Kungsbacka

Regler för AI

Dokumentegenskaper:	Titel: Regler för AI,
Beslutad av:	Digitaliseringschef, 2024-03-20
Gäller från:	2024-03-20
Ansvarig förvaltning:	Kommunstyrelsens förvaltning
Kontakt:	Kungsbacka direkt 0300-83 40 00 | info@kungsbacka.se Kungsbacka kommun, 434 81 Kungsbacka www.kungsbacka.se

Regler för AI

1 Inledning

Artificiell intelligens (AI) är ett område som väcker allt större intresse och påverkar vårt samhälle på många sätt. År 2023 blev ett viktigt år för AI då generativ AI på kort tid blev tillgängligt för allmänheten, bland annat tack vare lanseringen av ChatGPT från OpenAI i november 2022.

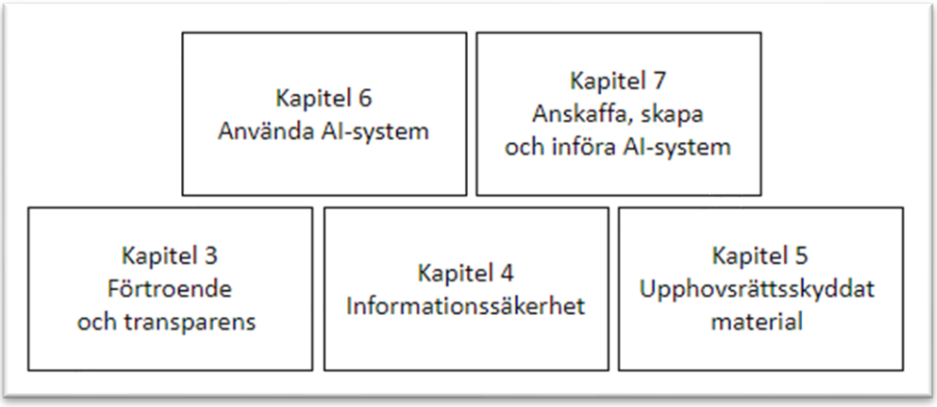
I Kungsbacka kommun är vi nyfikna och vågar effektivisera och utveckla våra arbetssätt med hjälp av ny teknik. Samtidigt måste vi vara ansvarsfulla när vi utforskar de nya möjligheter som tekniken ger oss. För alla nya teknologier måste vi vara medvetna om riskerna, men också ha insikt i de möjligheter de erbjuder oss.

Du uppmuntras att vara nyfiken på den ny tekniken och att prova den på ett klokt sätt. På samma sätt som med alla digitala system är du ansvarig för ditt eget handlande när du använder AI-system.

- Du har ansvar för den information som du matar in i ett AI-system
- Du har ansvar för hur du använder resultatet från ett AI-system

Reglerna i det här dokumentet syftat till att göra dig som medarbetare trygg i hur du kan använda AI i ditt arbete. Reglerna grundar sig till stor del i de lagar, förordningar och regler som kommunen, dess chefer och medarbetare redan lyder under. Förtroende, transparens och informationssäkerhet är några viktiga aspekter. EU:s AI-förordning lägger ytterligare ett lager ovanpå detta.

Reglerna gäller för alla medarbetare i Kungsbacka kommun.



Kapitel 3,4 och 5 utgör grunden för reglerna i kapitel 6 och 7

Kapitel 2 innehåller definitioner av begrepp som används i dokumentet.

Kapitel 3, 4 och 5 handlar om förtroende, transparens, informationssäkerhet och upphovsrätt, områden som utgör grunden för reglerna i de efterföljande kapitlen.

Kapitel 6 innehåller regler kring hur du får använda ett AI-system i ditt dagliga arbete

Kapitel 7 innehåller regler som handlar om att anskaffa, skapa eller införa ett AI-system.

2 Definitioner

Artificiell intelligens (AI) - Ett AI-system är ett maskinbaserat system som, för uttryckliga eller underförstådda ändamål, utifrån de indata det tar emot drar slutsatser om hur man genererar utdata, t.ex. förutsägelser, innehåll, rekommendationer eller beslut som kan påverka fysiska eller virtuella miljöer.

AI-system – med ordet “system” avses appar, programvaror samt andra digitala tjänster och verktyg. Med “AI-system” avses system innehållande större eller mindre inslag av artificiell intelligens (AI). Om det specifikt handlar om generativ AI är detta utskrivet.

Generativ (skapande) AI – AI-system som kan skapa nytt material baserat på enkla instruktioner eller exempel kallas för generativ AI, eller skapande AI. Det finns en stor mängd AI-baserade datorprogram, system och IT-tjänster som genererar nytt material i form av till exempel text, bild, video och programmeringskod. Många av dem är tillgängliga via internet för allmänheten. Detta dokument har inte ambition att försöka namnge dessa AI-system. Några exempel är Copilot (tidigare Bing Chat Enterprise) och ChatGPT (AI-baserade chattbotar) som kan skapa text utifrån dina instruktioner samt Dall-E som kan skapa bilder utifrån dina instruktioner.

Allmänt tillgängliga AI-system – AI-system som är tillgängliga för vem som helst att använda t.ex. via webbsida eller via en app. Detta dokument har inte ambition att försöka namnge dessa AI-system. Kända exempel är ChatGPT och DALL-E från OpenAI. Allmänt tillgängliga AI-system är inte anskaffade, skapade eller införda specifikt för kommunens behov och är därför inte heller granskade av kommunen.

Öppen information – information som är tillgänglig för alla utan några begränsningar.

Känsliga personuppgifter – uppgifter om ras eller etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse, medlemskap i en fackförening, hälsa, en persons sexualliv eller sexuella läggning, genetiska uppgifter och biometriska uppgifter som entydigt identifierar en person.

Särskilt skyddsvärda personuppgifter – uppgifter som på grund av sin karaktär är extra viktiga att skydda, även om de inte definieras som känsliga personuppgifter i Dataskyddsförordningen. Personnummer, vissa uppgifter om ekonomiska förhållanden, omdömen och värderingar av en person såsom social förmåga, inlärningsförmåga och liknande, provresultat, resultat av personlighetstester och annan information som ligger nära den privata sfären är exempel på extra skyddsvärda personuppgifter.

3 Förtroende och transparens

I Sverige finns det ett högt förtroende för offentlig sektor. För att bevara eller stärka förtroendet när offentlig sektor använder AI måste AI användas på ett ansvarfullt sätt, med stöd av välutvecklade processer, ett systematiserat arbetssätt och med god dokumentation.

Kommunens verksamheter måste alltid följa den lagstiftning som gäller. I verksamhet som är lagstyrd behöver man analysera om det man tänkt göra är lagligt innan man börjar använda AI-systemet. Det behöver också övervägas om det skapas nya allmänna handlingar genom användning av AI-systemet och hur de i så fall ska lagras och diarieföras.

En viktig aspekt av förtroende är transparens. Sverige har med offentlighetsprincipen ett väl utvecklat system för transparens inom offentlig sektor.

Enligt förvaltningslagen är det ett krav att en myndighet ska motivera sina beslut och förklara vad som gjort att myndigheten nått sin slutsats. Därför behöver varje verksamhet som använder ett AI-system på ett sätt som påverkar kommunens invånare eller brukare kunna förstå och förklara sitt AI-systems inre funktioner på en lämplig nivå. Utmaningen ligger i att ett AI-system kan bestå av tusentals, eller till och med miljontals, numeriska värden, värden som AI-systemet lär sig under dess träningsfas. Det är därför ofta inte möjligt, ens för den som utvecklat systemet, att förklara exakt hur systemet når en viss slutsats. Därför kan det i vissa sammanhang vara olämpligt att använda AI av det skälet.

En verksamhet som använder ett AI-system behöver veta om AI-systemet skapar nya allmänna handlingar och hur de i så fall ska lagras och diarieföras.

En verksamhet som använder ett AI-system på ett sätt som påverkar kommunens invånare eller brukare behöver kunna förstå och förklara sitt AI-systems inre funktioner.

Se även ”Automatisera lagligt – handbok i automatisering av ärendehandläggning och beslut”

4 Informationssäkerhet

I de flesta fall kräver användningen av ett AI-system att du matar in någon form av information i AI-systemet. Kommunens information är en av våra viktigaste tillgångar. Vår information omfattas av lagstiftning som rör bland annat personuppgiftsbehandling och sekretesskydd. Kommunens information ska vara informationsklassad och hanteras i system som är godkända utifrån informationsklassen. Det gäller även AI-system.

När kommunen anskaffar eller utvecklar system kan vi ställa krav på systemet utifrån den information som ska hanteras i det. Det kan inte kommunen när det gäller allmänt tillgängliga system som var och en kan använda. Där gäller i stället de användarvillkor du accepterar när du använder systemet.

Du bör vara medveten om *hur* AI-systemet och dess leverantör kan komma åt och använda den information som du matar in i systemet. Du måste veta om leverantören garanterar att den information som du matar in i AI-systemet inte används för att träna systemet, eller om det finns risk att den information som du matar in återanvändas i svar som ges till andra användare, vilket innebär att den kan spridas helt öppet.

Det finns även annan information som kan leda till skada om den sprids på fel sätt. Nedan kan du läsa om vad som gäller för behandling av personuppgifter, sekretessbelagd information, annan känslig information och upphovsrättsskyddat material i AI-system.

Allmänt tillgängliga AI-system får enbart användas för öppen information.

Hanteringen av personuppgifter och sekretessbelagd information omfattas av lagregler som måste följas. Det gäller även vid användning av AI-system.

Se även “Regler för informationssäkerhet för medarbetare och förtroendevalda”

4.1 Personuppgifter

Att mata in personuppgifter i ett AI-system är att behandla personuppgifter. All behandling av personuppgifter ska följa de grundläggande principerna i GDPR. Om du inte känner till de grundläggande principerna i GDPR kan du läsa om dem på Integritetsskyddsmyndighetens hemsida. Om och hur principerna uppfylls måste analyseras *innan* AI-systemet används för den tänkta behandlingen. Eftersom AI är en ny teknik ska det i de flesta fall också göras en konsekvensbedömning enligt GDPR innan personuppgifter används i ett AI-system.

Om det inte har gjorts en analys av behandlingen utifrån GDPR är det olagligt att mata in personuppgifter i ett AI-system.

När det gäller allmänt tillgängliga AI-system saknas tillräckliga möjligheter för kommunen att kontrollera leverantörens behandling eller säkerhet.

Allmänt tillgängliga AI-system levereras ofta som molntjänster av utländska företag. Samma lagar och regler gäller vid användning av AI-system i molnet som för andra typer av molntjänster.

Innan personuppgifter matas in i ett AI-system måste det göras en analys av behandlingen enligt GDPR.

Personuppgifter får inte matas in i allmänt tillgängliga AI-system.

4.2 Sekretess

Enligt offentlighets och sekretesslagen får information som omfattas av sekretess inte *röjas* för utomstående. När information delas med en IT-leverantör, t ex genom att laddas upp i en molntjänst, är informationen röjd. Om informationen är krypterad på ett sådant sätt att leverantören inte kan göra den läsbar är den dock inte röjd. Att röja sekretessbelagd information är bara tillåtet om det finns en sekretessbrytande bestämmelse.

Om du röjer sekretessbelagda uppgifter genom att mata in dem i ett AI-system kan det vara straffbart enligt brottsbalken som *brott mot tystnadsplikten*.

För att avgöra om det är tillåtet att mata in sekretessbelagda uppgifter i ett AI-system behöver det göras en teknisk och juridisk analys.

Mata aldrig in sekretessbelagd information i ett AI-verktyg som inte har analyserats och konstaterats vara säkert för sådan information.

Sekretessbelagd information får inte matas in i allmänt tillgängliga AI-system.

4.3 Annan känslig information

Annan känslig information är till exempel information som, om den kommer i fel händer, riskerar skada individer, grupper av individer eller kommunen i allmänhet. Det kan röra sig om information från interna möten eller preliminära bedömningar i olika ärenden. Allmänt tillgängliga AI-system som använder inmatad information för att träna systemet, kan ta med din information i svar till andra användare. Det vill säga den information som du matar in riskerar att göras tillgänglig för andra.

Mata aldrig in annan känslig information i ett AI-verktyg som inte har analyserats och konstaterats vara säkert för sådan information.

Annan känslig information får inte matas in i allmänt tillgängliga AI-system.

5 Upphovsrättsskyddat material

Det är inte tydligt i lagstiftning och praxis om det är tillåtet eller otillåtet att mata in upphovsrättsskyddat material i ett AI-system. Det kan vara tillåtet enligt ett undantag i 2 kap 15a § upphovsrättslagen, men det är inte rättsligt prövat.

I de användarvillkor som man godkänner när man använder ChatGPT och Copilot anges att man själv ansvarar för att inte bryta mot någon annans upphovsrätt. Det innebär att det är man 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 tänk igenom om du verkligen behöver mata in upphovsrättsskyddat material. Andra system kan ha andra villkor.

Det kan vara bra att veta att författningar (som lagar, förordningar och föreskrifter), beslut av myndigheter och yttranden av myndigheter inte omfattas av upphovsrätt.

Det är oklart om bilder och texter som genererats av AI skyddas av upphovsrätt eller ej. I regel gör dock tillhandahållare av generativ AI inte anspråk på upphovsrätt till innehåll som genererats, utan det står dig fritt att använda materialet som du vill.

Tillhandahållare av generativa AI-system garanterar dock i regel inte att det resultat som generas inte är likt tidigare upphovsrättsskyddat material. Eftersom modellerna är tränade på material från internet kan resultaten bli väldigt likt en bild eller text som är vanlig på internet. Av användarvillkoren till både Copilot och till OpenAI (som tillhandahåller ChatGPT och Dall-E) framgår att du som användare är ansvarig om du använder AI-genererat material på ett sätt som strider mot någon annans upphovsrätt. Det innebär att om du sprider en AI-genererad bild eller text som är för lik någons upphovsrättsskyddade verk kan du begå ett upphovsrättsbrott och bli skadeståndsskyldig.

6 Att använda AI-system

Oavsett om AI-systemet som du använder är tillgängligt för allmänheten, är av typen generativ (skapande) AI eller om AI-systemet är anskaffat, skapat eller infört specifikt för kommunens behov så gäller två grundläggande saker:

Du har ansvar för den information som du matar in i AI-systemet (se kapitlet om informationssäkerhet).

Du har ansvar för hur du använder resultatet från AI-systemet.

6.1 Allmänt tillgängliga AI-system

I takt med teknikutvecklingen lanseras nya AI-system som görs allmänt tillgängliga för användning. Allmänt tillgängliga AI-system kan tillföra värde för kommunen genom att underlätta vissa arbetsuppgifter och effektivisera arbetet.

Utöver det som framgår i kapitlet om informationssäkerhet och upphovsrätt måste du även tänka på följande innan du använder ett allmänt tillgängligt AI-system:

Du som använder systemet accepterar de användarvillkor som gäller för systemet. Du personligen är ansvarig för att känna till och följa villkoren. Vid osäkerhet om användningen, fråga din närmaste chef.

Eftersom kommunen inte har något avtal med leverantören av allmänt tillgängliga AI-system finns det ingen möjlighet för kommunen att påverka eller få insyn i leverantörens användning av informationen i AI-systemet. Det är du personligen som godkänner de villkor som gäller för systemet. Det är därför du personligen som tar ansvar för den information du lämnar ifrån dig och att leverantören använder informationen på det sätt som framgår av villkoren. Du använder alltså ett allmänt tillgängligt AI-system på egen risk och på eget ansvar.

Tänk på att det är svårt att veta exakt vad ett AI-system gör. Vad vet du om det AI-systemet som du tänkt att använda? Gör systemet det som du tänkt dig och uppfyller det de krav på kvalitet som krävs för uppgiften? Du personligen är ansvarig för resultatet. Ju mindre du vet om systemet, desto större risk tar du.

AI-baserade chattbotar levereras ofta av utländska företag som molntjänster. Samma lagar och regler gäller vid användning av AI-system i molnet som för andra typer av molntjänster.

Du som använder systemet accepterar de användarvillkor som gäller för systemet. Du personligen är ansvarig för att känna till och följa villkoren.

Använd aldrig ett allmänt tillgängligt AI-system som stöd för beslutsfattande eller som en del av ärendehandläggningen.

Använd inte ett AI-system som är helt nytt eller kommer från en okänd avsändare i ditt arbete.

Ge aldrig ett allmänt tillgängligt AI-system åtkomst till information på dina enheter.

Allmänt tillgängliga AI-system får enbart användas för öppen information.

Du får under inga omständigheter mata in personuppgifter, sekretessbelagd information eller annan känslig information i AI-system tillgängliga för allmänheten.

6.2 Copilot

Copilot är Microsofts AI-system och en del av Microsofts 365-miljö (M365) som vi har i kommunen. Det innebär att den information du matar in hanteras i Microsofts molntjänst på samma sätt som informationen i OneDrive eller Outlook. Precis som med övriga produkter i M365 får känsliga personuppgifter, särskilt skyddsvärda personuppgifter och sekretessbelagda uppgifter inte behandlas i Copilot.

När du är inloggad med ditt Kungsbacka kommun-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 ett allmänt tillgängligt AI-system 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.

Den tekniska säkerheten för M365, och därmed också för Copilot, är tillräcklig för att behandla personuppgifter som inte är känsliga eller särskilt skyddsvärda.

Varje förvaltning måste ta ställning till i vilken mån det finns laglig grund för att behandla personuppgifter i Copilot, och i så fall för vilka ändamål.

6.3 Använda resultatet från generativ (skapande) AI

AI-system som kan skapa nytt material baserat på enkla instruktioner eller exempel kallas för generativ AI, eller skapande AI. Du ansvarar för resultatet som AI-systemet har skapat, därmed måste du också kontrollera det.

Du behöver vara medveten om hur generativ AI skapar sitt resultat och att resultatet kan vilseleda. Det vill säga du behöver vara medveten om potentialen för felaktig information från dessa generativa AI-system.

De generativa AI-systemen är inte sökmotorer, de är däremot väldigt duktiga på att skapa nytt trovärdigt material utifrån dina instruktioner, till exempel att skapa ny text. För att skapa ny text använder sig det generativa AI-systemet av en språkmodell som, baserat på sannolikhet, beräknar vilket kommande ord som är mest troligt. De generativa AI-systemen kan också erbjuda olika svar på samma fråga om den ställs mer än en gång, och de kan hämta sina svar från källor som du inte skulle lita på i andra sammanhang.

Att texten som det generativa AI-systemet har skapat ser trovärdig ut är inte samma sak som att den är faktamässig korrekt, att den har rätt tonalitet eller att den är fri från fördomar eller annan bias. Behandla alltid de resultat som dessa generativa AI-system producerar som ett utkast.

Du är alltid ansvarig för att texten är faktamässigt korrekt, har rätt tonalitet och är fri från fördomar och annan bias. Du kan inte överlåta din yrkesmässiga bedömning till ett generativt AI-system. Du ansvarar för texten som om du själv hade skrivit den.

7 Att anskaffa, skapa och införa AI-system

Att anskaffa eller skapa nya AI-system, eller att införa befintliga AI-system i nya verksamheter är en del i Kungsbacka kommuns digitaliseringsarbete och integreras med övrig styrning och utveckling i kommunen. Genom att vi anskaffar, skapar och inför AI-system enligt befintliga processer och metodiker säkerställer vi också att alla relevanta aspekter utreds.

När kommunen anskaffar eller själv skapar system utvärderas systemets och leverantörens förmåga att skydda informationen. Det ställs också krav på funktioner och avtalsvillkor som säkrar kommunens kontroll över informationen.

En viktig del av processen är att säkerställa att AI-systemet är tillräckligt transparent och uppfyller de krav som behöver ställas utifrån tillämplig lagstiftning. Tillämplig lagstiftning kan vara den speciallagstiftning som gäller för verksamheten, offentlighets- och sekretesslagen, GDPR, förvaltningslagen, kommunallagen och lagstiftning om cybersäkerhet och EU:s kommande AI-förordning.

Befintliga processer och metodiker anpassas kontinuerligt utefter ny teknik, direktiv och lagstiftning.

> Befintliga processer och metodiker för digitaliserings- och innovationsarbete gäller och ska användas, oavsett om den tänkta lösningen innehåller AI eller inte.

7.1 Skapa AI-system

Det finns flera möjligheter att skapa eller utveckla egna AI-system. Power Platform som är en molnbaserad ”low-code” plattform från Microsoft och en del i vår M365-miljö 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 man ska skapa ett AI-system med hjälp av Power Platform är det extra viktigt att välja information, tillämpning och behandling som är förenlig med reglerna i detta dokument.

Tänk på att det AI-system som du har skapat kanske inte är lämplig att skalas upp och tas i bruk i Power Platform om den inte är förenlig med reglerna i detta dokument. 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.

> Du får inte använda verkliga personuppgifter, sekretessbelagd information eller annan känslig information när du laborerar med och testar AI-teknik i Power plattformen.
>
> De lösningar du bygger i Power Platform får inte tas i drift utan att en bedömning har gjorts av lösningen enligt det som nämns ovan.