Risikovurdering for Frischsenterets informasjonssystem

1 Innledning

Personvernforordningen (GDPR) gjelder som norsk lov etter *personopplysningsloven av 15. juni 2018 (POL)*. Artikkel 32 nr 1 pålegger den behandlingsansvarlige å «*… gjennomføre tiltak for å oppnå et sikkerhetsnivå som er egnet med hensyn til risikoen,…*» og i nr 2 heter det at «*Ved vurderingen av egnet sikkerhetsnivå skal det særlig tas hensyn til risikoene forbundet med behandlingen, ….*».

Dette notatet tar sikte på å vurdere risikobildet ved Frischsenterets behandling av personopplysninger i forskningen i lys av reglene i *GDPR* og *POL*. Risikovurderingen er opprinnelig utarbeidet på grunnlag av Datatilsynets veileder om *Risikovurdering av informasjonssystem med utgangspunkt i forskrift til personopplysningsloven (TV-506:2002, veiledningen)*, basert på dagjeldende lovgivning. Risikovurdering basert på GDPR er beskrevet i Datatilsynets veileder om «*Internkontroll og informasjonssikkerhet*» av 10/6-2018.

Sikkerhetssystemet er dokumentert i *Informasjonssikkerhetsinstruks for Frischsenteret* og *Sikkerhetsinstruks for Frischsenteret*. Oversikt over personopplysninger er dokumentert i det løpende registeret *Dataoversikt for Frischsenteret*.

GDPR artikkel 35 pålegger behandlingsansvarlig å utføre en vurdering av personvernkonsekvenser (DPIA) når behandlingen «*vil medføre en høy risiko for fysiske personers rettigheter og friheter*», og i alle tilfeller ved «*behandling i stor skala av*» sensitive opplysninger. Informasjonssikkerhetsinstruksen for Frischsenteret fastslår i punkt 2.9 at det skal utarbeides DPIA for alle nye prosjekter som gjør bruk av personopplysninger. GDPR artikkel 35 nr 7 e) angir at DPIA skal inneholde «en vurdering av risikoene for de registrertes rettigheter og friheter …». Dette dokumentet er ment som en overordnet risikovurdering av behandling av personopplysninger i forskningen med eksempler, og samtidig tjene som et rammeverk for den konkrete vurderingene som foretas i hver enkelt DPIA.

2 Begreper

2.1 Konfidensialitet, integritet og tilgjengelighet

Personvernet har som formål å ivareta hensynet til enkeltmennesker ved å sikre personopplysningers konfidensialitet, integritet og tilgjengelighet. Som forskningsinstitusjon har Frischsenteret ingen forvaltningsoppgaver der enkeltmennesker direkte berøres gjennom beslutninger basert på våre registreringer. Riktignok er endel forskning begrunnet i nytten som kunnskapsgrunnlag for beslutninger, men konsekvensene for den enkelte er her kun indirekte. Behovet for personopplysningenes integritet og tilgjengelighet ved Frischsenteret er derfor ikke direkte regulert av personopplysningsloven. Disse hensynene er selvsagt viktige for forskningens kvalitet, tidsmessig relevans og kostnader, men kommer her i annen rekke i forhold til konfidensialitet.

For enkelte datasett med opplysninger om bedrifter, organisasjoner, kommuner o.l. kan det tenkes at forskning vil få betydelig konsekvens for enheten, f.eks. ved krav til effektivitet, organisering og insentiver, men disse er ikke personer etter personopplysningsloven.

2.2 Konsekvens

Akseptabel risiko er beskrevet i sikkerhetsinstruksen. Risiko kan beskrives som produktet av konsekvens og sannsynlighet. Alvorligheten av konsekvensene for enkeltmenneskers personvern av en hendelse graderes etter en skala fra *TV-506:2002* med:

- K=4, hendelsen kan føre til tap av liv eller vedvarende helsetap, eller kan medføre betydelig og uopprettelig økonomisk tap, eller kan føre til alvorlig tap av anseelse eller integritet som påvirker liv, helse eller økonomi.

- K=3, hendelsen kan føre til tap av helse, eller kan medføre uopprettelig økonomisk tap, eller kan føre til alvorlig tap av anseelse og integritet.

- K=2, hendelsen kan medføre betydelig økonomisk tap – men som kan gjenopprettes, eller kan føre til tap av anseelse eller integritet (eksempelvis kompromittering av opplysninger den registrerte oppfatte som krenkende, eller som andre kan gjøre nytte av).

- K=1, hendelsen kan medføre økonomisk tap – men som kan gjenopprettes, eller kan føre til tap av anseelse eller integritet (eksempelvis kompromittering av opplysninger den registrerte oppfatter som følsomme).

I tillegg kan en hendelse ha konsekvenser for samfunnet mer generelt, eller for Frischsenteret.

2.3 Sannsynlighet

Sannsynligheten for at en hendelse inntreffer kan være svært vanskelig å kvantifisere. I stedet anviser TV-506:2002 en ordinal kategorisering av sannsynlighet på en skala fra 1 til 4 som en mulig fremgangsmåte. Veiledningen foreslår at denne kan beregnes ut fra historiske hendelser, ut fra letthetsbetraktning, eller ut fra motivering.

Historiske data gir ingen veiledning for Frischsenteret; så vidt vites har det ikke i Frischsenterets og forløperne SAF-Oslo og SNF-Oslo's historie ikke skjedd sikkerhetsbrudd med personvernkonsekvenser. Heller ikke motivering kan tenkes å spille en stor rolle i en sannsynlighetsvurdering, ettersom de forskningsdata som vi innehar neppe kan gi opphav til særlig økonomisk eller annen vinning. Dog er skillet mellom uaktsomhet, forsett og overlegg relevant for hvor lett opplysninger kan komme på avveie. Likevel er nok letthetsvurderingen mest relevant, slik det er definert i veiledningen som:

- S=4, sikkerhetstiltak er ikke etablert, eller kan omgås/brytes av egne medarbeidere og eksternt personell med små til normale resurser. Det er ikke nødvendig med kjennskap til tiltakene.

- S=3, sikkerhetstiltak er ikke fullt etablert, eller fungerer ikke etter hensikten. Egne medarbeidere trenger kun små til normale ressurser for å omgå/bryte tiltakene – det er ikke nødvendig med kjennskap til tiltakene. Eksternt personell trenger normal kjennskap til tiltakene (eksempelvis til hvilke rutiner som gjelder, eller hvordan sikkerhetsteknologi er implementert) – i tillegg til små/normale resurser.

- S=2, sikkerhetstiltak er etablert i forhold til sikkerhetsbehovet og fungere etter hensikten. Tiltakene kan likevel omgås/brytes av egne medarbeidere med små til normale ressurser, som i tillegg har normal kjennskap til tiltakene. Eksternt personell trenger gode ressurser, og god/fullstendig kjennskap til tiltakene for å omgå/bryte disse.

- S=1, sikkerhetstiltak er etablert i forhold til sikkerhetsbehovet og fungerer etter hensikten. Tiltakene kan kun omgås/brytes av egne medarbeidere med gode ressurser, og god/fullstendig kjennskap til tiltakene. Eksternt personell kan ikke omgå/bryte tiltakene.

2.4 Risiko

Risikoen knyttet til en uønsket hendelse er i denne sammenheng en forventningsverdi, dvs produktet av sannsynlighet og konsekvens. Risikoen knyttet til Frischsenterets samlede bruk av personopplysninger er litt forenklet summen av risiki knyttet til enkelthendelser. Den samfunnsmessige problemstillingen er å finne den riktige avveiningen mellom kostbare sikkerhetstiltak og samlet risiko. I veiledningen beskrives dette som toleransenivå eller akseptabel risiko, som skal fastsettes av virksomheten selv. Som eksempel settes opp en matrise av sannsynlighet og konsekvens der det skraverte området representerer trusler som ikke kan aksepteres, og gjennomgåtte enkelthendelser kan da plasseres i matrisen.

Tabell 1: Akseptabel risiko fra veiledningen

Konsevens ▶ ▼Sannsynlighet	Liten	Moderat	Stor	Katastrofal
Lav				
Moderat	(Hendelse A)		(Hendelse B)	
Høy				
Svært høy				

3 Risiko ved Frischsenteret

Risiko ved Frsichsenteret diskuteres i forhold til internkontrollinstruks og sikkerhetsinstruks gjeldende fra 1.1.2019. Det forutsettes at leseren har disse dokumentene tilgjengelig.

3.1 Personvernkonsekvenser

Det er strategisk viktig for Frischsenteret å ha strenge regler for akseptabel risiko. Forskningen ved Frischsenteret sikter mot høy vitenskapelig kvalitet og relevans for politiske beslutninger. En betydelig etterspørsel fra departementer og forskningsprogrammer og en lang rekke publikasjoner i internasjonale fagtidskrifter vitner om at denne målsettingen blir oppfylt. Grunnlaget for mye av forskningen er god utnyttelse av et solid empirisk datamateriale med individopplysninger. Konsekvensene av *ikke* å behandle personopplysninger ved Frischsenteret er knyttet til allmenne samfunnsinteresser som vil lide av mangel på forskning. Disse samfunnsinteressene kan ofte være av stor betydning og få konsekvenser for beslutninger som angår svært mange mennesker, men vil likevel ved en konkret avveining lett bli funnet mindre vesentlige enn personvernet til registrerte enkeltpersoner. For Frischsenteret kan det imidlertid være katastrofalt om forskningen ikke kan utføres fordi vi ikke har anledning til å lagre og behandle personopplysninger.

For den enkelte vil brudd på konfidensialitet for den type data som er lagret ved Frischsenteret neppe være av styrke K=4 eller K=3 i avsnitt 2.2 ovenfor, selv om enkelte kjennetegn er sensitive etter GDPR. Siden opplysningene ikke er lagret hos Frischsenteret som ledd i noe beslutningssystem, vil ikke sikkerhetsbrudd f.eks. bety feilbehandling i helsevesenet, feilutbetalinger av stønader, erstatninger e.l. Konsekvenser vil kun komme som følge av at uvedkommende får kjennskap til opplysninger de ellers ikke hadde hatt om enkeltindivider. Disse kan godt tenkes å være i kategoriene K=1 og K=2. Hvis f.eks. en person ikke får arbeid hos en arbeidsgiver fordi denne har fått kjennskap til diagnosen i en tidligere sykemelding, kan konsekvensen være et betydelig økonomisk tap (K=2) men ikke et uopprettelig økonomisk tap (K=3). Frischsenteret legger derfor stor vekt på å unngå konfidensialitetsbrudd.

3.2 Sannsynlighet

Det vil ofte være en avveining mellom tilgjengelighet og konfidensialitet. Forskningsaktiviteten er sjelden tidskritisk, og tilgjengelighet til en hver tid er ikke avgjørende for forskingens kvalitet og det er derfor mulig å ha god sikring av konfidensialiteten.

I tillegg er det kun i innsamlings- og kvalitetssikringsfasen nødvendig å vite hvilke enkeltpersoner som er registrert, og ikke i analysefasen av et forskningsprosjekt. Det er sjelden at Frischsenteret samler inn egne data, og når det har skjedd er det vært ikke-sensitive

opplysninger. Personopplysninger etter GDPR omfatter levende fysiske identifiserbare personer og er i sikkerhetsinstruksen konkretisert som sensitivtetsgruppene S og I og identifiserbarhetsgruppe 1-6. Direkte identifiserende kjennetegn er nesten alltid unødvendig, og skulle de være nødvendige bestemmer sikkerhetsinstruksen at identifikasjonen skal lagres separat (4.5.a, 4.6.b), og dataene alltid behandles i sikker sone (4.5.a og 4.6.a). Sikker sone er egne tjenermaskiner med høy sikkerhet, tilgang kun for spesifikt autoriserte brukere med dobbelt autentisering, og behandling av data kun i kryptert terminaltjenersesjon.

Vanligere enn direkte identifiserbare datasett er enten anonyme data, eller datasett der personidentifikasjonen er erstattet av et løpenummer som kun Statistisk sentralbyrå (SSB) kan tolke. Dette løpenummeret kan være nødvendig for å koble data fra ulike registre, men vil ikke kunne brukes til å direkte identifisere enkeltpersoner uten sikkerhetsbrudd også ved SSB.

Det kan være nødvendig å ha en del kjennetegn i en analyse som gjør det mulig med *indirekte* identifisering av et mindre eller et større antall enkeltpersoner. Det kan f.eks. være snakk om alder, kommunenummer, kjønn, utdanning o.l. som til sammen ville identifisere enkeltindivider, i det minste i mindre kommuner. Sjelden eller aldri vil det være mulig å identifisere hele eller størsteparten av befolkningen med slike metoder. Det vil dessuten være behov for betydelig ressurser og kompetanse for å identifisere enkeltindividene, i form av kjennskap til statistiske eller samfunnsvitenskapelige metoder, kjennskap til statistikkprogrammer og/eller deres filformater, samt kjennskap til variabellister og kodebøker.

Selv om det er vanskelig å identifisere enkeltindividene, kan konsekvensene være såpass store at det er nødvendig med klare tiltak for å forhindre tilgang til indirekte identifiserbare datasett. Sikkerhetsinstruksen bestemmer i pkt 4.1.d at data skal omformes slik at de kan klassifiseres i laveste sikkerhetsklasse forenelig med analysen.

Sensitive personopplysninger skal alltid behandles i sikker sone. Som hovedregel skal også andre indirekte identifiserbare personopplysninger behandles i sikker sone, men enkelte analyser krever maskinkraft eller spisskompetanse som ikke er tilgjengelig innenfor Frischsenteret. Derfor åpnes det i 4.4.d jfr 4.8.b at ikke-sensitive personopplysninger kan behandles utenfor sikker sone kun dersom datasettet er totalt omkodet både når det gjelder variabelnavn og variabelinnhold, og da kun for kortere tid og etter en konkret sikkerhetsvurdering. Koden for forståelse av variablenes innhold skal da forbli i sikker sone.

3.3 Akseptabel risiko

Akseptabel risiko ved Frischsenteret er fastlagt i avsnitt 2 i sikkerhetsinstruksen. I forhold til kravene i tabell 1 fra veiledningen er Frischsenterets krav strengere. I tillegg til katastrofale (K=4) personvernkonsekvenser skal heller ikke store (K=3) personvernkonsekvenser forekomme (pkt 2.1.a). Etter 2.1.b skal det være *svært usannsynlig* at *sensitive* personopplysninger kommer på avveie med negative konsekvenser for enkeltregistrerte, og punktet svarer omtrent til S=1 og K=2 i matrisen. Etter 2.1.c skal det være *usannsynlig* at *ikke-sensitive* personopplysninger kommer på avveie, svarende til S=2 og K=1. Akseptabel risiko svarer da til det uskraverte området i tabell 2.

Tabell 2: Akseptabel risiko ved Frischsenteret

Konsevens ► ▼Sannsynlighet	Liten	Moderat	Stor	Katastrofal
Lav				
Moderat				
Høy				
Svært høy				

4 Eksempel på hendelser

Veiledningen anfører identifikasjon av uønskede hendelser som metode for vurdering av risiko. Ideelt skulle en gjennomgå alle mulige uønskede hendelser, og spesifisere årsaker, hvilke konsekvenser hendelsen får og med hvilken sannsynlighet den opptrer. En full opplisting ville sprenge rammen for dette dokumentet og klart overgå kravene om omfang. I stedet vil en her skissere enkelte hendelser, som er blant de mest alvorlige eller mest sannsynlige som overhode er tenkelige ved Frischsenteret, samtidig som de viser noe av variasjonen i mulige datatyper, årsaksammenhenger og konsekvenser.

4.1 Hendelse I: Opplysninger om fødeland for de fleste utenlandsfødte innbyggerne i landet legges ut på internett.

Mulige årsaker:

a) Et norsk hacker med fremmedfiendtlige holdninger og brukerrettigheter ved Universitetet i Oslo vet at Frischsenteret forsker på arbeidsmarkedet for innvandrere. Han sender en epost med et vedlegg til en prosjektmedarbeider, som åpner vedlegget og dermed får installert et program som åpner en forbindelse gjennom ytre sikkerhetsbarriere og logger alle tastetrykk den betrodde medarbeideren gjør slik at hackeren får tilgang til hans passord mot sikker sone. Fordi sikker sone kun er tilgjengelig fra intern sone eller ved SMS til forskerens mobiltelefon ringer hackeren mobiloperatøren og ber om å få omrutet all trafikk til forskerens mobiltelefon til hackerens mobiltelefon i stedet. Mobiloperatøren har ikke tilfredsstillende sikkerhet og tror det er forskeren som legitimt ber om omruting. Hackeren logger seg på sikker sone med terminalsesjon og flytter datasett fra prosjektområde til brukerområde på sikker sone. Deretter logger hackeren seg på sikker sone med sftp (kryptert filoverføringsprogram) med bruk av sikkerhetssertifikat stjålet fra intern sone og laster data ned til egen maskin. Hackeren kontakter andre i sitt fremmedfiendtlige miljø, som har samfunnsvitenskapelig metodekompetanse og er i stand til å forstå innholdet i filene. Filene har kun løpenummer som identifikasjon, og en tallkode som for fødeland, så hackeren bryter seg på tilsvarende måte inn hos SSB for å hente ut filene som oversetter løpenummer til personnummer, deretter til andre filer hos SSB som kobler personnummer til navn. Endelig bearbeides alle opplysningene og legges ut på internett som lister.

Mulige konsekvenser:
I endel tilfeller vil folk oppfatte det som tap av anseelse eller integritet å få offentliggjort sitt fødeland.

Mulig sannsynlighet:
Lagring og bearbeiding av dette datasettet skal etter sikkerhetsinstruksen foretas i sikker sone. Det er brudd på sikkerhetsbestemmelsene om prosjektmedarbeideren ikke har et oppdatert virusprogram og siste sikkerhetspatcher til operativssystemet, eller at et vedlegg av tvilsomt opphav blir åpnet, jfr sikkerhetsinstruksens pkt 3.2.f. Driftsansvarlig har dessuten et ansvar etter pkt 6.1.f å sørge for drift og loggføring av sikkerhetsbarrieren til intern sone, noe som burde avdekket fjernkontroll av prosjektmedarbeiderens maskin. Mobiloperatøren må også bryte personvernregler. Dernest fordrer det ekstremt høy EDB-kompetanse og betydelig statistikkompetanse hos de utenforstående å forstå og trekke ut informasjonen. I tillegg til å bryte seg inn i Frischsenterets systemer må hackeren også bryte seg inn i flere av SSBs systemer. Det ville faktisk være betydelig enklere for hackeren å kun bryte seg inn i SSBs systemer, siden de aktuelle filene hos Frischsenteret er levert av SSB. Det vil også være mulig å blokkere internettsidene når sikkerhetsbruddet blir kjent, og å spore hvem som har lagt det ut slik at disse stilles til ansvar.

Konklusjon:
Hendelsen har moderat konsekvens (K=2) og er så usannsynlig at det grenser til umulig (S=1).

4.2 Hendelse II: Opplysninger om diagnose ved sykemelding for de fleste innbyggerne i en liten kommune legges ut på internett.

Mulige årsak:
En betrodd medarbeider blir oppsagt på et grunnlag han mener er urettferdig. Han går inn på dataområdet i sikker sone, tar et utdrag for en liten kommune, legger dataene over på sitt brukerområde, og derfra videre over på sin vanlige harddisk før de kobles mot diagnosekoder hentet fra standarder gitt ut av SSB og navn hentet fra ligningen. Han sender filen som epost til en kamerat som legger tabellen ut på internett.

Mulige konsekvenser:
I enkelte tilfeller kan arbeidsgiver ha interesse av å vite om en sykediagnose før eventuell ansettelse. I slike tilfeller kan noen derved unngå å få et arbeid han/hun ellers ville fått noe

som medforer et betydelig økonomisk tap. I et større antall tilfeller vil folk oppfatte det som tap av anseelse eller integritet å få offentliggjort sin sykediagnose.

Mulig sannsynlighet:
Lagring og bearbeiding av dette datasettet skal etter sikkerhetsinstruksen foretas i sikker sone. Her er det for det første et brudd på bestemmelsene i 3.3.k om fjerning av autorisasjon når medarbeidere slutter, enten fra direktøren etter 6.1.a og/eller av driftsansvarlig etter 6.1.f. Dernest er det både illojalt, et brudd på taushetserklæringen og brudd på sikkerhetsinstruksen fra den oppsagte medarbeideren selv, bl.a. ved lagring av data utenfor sikker sone. Medarbeiderens aktivitet vil bli logget slik at han kan stilles til ansvar, og internettsidene vil kunne blokkeres og den som har lagt dem ut vil også kunne stilles til ansvar. Selv om motivasjonen fra den oppsagtes side er klar, er det likevel tvilsomt om han ville ta sjansen gitt at han kan være sikker på å bli stilt til ansvar.

Konklusjon:
Hendelsen har moderat konsekvens (K=2) og er svært usannsynlig (S=1).

4.3 Hendelse III: Naboen til en arbeidsløs får vite om deltakelse på flere arbeidsmarkedstiltak uten at vedkommende har fått arbeid etterpå.

Mulige årsaker:
En betrodd medarbeider med tilgang til et indirekte identifiserbart datasett om arbeidsmarkedstiltak utforsker om han kan identifisere enkeltpersoner han kjenner i datasettet, gjenkjenner naboen til foreldrene hjemme, og forteller foreldrene om det.

Mulige konsekvenser:
Den arbeidsløse oppfatter det som tap av anseelse at hans nabo får vite at han ikke har fått jobb etter fullført arbeidsmarkedstiltak. Om dette blir videre kjent, f.eks. blant enkelte potensielle arbeidsgivere, kan det også medføre at han må søke noe lenger etter arbeid.

Mulig sannsynlighet:
Datasettet tilhører klasse C og skal lagres og behandles i sikker sone. Frischsenterets medarbeider bryter sin taushetserklæring, og det kreves en del metodisk kompetanse for å identifisere naboen, men det kreves ikke store ressurser ellers hos medarbeideren. Siden opplysningene er hentet fra sikker sone er det derimot sannsynlig at analysen har blitt logget,

slik at medarbeideren vil stilles til ansvar dersom naboen forfølger saken. Det er vanskelig å se at en betrodd medarbeider hadde tatt sjansen på et slikt brudd.

Konklusjon:
Hendelsen har liten konsekvens (K=1) og er usannsynlig (S=2).

5 Konklusjon

Gjennomgangen av mulige konsekvenser, sannsynligheter og enkelthendelser har ikke avdekket forhold som overstiger den akseptable risikoen som er beskrevet i sikkerhetsinstruksen.

Ytterligere sikkerhetstiltak utover gjeldende instrukser ville forringe forskningens kvalitet, samtidig som de ville være uforholdsmessig ressurskrevende i form av investeringer og tungvinte arbeidsrutiner.