioner från AI-förordningen artikel 3:

AI-system: ett maskinbaserat system som är utformat för att fungera med varierande grad av autonomi och som kan uppvisa anpassningsförmåga efter införande och som, för uttryckliga eller underförstådda mål, drar slutsatser härledda från den indata det tar emot, om hur utdata såsom förutsägelser, innehåll, rekommendationer eller beslut som kan påverka fysiska eller virtuella miljöer ska genereras.

AI-modell för allmänna ändamål: en AI-modell, även när en sådan AI-modell tränas med en stor mängd data med hjälp av självövervakning i stor skala, som uppvisar betydande generalitet och på ett kompetent sätt kan utföra ett brett spektrum av distinkta uppgifter oavsett hur modellen släppts ut på marknaden och som kan integreras i en rad system eller tillämpningar i efterföljande led, utom AI-modeller som används för forsknings-, utvecklings- eller prototypverksamhet innan de släpps ut på marknaden.

AI-system för allmänna ändamål: ett AI-system som bygger på en AI-modell för allmänna ändamål och som har kapacitet att tjäna en rad olika ändamål, både för direkt användning och för integrering i andra AI-system.

Risk: kombinationen av sannolikheten för skada och denna skadas allvarlighetsgrad.

Leverantör: en fysisk eller juridisk person, en offentlig myndighet, en byrå eller ett annat organ som utvecklar ett AI-system eller en AI-modell för allmänna ändamål eller som har ett AI-system eller en AI-modell för allmänna ändamål och släpper ut det eller den på marknaden eller tar AI-systemet i bruk i eget namn eller under eget varumärke, antingen mot betalning eller kostnadsfritt.

Avsett ändamål: den användning för vilken ett AI-system är avsett av leverantören, inbegripet det specifika användningssammanhanget och de specifika användnings-villkoren, enligt specifikationerna i de uppgifter som tillhandahålls av leverantören i bruksanvisningen, reklam- eller försäljningsmaterial och uttalanden samt i den tekniska dokumentationen.