Kommunal- og moderniseringsdepartementet

postmottak@kmd.no

Oslo, 22. juni 2021

Innspill til EU-kommisjonens nye forslag til regulering av kunstig intelligens

Det vises til EU-kommisjonens nye forslag til regulering av kunstig intelligens (KI) som ble lagt frem 21. april 2021, og departementets pågående arbeid med å samle inn innspill. Nedenfor følger Microsoft Norges skriftlige innspill.

Microsoft er en internasjonal IT-leverandør som er verdens ledende innen programvare og skyløsninger. I Norge er det rundt 600 ansatte fordelt på Microsoft Norge og Microsoft Development Center (tidligere Fast Search & Transfer) med utviklingsmiljøer lokalisert i Oslo, Trondheim (ved NTNU) og Tromsø (ved UiT). Vår virksomhet i Norge er sterkt knyttet til vårt partnernettverk bestående av utviklere og IT-leverandører over hele landet. Totalt har vi ca. 1.700 partnere i Norge. Vår visjon er "Empower every person and every organization on the planet to achieve more". I november 2019 åpnet Microsoft to datasentre i Norge – i Oslo og Stavanger.

Innledning

Microsoft setter pris på muligheten til å komme med innspill til dette nye forslaget til regulering av kunstig intelligens. I den forbindelse ønsker vi først og fremst å uttrykke vår støtte for EUs arbeid med å legge til rette for at utviklingen og bruken av kunstig intelligens (KI) i Europa respekterer viktige europeiske verdier. Det er en kjent problemstilling at selv om KI åpner opp for enorme muligheter for mennesker og samfunn i Europa, byr denne teknologien også på nye utfordringer.

Ved bruk av kunstig intelligens skaper mennesker hver dag nye innovative produkter og tjenester, gjør eksisterende tilbud bedre og sikrere, bidrar til å løse store samfunnsutfordringer, og rett og slett gjør livene våre bedre. Ved å gi datamaskiner evnen til å utføre oppgaver som tidligere kun ble utført av mennesker – slik som å innhente innsikt fra data, dra sluttsatser, og til og med gjennom å se, høre og å forstå – øker kunstig intelligens potensialet for hva vi kan oppnå, og fører til bedre resultater og bedre liv. Den nye hverdagen som kom med koronapandemien og de mange innovative løsningene på å løse utfordringene knyttet til dette, viste oss enda flere av de positive sidene og mulighetene med KI.

1. KI må være en del av løsningen – ikke problemet

Selv om kunstig intelligens kan være en løsning på mange av dagens utfordringer, må vi sikre at det ikke er en del av problemet. For å forsikre oss om at teknologien ikke fører til økte forskjeller og andre uønskede konsekvenser er vi nødt til å evaluere hvordan vi benytter KI og hva som må gjøres kollektivt for å sikre at KI utvikles og tas i bruk på en ansvarlig måte, slik at det viderefører våre felles samfunnsverdier og mål.

Det å sikre tillit til KI – at den er trygg, etisk og ansvarlig – krever samtidig et kulturelt skifte, både blant de som utvikler og leverer KI-teknologier, og de som tar den i bruk i hverdagen til ulike formål. Microsoft mener at virksomheter som utvikler teknologi også har et ansvar for å sikre dens ansvarlighet og levedyktighet. Av denne grunn ser vi derfor på tillitsvekkende KI mer som en reise enn en destinasjon. Dette er et viktig tema for oss. Vi er opptatt av å samarbeide med ulike aktører for å sikre at det å utvikle tillitsvekkende KI blir normen i tiden fremover, og ikke unntaket. Hovedutfordringen i dette arbeidet blir å løse utfordringene på en måte som ikke setter en stopper for bruk og utvikling av KI med det enorme potensialet KI faktisk har.

På bakgrunn av dette følger nedenfor konkrete innspill til EU-kommisjonens foreslåtte nye regulering av KI.

2. Viktigheten av klart språk og virkeområde

Slik teksten er utformet bærer nåværende språk i stor grad preg av tvetydighet, herunder at viktige begreper mangler klar definisjon og forståelse. I tillegg anses språket for å ikke være heldig i denne type lovregulerende tekst ved at det er for generelt og vidtrekkende. Dette vil kunne skape mye forvirring og usikkerhet blant nøkkelaktører i markedet. Eksempelvis skaper tvetydigheten tvil rundt betydningen av begreper slik som «placing on the market» og «making available on the market». Det er også usikkerhet rundt betydningen av visse vilkår, slik som «robustness». Ved å benytte denne type språk legger forslaget opp til urimelig vidtrekkende plikter, som ikke bare er vanskelig å etterleve, men også uklare.

Språkbruken er også problematisk fra et juridisk språklig perspektiv. Slik det geografiske virkeområdet er definert er det vanskelig for tilbydere å vite om deres systemer regnes for å ha innvirkning på individer som bor innenfor EU-området. Når det gjelder det materielle virkeområdet har kommisjonen forklart at selv om en stor variasjon av KI-baserte systemer faller innunder virkeområdet for forslaget, må definisjonen alltid ses i sammenheng med de fire opplistede risikokategoriene. Selv om dette til en viss grad er en rimelig og nyttig forklaring, og som i noe utstrekning reduserer bekymringen for den vide rekkevidden, må denne presiseringen inkluderes i selve lovforslaget.

3. Krav til overordnede systemer

I artikkel 9 innføres krav om et risikostyringgssystem, hvor skillet mellom dette og kvalitetssikringssystemet i artikkel 17 ikke er helt klart. Fra vårt ståsted oppstår det også spørsmål om hvilken relevans og nytteverdi det har å ha begge systemer.

Artikkel 10, som retter søkelys på data og data forvaltning (engelsk: «data governance») , innfører plikten om «error-free data sets» (norsk: «feilfrie datasett») som er umulig å etterleve i praksis. I tillegg kan hva som utgjør en «error» variere fra applikasjon til applikasjon. Artikkel 10 innfører også en plikt til å ta i betraktning «specific geographical, behavioral, and functional settings», som skaper utfordringer på flere måter. Blant annet fordi «data» er et komplekst tema som ikke bare er en representasjon basert på mennesker (geografi, språk, representasjon av mangfold), men også på andre type skiller slik som holdninger og miljø.

Kravene til rapportering, protokollføring og menneskelig kontroll er i tillegg for vidtrekkende, og skaper unødvendig stor byrde å etterleve.

4. Rettferdige algoritmer

Forslaget fokuserer på en forenklet tilnærming til rettferdige algoritmer. Kommisjonen gir uttrykk for at dersom alle kravene etterleves, vil det ikke oppstå diskriminering som følge av algoritmene eller andre utfordringer knyttet til rettferdighet. Etter vår erfaring gjenspeiler ikke dette realiteten.

Hva som er rettferdig er i seg selv en konkret vurdering som kan gi ulike resultat. I tillegg bidrar ikke kravet om at datasett må være «error-free» til å løse utfordringen med at underliggende og dype strukturelle skjevheter kan påvirke algoritmer og skape utfordringer knyttet til skjevhet og diskriminering.

5. Forholdet mellom sikkerhet og grunnleggende menneskerettigheter

Delvis på grunn av den formalistiske tilnærmingen, treffer ikke forslaget riktig når det gjelder faktisk og tilstrekkelig beskyttelse av fundamentale menneskerettigheter. Forslaget stiller opp en rekke krav som høy-risiko KI-systemer må etterleve, mens tilfellene der problemer og brudd oppstår selv om alle kravene er etterlevd ikke er behandlet i det hele tatt.

Det er heller ikke tatt med rettigheter og krav som er tilgjengelige for individer der disse har blitt innskrenket eller påvirket på annen måte. Forslaget henviser i liten grad til grunnleggende menneskerettigheter, noe som i stedet er henvist til i det forklarende notatet (Explanatory Memorandum) og i fortalepunktene.

Forslaget kombinerer to konsepter som er fundamentalt atskillelige: KI-systemer utgjør høy risiko i forbindelse med sikkerhet og helse, og høy risiko for beskyttelsen av fundamentale rettigheter. Det er grunn til å stille spørsmål ved om disse kravene og overvåkningen av markedet som er laget for sikring av produktsikkerhet, faktisk vil resultere i beskyttelse av disse rettighetene. Av natur er risikoene knyttet til design- og utviklingsvalg i KI-systemer avhengig av den konkrete bruken av disse systemene.

6. Relasjonen mellom tilbydere og brukere

Kommisjonen har forklart at hovedvekten ligger på det tiltenkte formålet bak et KI-system og på dens brukerveiledning i vurderingen av om en bruker benytter systemet på en måte som går lenger enn det tilbyder har tenkt.

Dette er til en viss grad nyttig for å klargjøre forholdet mellom tilbydere og brukere, men dynamikken mellom de to bør tydeliggjøres i enda større grad. Mer konkret ville det vært hensiktsmessig om kontraktuelle plikter nevnes uttrykkelig som gyldig tiltak som kan benyttes for å skape klare grenser i denne dynamikken.

7. Egenevaluering som norm er en styrke

Forslaget viser til egenevaluering som en hovedregel, med unntak kun for visse begrensede tilfeller. Dette anses som en riktig og hensiktsmessig tilnærming for å sikre at innovasjon ikke hindres unødig og muliggjør at SMB'er kan engasjere seg i markedet. Selv om noen har vært kritiske til denne tilnærmingen og har foreslått tredjepartsevaluering som norm og et utgangspunkt, er dette en tilnærming som kan gjøre mer skade enn nytte, og bør unngås.

8. Nytt område som krever nytenkning

Et annet aspekt ved forslaget som er viktig å reflektere over, spesielt sett i lys av håndhevingsregimet, er at dette er et nytt område. Det finnes ingen etablerte normer, harmoniserte regelsett eller felles hoveddimensjoner for å sikre rettferdighet, pålitelighet og trygghet.

Dette er kun ett av argumentene for at evalueringene av etterlevelse er problematiske. Femårssertifiseringen er uproporsjonalt lang tid med tanke på at hver enkelt vesentlig («substantial») modifisering krever en ny evaluering. Det er urealistisk å tro at et KI-system kan eksistere i fem år uten å gjennomgå vesentlige endringer, uavhengig av hvordan ordet «substantial» defineres.

Et annet poeng er at forslaget om å benytte CE-merking for software ikke treffer. Dette er en standard laget for hardware og det risikerer å skape forvirring – ikke minst med tanke på utfordringen som oppstår i forbindelse med det og rent praktisk feste CE-merket til software generelt og særlig KI-systemer.

9. Risiko for fragmentert tilnærming, uproporsjonale gebyrer og viktigheten av åpenhet

Etter artikkel 65 kan et medlemsland kreve at en "relevant operator" trekker tilbake sitt system fra markedet eller tilbakekaller dette, selv om systemet etterlever lovgivning, dersom det utgjør en risiko for personers helse eller sikkerhet, eller for lovgivning med formal å beskytte fundamentale menneskerettigheter. Denne tilnærmingen øker sannsynligheten for at det blir en fragmentert tilnærming, samtidig som den skaper rettslig usikkerhet.

I artikkel 71 nr. 3 er det gitt ramme på gebyr på 6% av verdensomspennende årlig omsetning for å ikke etterleve artikkel 10. Artikkel 10 i seg selv oppstiller vilkår som kan være subjektive og/eller vanskelig å evaluere, slik som krav om at opplæring, validering, og testing må være relevant, representativ, feilfri og fullstendig. Med tanke på utfordringen med denne vurderingen, oppfattes dette som at det vil kunne føre til uproporsjonale gebyrer.

Åpenhet rundt bruk av følelsesgjenkjennings-/biometriske kategoriseringssystemer ser ut til å være et viktig første steg og det bør diskuteres om det bør følge ytterligere plikter. Mye av debatten rundt gjenkjenning av følelser er fokusert rundt hva KI kan og ikke kan gjøre. Vi mener det er like viktig å være tydelig på muligheter og begrensninger i teknologien, slik at brukere er informerte når de tar egne valg.

10. Teknologinæringens rolle

Etter vår oppfatning har EU-kommisjonen gjort en god jobb med å introdusere og inkludere mekanismer for oppdatering av KI-systemer, for å sikre fremtidig tilpasningsdyktighet av disse og at sentrale europeiske verdier ivaretas. Samtidig tror vi det er hensiktsmessig om teknologiaktører og ulike beslutningstakere jobber enda tettere sammen i arbeidet videre med forslaget og regulering av kunstig intelligens.

Med vennlig hilsen,

Kristine Beitland,

Direktør for Myndighetskontakt, Microsoft Norge