Data & AI governance

A. INDLEDENDE OPLYSNINGER	
Aktivitetsområde	Indsatsområdet Digital sikkerhed, tillid og dataetik
Institut	Alexandra Instituttet
Titel *Dækker indholdet af aktiviteterne*	Data & AI governance
Nummerering *Af beskrivelsen*	2
Version	1
Periode *Forventet start og slut*	13.01.2023 – 31.12.2023
Kontaktperson	Kristian Krämer

B. ÆNDRINGER
Angiv her, hvis en planlagt aktivitet er ændret i forhold til den forudgående version af beskrivelsen.

C. BESKRIVELSE	
1. Mål *Hvorfor? Hvad er målet for aktiviteterne? Hvordan bidrager de til det overordnede mål for indsatsområdet?*	Mens kunstig intelligens-feltet udvikler sig hurtigt, er sikkerhedsaspektet for fremtidige løsninger baseret på kunstig intelligens mere uklart. Data udgør grundlaget for at udvikle kunstig intelligens og påvirker i høj grad aspekter som ydeevne, retfærdighed, robusthed og skalerbarhed. Det gælder både i positiv og negativ retning. Desuden kræver kunstig intelligens adgang til data – data som måske samtidigt skal beskyttes så det ikke bliver lækket eller misbrugt. Hvis der er fejl, mangler eller andre problemer i data eller algoritmer vil det i nogle tilfælde blot betyde at et system ikke løser en given opgave godt nok til at blive taget i brug. Men der er også sammenhænge, hvor problemer er kritiske og potentielt kan have fatale konsekvenser. Det gælder særligt indenfor sektorer som sundhed, finans, offentlig forvaltning, m.fl. hvor mange af de beslutninger der tages, kredser om eller har indvirkning på individniveau, og hvor data ofte er følsomme, og hvor beskyttelse af disse derfor er vigtigt. Hvis kunstig intelligens skal anvendes i disse sammenhænge, er det nødvendigt at der er foretages tilstrækkelige undersøgelser af det pågældende system. Paradoksalt nok er dataarbejde ofte det mindst prioriterede aspekt i udviklingen af kunstig intelligens systemer og indenfor feltet er der da også en tiltagende opmærksomhed på sikkerhedsaspektet.[1] I et nyligt survey med deltagelse af knap 4300 forskere og udviklere indenfor kunstig intelligens, peger 69% på at dette område fremadrettet bør prioriteres højere end det gøres i dag og at der bør udvikles tekniske løsninger der kan understøtte dette.[2] En stor udfordring er at kunstig intelligens er et paraplybegreb der dækker over en række subfelter, der hver har deres lokale udfordringer. Både i forhold til datatyper og

[1] https://research.google/pubs/pub49953/
[2] https://aiimpacts.org/what-do-ml-researchers-think-about-ai-in-2022/

	modelarkitekturer. Derudover øges kompleksiteten, da der typisk knytter sig forskellige udfordringer til forskellige industrier. Det gør det svært at udvikle generelle software løsninger indenfor temaer som *explainability*, *fairness* og *privacy*. Målet for aktiviteten er at sætte fokus på de overvejelser, valg og handlinger der foretages undervejs i forbindelse med udvikling og evaluering af ansvarlige kunstig intelligens systemer. Et fokus der skal gøre det tydeligt og hands-on for danske virksomheder hvordan de kan udføre dette dataarbejde på en forsvarlig og tillidsskabende måde. Et centralt element i aktiviteten vil derfor være at etablere og udføre to eller flere cases indenfor kritiske sektorer, der kan demonstrere ovenstående. Disse cases vil blive udarbejdet i tæt samarbejde med Alexandra Instituttets resultatkontrakt-indsatsområder ”Accelerering af digital sundhed og velfærd i Danmark” og ”Digitale teknologier til datadrevet, bæredygtig vækst”. Konkret bidrages der til de overordnede mål-indikatorer med to eller flere caseforløb sammen med virksomheder, videreudvikling af teknologisk service indenfor AI, kompetenceopbygning, udbygning af samarbejde med vidanspartnere, samt vidanspredning af resultater til dansk erhvervsliv og andre interesserede i form af eksempelvis indlæg på konferencer, webinarer, formidlingsrapporter, artikler, m.v.
2. Indhold **Hvad skal der ske?** **Hvilke(n) konkret€ aktiviteter udføres?**	Med afsæt i ovenstående målbeskrivelse udføres følgende aktiviteter i 2023: **Anvendelse af følsomme data** Der er brug for data i fremtidens løsninger, men data der er følsomme for de registrerede, kan dog være svære at få adgang til, dette kan også gælde forretningsfølsomme data for virksomheder. Der findes både tekniske løsninger, samt brugercentrerede løsninger, der gør beskyttelse og brug af data muligt på samme tid. I denne aktivitet vil vi have fokus på at få prøvet nogle af disse løsninger af i praksis. I tæt samarbejde med aktiviteten ”Bedre anvendelse af sundhedsdata” i Resultatkontrakten-indsatsområdet ”Accelerering af digital sundhed og velfærd i Danmark” udføres min. 1 case om brug af følsomme data og/eller maskinlæring indenfor sundhedssektoren. **Regulering af data håndtering og kunstig intelligens** I disse år er der en lang række lovtiltag på vej på Europæisk plan, der overordnet har som formål at skabe tydeligere juridiske rammer for virksomheders anvendelse af data, datadeling, samt udvikling og anvendelse af kunstig intelligens. For mange virksomheder, særligt SMV’er, er det komplekst at følge med. I denne delaktivitet vil særligt have fokus på udmøntningen af Data Governance Act, som er vedtaget, og AI Act, som forventes vedtaget i løbet af 2023. Omdrejningspunktet vil være på de nye rammer, som denne lovgivning sætter for danske virksomheder. I 2023 vil der blive indhentet og opbygget viden og kompetencer, så Alexandra Instituttet kan give danske virksomheder teknisk vejledning om indhold af begge forordninger, herunder også hvilke tiltag danske virksomheder kan gøre i forbindelse med forretningsudvikling baseret på følsomme data, samt udvikling af kunstig intelligens baserede softwareløsninger. Der vil blive arbejdet tæt sammen med danske virksomheder og det resterende økosystem i Danmark. **Dansk AI-sandkasse** Som en del af AI Act lægges der er op til at der skal etableres såkaldte regulatoriske sandkasser, hvor virksomheder kan få hjælp og vejledning til udvikling af kunstig intelligens løsninger. I Spanien og Norge er der etableret sandkasser der assisterer virksomheder med juridiske aspekter, og i Danmark findes der blandt andet en sandkasse for anvendelse af sundhedsdata, der ledes af KU. Denne delaktivitet vil have fokus på at indsamle viden og erfaringer fra de nuværende sandkasser, som skal bruges til baggrund for udvikling af tekniske løsninger, der kan hjælpe danske virksomheder med at leve op til de kommende krav fra AI Act. **Dataetik i praksis – videns- og kompetenceopbygning hos danske virksomheder**

	Baseret på viden fra den udførte state-of-the-industry analyse i 2022 udføres et eller flere case samarbejder om konkrete tiltag i virksomheder eller andre organisationer om hands-on at arbejde med dataetiske overvejelser i forbindelse med anvendelse af følsomme data og udvikling af kunstig intelligens.
3. Aktører *Hvem udfører aktiviteterne? Hvilken afdeling af instituttet? Evt. hvilke eksterne parter er med (videninstitutioner, virksomheder, erhvervsorganisationer, myndigheder, klynge-organisationer eller andre.)*	Alexandra Instituttets medarbejdere på tværs af fire afdelinger vil udføre aktiviteterne: *Artificial Intelligence & Data Analytics*, *Insights, Security lab.* og *Strategic Business & Governance.* Eksterne samarbejdspartnere: • Universiteterne: I dette aktivitetsområde vil vi samarbejde med flere af de danske universiteter. Vi har samarbejder i gang med både AU, ITU og AAU, og vil både forsøge at udbygge samarbejdet med disse og indlede nye. • Klynger: Hovedsageligt DigitalLead, men også ˚andre klynger hvor denne aktivitet kan støtte op med teknologiske løsninger • Dansk Industri
4. Sammenhæng med andre projekter *Indgår aktiviteten i andre eksternt finansierede projekter?*	Der er synergi med følgende andre projekter hos Alexandra Instituttet: • CRUCIAL, støttet af Innovationsfonden, formål i dette projekt er bl.a. at anvende forretningsfølsomme data fra kritisk infrastruktur på en sikker måde. • AI Denmark projektet, støttet af Industriens Fond, der har til formål at understøtte SMV'er med at komme hurtigere i gang med at udnytte data og AI-værktøjer i deres forretning.
5. Følgegruppe *Har følgegruppen forholdt sig til aktiviteten? I så fald hvordan?*	D. 9. november, 2022, blev der afholdt møde med følgegruppen, hvor aktiviteten og delaktiviteterne blev præsenteret og diskuteret. Aktivitetsbeskrivelsen er efterfølgende blevet tilpasset på baggrund af følgegruppens kommentarer. Eksempelvis er Data Governance Act blevet inkluderet som del af delaktiviteten der handler regulering grundet efterspørgsel blandt flere deltagere i følgegruppen.
6. Formidling af resultater *Hvordan/hvor kan interesserede virksomheder m.fl. få viden om resultaterne af aktiviteterne?* *Anføres/tilføjes hvis det ikke allerede fremgår af beskrivelsen ovenfor, f.eks. ved links til konferencer, hjemmeside, publikationer etc.*	Hovedparten af resultater og erfaringer i denne aktivitet vil blive formidlet gennem aktiviteten *Demonstration af anvendelighed og værdiskabelse* Ift. indsatsområdet *Digital sikkerhed, tillid og dataetik,* i form af videnskabelige artikler, white papers, blog posts- og oplæg på konferencer eller webinarer hvortil denne aktivitet leverer fagligt indhold. En del formidling vil blive gjort i samarbejder med andre aktører for at nå ud til en bredere målgruppe. En del af formidlingen vil ske direkte gennem standardiseringsarbejdet. De udviklede ydelser vil også stilles til rådighed for markedet gennem den digitale TDU.