



Policy

Reglemente

Föreskrift

➤ Bestämmelse

Riktlinje

Strategi

Handlingsplan

Plan

Rutin



**Trollhättans
Stad**

Bestämmelser för AI

Dokumentinformation

Dokumentbeteckning:	Bestämmelse
Antaget av:	Stadsdirektör
Diarienummer:	KS 2024/1062
Datum/Ersätter:	2024-12-02
Handlingen publiceras:	Trollhättans stads intranät
Syfte:	Syftet med bestämmelsen är att säkerställa säker hantering av information och beaktande av etiska principer i samband med användning av AI-system, samt att ge medarbetare trygghet att vara nyfikna och utforskande kring AI-teknikens möjligheter.
Gäller för:	Trollhättans Stad
Referensdokument:	KS 2024/273, dokument- och informationshanteringsplan.
Dokumentansvarig:	Digitaliseringschef
Uppföljning:	2026-03-01

Innehåll

Bestämmelser för AI	0
Dokumentinformation.....	1
1. Inledning.....	3
2. Vägledning.....	4
3. Bestämmelser.....	8
4. Framtida överväganden.....	10

1. Inledning

Artificiell intelligens (AI) förändrar vårt samhälle i ökande takt och skapar kontinuerligt nya möjligheter. Utifrån Trollhättans Stads ledord mod, nytänkande, delaktighet och handlingskraft är utforskandet av ny teknik och dess möjligheter en självklarhet.

Med dessa möjligheter följer också risker och ansvar, där användningen av AI väcker viktiga frågor kring etik, integritet och säkerhet. Det är avgörande att AI implementeras på ett sätt som respekterar medborgarnas rättigheter och säkerställer en balans mellan innovation och ansvarstagande.

Syftet med bestämmelsen är att säkerställa säker hantering av information och beaktande av etiska principer i samband med användning av AI-system, samt att ge medarbetare trygghet att vara nyfikna och utforskande kring AI-teknikens möjligheter.

Dokumentet inkluderar även en vägledning som ger Trollhättans Stads chefer och medarbetare ytterligare stöd i smart användande av AI-system, samt strategiska överväganden för framtiden.

Definitioner

Allmänt tillgängliga AI-system – AI-system som är tillgängliga för vem som helst att använda och som inte är anskaffade, utformade och/eller införda i Trollhättans Stads digitala infrastruktur.

2. Vägledning

Artificiell intelligens (AI) kan definieras som system eller mjukvara som kan utföra uppgifter självständigt och anpassar sitt beteende baserat på förändringar i sin omgivning. Systemen är designade att vara autonoma, vilket innebär att de kan fatta beslut och agera utan mänsklig inblandning. De är också adaptiva, så att de kan lära sig från erfarenheter och därmed förbättra sin prestation över tid. Denna förmåga till självlärande och anpassning är central för modern AI.

Du uppmuntras att vara nyfiken på den nya tekniken och att prova den på ett klokt sätt, med följande vägledning och bestämmelser som ett stöd.

Olika sätt att använda AI

AI-teknik implementeras idag på allt fler ställen och de flesta av oss använder AI varje dag utan att vi är medvetna om det. Det kan vara allt ifrån en rättstavningsfunktion till en navigationstjänst som föreslår den optimala ruten. I denna vägledning skiljer vi på inbyggd AI och AI-assistenter.

Inbyggd AI

AI kan vara en inbyggd del i en digital lösning. Det kan vara olika slags AI som tex en chattfunktion eller en prognosfunktion. Då används tekniken som en del i en lösning tillsammans med traditionella program och tjänster. Genom att använda smarta AI-lösningar i digitaliseringsprojekt kan staden förbättra effektivitet och anpassningsförmåga till medborgarnas behov. Många verksamhetssystem kommer på sikt att innehålla AI som delkomponent.

AI som assistent

AI som assistent är ett AI-system som både förstår och kan skapa texter och andra medier. AI-assistenter används ofta via ett chattfönster och erbjuder en rad möjligheter att få stöd i olika arbetsuppgifter. Ett tydligt exempel på detta är ChatGPT. Genom att använda ChatGPT kan användare få hjälp med att sammanställa dokument, generera kreativa idéer, och till och med skapa kunskapsunderlag genom att förse systemet med enkla instruktioner, så kallade "prompter". ChatGPT kan analysera och bearbeta stora mängder text och dokument snabbt, vilket gör den till en ovärderlig assistent i många olika sammanhang.

Ett exempel på hur ChatGPT kan användas är att be om en sammanfattning av en komplex rapport eller att generera en presentation utifrån specifika krav. På så vis kan AI bidra till att spara tid och öka effektiviteten i olika arbetsuppgifter.

Smart användning av AI-assistenter

AI-assistenter kan vara en stor hjälp i att utföra vissa arbetsuppgifter. För att få ut så mycket som möjligt finns det några saker att vara uppmärksam på.

Att få ett bra resultat

Att skriva instruktioner till en AI-assistent kallas för att prompta den. Det är viktigt att tänka på att ju mer specifika instruktionerna är, desto bättre blir resultatet. Så här kan ett exempel se ut:

”Jag har i uppgift att sammanställa ett dokument om kvalitetsledningssystem till ledningsgruppen. Vi kommer att ha en diskussion efteråt om innehållet. De har bara grundläggande kunskaper i hur man arbetar med kvalitetsledning. Det jag vill att du ska göra är att göra en sammanfattning som ledningsgruppen kan läsa innan vårt möte. Jag vill också ha ett underlag till en Power Point-presentation som innehåller de viktigaste punkterna att ta upp till diskussion. Sammanfattningen ska vara ca två A4-sidor och presentationen ska innehålla lite teori om kvalitetsledning, varför det är bra att jobba med och hur vi kan gå vidare i arbetet nu när vi känner bakgrunden.”

Ställ följdfrågor om du inte är nöjd med resultatet och tänk återigen på att ju mer specifikt du skriver desto bättre blir sannolikt svaren. För ännu träffsäkrare resultat kan prompting göras på engelska eftersom AI-assistenter till största del är tränade på engelsk information.

En bra grundtanke är att skriva med en AI-assistent som om det vore en med en kollega. Skriv utförligt med hela meningar och tydliggör vad som är viktigt.

Ju mer specifika instruktionerna är, desto bättre blir resultatet.

Att använda resultatet

Informationen som fås tillbaka ifrån en AI-assistent behöver värderas. Det är alltid viktigt att kontrollera och kvalitetsgranska resultatet. Det är också viktigt att tänka på att en AI-assistent enbart kan det den har läst och saknar kontextuell förståelse som kan vara viktig för resultatet.

Det är också viktigt att förstå begreppet ”hallucinationer”. Hallucinationer innebär i detta sammanhang att AI-assistenter kan svara med påhittade fakta och information som inte finns. Hallucinationer är sällsynta men då de förekommer är det ytterligare ett skäl att kontrollera resultatet.

En bra grundtanke är att se på resultatet som om uppgiften givits till en medarbetare som är ny på jobbet, även om svaret är utförligt behöver informationen kontrolleras.

Säker användning av allmänt tillgängliga AI-assistenter

Vid användning av allmänt tillgängliga AI-assistenter är det viktigt att vara uppmärksam på vilken information som delas. Även om känsliga uppgifter utelämnas kan det finnas utrymme för att utläsa information "mellan raderna" som trots maskning gör informationen olämplig att dela med en allmänt tillgänglig AI-assistent. Detta fenomen, att information kan utläsas "mellan raderna", kallas inferens.

AI-assistenter kan också använda inmatad information som träningsdata för vidareutveckling av tjänsten. Om vi från olika håll delar med oss av uppgifter som är harmlösa när de är åtskilda, kan informationen läggas samman av AI-assistenten och sammanlagt avslöja mer om vår organisation än vi önskar. Det kan således leda till att vi synliggör brister i vår organisation som kan utnyttjas.

Ett bra verktyg för att bedöma vilken information som är lämplig att dela med allmänt tillgängliga AI-assistenter är att göra en riskanalys innan den introduceras i verksamheten. Riskanalysen går ut på att identifiera vilka risker som användningen kan medföra samt hur allvarliga riskerna är utifrån sannolikhet och konsekvens.

En tydlig minnesregel för att använda allmänt tillgängliga AI-assistenter är "Skulle informationen jag vill dela med AI-assistenten kunna publiceras öppet på trollhattan.se?"

Är svaret ja, kan du troligen dela den med en allmänt tillgänglig AI-assistent.

Rekommenderade kurser

Olika roller inom Trollhättans Stad har olika kompetensutvecklingsbehov kopplat till AI. Nedan ges sammanfattade rekommendationer för utvalda målgrupper.

- För den breda målgruppen rekommenderas filmserien *Starta din AI-resa* från AI Sweden, vårt nationella center för tillämpad AI. Den riktar sig till alla som vill få en större förståelse för vad AI är och består av sex moduler som är mellan 5-10 minuter långa.

[AI Sweden – starta din AI-resa](#)

- För chefer rekommenderas filmserien *AI för offentlig sektor* från AI Sweden. Det är en filmserie i tre delar mellan 5-10 minuter långa som ger en inblick i vad ledare och beslutsfattare behöver tänka på kopplat till AI.

[AI Sweden – AI för offentlig sektor](#)

- För roller som har behov av en fördjupad en fördjupad kompetens inom AI och dess bakomliggande logik rekommenderas utbildningen Elements of AI. Utbildningen är bland annat framtagen av Helsingfors universitet och Linköpings universitet och finns tillgänglig gratis på flera språk.

[Elements of AI](#)

Tänk på att

- 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.
- Du kan alltid kontakta din informationssäkerhetssamordnare för vägledning samt stöd i hantering av information.
- Du kan alltid kontakta Kontoret för Digitalisering och IT för vägledning samt stöd i anskaffande av AI-system.
- Du kan alltid kontakta Personalkontoret för stöd i strategisk kompetensförsörjning.

3. Bestämmelser

Nedan följer bestämmelser för användning av AI inom Trollhättans Stad.

Informationssäkerhet

Trollhättans Stad ansvarar för sin informationshantering och har ett flertal krav på sig, exempelvis att skydda informationen mot obehörig åtkomst och att dokumentera personuppgiftsbehandlingen i en registerförteckning. Detta gäller både den interna hanteringen och våra leverantörers hantering.

Många AI-system tillhandahålls av stora leverantörer som har sitt huvudsakliga verksamhetsställe utanför EU/EES. Eftersom vi omfattas av kraven i dataskyddsförordningen (GDPR) så behöver vi säkerställa att även våra leverantörer lever upp till dessa krav, vilket kan vara utmanande när vi nyttjar tjänster från stora leverantörer.

Bestämmelser för användning av allmänt tillgängliga AI-system

Till stöd för bedömningen av vilken information som får hanteras i allmänt tillgängliga AI-system finns nedanstående sammanställning. Uppgifter som inte får hanteras i allmänt tillgängliga AI-system är:

- Sekretessuppgifter
- Personnummer och samordningsnummer
- Skyddsobjekt och säkerhetsklassade data
- Känsliga personuppgifter, vilket definieras som
 - etniskt ursprung
 - politiska åsikter
 - religiös eller filosofisk övertygelse
 - medlemskap i en fackförening
 - hälsouppgifter
 - en persons sexualliv eller sexuella läggning
 - genetiska uppgifter
- biometriska uppgifter som används för att entydigt identifiera en person
- Särskilt skyddsvärda personuppgifter
 - Löneuppgifter
 - Uppgifter om lagöverträdelse
 - värderande uppgifter, till exempel uppgifter från utvecklingssamtal, uppgifter om resultat från personlighetstester eller personlighetsprofiler
 - information som rör någons privata sfär
- uppgifter om sociala förhållanden

Ytterligare information och förtydliganden finns i KS 2024/273, dokument- och informationshanteringsplan.

Myndighetsbeslut

I vår myndighetsutövning måste samtliga beslut som fattas kunna motiveras. I det fall ett AI-system har använts som en del av beslutsfattandet måste spårbarhet och kännedom finnas om AI-systemets resonemang och regelverk, vilket i stor utsträckning inte är möjligt. Allmänt tillgängliga AI-system ska av denna anledning inte användas som en del av beslutsprocessen vid myndighetsutövning. Vid användning av för staden anskaffade AI-system vid beslutsfattande måste spårbarheten i beslutet kunna säkerställas samt lagstiftning kring automatiskt beslutsfattande beaktas.

Etiska bestämmelser

Vid utveckling, inköp och användning av ett AI-system i Trollhättans Stad ska förutom gällande lagar och regler även nedanstående etiska bestämmelser beaktas och följas. Detta är särskilt viktigt i AI-system som används för att göra bedömningar eller beslut som påverkar en medborgares, brukares, elevs med fleras rätt till tjänster, förmåner eller andra rättigheter.

Transparens

AI-system ska utformas och införas på ett sätt som innebär att systemets möjligheter och begränsningar kommuniceras på ett begripligt sätt till alla intressenter.

Etiskt och rättvist

AI-system ska utformas och användas på ett sätt som respekterar mänskliga rättigheter, värderingar och kulturell mångfald. Detta inkluderar att minska riskerna med AI-algoritmer i förhållande till partiskhet och diskriminering. Det är viktigt att beakta etiska konsekvenser i varje skede av AI-utveckling och driftsättning.

Säkerhet och robusthet

AI-system ska utformas att vara tillförlitliga, kontrollerbara och säkra. De ska vara motståndskraftiga mot manipulation och fel. Detta innefattar säkerställande av att AI-systemet inte orsakar oavsiktlig skada på människor eller miljö.

4. Framtida överväganden

AI som strategisk fråga

AI har potentialen att förändra och förbättra våra arbetssätt i grunden, men för att detta ska bli verklighet måste vi betrakta AI som ett strategiskt utvecklingsområde och en del av stadens övergripande riktning. Det kräver ett aktivt engagemang och en vilja att investera i nya lösningar, även med vetskap om att risken för misslyckande är högre än vid traditionella utvecklingsprojekt.

För att lyckas med AI är det avgörande att betrakta data som en strategisk resurs. AI-lösningar är beroende av tillgång till korrekta, relevanta och uppdaterade data. Detta kan innebära att vi måste fatta beslut kring standardisering av datamängder och säkerställande av att data är tillgänglig och användbar över olika system och verksamheter.

Reglering och lagstiftning

Utvecklingen inom AI-teknik och dess tillämpningar går snabbt, samtidigt som arbete med reglering och lagstiftning sker på flera håll. Det medför att löpande utvärdering och revidering av denna vägledning kommer att vara nödvändig.

EU förordningen AI Act

EU har nyligen antagit AI-förordningen (AI Act), en lagstiftning som syftar till att reglera användningen av artificiell intelligens i medlemsländerna. Den fokuserar på att skydda medborgarnas rättigheter genom ökad transparens, krav på ansvariga och reglering av AI-system.

AI-förordningen innebär att vi måste säkerställa att våra AI-system används etiskt och pålitligt, utan att inskränka medborgarnas rättigheter eller bidra till diskriminering. Lagen kommer att komplettera GDPR och ställa krav på ökad informationssäkerhet och transparens i beslut som fattas med hjälp av AI. Förordningen kommer att integreras i svensk lagstiftning och kräva att vi anpassar våra rutiner och system för att följa de nya reglerna.

Regeringen har tillsatt en utredning för att se över behovet av nationella anpassningar till följd av AI-förordningen. Denna bestämmelse kan komma att revideras utifrån den kommande tillämpningen av AI-förordningens regler samt kompletterande bestämmelser i svensk lag.

[SKR – AI Act](#)

[Regeringen – Utredning kring AI-förordningen](#)

Regeringsuppdrag kring riktlinjer för generativ AI

Regeringen har gett Myndigheten för digital förvaltning (Digg) och Integritetsskyddsmyndigheten (IMY) i uppdrag att ta fram vägledande riktlinjer för att stödja offentlig sektor i användandet av AI. Riktlinjerna ska ge stöd till anställda på myndigheter, kommuner och regioner. I uppdraget kommer Digg också erbjuda ett kompetenslyft för att stödja offentliga organisationer i användandet av AI.

Denna bestämmelse kan komma att revideras utifrån de vägledande riktlinjerna från regeringen.

[Regeringen - uppdrag till Digg och IMY](#)

Trollhättans stad
Gårdhemsvägen 9, 461 83 Trollhättan

Telefon: 0520-49 50 00

Epost: kontaktcenter@trollhattan.se

Webb: www.trollhattan.se



Trollhättans Stad