Kunstig intelligens:

Kan Europa både regulere og innovere?

EUs AI Act, vedtatt våren 2024, representerer det første store forsøket på å etablere et omfattende regulatorisk rammeverk for AI, med mål om å beskytte grunnleggende rettigheter og samfunnsinteresser. Regulering av teknologi er aldri enkelt, spesielt ikke når utviklingen går så raskt som den gjør nå på AI-fronten. Metodikk og erfaring fra revisjon kan spille en viktig rolle i operasjonaliseringen av EUs kunstig intelligens-forordning.

Ph.D.
Lars Erlend Leganger
Direktør
PwC

Introduksjon

På midten av 2000-tallet skapte en kombinasjon av faktorer[1] en bølge av optimisme og bruk av kunstig intelligens («AI» på folkemunne – fra engelsk «artificial intelligence»). Mer utstrakt og sofistikert bruk av AI har ført til mer oppmerksomhet rundt – og bekymring for – risikoer ved utvikling og bruk av AI i beslutningsstøtte og automatisering. EUs AI Act er det første store forsøket på å etablere et omfattende regulatorisk rammeverk for kunstig intelligens. Forordningen skal primært beskytte forbrukere og samfunnsinteresser, men den kan – avhengig av hvordan den blir operasjonalisert – få store konsekvenser for næringslivet, og kanskje særlig i bransjer som er leverandører av tillit. Denne artikkelen beskriver EUs AI Act og potensielle konsekvenser av forordningens ikrafttredelse, med særlig fokus på regnskap og revisjon.

EUs AI Act skal beskytte samfunnet og innbyggernes grunnleggende rettigheter

EUs AI Act ble foreslått av Europakommisjonen i april 2021 og endelig vedtatt våren 2024. Forordningen skal sikre at utvikling og bruk av AI i EU skjer på en måte som respekterer innbyggernes grunnleggende rettigheter, og ikke medfører unødvendig samfunnsrisiko. Reglene er bransjenøytrale, og gjelder for alle som tilbyr, produserer, bruker, importerer eller distribuerer AI-systemer i EU, samt tilbydere og brukere av AI-systemer som er lokalisert i et tredjeland, men hvor resultater fra systemet benyttes i EU.

AI Act har i utgangspunktet en risikobasert tilnærming til regulering, hvor AI-anvendelser deles i fire risikonivåer: Uakseptabel risiko, høyrisiko, begrenset risiko, og minimal risiko. Kravene for de forskjellige risikonivåene fases gradvis inn, fra «uakseptabel risiko»- forbud fra og med februar 2025, til alle krav for alle risikonivåer er på plass fra august 2027.

«**Uakseptabel risiko**»-forbudene gjelder forholdsvis få og særdeles inngripende anvendelser, som statlig «social scoring» av innbyggere, og prediksjon av individers risiko for å begå fremtidige kriminelle handlinger.

«Høyrisiko»-kategorien treffer spesielt bransjer som ligger nær individ- og samfunnsinteresser, som helse, utdanning, essensiell infrastruktur, og enkelte finansielle tjenester. Men også personalledelse/HR anses som et «høyrisiko»-område: Selskap med digitalt fremoverlente personalavdelinger som jobber datadrevet og tar nye verktøy i bruk raskt, bør sette seg godt inn i AI Acts kapittel 3: AI for prioritering/screening av jobbsøknader, ressursplanlegging basert på individuelle ferdigheter, og leveranse-monitorering er eksempler på systemer som fra 2027 må svare ut en smørbrødliste «høyrisiko»-krav som skal sikre transparens, sikkerhet, og plassering av ansvar ved bruk.

Etter gjennombruddet til OpenAI og ChatGPT i 2023 kom det også inn krav til leverandører av «**general purpose AI**» (GenAI) i AI Act. Leverandører av slike løsninger får relativt omfattende tekniske dokumentasjonskrav, og leverandører av GenAI-løsninger med «systemisk risiko» må videre dokumentere at modellene er robuste mot manipulering, cyberangrep, og de må ha prosesser på plass for løpende å vurdere og mitigere risiko og for å håndtere og rapportere uønskede hendelser.

[1] I korte trekk: Billigere regnekraft: Bedre og billigere grafikkprosessorer (GPUer) gjorde det mulig å bygge komplekse AI-modeller mye raskere enn tidligere. Større datasett: Fremveksten av internett og digitale tjenester førte til en eksplosjon i mengden tilgjengelige data for å bygge AI-modeller. Bedre algoritmer: Nye maskinlæringsalgoritmer og teknikker gjorde det mulig å løse problemer som tidligere var utenfor rekkevidde. Praktiske anvendelsesområder: Gjennombrudd på et bredt spekter praktiske anvendelsesområder (talegjenkjenning, bildegjenkjenning, språkforståelse, mm.) har regelmessig fornyet/forsterket interessen for – og investeringsviljen i – kunstig intelligens. Enklere implementering: Nye sky- og API-baserte fagsystemer gjorde det enklere å integrere skreddersydde AI-løsninger for spesifikke oppgaver.

Innlemmingen av GenAI-krav i AI Act kom sent i forordningens tilblivelse, og har et visst preg av hastverksarbeid: Her avviker AI Act fra den risikobaserte systematikken, de GenAI-spesifikke kravene er de samme uansett anvendelsesområdets risikokategori.

AI-tilbydere og -brukere som ikke treffes av «uakseptabel risiko», «høyrisiko», eller «GenAI»-kategoriene blir kun i liten grad påvirket av AI Act. Det kommer et gjennomsiktighetskrav om at det alltid skal være åpenbart for deg som individ at noe er AI-generert: En AI-drevet kundesupport-chatbot får ikke utgi seg for å være et levende menneske, og AI-generert bilde og lyd skal merkes som sådan, og alle GenAI-tilbydere blir pålagt noen utvidede rapporterings- og dokumentasjonskrav.

Finansiell rapportering, bærekraftsrapportering – og AI-rapportering?

I revisjon og regnskap benyttes AI i dag dels til å automatisere rutineoppgaver – automatisk datainnsamling/bearbeiding, klassifisering av transaksjoner, og bilagsbehandling – og dels som beslutningsstøtte – generere estimater, gjøre benchmarking, og anomalideteksjon[2]. I utgangspunktet er dette anvendelser som ligger utenfor AI Acts «uakseptabel»- og «høyrisiko»-definisjoner. Regnskap og revisjon er tungt regulerte bransjer, og de allerede eksisterende lover og regler for regnskap og revisjon vil normalt være første skranke for hvor, og hvordan en tar AI-systemer i bruk i disse verdikjedene, lenge før AI Act eventuelt kommer på banen.

På den andre siden vil dagens og fremtidige versjoner av AI Act drastisk utvide hvilke krav som stilles til revisors kompetanse. Bruk av AI kan gjøre finansiell rapportering mer effektiv og presis, men kan også åpne for nye feilkilder og angrepsvektorer. AI-genererte vurderinger og analyser kan på overflaten og for den uerfarne fremstå som svært troverdige selv der det underliggende datagrunnlaget er mangelfullt, eller der AI-verktøy er feilaktig brukt utenfor sine egentlig tiltenkte anvendelsesområder.

Det er også interessant hvordan AI Acts nåværende og fremtidige krav vil operasjonaliseres. AI Acts definisjon av hvilke systemer som anses som «AI» (og som dermed er in-scope for forordningen), gir et visst tolkningsrom, der mange automatiserings- og beslutningsstøtte-løsninger kan tolkes inn eller ut av definisjonen, avhengig av hvilket resultat fortolkeren ønsker seg[3]. På samme måte kan AI-tilbydere forsøke seg på produkttilpasninger for å komme inn under unntakene for høyrisikokrav i AI Act artikkel 6 pkt. 3 – som fritar fra høyrisiko-kravene dersom AI-systemet kun utfører en «narrow procedural task», «preparatory task to an assessment», o.l. – og slik potensielt oppnå etterlevelse (på papiret) billigere enn ved å svare ut alle «high risk»-kravene, se figur 1 nedenfor.

I AI Act er det nasjonale «competent authorities» og «market surveillance authorities» som får jobben med å sikre at etterlevelse skjer i tråd med forordningens formål. Avhengig av hvordan disse velger å løse oppdraget, kan det fort oppstå et marked for uavhengige tredjeparter som bekrefter at selskapers klassifiseringer, risikovurderinger, rapportering og tiltak er gjennomført i tråd med AI Act og andre fremtidige AI-krav.

Behovet for å kunne stole på selskapers AI-rapportering har mange fellestrekk med behovet for å stole på selskapers finansielle rapportering og bærekraftsrapportering. De konkrete, materielle, faglige vurderingene som skal gjøres, vil være annerledes for AI, men generell revisjonsmetodikk, og ikke minst forståelsen for hvordan selskapets AI-regulatoriske etterlevelse henger sammen med selskapets øvrige regulatoriske plikter, vil fortsatt være høyst relevant. Aktører i revisjonsbransjen har et verdifullt forsprang dersom de ønsker å gå inn i et fremtidig AI-bekreftelser-marked, og bør med stor interesse følge med på hvordan AI Act operasjonaliseres.

EU regulerer mens USA innoverer?

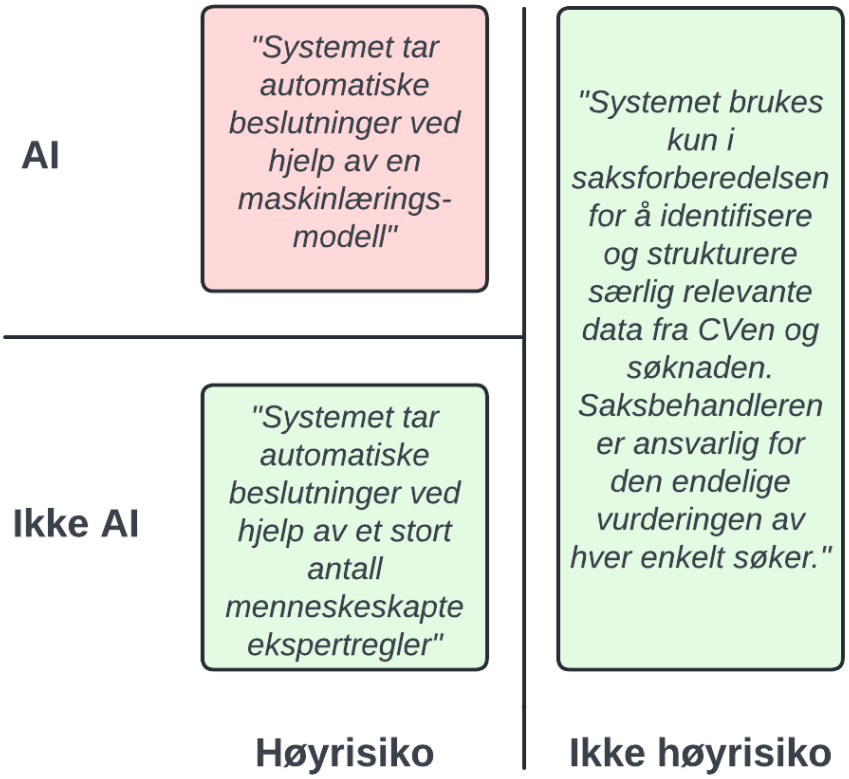
For regnskaps- og revisjonsbransjens egen bruk av AI vil etterlevelse av AI Act altså neppe medføre store utfordringer, og for sistnevnte kanskje også nye inntektskilder. Men den enes inntekt er den andres utgift: I oktober i 2023, før den endelige AI Act teksten var landet, meldte Cédric O fra den franske GenAI-utfordreren Mistral at «EU's AI act could kill our company»[4], og i en studie utført av Copenhagen Economics for CCIA Europe fremheves usikkerhet og frykt for økende regulatorisk byrde som et signifikant hinder for europeisk innovasjon innen GenAI[5].

2 Andrusko & Amble gir en grei og (relativt) fersk drøfting av mulighetsrommet for GenAI i regnskap i Death, Taxes, and AI: How Generative AI Will Change Accounting: https://a16z.com/generative-ai-in-accounting/.

3 Det er nok en tendens til at systemer som omtales som cutting-edge AI i salgsmateriale, brått blir "vanlige" IT-løsninger i møte med AI-spesifikke regulatoriske krav.

4 https://sifted.eu/articles/eu-ai-act-kill-mistral-cedric-o

5 https://copenhageneconomics.com/wp-content/uploads/2024/03/Copenhagen-Economics-Generative-Artificial-Intelligence-The-Competitive-Landscape.pdf



Figur 1: Et (hypotetisk) AI-system for å vurdere jobbsøkere basert på data i CV og søknad kan med relativt små justeringer havne innenfor (rød) eller utenfor (grønn) definisjonen av høyrisiko bruk av AI. Kostnaden ved å etterleve AI Acts høyrisiko-krav kan gi opphav til kreative "anti-patterns" – Hva om man f.eks. formelt deler et automatisk-vurdere-jobbsøkere-system opp i to separate løsninger – en maskinlæringdrevet saksforberedende AI-men-ikke-høyrisiko-løsning som henter ut og strukturerer beslutningsrelevante data, og en ekspertregelbasert høyrisiko-men-ikke-AI-løsning som fatter automatiske beslutninger basert på de strukturerte dataene?

Etablerte aktører som Microsoft, Amazon, og Meta/Facebook, med sine allerede enorme compliance-funksjoner, kan etterleve AI Acts nye krav uten å blunke. For små utfordrere som Mistral – i oktober var de 20 ansatte – kan regulatoriske krav innebære tap av momentum i en kritisk etableringsfase. Da kan det være et fristende alternativ å starte opp i – eller samarbeide med noen i – en mer innovasjonsvennlig jurisdiksjon, og heller bevege seg til det europeiske markedet først når selskapet er stort nok til å bære den regulatoriske ekstrabyrden.

Frankrikes president Macron oppsummerte tankene til mange i det europeiske GenAI-miljøet da AI Act ble vedtatt: «We can decide to regulate much faster and much stronger than our major competitors. But we will regulate things that we will no longer produce or invent. This is never a good idea»[6]. På AI-feltet er det enn så lenge slik at EU leder an på regulering, mens USA leder an på innovasjon og verdiskapning. I Mistrals tilfelle endte de med å inngå i et strategisk partnerskap med Microsoft – til enkelte EU-byråkraters store frustrasjon, som mente Mistral burde vist takknemlighet for at EU tross alt hadde forsøkt å hensynta innovasjon i AI Act, og ikke «flagget ut» like etterpå[7]. Revisjonsbransjen kan spille en viktig rolle i operasjonaliseringen av AI Act ved proaktivt å dele relevante innspill og erfaringer som kan bidra til at en ikke bare sikrer etterlevelsen av de regulatoriske kravene, men også fremmer AI-innovasjon og -konkurranseevne i Europa.

6 https://www.euronews.com/next/2023/12/15/potentially-disastrous-for-innovation-tech-sector-says-eu-ai-act-goes-too-far
7 https://www.euronews.com/next/2024/02/27/furious-critics-question-microsofts-deal-with-mistral-ai-as-eu-set-to-look-into-it

Dagens «akseptabel risiko» kan være morgendagens «høy risiko»

Historisk har de store byksene i regulatoriske krav ofte kommet etter kriser – krakket på Wall Street i 1929, Enron-skandalen i 2001, og finanskrisen i 2008 har på hver sin måte ansporet tettere oppfølging av finansiell rapportering. Den stadig pågående klimakrisen har drevet frem krav om bærekraftsrapportering. Selv om mange ser på AI-utviklingen med stor fremtidsfrykt, har vi ikke hatt noen virkelig store kriser forårsaket av bruk av kunstig intelligens – ennå. AI Act er designet for å være dynamisk og tilpasningsdyktig: Dagen den første store AI-skandalen smeller, kan de impliserte bransjene fort havne på «high risk»-listen – inkludert regnskap og revisjon.

Oppsummering

EUs AI Act tar sikte på å beskytte europeiske samfunnsinteresser og innbyggernes grunnleggende rettigheter mot (mis)bruk av kunstig intelligens. Forordningen har en risikobasert tilnærming, der anvendelser med uakseptabel og høy risiko blir hhv. forbudt og strengt regulert. Tilbydere av GenAI-løsninger blir pålagt krav til rapportering og tiltak som skal redusere systemisk risiko ved bred anvendelse. De fleste av dagens anvendelser i regnskap og revisjon faller utenfor de strengeste AI Act- reguleringene. Økende bruk av AI i utarbeiding av finansiell rapportering vil imidlertid kreve økt kompetanse og tilpasning fra revisjonsbransjen, som også kan finne nye inntektsmuligheter i å bekrefte selskapers AI-rapportering. Ikrafttredelsen av AI Act setter EU i førersetet på AI-regulering, men økt regulatorisk byrde kan fort gå ut over innovasjon og svekke europeisk konkurranseevne i forhold til USA. Revisjonsbransjen har en viktig rolle å spille i utformingen av den riktige balansen mellom regulert og innovativ AI-utvikling og -bruk i Europa.