

Att använda AI i Göteborgs Stad

Det här är ett stöddokument som konkretiserar styrningen av hur vi ska arbeta med AI i Göteborgs Stad. Dokumentet vänder sig till medarbetare och chefer och gäller för all användning av AI-system i verksamheten, både publika och stadens egna. Dokumentet omfattar dock inte utveckling, upphandling eller automatiserat beslutsfattande.

AI kan vara ett bra hjälpmedel i arbetet men att använda AI är också förenat med risker. Dokumentet syftar till att stödja en laglig, etisk och ansvarsfull användning av AI i linje med gällande styrdokument.

Dokumentet beskriver vad den som vill använda AI i arbetet behöver tänka på. Det är verksamhetens ansvar att skapa förutsättningar för medarbetarna att kunna använda AI på ett korrekt och säkert sätt.

Med AI-system avses i det här dokumentet både fristående verktyg och AI som är inbyggd i till exempel en tjänst eller plattform. I vissa fall är det inte uppenbart att en digital tjänst använder AI men genom att följa förhållningsreglerna i dokumentet kan riskerna med AI-användning minimeras.

Dokumentet innehåller ett antal olika **förhållningsregler** som du som använder AI när du utför arbetsuppgifter behöver tänka på, en **fördjupningsdel** som du kan läsa om du behöver få bättre förståelse för vad förhållningsreglerna innebär och en **lista med förklaringar** av förkortningar och begrepp som används i dokumentet.

Du som vill använda ett AI-system i arbetet behöver tänka på att:

1. Du ansvarar för att du använder AI på ett etiskt och ansvarsfullt sätt.
2. Innan du använder ett AI-system behöver du vara säker på att systemet uppfyller säkerhetskraven för den typ av information som du ska ge det tillgång till.
3. Det finns särskilda krav på AI-användning som innebär höga risker och en del AI-system är helt förbjudna.
4. Du ansvarar för att sekretessuppgifter och skyddsvärd information inte röjs för obehöriga.
5. Du får bara behandla personuppgifter om det sker i enlighet med dataskyddslagstiftningen.
6. När du använder ett AI-system måste du kontrollera resultatet som systemet har skapat och du är ansvarig för hur resultatet används.
7. Du ansvarar för att ditt AI-genererade material används på ett transparent sätt som bidrar till förtroendet för stadens verksamhet.
8. Information du lägger in i systemet och resultat som du får ut kan bli allmänna handlingar.
9. Du ansvarar för att upphovsrätten respekteras när du använder ett AI-system.

1. Du ansvarar för att använda AI på ett etiskt och ansvarsfullt sätt

Tänk på att

- Det alltid är du som människa som ska ha kontrollen och säkerställa att AI-systemet används på ett etiskt och ansvarsfullt sätt.
- Ha en nyfiken men försiktig inställning när du testar nya system och funktioner.

Fördjupning

AI-system är digitala arbetsverktyg. Att använda ett AI-system i arbetet innebär att ta stöd av en maskin för att utföra sina arbetsuppgifter. Precis som när vi använder andra maskiner i tjänsten är det viktigt att använda dem på ett ansvarsfullt sätt så att de inte orsakar någon skada.

Det är lätt att uppfatta ett AI-system som intelligent men det är viktigt att vara medveten om att de inte har någon egen intelligens. Ett AI-system har inte heller moral eller etik. Slutsatser som för en maskin kan framstå som logiska kan få helt fel konsekvenser i verkligheten. Det är viktigt att det alltid är du som människa som har kontrollen och säkerställer att AI-systemet används på ett etiskt och ansvarsfullt sätt.

Förutom att följa gällande lagar ska vi använda AI-system på ett sätt som respekterar de etiska principerna om *respekt för människans autonomi, skadeförebyggande, rättvisa och förklarbarhet*¹. De här principerna kan aktualiseras på olika sätt när vi använder AI för att utföra arbetsuppgifter.

Respekt för människans autonomi

Med människans autonomi menas att människor som interagerar med AI-system ska kunna behålla makten att bestämma över sina egna liv och kunna delta i den demokratiska processen. AI-system bör utformas för att förhöja, komplettera och stärka människors kognitiva, sociala och kulturella förmågor. Fördelningen av funktioner mellan människor och AI-system bör utgå från en människocentrerad utformning.

Skadeförebyggande

AI-system ska inte användas på ett sätt som innebär att de orsakar skada eller påverkar människor negativt på annat sätt. Människans värdighet och integritet ska skyddas. Sårbara personer bör uppmärksammas särskilt och inkluderas i utvecklingen och spridningen av AI-system. Situationer där AI-system kan orsaka negativa effekter på grund av olika tillgång till inflytande eller information bör också uppmärksammas särskilt, till exempel mellan arbetsgivare och arbetstagare eller kommunen och dess medborgare.

¹ Etiska riktlinjer för tillförlitlig AI, Europeiska kommissionens expertgrupp på hög nivå för AI-frågor, 2019, s. 13-15

Rättvisa

Principen innebär ett åtagande om att säkerställa en jämlik och rättvis fördelning av både fördelar och kostnader samt att säkerställa att individer och grupper inte utsätts för orättvis snedvridning, diskriminering och stigmatisering. Lika möjligheter i fråga om tillgång till utbildning, varor, tjänster och teknik bör också främjas. Dessutom får vår användning av AI-system aldrig leda till att någon vilseleds eller hindras i sin valfrihet. Rättvisa innebär också att noggrant överväga hur avvägningen ska göras mellan motstridiga intressen och mål.

Förklarbarhet

Förklarbarhet är avgörande för att bygga och bibehålla medborgarnas förtroende för vår verksamhet även när vi tar hjälp av AI-system för att utföra den. Det är inte alltid möjligt att förklara varför en modell har gett ett visst resultat eller beslut (och vilken kombination av inputfaktorer som bidrog till det). Det kallas för svarta lådan-problematiken och behöver uppmärksammas särskilt. I sådana fall kan det behövas andra åtgärder för att skapa förklarbarhet. I vilken grad förklarbarhet är nödvändig beror i stor utsträckning på sammanhanget och på hur svåra konsekvenserna blir om resultatet är felaktigt eller missvisande på annat sätt.

Balans mellan principerna

Det kan uppstå spänningar mellan de här olika principerna. Det finns ingen färdig lösning för hur det ska hanteras men det är viktigt att inte ignorera sådana spänningar utan att uppmärksamma och hantera dem. Var särskilt uppmärksam på situationer som rör mer sårbara grupper, till exempel barn, personer med funktionsnedsättning och andra som historiskt sett har varit missgynnade eller riskerar utanförskap, och på situationer som utmärks av bristande balans i fråga om inflytande eller information, till exempel mellan arbetsgivare och arbetstagare eller mellan myndighet och enskilda.

Samtidigt som det finns risker med att använda AI kan det finnas stora fördelar med det. De senaste årens snabba utveckling visar att det kan finnas stora effektiviseringsvinster med att använda AI i offentlig sektor. Det är därför viktigt att du har en nyfiken men försiktig inställning när du testar nya system och funktioner.

Ytterligare information

- Här är en länk till Etiska riktlinjer för tillförlitlig AI, Europeiska kommissionens expertgrupp på hög nivå för AI-frågor: [Ethics-guidelines-AI_SV.pdf](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du prata med din chef.

2. Innan du använder ett AI-system behöver du vara säker på att systemet uppfyller säkerhetskraven för den typ av information som du ska ge det tillgång till

Tänk på att

- Du aldrig kan ångra eller ta bort en inmatning i ett AI-system.
- Alltid utgå från att den information som du matar in eller ger ett AI-system tillgång till kan sparas av leverantören och att den kan delas med andra utomstående utan att du har någon kontroll över det.
- Innan du använder ett AI-system, behöver du ta reda på om en riskbedömning av systemet är gjord av systemägaren och att den information du tänker använda är klassificerad av din verksamhet för att veta hur systemet får användas
- Du får inte ladda ner ett AI-system till din dator eller mobila enhet innan det har riskbedömts.
- Om AI-systemet inte kräver nedladdning kan du använda information med konsekvensnivå 0 i konfidentialitet i det innan systemet har riskbedömts.

Fördjupning

Informationssäkerhet kan kortfattat förklaras som arbetet med att förhindra att information läcker ut, förvanskas eller förstörs och att rätt information ska finnas tillgänglig för rätt personer i rätt tid.

Om du använder AI-system i ditt arbete är det viktigt att vara medveten om att AI-användning innebär särskilda risker för informationssäkerheten. Det kan bland annat handla om dataläckage genom att information överförs till externa tjänster och kan komma att hamna i orätta händer. Det finns även en risk för att en övertro på systemet och en olämplig användning kan leda till att tekniken används på ett sätt som är olämpligt för informationssäkerheten. Du ansvarar för att du använder AI-system på rätt sätt och för den information som du lägger in i ett AI-system eller på annat sätt ger systemet tillgång till.

En av de stora utmaningarna med att använda AI idag, särskilt när det gäller publika AI-system, är att det ofta är svårt för oss som enskilda användare att veta hur den information som vi vill mata in i ett AI-system kommer att hanteras av systemet. Innan du har försäkrat dig om att ett AI-system är säkert för den information du vill använda för att utföra en arbetsuppgift behöver du utgå från att all information som du matar in i ett AI-system sparas av systemet och att informationen kan komma att delas med andra utomstående utan att du har någon kontroll över det.

Alla verksamheter inom Göteborgs Stad ska bedriva ett systematiskt och långsiktigt informationssäkerhetsarbete. Det görs genom att klassificera information, genomföra en riskanalys för att därefter vidta lämpliga organisatoriska, personrelaterade, fysiska och tekniska säkerhetsåtgärder. De olika stegen beskrivs i Göteborgs Stads riktlinje för informationssäkerhet.

Enligt riktlinjen för informationssäkerhet är det systemägaren som ansvarar för att säkerställa att **systemet är riskbedömt**. Det är nämnden eller styrelsen som ansvarar för att **klassificera informationen** som hanteras i verksamheten. Du som användare ansvarar för att hantera stadens information utifrån hur den har klassificerats.

Genom att stadens verksamheter kontrollerar de informationssäkerhetsrisker som kommer med användningen av ett AI-system och överväger hur dessa kan hanteras kan vi minska risken för incidenter och negativa konsekvenser för våra verksamheter.

Eftersom användningen av AI-system kan medföra nya informationssäkerhetsrisker och även kan bidra till att förstärka befintliga brister är det viktigt att alla AI-system riskbedöms innan de börjar användas. Även under användning behöver riskerna följas upp och om en användning anses vara allt för riskfylld ska den upphöra. Du behöver därför kontakta din informationssäkerhetssamordnare innan du börjar använda ett AI-system för att kontrollera om en riskbedömning av systemet har genomförts och hur det får användas.

Om ett AI-system inte har riskbedömts får endast information med konsekvensnivå 0 i konfidentialitet användas i systemet. Det är sådan information som, om den tillgängliggörs, röjs eller sprids till obehöriga, inte medför någon negativ skada för Göteborgs Stads verksamhet, tillgångar, annan organisation eller individer.²

Om du behöver ladda ned AI-systemet till din dator eller mobila enhet för att kunna använda det måste en riskbedömning alltid göras först. Innan det har gjorts kan du inte veta vilken information systemet får tillgång till på din dator eller mobila enhet.

Ytterligare information

- Här hittar du länk till [Göteborgs Stads riktlinje för informationssäkerhet](#), stadens [sida för informationssäkerhet](#) och [Göteborgs Stads regel för IT-användare](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel informationssäkerhetssamordnare.

3. Det finns särskilda krav på AI-användning som innebär höga risker och en del AI-system är helt förbjudna

Tänk på att

- Alltid överväga om din tänkta användning av ett AI-system kan utgöra en högrisk-användning.

² Göteborgs Stads riktlinje för informationssäkerhet, Bilaga 1 Göteborg Stads klassificeringsmodell

- Innan en högrisk-användning av ett AI-system behöver verksamheten ta ställning till användningen och utreda vilka åtgärder som behöver vidtas.

Fördjupning

Förbjuden AI

En del AI-system är enligt AI-förordningen helt förbjudna inom EU. Det beror på att de anses utgöra ett tydligt hot mot människors säkerhet, försörjningsmöjligheter och rättigheter. Sådana AI-system får överhuvudtaget inte användas i stadens verksamheter. Det gäller AI-system för:³

- Kognitiv beteendemanipulation av personer eller specifika sårbara grupper (till exempel röstaktiverade leksaker som kan uppmuntra till farligt beteende hos barn)
- Social poängsättning och klassificering av människor baserat på beteende, socioekonomisk status eller personliga egenskaper
- Biometrisk identifiering och kategorisering av fysiska personer
- Biometrisk identifieringssystem i realtid och på distans (till exempel ansiktsgenkänning i offentliga utrymmen)

Högriskområden

Utöver dessa finns det en del användning av AI-system som bedöms kunna utgöra allvarliga risker för hälsa, säkerhet eller grundläggande rättigheter. Sådan användning kategoriseras som högrisk-AI. Reglerna om högrisk-AI i AI-förordningen träder i kraft först den 2 augusti 2026, men det är viktigt att vara medveten om dem redan nu. Det gäller AI-system som används inom följande områden:⁴

- Produkter som faller under EU:s produktsäkerhetslagstiftning (till exempel leksaker, flyg, bilar, medicintekniska produkter och hissar)
- Biometrisk identifiering och kategorisering (till exempel för att läsa av personers känslotillstånd)
- Kritisk infrastruktur (till exempel försörjning av vatten, gas, värme, el eller digitala kommunikationsnät)
- Utbildning och yrkesutbildning (till exempel poängsättning av prov)
- Anställning (till exempel vid rekrytering eller urval av individer, för att annonsera lediga jobb, granska ansökningar, utvärdera kandidater i intervjuer eller tester, eller för att övervaka eller utvärdera anställdas prestationer eller beteende i arbetsrelaterade förhållanden)
- Viktiga privata och offentliga tjänster (till exempel för att bedöma rätt till ekonomiskt bistånd eller hälso- och sjukvård)
- Brottsbekämpning (till exempel vid profilering av individer eller för att förutsäga, förebygga, upptäcka eller utreda brott)
- Migration, asyl och gränskontroll (till exempel automatiserad prövning av viseringsansökningar eller för att verifiera äktheten av resehandlingar)

³ EU:s AI-förordning Artikel 5

⁴ EU:s AI-förordning Artikel 6 och Bilaga I och III

- Rättskipning och demokratiska processer (till exempel AI-lösningar för att påverka resultatet av ett val eller enskildas deltagande i dessa)

Om en användning av ett AI-system bedöms utgöra en högrisk-användning så finns det vissa lagkrav som ställs på användare av systemet för att säkerställa att högrisk-AI-systemet används på ett säkert och etiskt ansvarsfullt sätt, med respekt för grundläggande rättigheter.

Bland annat behöver du som användare ha nödvändig kompetens och utbildning för att kunna använda ett högrisk-AI-system. Det är viktigt för att kunna ta ansvar för att du använder det i enlighet med det avsedda ändamålet och enligt leverantörens instruktioner.

Det är varje verksamhets ansvar att på förhand ta ställning till om en viss användning kan komma att bedömas som en högrisk-användning och att följa rättsutvecklingen på området. En användning av högrisk-AI måste ha föregåtts av en noggrann analys och ställningstagande av verksamheten om vilka åtgärder som behöver vidtas innan systemet kan börja användas.

Som medarbetare behöver du vara medveten om att det finns särskilda krav på AI-användning som innebär höga risker och att en del AI-system är helt förbjudna. Om du överväger att använda dig av AI i tjänsten för något som skulle kunna omfattas av de ovan uppräknade användningsområdena är det därför viktigt att uppmärksamma frågan så att din verksamhet kan ta ställning till om det är en högrisk-användning.

Ytterligare information

- Här hittar du AI-förordningen i fulltext: [Regulation - EU - 2024/1689 - EN - EUR-Lex](#)
- IMY, Digg och SKR har tagit fram viss vägledning gällande förordningen:
 - [Frågor och svar om AI-förordningen | IMY](#)
 - [AI-förordningen | Digg](#)
 - [AI Act, AI-förordningen | SKR](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel en förvaltnings- eller bolagsjurist.

4. Du ansvarar för att sekretessuppgifter och skyddsvärd information inte röjs för obehöriga

Tänk på att

- Alltid utgå från att den information som du matar in i AI-systemet när du promptar lämnar myndigheten och överförs till leverantören av AI-systemet.
- Sekretessbelagda uppgifter inte får matas in i ett AI-system. Det enda undantaget är när systemet har riskbedömts för att vara säkert för sådan hantering.

- Utöver sekretessbelagda uppgifter kan även andra uppgifter vara känsliga.
- Se helhetsbilden av den information som du hanterar, eftersom den kan vara mer känslig när den sammanställs med en stor mängd annan information.

Fördjupning

Bestämmelserna i offentlighets- och sekretesslagen innebär att det finns sekretessgränser runt en myndighet. Varje kommunal nämnd med tillhörande förvaltning är att se som en egen myndighet i det här sammanhanget. Ett bolag är inte en myndighet men stadens bolag ska jämföras med myndigheter när man tillämpar offentlighets- och sekretesslagen.⁵

Sekretessuppgifter får inte lämnas ut till någon utanför myndighetens sekretessgräns utan lagstöd. Detta gäller oavsett på vilket sätt uppgifterna lämnas ut och även om det sker när du använder ett AI-system.

Sekretessuppgifter

Det finns många olika slags uppgifter som omfattas av sekretess i stadens verksamheter. Det kan till exempel gälla uppgifter om enskilda individer, om medarbetare eller om verksamheten. Ofta är det självklart om de uppgifter vi hanterar omfattas av sekretess eller inte men ibland kan det vara svårare att bedöma.

Sekretessen gäller för myndigheter och bland annat för oss som arbetar i myndighetens verksamhet. Det är viktigt att vara medveten om att det kan vara en brottslig handling att med uppsåt eller av grov oaktsamhet röja en sekretessbelagd uppgift när det inte finns stöd i lag (brott mot tystnadsplikten).⁶

För att säkerställa att sekretesskyddet upprätthålls när du använder AI måste du dels veta om uppgifterna som du vill använda i ett AI-system omfattas av sekretess, dels om användningen innebär att uppgifter riskerar att röjas för obehöriga. Det här är viktigt för att du ska kunna veta vilka arbetsuppgifter som ett AI-system kan användas för.

De sekretessuppgifter som finns i kommunen klassificeras som konsekvensnivå 2 eller 3 i konfidentialitet,⁷ vilket innebär att systemet måste vara riskbedömt för att sekretessuppgifter på ett säkert sätt ska få användas i det. Det innebär att det kan medföra betydande negativ påverkan eller allvarliga konsekvenser för Göteborgs Stads verksamhet, tillgångar, annan organisation eller individer om informationen tillgängliggörs, röjs eller sprids till obehöriga.⁸

Annan skyddsvärd information

Utöver uppgifter som omfattas av sekretess är det också bra att tänka på att det kan finnas information som bör skyddas för att den är känslig av andra skäl och som därför inte bör spridas obehindrat. Olika sorters intern information kan göra kommunen sårbar för

⁵ 2 kap. 3 § offentlighets- och sekretesslagen

⁶ 20 kap. 3 § brottsbalken

⁷ [SKRs stödmaterial för KLASSA, Information som faller under OSL](#)

⁸ Göteborgs Stads riktlinje för informationssäkerhet, Bilaga 1 Göteborg Stads klassificeringsmodell

angrepp (digitala eller fysiska) eller ge en felaktig bild av kommunen och dess verksamhet om den sprids.

Om du använder information i ditt arbete som inte omfattas av sekretess men som inte har klassificerats ännu, men som du misstänker att det finns skäl att vara försiktig med när det gäller just användning av AI-system så är rekommendationen att ta upp frågan för diskussion i din verksamhet.

Exempel: *Ekonomiska brottslingar kan använda information om ekonomi för att planera angrepp. Information om fysisk och digital infrastruktur kan användas för att rikta attacker eller hitta sårbarheter. Listor med adresser på kommunala boenden skulle kunna användas för att rikta fysiska attacker. Listor med datorprogram kan användas för att hitta sårbarheter som kan utnyttjas vid cyberangrepp. Kartor kan innehålla information som kan användas för att planera fysiska angrepp.*

Sammanställningar av information kan vara skyddsvärda

Ytterligare en aspekt är att vi med hjälp av AI kan göra sammanställningar utifrån stora mängder information. Uppgifter som var för sig inte klassificeras som känslig kan komma att vara det om de sammanställs. Om du använder information i ditt arbete som du misstänker att det finns skäl att vara försiktig med när det gäller just användning av AI-system så är rekommendationen att ta upp frågan för diskussion i din verksamhet.

Exempel: *Ett större antal kartor kan sammantaget vara känslig information även om kartorna var för sig inte är det.*

Ytterligare information

- Hos Digg och IMY kan du ta del av deras riktlinjer och vad de skriver om sekretess: [Säkerställ sekretess vid användning av generativ AI | Digg](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel informationssäkerhetssamordnare eller en förvaltnings- eller bolagsjurist.

5. Du får bara behandla personuppgifter om det sker i enlighet med dataskyddslagstiftningen

Tänk på att

- Personuppgifter är alla slags uppgifter som direkt eller indirekt kan kopplas till en levande fysisk person. Det är därför många olika slags uppgifter som anses vara personuppgifter och då blir också dataskyddslagstiftningen tillämplig.
- När du behandlar personuppgifter i ett AI-system så behöver du veta vem som är personuppgiftsansvarig för behandlingen och vilken rättslig grund som är tillämplig. Ju mer integritetskänsliga personuppgifter som du hanterar, desto

tydligare behöver det rättsliga stödet för behandlingen vara och desto viktigare är det att personuppgifterna hanteras på ett säkert sätt.

- Du får bara behandla personuppgifter för de ändamål som de samlades in för och du får inte behandla fler personuppgifter än vad som behövs för ändamålen
- Håll dig uppdaterad om de riskbedömningar din organisation gör avseende personuppgiftsbehandlingar där AI används för behandlingen.

Fördjupning

När du använder personuppgifter i ditt arbete måste du se till så att det sker på ett sätt som är förenligt med dataskyddslagstiftningen, det vill säga EU:s dataskyddsförordning (GDPR) och kompletterande nationell lagstiftning.⁹ Lagstiftningen kring hur personuppgifter får hanteras är komplex och fokus i detta avsnitt är de viktigaste delarna att ha med sig när man använder AI.

Det är viktigt att förstå att personuppgifter är alla slags uppgifter som direkt eller indirekt kan kopplas till en levande fysisk person. Även indirekta uppgifter, bilder och data som i kombination med varandra kan användas för att identifiera en person anses utgöra personuppgifter.

Känsliga och extra skyddsvärda personuppgifter är uppgifter som kräver en särskild hantering, oavsett om du vill använda uppgifterna i ett AI-system eller på annat sätt. Känsliga personuppgifter är uppgifter som avslöjar etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse, medlemskap i fackförening, genetiska uppgifter, biometriska uppgifter, uppgifter om hälsa eller uppgifter om en persons sexualliv eller sexuella läggning. Det är förbjudet att behandla känsliga personuppgifter förutom när det finns ett tillämpligt undantag i lagstiftningen. Exempel på extra skyddsvärda personuppgifter är personnummer, uppgifter om barn, uppgifter om lagöverträdelse, löneuppgifter, värderande uppgifter och uppgifter om sociala förhållanden.

Alla olika slags åtgärder som görs med personuppgifter är en behandling av personuppgifter. Exempel på personuppgiftsbehandlingar är när personuppgifter matas in i ett AI-system och när AI-systemet instrueras på ett sätt som får det att använda sig av personuppgifter som finns inlärd sedan tidigare.

När personuppgifter behandlas är det dessutom viktigt att du vet vem som ansvarar för behandlingen, om det är den organisation där du arbetar¹⁰ eller någon annan aktör. Finns inte denna information tillgänglig för dig är det lämpligt att fråga till exempel närmsta chef eller organisationens dataskyddskontakt. Det är den verksamhet som ansvarar för personuppgiftsbehandlingen som är personuppgiftsansvarig och ska se till att behandlingen lever upp till kraven i dataskyddslagstiftningen.

För att en personuppgiftsbehandling ska vara tillåten ska det finnas en rättslig grund och annat rättsligt stöd för behandlingen. Ju större risker det är med att behandla

⁹ Observera att förutom den svenska dataskyddslagen finns det även nationell speciallagstiftning, så kallad registerlagstiftning, som reglerar hur personuppgifter får behandlas inom specifika sektorer inom samhället. Registerlagstiftning finns bland annat för hälso- och sjukvården och socialtjänsten.

¹⁰ Inom Göteborgs Stad är det nämnder och bolag som är personuppgiftsansvariga, det vill säga ytterst ansvariga för att personuppgifter behandlas i enlighet med dataskyddsförordningen.

personuppgifter i ett AI-system desto högre krav ställs det på den rättsliga grunden och det rättsliga stödet för behandlingen. Du ska därför enbart behandla personuppgifter för behandlingar där din organisation fastställt den rättsliga grunden.

Något av det viktigaste du som medarbetare inom Göteborgs Stad därutöver kan göra när du behandlar personuppgifter i arbetet är att alltid se till så att de grundläggande dataskyddsrättsliga principerna för dataskydd är uppfyllda, både innan du börjar använda dig av personuppgifter och under tiden du använder dem. Det gäller även när du vill använda dig av någon form av AI.

De grundläggande dataskyddsrättsliga principerna innebär bland annat att personuppgifter bara får behandlas för specifika, särskilt angivna och berättigade ändamål.

Personuppgifter som redan har behandlats för ett visst ändamål får inte vidarebehandlas för nya ändamål som är oförenliga med de ursprungliga ändamålen. Du får inte heller behandla fler personuppgifter än vad som är nödvändigt och som behövs för ändamålen. Du ska sedan alltid se till så att det är riktiga uppgifter som behandlas och att uppgifterna inte heller behandlas längre än vad som behövs. Av de grundläggande principerna framgår också att personuppgifter behöver skyddas på ett lämpligt sätt genom olika säkerhetsåtgärder och att den personuppgiftsansvariga ska kunna visa att verksamheten lever upp till dataskyddsförordningen och hur det görs.

Det är viktigt att vara medveten om att det finns många utmaningar och svårigheter med att efterleva de grundläggande principerna i dataskyddsförordningen när personuppgifter behandlas i samband med användningen av AI. Eftersom ett AI-system ofta kräver mycket stora mängder data är det exempelvis svårt att säkerställa att personuppgifter i ett AI-system enbart behandlas för det ändamål för vilket de ursprungligen samlades in och att behandlingen begränsas till vad som är nödvändigt och inte blir för omfattande. Då det är svårt att få inblick i och förklara hur en AI-modell fungerar är det även svårt för den personuppgiftsansvariga att uppfylla de krav som finns på att ge information om behandlingen till de individer vars personuppgifter behandlas. Det innebär ett sämre skydd och risker för enskildas rättigheter.

Alla personuppgifter som behandlas måste skyddas så att inte obehöriga får tillgång till dem eller att de förstörs till exempel genom en olyckshändelse eller genom en cyberattack. För att förhindra sådana risker ska den personuppgiftsansvariga vidta lämpliga säkerhetsåtgärder. Exempel på säkerhetsåtgärder är kryptering och säkerhetskopiering men vilka åtgärder som behöver vidtas beror på risken med behandlingen. För att bedöma riskerna behöver verksamheten, den personuppgiftsansvariga, genomföra en konsekvensbedömning enligt dataskyddsförordningen om det är sannolikt att behandlingen medför höga risker för enskilda. Tekniken för AI-system utvecklas snabbt och du som medarbetare behöver därför hålla dig uppdaterad om de riskbedömningar din organisation gör avseende personuppgiftsbehandlingar där AI används för behandlingen.

Ytterligare information

- För ytterligare information om behandling av personuppgifter och dataskyddslagstiftningen se [Integritetsskyddsmyndighetens \(IMY\) webbsida, imy.se](https://www.imy.se)

- [IMY har en Vägledning om GDPR och AI | IMY](#) och information specifikt om de grundläggande dataskyddsrättsliga principerna hittar du här [AI och grundläggande principer – Vägledning om GDPR och AI | IMY](#)
- IMY har tagit fram riktlinjer för hanteringen av personuppgifter vid användningen av generativ AI: [GDPR vid användning av generativ AI | IMY](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel dataskyddskontakt. Det finns också möjlighet att vända sig till dataskyddsenheten på Intraservice som är Göteborgs Stads dataskyddsombud.

6. När du använder ett AI-system är du alltid ansvarig för resultatet och hur det används

Tänk på att

- Även om texten som ett AI-system har skapat verkar trovärdig är det inte samma sak som att den är faktamässigt korrekt. Du ska alltid behandla de resultat generativ AI skapar som ett utkast som kräver ytterligare bearbetning.
- Det är viktigt att du kontrollerar resultaten när det gäller till exempel beräkningar, analyser av siffror och sammanställning av statistik.
- Det är viktigt att du kontrollerar så att alla AI-genererade texter som du använder har en professionell ton och är sakliga och opartiska.

Fördjupning

AI är ett digitalt arbetsverktyg vars resultat bygger på sannolikheter och mönster i data, vilket innebär att de kan innehålla fel. Du ansvarar själv för hur du använder AI-genererat innehåll och får aldrig överlåta din yrkesmässiga bedömning till ett AI-system. Resultat måste alltid granskas noga.

AI-genererade texter, beräkningar och bilder kan framstå som trovärdiga, men ska ses som utkast som ofta kräver bearbetning. Kritisk granskning och faktakontroll är avgörande – att något låter korrekt innebär inte att det är det.

Forskning visar att vi lätt överskattar AI:s pålitlighet, ett fenomen kallat automation bias. Vi tenderar att se AI som objektivt och vetenskapligt, vilket ökar risken att vi okritiskt accepterar dess resultat som sanningar.¹¹

När du använder en AI-genererad text måste du förstå innehållet och kunna förklara det. Använd helst AI inom områden där du har kunskap, så att du lättare kan upptäcka fel. Om

¹¹ Se Diskrimineringsombudsmannens rapport AI och risker för diskriminering i arbetslivet, s. 10 och 56

du är osäker på begrepp eller formuleringar – ta bort dem eller rådgör med en kollega som är mer insatt.

Kontrollera alltid AI-genererade beräkningar, analyser och diagram. Säkerställ att data presenteras korrekt och inte ger en missvisande bild. AI:s tolkning av data kan vara felaktig, särskilt om källdatan inte speglas på rätt sätt.

Var också uppmärksam på att AI kan producera innehåll som är partiskt, osakligt, olämpligt eller diskriminerande. Som anställda i Göteborgs Stad har vi ett ansvar att kommunicera professionellt, opartiskt och sakligt – AI-genererade texter måste granskas med det i åtanke.

Slutligen kan det vara svårt att förstå hur AI-system kommer fram till sina resultat, den så kallade ”svarta lådan”-problematiken. Brist på transparens kan innebära både säkerhetsrisker och risk för felaktigheter – något du som användare behöver ta hänsyn till.

Krav på AI-kunnighet vid användning av AI som beslutsstöd

Kravet på AI-kunnighet i AI-förordningen innebär bland annat att medarbetare som använder AI-system som beslutsstöd i sitt dagliga arbete behöver ha den kunskap och kompetens som krävs för att utvärdera och ifrågasätta resultaten som produceras av AI-systemet.

Om ett AI-system används som stöd för att ta fram beslutsunderlag finns det risk för såväl direkt som indirekt diskriminering. Risker för diskriminering kan uppstå på grund av bristande kvalitet i data, bristande träning och bristande transparens. Det är därför viktigt att de som använder AI-system har gedigna och relevanta kunskaper för att kontrollera resultaten som produceras av AI-systemet, så att grundläggande rättigheter skyddas.

Ytterligare information

- För ytterligare information om vad du behöver tänka på när AI-system används som stöd för att ta fram beslutsunderlag kan du hitta mer stöd här: [Beslut med stöd av generativ AI bör ha mänsklig kontroll | Digg](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel specialister som kan hjälpa dig faktagranska texter som handlar om ämnen du inte själv är bekant med.

7. Du ansvarar för att ditt AI-genererade material används på ett ansvarsfullt och transparent sätt som bidrar till förtroendet för stadens verksamhet

Tänk på att

- Du inte ska använda AI för sakens skull. Det är alltid viktigt att du överväger om det är lämpligt utifrån vad saken gäller och dess sammanhang att använda AI-genererat material eller om det kan minska förtroendet för staden.
- Det är viktigt att du alltid anger tydligt om en bild, film eller ljudinspelning är helt eller delvis AI-genererad. Detsamma gäller även helt AI-genererad text som inte har bearbetats av en människa i efterhand.

Fördjupning

Grunden för en god förvaltning är att kommunen är saklig och opartisk i sin verksamhet.¹² Det är viktigt att medborgarna kan lita på att den information som kommunen kommunicerar är saklig och sanningsenlig.

Det är även viktigt att AI-genererat material används i enlighet med stadens riktlinje för kommunikation och stadens grafiska profil. Stadens information och kommunikation ska vara faktabaserad och grundad på fattade beslut så att trovärdigheten för stadens olika verksamheter och för staden som helhet stärks. Av det följer även att vi inte ska använda metoder eller kanaler för information och kommunikation som kan innebära att förtroendet för staden på något sätt minskas. Det är inte heller säkert att AI-genererat material är representativt för medborgarna i staden, eller beaktar normkritiska perspektiv på det sätt som garanterar att vår kommunikation är tillgänglig för våra målgrupper utifrån deras olika behov och förutsättningar.

Exempel: Du ber ett AI-system att skapa en bild av Järntorget. Om den AI-genererade bilden går att misstas för ett fotografi kan en betraktare luras att tro att det är en korrekt representation av verkligheten. Om betraktaren sedan upptäcker att bilden inte stämmer, för att till exempel en butik som syns på bilden inte finns i verkligheten, kan tilltron till information från Göteborgs Stad skadas.

Det är viktigt att du överväger om det är lämpligt utifrån vad saken gäller och dess sammanhang att använda AI-genererat material överhuvudtaget eller om det kan minska förtroendet för staden.

AI-genererade bilder, filmer och ljudinspelningar ska användas med särskild försiktighet och på ett sätt som gör att det inte finns någon risk för att de felaktigt uppfattas som verklighetsbaserade.

Det är även viktigt att du alltid anger tydligt om en bild, film eller ljudinspelning är såväl helt som delvis AI-genererad. Detsamma gäller även helt AI-genererad text som inte har bearbetats av en människa i efterhand. Detta görs lämpligen när verkets källa anges, till

¹² 5 § Förvaltningslagen

exempel "Bild genererad av AI" eller liknande. Om det går att lägga in metadata i verket ska det även framgå där att den är AI-genererad. Lägg till metadata i programmet som bilden, filmen eller ljudinspelningen tas fram i eller tagga den som AI-genererad om du lagrar den i Göteborgs Stads mediabank. Det ska även framgå av eventuell alternativtext till en bild, film eller ljudinspelning att den är AI-genererad. Det behövs för att informationen ska kunna syntolkas i digitala gränssnitt.

Ytterligare information

- Det finns ytterligare information om hur du ska [arbeta med AI-genererade bilder i den grafiska profilen](#).

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel kommunikatörer och förvaltnings- eller bolagsjurister.

8. Information som du lägger in i systemet och resultatet som du får ut kan bli allmänna handlingar

Tänk på att

- Rättsläget är oklart gällande vilka allmänna handlingar som skapas vid användning av AI-system. Det är mer sannolikt att resultaten är allmänna handlingar än att promptarna är det.
- Det är möjligt att en domstol bedömer att både promptar och resultat är allmänna handlingar. Utgå därför från att båda dessa kan behöva lämnas ut vid en begäran.
- Verksamheten behöver på förhand ta ställning till om det uppstår allmänna handlingar vid medarbetarnas AI-användning och att se till att sådana handlingar hanteras på ett korrekt sätt så att både offentlighetsprincipen och arkivbestämmelserna följs.
- En ny, självständig bedömning alltid måste göras vid varje begäran från allmänheten om ett utlämnande.

Fördjupning

En fråga som uppkommer vid användning av AI-system är om och i vilka situationer det uppstår allmänna handlingar. Bestämmelserna om allmänna handlingar finns i tryckfrihetsförordningen. Varje kommunal nämnd med tillhörande förvaltning är att se som en egen myndighet i det här sammanhanget. Ett bolag är inte en myndighet men stadens bolag ska jämföras med myndigheter när man tillämpar tryckfrihetsförordningen.

Det är ett ansvar för verksamheterna att på förhand ta ställning till om det uppstår allmänna handlingar vid verksamhetens AI-användning, att dessa hanteras på ett korrekt sätt och att det finns rutiner för till exempel diarieföring och gallring. Som medarbetare bör du vara medveten om att även ett AI-system kan generera allmänna handlingar och att en begäran om att ta del av allmänna handlingar ska hanteras i en särskild ordning.

Allmänna handlingar

Enligt tryckfrihetsförordningen är alla handlingar som är antingen inkomna eller upprättade samt förvarade hos en myndighet allmänna. En handling är **inkommen** när den har kommit till myndigheten. Handlingar som kommer in genom olika digitala verktyg som vi använder i tjänsten, till exempel e-postmeddelanden eller sms, anses ha kommit in till myndigheten när handlingen är tillgänglig med tekniskt hjälpmedel, det vill säga när informationen är läsbar med hjälp av det digitala verktyget.

Som huvudregel är en handling **upprättad** när den har expedierats, det vill säga skickats utanför myndighetens gränser. Slutligen ska handlingen också vara **förvarad** hos myndigheten för att vara en allmän handling där. Som förvarad räknas även handlingar som går att nå via vanliga digitala hjälpmedel (exempelvis externa servrar och/eller molntjänster) så handlingen behöver inte fysiskt finnas hos myndigheten.

Det kan vara komplicerat att ta ställning till på vilket sätt det skapas allmänna handlingar när vi använder olika AI-system, inte minst eftersom lagstiftningen inte är anpassad utifrån den tekniska utvecklingen. Nedan beskrivs några olika exempel på handlingar som kan skapas vid användning, utveckling och träning av ett AI-system. Tänk på att verksamheten alltid ska göra en ny självständig bedömning om någon av de nedan beskrivna handlingarna begärs ut.

Handlingar som kan skapas vid användning, utveckling och träning av ett AI-system

I Riktlinjer för generativ AI inom offentlig förvaltning gör Digg och IMY bedömningen att **prompten** och den information som matas in i ett AI-system som bygger på generativ AI sannolikt kommer kunna betraktas som någon form av utkast och därför inte nödvändigtvis ska anses vara en allmän handling. I riktlinjerna beskrivs att samma överväganden, om det skapas allmänna handlingar eller inte, behöver göras av det **resultat** som ett AI-system producerar och att resultatet förmodligen i högre grad kommer att betraktas som en allmän handling.

Rättsläget är emellertid oklart och det är möjligt att en domstol bedömer att både promptar och resultat är allmänna handlingar. Utgå därför från att båda dessa kan behöva lämnas ut vid en begäran.

Frågor om i vilka situationer det uppstår allmänna handlingar kanske kommer bli ännu mer relevanta vid utveckling eller träning av AI-system. Exempelvis behöver det bedömas om **tränings- och valideringsdata** blir nya allmänna handlingar eller om de utgör kopior av tidigare allmänna handlingar. Ett AI-systems **källkod** kan i sig vara en allmän handling. Förändringar av källkod kan innebära att det skapas nya allmänna handlingar och att det anses ske en gallring av tidigare version. Vid användningen av ett

AI-system kan det också komma att skapas olika slags **loggar** som kan komma att ses som allmänna handlingar.

Begäran om att ta del av allmän handling

Huvudregeln är att allmänna handlingar är offentliga, vilket innebär att de ska lämnas ut på begäran om det inte finns hinder mot det (sekretess). Det gäller även handlingar som skapas vid användning av olika digitala verktyg. Om du får en begäran om att ta del av uppgifter eller handlingar från ett AI-system som du använder är det viktigt att hantera den skyndsamt. Vid en begäran behöver verksamheten ta ställning till om det som begärts ut är en allmän handling och om det finns uppgifter som omfattas av sekretess.

Ytterligare information

- Det finns ytterligare information om allmänna handlingar och hur du hanterar en begäran om att ta del av allmänna handlingar i [Göteborgs Stads rutin för hantering av begäran om utlämnande av allmänna handlingar](#)
- Hos Digg eller IMY kan du ta del av deras riktlinjer och vad de skriver om allmänna handlingar: [Säkerställ sekretess vid användning av generativ AI | Digg](#)

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar eller bolag experter som kan hjälpa till, till exempel registrator, arkivarie eller annan tjänsteperson som ansvarar för diariet i din verksamhet eller en förvaltnings- eller bolagsjurist.

9. Du ansvarar för att upphovsrätten respekteras när du använder ett AI-system

Tänk på att

- Låta ett AI-system bearbeta ett verk kan strida mot upphovsrätten i sig.
- Aldrig instruera ett AI-system att bryta mot upphovsrätten.
- Du genom att mata in ett upphovsrättsligt skyddat verk i ett AI-system kan göra så att verket tillgängliggörs för allmänheten.
- Vara försiktig med att sprida eller offentliggöra AI-genererade bilder, texter, ljud, musik, datorprogram med mera.
- Kontrollera så att ett AI-genererat verk inte inkräktar på någon annans upphovsrätt, genom att exempelvis söka efter verket på internet eller låta kollegor bedöma det.

Fördjupning

Upphovsrätten är det regelverk som styr vem som får göra vad med ett litterärt eller konstnärligt verk. Med **verk** avses exempelvis bilder, texter, ljud, musik, datorprogram

med mera. Dessa omfattas oftast av upphovsrätt och då är utgångspunkten att det är den som har skapat ett verk som har ensamrätt att bestämma hur det ska användas, vilket innebär att andra inte får använda det hur som helst. Den som bryter mot upphovsrätten kan bli skadeståndsskyldig för upphovsrättsintrång.

Utöver upphovsrätten som skyddar konstnärliga och litterära verk finns det även skydd för uppfinningar, varumärken och design-/mönster. Reglerna för dessa skiljer sig lite åt men det som står om skydd för verk i det här avsnittet gäller på det stora hela även för dessa.

Även om rättsläget ännu är ganska osäkert, och det saknas vägledning för vad det innebär för användarna, är det med stor sannolikhet så att redan träningen av AI-system på stora mängder data har innefattat upphovsrättsintrång.

Eftersom i princip all användning av verk kan vara i strid med upphovsrätten kan reglerna aktualiseras genom att ge ett AI-system tillgång till ett verk genom exempelvis promptning. De kan även aktualiseras genom de instruktioner som en användare ger AI-systemet genom att prompta. Slutligen kan reglerna aktualiseras genom det som skapas av AI-systemet. Vid ett upphovsrättsintrång kan den som utsatts välja vem den ska rikta krav mot; den som har utvecklat eller äger systemet, den som använder systemet för att skapa eller den som använder resultatet.

Ge AI-system tillgång till ett verk

Om du matar in ett verk i en AI-system innebär det framför allt två saker. Verket kan komma att omarbetas och verket kan komma att delas av systemet eller dess leverantör med andra. Båda dessa situationer kan innebära ett brott mot upphovsrätten. Det kan också innebära att den som har skapat det ursprungliga verket har upphovsrätt till den omarbetade version som AI-systemet skapar.

Exempel: *Du ber ett AI-system att omarbeta en bild som den laddar upp i AI-systemet. Genom att ge AI-systemet tillgång till bilden skapas en kopia av den vilket sannolikt strider mot upphovsrätten.*

Det är därför som utgångspunkt inte tillåtet att mata in andras bilder, texter, ljud, musik, datorprogram med mera i ett AI-system. Det enda undantaget är om verket inte omfattas av upphovsrätt eller om du vet att du har tillåtelse både att omarbeta det och att sprida det till andra. Om du är osäker så utgå från att alla verk omfattas av upphovsrätt.

Instruera AI-system på ett sätt som bryter mot upphovsrätten

Om du ger ett AI-system instruktioner som innebär att AI-systemet tar del av eller omarbetar upphovsrättsligt skyddade verk, kan själva instruktionerna innebära ett brott mot upphovsrätten. Detta har ännu inte prövats i domstol men skulle kunna vara en möjlig följd av lagstiftningen.

Exempel: *Du ber ett AI-system att skapa en bild av en tecknad anka som har en ljusblå sjömansblus och blå studentmössa. Om bilden som skapas är väldigt lik en annan känd anka som är upphovsrättsskyddad, skulle den kunna ha skapats i strid med upphovsrätten.*

Det är därför inte tillåtet att avsiktligt ge ett AI-system instruktioner att använda sig av, bearbeta eller försöka att efterlikna andras upphovsrättsskyddade verk.

Använda AI-genererade verk

Om ett AI-system genererar ett verk finns det alltid en risk för att det verket är väldigt likt ett befintligt upphovsrättsligt skyddat verk. Det kan bero på att AI-systemet, som har tränats på material från internet, använder sig av ett upphovsrättsligt skyddat verk för att generera sitt resultat. Om du sprider ett AI-genererat verk som är för likt någons upphovsrättsskyddade verk kan du begå ett upphovsrättsbrott.

Det kan vara bra att känna till att det redan idag finns privata aktörer som är specialiserade på att hitta upphovsrättsintrång på nätet och driva ersättningsanspråk för dessa. Den typen av tjänster kan komma att bli vanligare i framtiden i och med användningen av AI.

Innan du offentliggör eller sprider ett AI-genererat verk är det viktigt att du är försiktig och försöker kontrollera så att det inte är för likt någon annans verk. Ett sätt att undersöka detta är att ta delar av en AI-genererad text och söka i en sökmotor. Gäller det en bild finns sökmotorer för bildsökning som letar efter liknande bilder på nätet. En annan försiktighetsåtgärd är att låta några kollegor hjälpa till att bedöma om de ser någon risk med att sprida materialet.

Ytterligare information

- Läs mer om [upphovsrätt och AI på Digg:s hemsida](#).
- På sidan har Digg även en [lista med ett antal frågor som är bra att ställa sig vid användning av verk som bilder, texter, ljud, musik, datorprogram med mera](#).

Behöver du stöd?

- Om du har frågor eller behöver stöd ska du i första hand prata med din chef. I många fall har förvaltningar och bolag experter som kan hjälpa till, till exempel kommunikatörer och förvaltnings- eller bolagsjurister.

Förkortningar och begrepp

Artificiell intelligens (AI): Förmågan hos datorprogram och robotar att efterlikna (mänsklig) intelligens.

AI-modell: AI-modeller är väsentliga komponenter i AI-system men utgör inte AI-system i sig. En AI-modell är ett program som har tränats på data för att till exempel kunna identifiera mönster, generera text eller fatta beslut utan mänsklig intervention.

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.

Bias: Om informationen som AI-modellen tränats på är partisk eller innehåller fördomar kan AI ge oriktiga och orättvisa svar. Det kan uppstå om vissa grupper av människor är underrepresenterade eller felaktig information förekommer i träningsdatan. I dessa fall kan modellen fatta beslut eller ge svar som upprepar och överdriver stereotyper.

Digg: Myndigheten för digital förvaltning

Extra skyddsvärda personuppgifter: Uppgifter som på grund av sin karaktär är särskilt skyddsvärda. Det kan till exempel vara personnummer och samordningsnummer, uppgifter om barn, uppgifter om lagöverträdelser, uppgifter om ekonomiska förhållanden, värderande uppgifter och annan information som ligger nära den privata sfären.

Generativ AI: AI-system som kan skapa nytt material baserat på befintliga data och mönster utifrån enkla instruktioner. 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. Exempel är Copilot i Edge som kan skapa text utifrån dina instruktioner, Chat GPT och DeepSeek.

Hallucination: AI-modeller kan generera svar som verkar rimliga men som är felaktiga.

IMY: Integritetsskyddsmyndigheten

Indata: Data som matas in eller hämtas direkt av ett AI-system och som är grunden för den data som matas ut/produceras (utdata) av systemet.

Källkod: Den läsbara versionen av ett datorprogram, skriven i ett programmeringsspråk.

Känsliga personuppgifter: Det är uppgifter som avslöjar etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse, medlemskap i fackförening, genetiska uppgifter, biometrisk uppgifter, uppgifter om hälsa eller uppgifter om en persons sexualliv eller sexuella läggning.

Loggar: Information i textfiler om händelser när ett program körs på en dator eller server.

Publika AI-system: AI-system som är tillgängliga för vem som helst att använda till exempel via webbsida eller via en app, till exempel Chat GPT och Dall-E från OpenAI. Publikt eller allmänt tillgängliga AI-system är inte inskaffade, skapade eller införda specifikt för kommunens behov.

Personuppgift: Alla uppgifter som direkt eller indirekt kan knytas till en identifierad eller identifierbar fysisk levande person.

Personuppgiftsansvarig: Bestämmer ändamålen och medlen för behandlingen av personuppgifter, det vill säga hur och varför en behandling av personuppgifter ska ske.

SKR: Sveriges kommuner och regioner

Svarta lådan-problematiken: Problemet att det inte går att förklara alla detaljer av hur en AI-modell producerar sina resultat då de ofta består av mycket komplexa AI-modeller. Detta kan leda till säkerhetsrisker, när användaren av AI-systemet inte fullt ut har förståelse för hur AI:n arbetar. På engelska kallas detta för the Black Box problem.

Tränings- och valideringsdata: Träningsdata är data som används för att träna ett AI-system och valideringsdata data som används för att tillhandahålla en utvärdering av det tränade AI-systemet.