**Anskaffelse av KI-system eller IT-system med KI-funksjonalitet**

**Generelle krav:**

Det skal stilles krav til at produsenten fremlegger dokumentasjon slik at KI-systemets ytelse, tillitsverdighet og sikkerhet kan vurderes grundig før kontraktsinngåelsen. Den kommende KI-forordningen stiller krav til KI-systemer med begrenset og høy risiko, som det skal tas høyde for i kravsettingen, i tillegg til andre regulatoriske krav.

KI-systemer forbruker betydelig mengder energi, både til lagring og bruk av store datamengder til trening og bruk av KI-modeller. Alle KI-systemer bør bidra til å oppfylle FNs bærekraftsmål.

**Funksjonelle krav:**

Datakvalitet er noe man spesielt må vurdere i en KI-anskaffelse, da dataene som brukes til å trene KI-modeller påvirker systemets ytelse, pålitelighet og upartiskhet. Det bør også stilles krav om at leverandøren kan redegjøre for hvilke data de har brukt i utviklingen av KI-systemet, både i trening og valideringen, og hvilket behandlingsgrunnlag etter personvernforordningen.

**Tekniske krav:**

KI-systemet skal passe i virksomhetens infrastruktur. Det stilles derfor krav til beskrivelser av dataflyt og dataformater som skal håndteres og utveksles mellom andre systemer og KI-systemet.

**Krav til sikkerhet:**

De tiltakene som er beskrevet i Normen og er viktige for å sikre digitale systemer generelt, er også viktige i beskyttelsen av KI-systemer. KI kan innebære særlig sikkerhetsrisikoer, som det er viktig at anskaffelsen ivaretar. [ENISA peker på følgende trusler i KI-livsløpet[[1]](#footnote-1),](https://www.enisa.europa.eu/publications/multilayer-framework-for-good-cybersecurity-practices-for-ai) som anskaffelsen skal ivareta:

* **Unnvikelse**(eng: evasion): KI-systemer kan ofte lures på nye måter, eksempel ved at angriper benytter spesielle inngangsverdier (eng: *adversarial examples, prompt injections*).
* **Forgiftning** (eng: *poisoning*): Angriper påvirker treningsdata eller KI-modellen slik at oppførselen til systemet blir påvirket i en bestemt retning.
* **Lekket informasjon om treningsdata eller KI-modell**: En angriper får hentet ut informasjon om KI-modellen, slik som konfigurasjon, parametre og treningsdata
* **Kompromitterte komponenter i KI-systemet**: Angriper kompromitterer en komponent, for eksempel ved å utnytte sårbarheter i åpne kilder som KI-systemet benytter.
* **KI-systemet feiler**: Hele KI-systemet feiler, for eksempel ved en feil eller ved at en angriper klarer å ta ned KI-systemet.

1. European Union Agency for Network and Information Security, ENISA, er EUs ekspertisesenter for internettsikkerhet i Europa. [↑](#footnote-ref-1)