Pilot AI-regulatorisk sandlåda

Linda Lindström, eSam

Kort om eSam

Syftet med samverkan: underlätta och påskynda medlemmarnas digitala transformation.

- Frivilligt samverkansprogram mellan 41 myndigheter.
- Samverkan för digitala lösningar som t.ex. gemensamma tjänster.
- Arbetar för medlemmarnas digitala transformation och effektivisering.
- Återbruka lösningar som redan utvecklats och utveckla gemensamma lösningar för att underlätta vardagen för individer och företag.
- Medlemmarna vill använda sina gemensamma resurser på ett ansvarsfullt och effektivt sätt.

Vad är en AI-regulatorisk sandlåda?

- Författningsreglerad i AI-förordningen.

- Kontrollerad miljö som erbjuder leverantörer eller potentiella leverantörer av AI-system en möjlighet att utveckla, träna, validera och testa innovativa AI-system under en begränsad tid innan de släpps ut på marknaden eller på annat sätt tas i bruk.

- Genomförandet ska ske enligt en särskild sandlådeplan som beskriver målen, villkoren, tidsplanen, metoden och kraven för den verksamhet som ska bedrivas i sandlådan.

- Medlemsstaterna ska säkerställa att deras behöriga myndigheter inrättar minst en AI-regulatorisk sandlåda på nationell nivå.

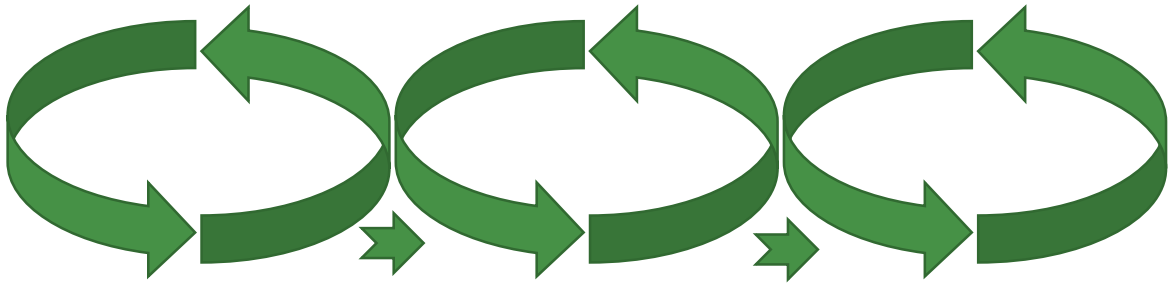
Mål med AI-regulatoriska sandlådor:
Främja innovation och öka rättssäkerheten.

Pilot AI-regulatorisk sandlåda

Initiativ inom eSam med Bolagsverket, Skatteverket och Arbetsförmedlingen tillsammans med Integritetsskyddsmyndigheten. Ekonomistyrningsverket deltog i andra iterationen.

Utgångspunkt :

I flera iterationer utvärdera ett eller flera AI-system utifrån de krav som förordningen ställer för att förstå kraven på AI-system och göra en analys av vad som krävs för att upprätta en AI-regulatorisk sandlåda inbegripet kompetenser, resurser, dokumentation och eventuellt teknisk infrastruktur.



Pilot AI-regulatorisk sandlåda forts

Förhoppningen är att arbetet kan bidra till kunskap om hur regulatoriska sandlådor för AI bör inrättas och fungera i Sverige och vilka förutsättningar som krävs för att aktörer ska kunna nyttja sådana sandlådor.

En pilot av en AI-regulatorisk sandlåda kan

- **möjliggöra** för att AI-regulatoriska sandlådor är användbara från AI-förordningens tillämpningsdatum
- **främja** svenska intressen i EU
- preventivt och gemensamt **utvärdera eller utveckla** ett AI-system i enligt med AI-förordningen och
- ge en fördjupad **kunskap** gällande AI-förordningen.

Iterationer

1. Breddad matchning - AI-system för att hjälpa arbetssökanden att se nya yrken inom ramen för befintlig kompetens och därigenom bredda deras arbetsmarknad

2. Remissammanställningar – AI-system för RK för sammanställning av inkomna remissvar, hitta relevanta stycken i remissvar och skapar en remissammanställning av dessa

3. Brottsutredning – stöd för handläggning av brottsutredningar, hjälpa till med vilken information som ska hämtas in, hitta avvikande mönster, tidsuppskatta, förhörsjämförelse, kontrollera uppgifter i realtid, analysera bevis och översätta

Genomförande

- En arbetsgrupp bestående av tvärfunktionell kompetens
- Sex halvdagsmöten, varav två fysiska möten.
- En styrgrupp

- Projektplan för planering, utformning av sandlådan, genomförande, utvärdering och kommunikation.
- Erfarenhetsinhämtning IMY:s erfarenheter
- Frågematris utifrån perspektiven: AI-förordningen, regulatorisk sandlåda, verksamhetsfrågor och nationella frågor

- Arbetsmetod: arbetsgruppen har fått en beskrivning av AI-systemet och utifrån beskrivningen diskuterat frågeställningar utifrån frågematrisen.

Några lärdomar

- Om AI-förordningen
- Om sandlådans innehåll
- Om arbetssättet

Lärdomar

Om AI-förordningen – definitionen

- Flera av rekvisiten är deskriptiva eller sådana som i princip föreligger som utgångspunkt. Dvs. det är sannolikt ganska vanligt att de uppfylls.
- Det kommer sannolikt vara rekvisitet om inferens som påverkar beslutet om det är ett AI-system eller inte.

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**.

Lärdomar

Om AI-förordningen

- **Många it-system kommer att utgöra ett AI-system.** Sannolikt kommer det mesta som inte är helt regelbaserad it att kunna omfattas av AI-förordningens definition av AI-system.

- Gränsdragningen av omfattningen av systemet, dvs. vad som ska ingå i systemet och vad som inte är en del av systemet, är komplicerad. Det **avsedda ändamålet** bör stå i centrum. Detta sätter ramen för systemet och gör att ett AI-system kan innehålla flera komponenter, modeller och data som behöver beaktas i bedömningen av systemets omfattning.

- För att göra bedömningen om vad som innefattas i ett AI-system krävs en **god helhetsbild** över it- och AI-arkitekturen.

- Även om det är komplicerat att i sig avgöra om det är ett AI-system är det sannolikt inte denna frågeställning som får mest betydelse. Frågan om vilken **risknivå** systemet tillhör och särskilt om systemet är ett **högrisksystem** bör vara mest central utifrån de konsekvenser som följer med kraven på högrisksystem.

Lärdomar

Om AI-förordningen

- Även om det vid system med minimal/ingen risk inte finns några krav på aktörerna enligt AI-förordningen, kan det för en myndighet ändå finnas anledning att överväga att ändå informera om AI-systemet, med då på **frivillig basis eller utifrån frivilliga uppförandekoder**.

- Det är inte alltid självklart vilken **roll** respektive aktör har enligt AI-förordningen, tanken verkar ändå vara att det ska finnas en utpekad leverantör. Sannolikt kan en aktör ha flera roller enligt AI-förordningen.

- Att det ingår en **AI-modell för allmänna ändamål** i AI-systemet föranleder **inte några ytterligare krav** på leverantören eller tillhandahållaren. Kraven baseras istället på AI-systemets risknivå.

Lärdomar

Sandlådans innehåll

- Det kan vara **svårt att avgöra på förhand** om ett projekt platsar i en AI-regulatorisk sandlåda. Vid ansökan till en AI-regulatorisk sandlåda kommer den första frågan att vara om det är ett AI-system. Sannolikt kommer dock tröskeln vara låg för vad som är ett AI-system.

- Syftet med de AI-regulatoriska sandlådorna är att det ska vara en slags **ventil** som underlättar att system snabbt kan komma ut på marknaden. Sandlådan bör **inte endast vara till för högrisksystem** utan även för system med lägre risk, så de kommer ut på marknaden snabbare.

- Beskrivningen av AI-systemet behöver vara tillräckligt definierad för att kunna bedömas. Eftersom det initialt kan vara svårt att veta var gränsen går för vad som utgör AI-systemet är det bättre att börja **brett i beskrivningen** och snäva in vartefter.

- I diskussionerna om AI-system kan det lätt ske en **sammanblandning** av bedömningarna. Det är därmed viktigt att tydliggöra att analysen av om systemet utgör ett AI-system alltid måste vara första steget. Därefter tas ställning till frågan om vilken risknivå AI-systemet tillhör.

Lärdomar

Sandlådans innehåll

- Det kan finnas **utmaningar** att **få fram all information** för en bedömning av systemet, bl.a. för att systemet är under utveckling och den dokumentation som kan behövas för beskrivning ännu inte finns tillgänglig.

- Det är endast **leverantörer och potentiella leverantörer** av AI-system som får **ansöka** till en AI-regulatorisk sandlåda. En ansökan kan dock ske i partnerskap med t.ex. tillhandahållare eller leverantör av en AI-modell för allmänna ändamål.

- **Harmonisering** av de AI-regulatoriska sandlådorna – det kan finnas en **risk för diskrepans** utifrån medlemsländernas kapacitet och fokus.

- **Samordnade sandlådor –** AI-sandlådan besvarar bara frågor enligt AI-förordningen. En aktör har behov av svar inom flera rättsområden. Sannolikt kommer det i så fall handla om olika sandlådor som **samordnas** snarare än en gemensam sandlåda.

Lärdomar

Om arbetssättet

- Det är viktigt att de som deltar besitter både kompetens, mandat och förmåga till dialog och samma reflektion görs i denna sandlåda. **Tvärfunktionellt arbete** i nära samarbete är helt avgörande för framgång.
- Det är viktigt med en **gemensam begreppsbild** då det annars finns risk för olika tolkningar av begreppens betydelse.
- En viktig framgångsfaktor är att arbetet bedrivs i en mindre gruppering med **aktivt deltagande** och att deltagarna är med från början till slut.
- **Tydliga förväntningar** på deltagarnas bidrag vid arbetsmötena underlättar och effektiviserar arbetet i sandlådan. Det kan också vara värdefullt med en **informerad ledning** som har insikt i förväntningarna på deltagandet i en sandlåda.
- Det är en framgångsfaktor att **dokumentera** och producera skriven text under hela arbetet. Genom att dokumentera vad gruppen är överens om uppmärksammas tidigt om det föreligger olika bilder om vad gruppen kommit fram till.
- I den AI-regulatoriska sandlådan kommer liknande frågor (såsom vad som är AI-system) återkomma varje gång. Det bör vara positivt med **återkommande frågor**, då processen bör kunna struktureras mycket mer.

Lärdomar

Om arbetssättet

- Behov av **tvärfunktionell kompetens**, såsom jurister, AI-utvecklare, teknisk kompetens och projektledningskompetens. Sådan kompetens bör tillhandahållas av den som tillhandahåller sandlådan för objektivitet.

- **Kärnteam** en fördel då skapar trygghet och effektivitet i arbetet.

- I ett utforskande arbetssätt är det nödvändigt att vara **öppensinnad och lösningsfokuserad**. Teamet behöver ha förmåga att förklara sin specialistkompetens på ett enkelt och pedagogiskt sätt.

- Ett arbetssätt där tolkningssvårigheter identifieras utifrån ett **praktiskt fall** ger ett större lärande.

- En aktör behöver bidra med visst **eget deltagande och kompetens**, särskilt vad gäller beskrivning av tilltänkt system.

- Behov av **stöd från beslutsfattare** för framgångsrikt deltagande i en AI-regulatorisk sandlåda.

Länkar

AI-regulatorisk sandlåda – en kontrollerad miljö för test av AI

Nyttor med att delta i en AI-regulatorisk sandlåda

Delrapport AI-regulatorisk sandlåda – en första iteration

Delrapport AI-regulatorisk sandlåda – den andra iterationen

TACK!