En case-studie om PDS-økosystemer.

Hvordan kan PDS-økosystemer løse utfordringene ved dagens persondata håndtering.

Sindre Kvålsgard

Masteroppgave ved institutt for informatikk

UNIVERSITETET I OSLO

20. mai 2016

En case-studie om PDS-økosystemer

Hvordan kan PDS-økosystemer løse utfordringene ved dagens håndtering, prosessering og lagring av persondata.

2016

En case-studie om PDS-økosystemer

Sindre Kvålsgard

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Motivasjon: Ny teknologi og endret lovgiving, gjør det relevant å se på nye modeller for håndtering, lagring og prosessering av persondata. **Problemet**: Tjenestetilbydernes tilnærming til innsamling, lagring og bruk av persondata er ikke tilfredsstillende. Vi observerer en tvetydig regulering, manglende databasesikkerhet og brukere som mangler kunnskap om hvordan de best kan håndtere sin egen persondata. **Tilnærming**: Eksplorerende case-studie med kvalitative intervjuer. To runder intervjuer, hvor eksperter på området («privacy») ble stilt spørsmål om fenomenet (håndtering av persondata) og en modell. **Funn**: Tiden er moden for en forandring i håndteringen av persondatabasert på ny regulering i EU og at brukere krever bedre beskyttelse av sine data. Personal Data Storage (PDS) økosystem er et mer brukersentrisk alternativ som gir brukeren kontroll over egne persondata og samsvarer bedre med den nye reguleringen. **Bidrag**: Denne studien bidrar med en abstrakt brukersentrisk modell for et PDS-økosystem der brukeren har utvidede muligheter for kontroll, samtykke og konfidensialitet for sine persondata.

Forord

Denne studien ble utført i tett dialog mellom forfatter og veiledere, Petter Nielsen og Kjetil Hustveit. Jeg samarbeidet også en del med en annen masterstudent, Vivek Kaul. Sammen definerte vi hva vi skulle legge i begrepet PDS. Dette var nyttig da vi begge skrev om PDS'er.

Innholdsfortegnelse

En case-studie om PDS-økosystemer III
Sammendrag V
Forord VII
1 En case-studie av PDS-økosystemer 1
2 Bakgrunn 9
1995 Data Protection Directive 9
GDPR (General Data Protection Regulation) 11
Platform for Privacy Preferences (P3P) 13
PDS (Personal Data Store) 14
Mydex CIC 15
OpenPDS 18
Enigma 20
Privacy By Design 21
3 Metode 22
Case-studie delt opp i fire faser 22
Metodens rammeverk 24
Beskrivelse av selve prosessen 26
Tilnærmingens styrker og svakheter 32
4 Sentrale aspekter ved persondata 34
GDPR 34
Lagring av persondata 36
Identifisere samme bruker på tvers av ulike systemer 37
SafeAnswers o.l. 38
Aspektene drar i samme retning 39
5 Analyse av tre utvalgte PDS-økosystemer 40
Evalueringskriteriene 40
OpenPDS 42
Mydex 43
Enigma 44

6 Diskusjon: krav til et PDS-økosystem 48
Lagring 48
Tilgang til data og samtykke 51
Sikkerhet 54
Identitetstilbyder 56
Gjør PDS-økosystemet adopterbart for markedet 58
7 Modellen 60
8 Studiens begrensninger og framtidig arbeid 65
Framtidig arbeid 66
Litteraturliste 68
Bibliografi 68
Vedlegg 1 71
Vedlegg 2 75

Figurer:

Figur 1: Mydex sikkerhetsmodell 16
Figur 2: Mydex informasjonsflyt 17
Figur 3: OpenPDS arkitektur 19
Figur 4: Fasene til case-studien 22
Figur 5: PDS-økosystem modellen, overordnet 60
Figur 6: PDS-økosystemets rådata håndtering/lagring 62
Figur 7: PDS-økosystemets plassering av applikasjoner/plattformer 64

Tabeller:

Intervjuoppsummering……..31
Analyseoppsummering……..46

1 En case-studie av PDS-økosystemer

Denne oppgaven diskuterer hvordan brukere kan få kontroll over egne persondata. Kontroll i denne sammenheng betyr at brukere av tjenester har:

- Oversikt over hvilke persondata som lagres av tjenestetilbydere;
- Oversikt over hvordan tjenestetilbyderne bruker persondata;
- Mulighet til å begrense og hindre tjenestetilbyderens tilgang og bruk av persondata.

Fokus er på hva regulering krever av tjenestetilbydere og hvordan nye modeller for håndtering, prosessering og lagring av persondata kan håndtere dagens utfordringer og flytte kontroll fra tjenestetilbyderne og over til brukerne. Eksempler på disse utfordringene er identitetstyveri og misbruk av persondata. Oppgavens viktigste bidrag er beskrivelsen av en modell for hva et Personal Data Store (PDS)-økosystem for håndtering, prosessering og lagring av private data bør inkludere av funksjoner og begrensninger.

Et økosystem refererer i denne oppgaven til «software ecosystem» (Hansen & Manikas, 2012) som beskriver et digitalt miljø hvor et sett av software-løsninger fungerer som en enhet. «Enheten» i denne oppgaven er et miljø for håndtering, prosessering og lagring av private data.

Motivasjon

Det er behov for nye perspektiver på hvordan håndtering, lagring og prosessering av private data skjer. Dagens situasjon har utviklet seg på tjenestetilbydernes premisser og på bekostning av brukerens rettigheter. Nye perspektiver er spesielt relevant nå som nye metoder for «BigData» kan anonymisere brukerne, og reguleringer åpner for nye muligheter og drivkraft i markedet. Det vil også bli mer ressurskrevende for tjenestetilbydere å sikre brukerens persondata, grunnet strengere krav til databasesikkerhet.

Den teknologiske utviklingen er raskt, i motsetning til politiske reformer og samfunnsforståelsen av nye fenomener. Byråkratiske offentlige organisasjoner bruker lang tid på å sette seg inn i nye fenomener og regulere de, i motsetning til små og store tek-bedrifter. Lovgivningen har begynt å ta igjen den teknologiske utviklingen når det gjelder eierskap og

håndtering av persondata, noe vi ser ta form i GDPR (General Data Protection Regulation, Des 2015).

Dagens standardkontrakt for samtykke er å skrive ned alt en tjenestetilbyder har tenkt til å bruke data til gjennom et veldig langt «pop-up agreement form». Få leser disse, og hvorfor skal man det når de fleste brukerne synes det er greie vilkår. Da kan det vel ikke være for ille? Samtidig vet vi at flere tjenestetilbydere har designet samtykkekontrakten slik at den skal være slitsom å lese i gjennom (Stone Business Law, P.C.). Et ofte brukt argument (Boitnott, 2014) er at brukerne faktisk ikke bryr seg hvorvidt det samles og lagres data om en. Her ser vi en mulig endring, grunnet økt fokus på beskyttelse av privatlivet og økningen i ID-tyverier.

Informasjonssikkerhet blir stadig bedre, ved bruk at avansert programvare og bedre treffende modeller. Samtidig er normen om å sanke så mye persondata som mulig fra brukere, utfordrende for små bedrifter som ikke har kunnskap eller ressurser til å sikre persondatas konfidensialitet og integritet (Hardekopf, 2015).

Det eksisterer på mange måter et vakuum i håndtering av persondata. På den ene siden har vi metoden kommersielle aktører alltid har brukt, på den andre strengere regulering og økende fokus på personvern. GDPR åpner for en ny gjennomgang av hvordan persondata håndteres, lagres og prosesseres. Min oppfatning er at tjenestetilbyderne så langt har hatt for mye spillerom, noe som har gått utover brukernes rettigheter og trygghet. Et PDS-økosystem kan ivareta brukernes rettigheter på en måte som er mer synkront med den nye reguleringen.

Dagens håndtering av persondata er ikke tilfredsstillende

I dette avsnittet argumenterer jeg for hvorfor måten tjenestetilbydere håndterer persondata ikke er tilfredsstillende med hensyn til brukerens rettigheter. Dette inkluderer hvordan tjenestetilbydere lagrer, håndterer og prosesserer persondata fra sine brukere.

Typen data og metadata det refereres til i denne studien er alle former for persondata som;

- GPS posisjon med tidsavtrykk
- Hvem meldinger ble sendt til og meldingens innhold

- Nettleserhistorikk
- IP-adresser.

Og mer sensitive data, som;

- Personnummer
- Kredittopplysninger
- Data om nåværende og tidligere helsesituasjon

Disse dataene (eller hva vi kan kalle «produktet») har en høy markedsverdi for enkelte aktører. For eksempel, verdien av helsedata er betydelig for forsikringsselskaper som skal bestemme premien for en livsforsikring.

Det kan ikke forventes at den gjennomsnittlige brukeren har tilstrekkelig juridisk kompetanse til å vite hvem som eier data registrert om dem, og hvilke rettigheter hun eller han har over disse. Vi må derfor forvente at mange er forvirret da feltet stadig utvikler seg og nye gråsoner oppdages av tjenestetilbyderne. Med gråsoner menes at vi ikke alltid har juridiske bestemmelser som følger utviklingen og nye muligheter teknologien gir. Så lenge tjenestetilbyderne oppgir hva data skal brukes til under «terms of service», står de rimelig fritt til å gjøre hva de vil (Grothaus).

Nadezhda Purtova beskriver denne praksisen slik: "*... if property rights are not assigned by a legislative action, personal data will be appropriated in proportion to the de facto power of the data market participants to exclude others. It follows that, so long as personal data bears high economic value, the real question is not whether there should be property rights in personal data but whose rights they should be.*" (Purtova, 2015) Purtova er klart inne på noe her. Uten strekkelig regulering rundt persondata, som har høy økonomisk verdi, vil sterke aktører utnytte dette. Tjenestetilbyderne vil kapitalisere på persondata innenfor mindre klare reguleringer.

Store internasjonale selskaper har lobbyister som jobber mot EU og regjeringer for å fremme sine interesser (Davies & Marks, 2015). Det krever ikke mye å se for seg at å beholde «status quo» og fortsette å behandle persondata som «nobody's property» etter dagens lovgiving. Dette vil gagne tjenestetilbydere og være på bekostning av brukerne. En annen viktig dimensjon i forhold til det juridiske er «default entitlement» (Purtova, 2015) til persondata. Hvem har ultimat kontroll over persondata og kan nekte andre tilgang til persondata? For å

fjerne all tvil om hvem det er som eier persondata, er det enklere å ikke fokusere på hvem som har tilgang til den, men heller hvem som kan nekte andre tilgang til den. Den er ifølge Purtova den virkelige eieren.

Fra et teknisk perspektiv er det viktig å se på følgene av manglende regulering av persondata og hvordan persondata håndteres. Der det er manglende regulering kan det oppstå manglede og usikker håndtering av persondata. Mange tjenestetilbydere ønsker å kapitalisere på persondata og samler inn så mye persondata de kan ved å ha en lang liste med ønsket persondata for at applikasjonen skal fungere «optimalt». Dette gir rom for mye «lovlig kaos» og er ikke tilfredsstillende fra et brukerperspektiv. Mengder av persondata samles inn uten grunn, noe som er tidkrevende der brukerne må registrere data selv og i forhold til sikkerhet og personvern.

Sett bort i fra den etiske og juridiske plikten å sørge for at andres persondata behandles konfidensielt og sikkert med tanke på integriteten til data, er en mer direkte trussel i dagen samfunn er økningen av identitetstyveri (Sandland, 2016).

Sentrale punkter fra dette avsnittet er at tjenestetilbyderne har for mye makt og samler inn mye persondata, selv om det innebærer en risiko for brukerne. Dette kan de gjøre fordi lovgivingen på dette området ikke regulerer det tilstrekkelig. Slik det fungerer i dag har brukerne manglende kunnskap og muligheter til å kontrollere sine egne persondata.

PDS-økosystemer kan løse utfordringene

Denne oppgaven har som ambisjon å finne fram til en modell for håndtering av persondata som gir økt kontroll og beskyttelse av persondata. Brukerne får makt til bare å dele det han/hun vil dele, og får full oversikt over hvilke data som går hvor.

For å kunne utvikle dette økosystemet trengs en forståelse av hvordan man kan designe et slikt økosystem. Denne oppgaven tar sikte på å beskrive en modell/økosystem som tåler tidens tann ved at den ikke baserer seg på noen varige aktører i særlig grad. Med dette menes at selv om verden/konteksten rundt modellen forandrer seg vil ikke dette ha noen innvirkning på integriteten og konfidensialiteten på persondata. Insentiver vektlegges sterkt i denne oppgaven, og spesielt hvordan insentiver vanskeligjør en endring fra dagens situasjon. Et

eksempel på insentiver er forretningsmodeller og hvordan dagens tjenestetilbydere tjener penger på kundens/brukerens persondata.

En Personal Data Store (PDS) er et konsept der brukeren lagrer alle sine persondata og metadata ett og bare ett sted. Brukeren kan så velge å dele disse persondata med tjenestetilbydere gjennom samtykkekontrakter. Disse kontraktene beskriver hvilke data som skal utveksles og til hvilket formål. PDSen gir brukeren kontroll over persondata, i form av hvor data er lagret, oversikt over samtykkekontrakter og muligheten til å redigere både persondata og samtykkekontrakter. Det omkringliggende økosystemet gir mulighet til persondataanalyse, anonyme svar på spørringer og generell tilpasning av PDSen mot tjenestetilbyderne.

En PDS og det omkringliggende økosystemet bør være «open-source». Dette standpunktet deles med utviklerne av Mydex CIC og OpenPDS, to systemer som prøver å puste liv inn i PDS konseptet fra to litt forskjellige vinkler. Med «open-source» sikrer man i større grad gjennomsiktigheten og tilliten til økosystemet.

Fra et sikkerhets-perspektiv er det flere som deler synet på at systemet skal være transparent og åpent. Dette vil si at brukere og utviklere selv kan se hele livssyklusen til persondata i økosystemet. Dette for å hindre «lytteposter» som overhører datatrafikken og sender den videre ut av økosystemet. Selv om «kriminelle» kan se mulige utnyttelser, vil det være flere altruister som vil melde ifra om disse hullene, slik at de kan tettes igjen. Eller så tetter de igjen hullene selv.

Det er også viktig å ta for seg den daglige bruken og de praktiske løsningene rundt et slikt økosystem. Slik at det både er brukervennlig, men også tilbyr sikre løsninger. Et viktig prinsipp er at det bare skal være én administrator til dette økosystemet og det er brukeren. Med dette menes at det er bare brukeren som har veto for godkjenning av samtykkekontraker eller endring av eksisterende rutiner i økosystemet.

Det er flere måter for brukerne å dele persondata på. En metode er standardiserte kontrakter: her vil tjenestetilbyderen og brukeren inngå en avtale om hvilke data som kan brukes og hvordan den kan brukes. Dette vil være en form for deling som er lik dagens standard sett bort i fra at data lagres av brukeren i stedet for tjenestetilbyderen. En annen metode er at brukerne ikke gir fra seg rådata. Rådata er persondata som ikke har blitt aggregert. Om man legger inn et prosesserings lag mellom rådata og tjenestetilbyderen, vil dette kunne sikre at brukeren

lettere kan anonymiseres, da aggregerte persondata vanskeliggjør identifisering om det er ønskelig. For eksempel, i stedet for at tjenestetilbyderen får alle filmene en bruker har sett, kan favorittsjanger vises til tjenestetilbyder. Kalkuleringen som vanligvis ville ha skjedd hos tilbyder, skjer nå hos bruker. En del av OpenPDS' SafeAnswers tilbyr denne funksjonen, hvor ferdige svar serveres rett til tjenestetilbyder.

Denne oppgaven vil argumentere for et PDS-økosystem der all tilgang til persondata er regulert gjennom overvåkede porter/innganger. Identitetstyveri vil med dette vanskeliggjøres siden fokus på sikkerhet kan konsentreres om disse portene, i stedet for mange med varierende grad av sikkerhet, som det er i dag.

Oppsummert skal et PDS-økosystem gi makten over persondata til brukeren, ved at brukeren bestemmer hvilke persondata som skal deles. Persondata som ikke er aggregert eller prosessert, kalles for rådata og er satt under brukerens kontroll i et PDS konsept. Deling av aggregerte persondata bør prioriteres fremfor rådata, da rådata er mer sensitivt. PDS-økosystemet bør være «open-source» da dette gir en større grad av gjennomsiktighet. Gjennomsiktighet gjør det mulig å spore persondata-transaksjoner som vil øke sikkerheten og vanskeliggjøre misbruk.

Hypotese og forskningsspørsmål

Hypotese: Denne oppgaven legger til grunn en hypotese om at det er mulig å lage et PDS-økosystem som oppfyller særdeles sterke krav til brukernes kontroll over egne data. Med dette menes; muligheten til anonymitet når brukeren skulle ønske det, full kontroll og oversikt over persondata, samt mulighet til å endre tjenestetilbydernes tilgang på den.

Forskningsspørsmål: Hvordan kan et PDS-økosystem løse dagens utfordringer ved håndtering, prosessering og lagring av persondata? Med dagens utfordringer menes hovedsakelig brukerens manglende kontroll over egen persondata, som kan føre til misbruk og tyveri av persondata. Under kontroll ligger for eksempel brukerens mulighet til å være anonym når han/hun ønsker det, samt å vite hvem brukeren har delt persondata med.

Metode

I denne oppgaven ble case-studie valgt. Typen av case-studie som ble benyttet er en form for eksplorerende studie. Kvalitative intervjuer ble gjort av relevante eksperter fra ulike disipliner. Prosessen startet med begrenset kunnskap om området i starten, noe som påvirket valg av forskningsspørsmål, metodologi, og fremgangsmåte. Etter hvert som mer fakta og informasjon på feltet ble gjort tilgjengelig endret dette seg.

Målet med denne oppgaven er å først utforske de ulike aspektene for lagring, håndtering og prosessering av persondata. Deretter evalueres tre PDS-økosystemer og deres funksjoner for deling av persondata med tjenestetilbydere og andre tredjeparts applikasjoner. Til slutt presentere en modell av et PDS-økosystem som tar tilstrekkelig hensyn til de ulike aspektene tidligere belyst. I den første intervjurunden var fokuset på at intervjuobjektene skulle hjelpe til med å finne de sentrale aspektene til fenomenet lagring, prosessering og håndtering av persondata. De hadde også en indirekte innflytelse på hvilke evalueringskriterier som ble valgt for å analysere PDS-økosystemene. Andre intervjurunde ble hovedsakelig brukt til en kvalitetssikring av modellen og evalueringskriteriene.

Perspektiv

Perspektivet i denne oppgaven er hovedsakelig brukersentrisk, men ikke utelukkende brukersentrisk. Det er også viktig å få med seg markedet og tjenestetilbyderne for å få til et slikt skifte. De etablerte tjenestetilbyderne bør enkelt se hvilke mulige insentiver som finnes.

Et av målene modellen foreslått i denne oppgaven er å sikre at brukeren har teknisk eierskap i tillegg til juridisk eierskap. Det legges også vekt på å finne modeller/funksjoner som favoriserer bruken av «no-trust» og modeller som har den egenskapen at den begrenser tredjeparts-autoritet og -innflytelse over data. Ved bruk av «no-trust» modeller blir ikke persondata tapt eller kompromittert om en av lagringstilbyderne blir kjøpt opp av noen med andre insentiver, eller om bedriften går konkurs. Med «no-trust» fjernes det menneskelige aspektet fra ligningen, altså muligheten til å gjøre feil.

Funn og konklusjon

Denne oppgaven tar for seg en kompleks problemstilling med flere interessenter (hovedsakelig tjenestetilbydere, brukere og styresmakter), mulige modeller og juridiske aspekter. Her er det samtidig ikke rom for å inngå mange kompromisser. Med det mener jeg at om man ikke vektlegger brukersentrismen mer enn tjenestetilbydernes interesser vil løsningen bli for lik dagens modell, og da har man ikke oppnådd nok for å sørge for at brukeren får kontroll over egne data.

Det viktigste bidraget i denne oppgaven er en modell som sikrer brukerens rettigheter tilstrekkelig, samt at den er relevant for tjenestetilbyderne. Denne modellen tar samtidig høyde for at det er viktig å sikre at tjenestetilbyderne får sine behov dekket og fortsatt kan tjene penger. De vil miste noen inntekter fra «bigdata»-informasjon, men de kan tjene litt inn igjen på lavere backend kostnader. Samtidig gir modellen/PDS-økosystemet tjenestetilbydere mulighet til å tjene penger på multiparty-computations (forklart senere). Modellen tilbyr også flere kontrollmekanismer for å sikre brukerens makt over egne persondata.

Oppgavens omfang

Denne oppgaven omfatter PDS-økosystemer og aspekter og handlinger som har en direkte innflytelse på utviklingen av disse. Mer detaljert: Lover og reguleringer, hovedsakelig i EU, tekniske spesifikasjoner som kryptografi og protokoller (veldig overordnet), samtykke og gjennomsiktighet, og minimering av begrensninger slik at markedet kan absorbere løsningen.

Retningslinjer for bedrifter eller andre tiltak for å hindre at bedrifter misbruker data gitt til dem av PDS'en vil ikke falle innenfor denne oppgaven. PDS-økosystemet i denne oppgaven vil heller fungere proaktivt og muligens reaktivt på misbruk av data, som igjen vil være et brudd på tillit/samtykkekontrakt. Fysiske begrensninger på hardware og andre fysiske barrierer for beskyttelse av data vil også falle utenfor.

For å oppsummere skal denne oppgaven svare på forskningsspørsmålet: *hvordan kan et PDS-økosystem løse dagens utfordringer ved håndtering, prosessering og lagring av persondata?* Ved å se på tre caser av PDS-økosystemer (Mydex, OpenPDS og Enigma) og til slutt benytte favoriserte funksjoner i vår egen PDS-økosystem modell.

2 Bakgrunn

I dette kapittelet skal jeg gå igjennom bakgrunnen til reguleringsaspektet utforsket i denne oppgaven. Etter det blir eksisterende modeller som tilbyr interessante funksjoner og løsninger til persondata-utfordringen sett på.

Under reguleringen av persondata ser vi på manglene i dagens direktiv og fordelene med den nye reguleringen som blir gjeldene våren 2018. Så skal vi raskt se på et tidligere prosjekt (P3P) for mer brukersentrisk håndtering av persondata, og se på hvorfor det ikke slo igjennom. Etter det beskrives oppgavens definisjon av PDS/PDS-økosystem og tre eksempler på PDS-økosystemer diskuteres. Vi vil avslutningsvis å se på design prinsipper for håndtering av persondata. Dette er relevant siden vi skal utvikle en modell for et PDS-økosystem.

1995 Data Protection Directive

Datalagringsdirektivet fra 1995 har gitt rom for dagens modell. Direktivet har mangler som kan oppsummeres med for få krav til tjenestetilbyderne når det kommer til håndtering og lagring av persondata. En av grunnene til dette er at direktivet ble utviklet for en annen teknologisk tidsalder (Hustinx, 2014). I tillegg er direktivet kun et veiledende rammeverk for medlemsland av EU. De må innen en gitt tidsfrist sørge for å nå målene gitt i direktivet. Innen våren 2018 vil den bli fullstendig erstattet av General Data Protection Regulation (GDPR).

Datatransport begrensninger

EUs databeskyttelse-direktiv (EU, data-transfer) inkorporert i den norske personopplysningsloven sier at transportering av personlig data utenfor grensene til EU, krever tilstrekkelig grad av databeskyttelse fra mottakerens side. Unntak kan gis om kontrolløren (entiteten som bestemmer formålet og meningen med prosesseringen av persondata) kan sikre at data blir håndtert på lik linje som om det var innenfor grensene til EU. Denne bestemmelsen er relevant fordi dagens digitale marked er globalt. De fleste store selskapene opererer på tvers av landegrensene. Uansett skal brukere, i hvert fall innenfor EU være sikre på at deres rettigheter blir ivaretatt gjennom reguleringer selv om data prosesseres utenfor EUs yttergrense. Modellen beskrevet senere i denne oppgaven vil gjøre grep for å gi

brukerne mer makt, og dermed gi tjenestetilbyderne mindre. Dette gjøres for å begrense risiko ved å redusere mengden tillit til andre parter utenfor PDS-økosystemet. Denne tillits-begrensningen sammen med reguleringen vil gjøre det vanskeligere å misbruke persondata.

Kritikk av direktivet

Direktivet er blitt kritisert for å være uklart. Det er ikke klart nok definert hvem som har «default entitlement»:

> *"Property rights in personal data under the current European data protection regime are ill-defined in that the 1995 Data Protection Directive does not assign default entitlements in personal data clearly to the data subject or to the data controller."* (Purtova, 2015).

Dette viser til at det ikke er klarhet i hvem det er som formelt eier persondata og at det bør spesifiseres klarere hvem som eier persondata i utgangspunktet (by default).

Et annet relevant område som har fått en del kritikk er «Safe harbor principles» som regulerer utveksling av persondata og data mellom USA og EU. «Safe harbor principles» er 7 prinsipper som gir noen amerikanske selskaper mulighet til å lagre persondata fra EUs innbyggere, så lenge de følger EUs reguleringer (oversatt til de 7 prinsippene). Kort fortalt går kritikken ut på at «Safe harbor principles» ikke gir en tilfredsstillende grad av beskyttelse av persondata (Carson, 2013). Et av punktene det er gitt kritikk for, er at de amerikanske selskapene som får lov til å lagre persondata om EU medlemmer kan sertifisere seg selv. Dette går direkte ut over EU-innbyggeres tillit til at tjenestetilbydere og reguleringer sikrer deres persondata fra misbruk. Det bør ikke være tvil om persondata blir håndtert etter EU-reguleringer.

GDPR (General Data Protection Regulation)

Her er et kort innblikk i de viktigste relevante punktene i denne nye EU reguleringen. Siden dette er en regulering og ikke et direktiv, plikter alle 28 medlemsland å følge bestemmelsene.

GDPR: Samtykke

Innhenting og prosessering av persondata er strengere regulert i GDPR, enn i 95-direktivet. Brukeren må ha utvetydig gitt samtykke til bruk av persondata etter å ha blitt tilstrekkelig informert (Purtova, 2015). Hvordan en velger å tolke denne bestemmelsen vil trolig bli et fokuspunkt i de neste to overgangs årene til GDPR. Brukeren skal også ha muligheten til å trekke tilbake samtykket som tidligere har blitt gitt.

Interessene til tjenestetilbyderen blir her tilsidesatt av de fundamentale rettighetene til brukeren. Brukeren har rett til å bli slettet fra datasamlingen til tjenestetilbyderen om han eller hun ønsker det (EU, 2015).

Et annet viktig punkt ved GDPR er brukernes rett til å innhente data som er samlet om en. Brukeren har rett til å rekvirere en kopi av persondata i både maskinleselig format og menneskelig lesbart format (EU, 2015). Dette åpner for å fylle en PDS med tidligere innsamlet data, samt muligheten for å flytte data mellom ulike tjenestetilbydere dersom dette skulle være ønskelig.

GDPR: Persondata

Dette er hvordan 95-direktivet definerer persondata;

> *«Personal data is any information relating to an individual, whether it relates to his or her private, professional or public life. It can be anything from a name, a photo, an email address, bank details, your posts on social networking websites, your medical information, or your computer's IP address.»* (EU, 2012)

GDPR vil komme med en mer omfattende definisjon av persondata som inkludere: fysisk, psykologisk, genetiske, mentale, økonomisk kulturell og sosiale identitet (Hunton & Williams, 2015). Dette vil innebære at persondata blir et bredere begrep som omfatter flere

typer data. Dette gjør håndtering, prosessering og lagring av data mer ressurskrevende for tjenestetilbyderne, siden persondata har strengere regulering enn annen data.

GDPR: Databeskyttelse

Kravene EU gir til tjenestetilbydere for beskyttelse av persondata blir flere og strengere i GDPR. Dette vil i større grad sikre at kun store tjenestetilbydere som klarer å sikre persondata tilstrekkelig vil få mulighet til å lagre persondata. Da dette ikke vil bli lønnsomt nok for mindre aktører, grunnet at det vil koste for mye å sikre tilfreds implementasjon av de nye kravene.

> *«With new obligations on such matters as data subject consent, data anonymization, breach notification, trans-border data transfers, and appointment of data protection officers, to name a few, the GDPR requires companies handling EU citizens' data to undertake major operational reform.»* (Heimes, 2016).

Kravet til databeskyttelse blir strengere overholdt i GDPR. Store bøter og ansettelser av databeskyttelses-eksperter er noen tiltak. Dette vil også gjøre det mer ressurskrevende for tjenestetilbydere å håndtere persondata.

GDPR: Sammendrag

Den nye reguleringen vil sannsynlig få bestemmelsene prøvd i retten i de nærmeste årene, og det er ikke sikkert at bestemmelsene vil stå som de gjør om noen år. Uansett gir GDPR en bedre beskyttelse av brukernes rettigheter og legger strengere krav på håndtering og lagring av deres persondata, som tjenestetilbyderne må følge. Det vil bli mer ressurskrevende for tjenestetilbyderne å prosessere og lagre persondata. Dette styrker motivasjonen for å se på hvordan PDS-økosystemer kan håndtere oppgaven med å lagre og prosessere persondata, framfor å la tjenestetilbyderne gjøre det.

Platform for Privacy Preferences (P3P)

P3P prosjektet ble utviklet og lagt ned for flere år siden. De hadde slagordet: «Enabling smarter Privacy Tools for the web». Dette prosjektet prøvde å implementere bedre personvern og oversikt over hvilke data tjenestetilbydere på nett brukte og hva de brukte dem til. Det er relevant å se hva som tidligere har blitt prøvd uten suksess. P3P baserte seg på frivillighet og håpet på at tjenestetilbydere skulle omfavne denne funksjonaliteten. Under blir det forklart hvordan P3P fungerte og kritikk/ideer på hvorfor det ikke slo igjennom.

Hvordan fungerte P3P

P3P har fortsatt en nettside med en konseptforklaring. P3P gir nettsidene en mulighet til å uttrykke deres persondata håndtering gjennom et standard format. Brukerne har en liten software-modul (mest sannsynlig integrert i nettleseren) som beskriver nettsidens persondata-praksis. For at dette skal fungere må også nettsiden ha integrert P3P slik at begge sider (bruker og nettside) kan utveksle informasjon. Både menneskelig- og maskin-leselig format skulle benyttes slik at man kunne automatisere beslutninger. Et eksempel på dette kunne være at brukeren bestemte seg for å aldri gi ut personnummer til nettsider/tjenestetilbydere. Dette ville gjøre det enklere for brukere å forholde seg til nettsidenes praksis, slik at de ikke trengte å lese nettsidens retningslinjer (P3P, 2006).

Hvorfor ble det avsluttet

Noen sier at P3P var for komplisert å jobbe med for tjenestetilbydere (Thylmann, 2011) , mens andre peker på at håndtering av persondata trenger regulering, altså tvang (Cranor, 2012). Årsaken til at P3P ble avsluttet kan være en kombinasjon av begge disse årsakene, eller andre grunner som er vanskeligere å oppdage. Mitt syn er at det generelt var for lite fokus på personvern og håndtering av persondata på den tiden, samt mangelfull regulering. Det er også vanskelig å skille de to fra hverandre, da økt fokus fra store folkemengder «tvinger» frem politisk handling.

Direktivet fra 1995 vil bli erstattet av GDPR innen våren 2018. GDPR beskytter og styrker brukernes rettigheter når det gjelder tjenestetilbydernes håndtering av persondata. P3P manglet gjennomslagskraft, og en av de var en regulering tilsvarende GDPR.

PDS er på mange måter en videreutvikling av P3P konseptet. Og et PDS-økosystem er en videreutvikling av PDS, som passer bedre med den nåværende teknologiske utviklingen. Med dette menes at det fokuseres sterkere på plattformutvikling som andre mindre aktører kan benytte seg av. Plattform-konseptet (som et PDS-økosystem bør være) vil gjøre det mer oversiktlig for brukerens persondata trafikk og raskere for tjenestetilbyderen å tilpasse applikasjonen til brukeren.

PDS (Personal Data Store)

PDS beskriver et system som lagrer all data og metadata om en bruker. Det finnes mange forskjellige definisjoner av PDS. Noen gir begrepet flere egenskaper, andre færre. Her bruker vi begrepet i en ganske grunnleggende form hvor vi utvider begrepet til «PDS-økosystem» for å legge til egenskaper som vil gjøre modellen attraktiv. Andre termer som beskriver det samme som PDS er: personal data vault og personal data locker.

En grunnleggende PDS innehar en enkel protokoll for deling og innhenting av data. Flere abstraksjonslag som analyseverktøy og mer komplekse handlinger vil gjøre at konseptet faller inn under PDS-økosystem. PDS definisjonen jeg viser til under er kommet fram til av forfatter og Vivek Kaul.

Funksjoner hvor «grad 1» tilhører bare PDS og «grad 2» tilhører PDS-økosystem kan oppsummeres som:

Grad 1: - Sikker lagring av persondata.

- Lagring av ID-kreditter (navn, personnummer, adresse o.l.)
- Se hvem du deler hvilke data med, og mulighet for endring
- Standard samtykkekontrakter for bruk at data.
- Muligheten til enkel prosessering og innhenting av data.

Grad 2: - Analyse og mer avansert prosessering av data.

- Komplekse strukturer for deling av data.
- Multi-party-computation. Muliggjør avansert form for anonymisering.

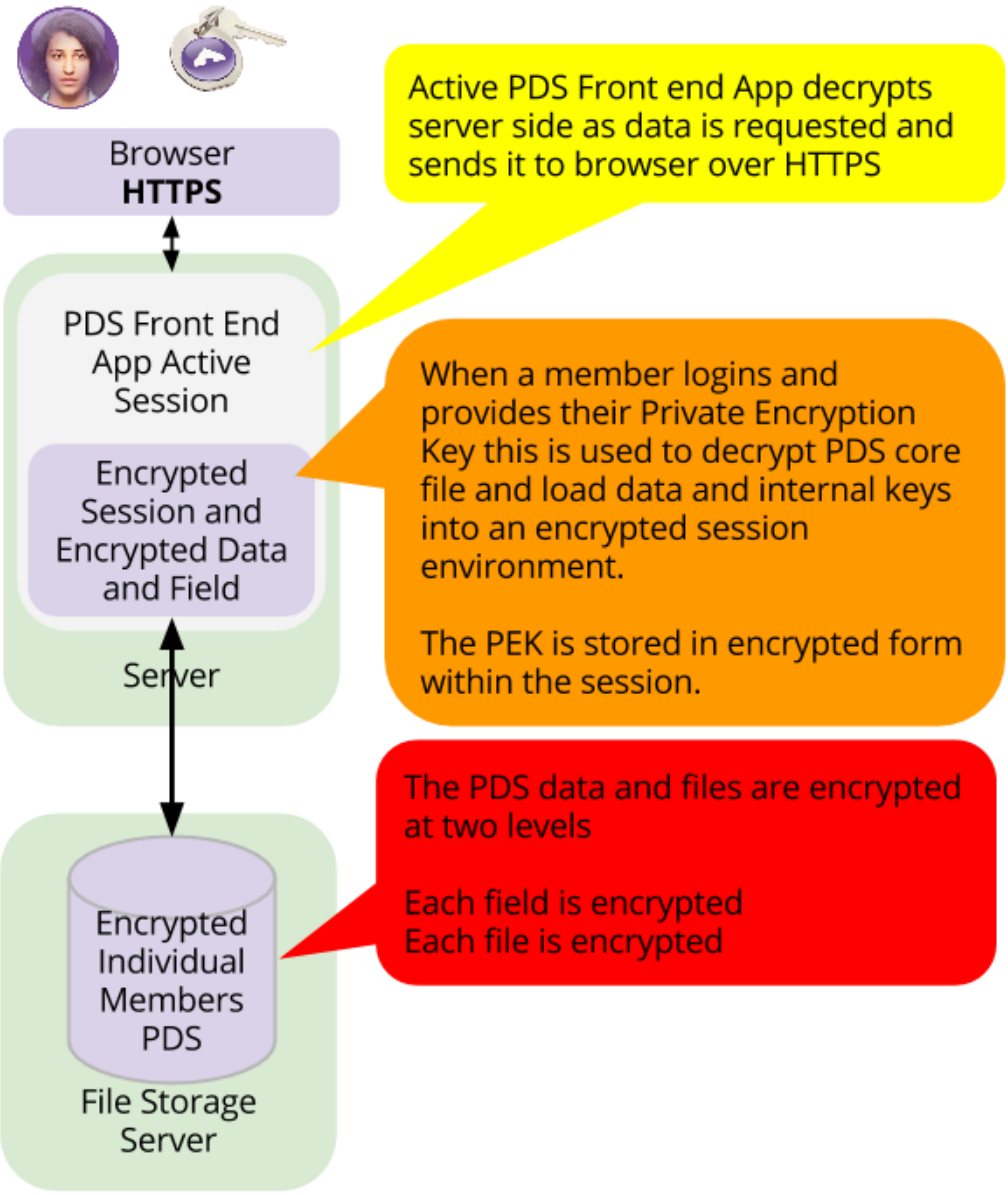
Det finnes flere elementer man kan legge til i «grad 2», men dette viser til mulighetene og en kort beskrivelse av hva en PDS/PDS-økosystem er. Videre beskrives Mydex CIC, OpenPDS og Enigma. Alle er eksempler på PDS-økosystemer.

Mydex CIC

Mydex CIC (Community Interest Company) har laget en PDS-økosystem for sikker lagring og deling av persondata. De beskriver hvordan de etisk sett er forskjellige fra dagens persondata standard, men jeg syntes det var vanskelig å finne eksplisitte beskrivelser på hvordan deres modell faktisk fungerer. For eksempel så beskriver de «hva de ikke gjør» i avsnittet «businessmodell», men ikke hvordan de faktisk tjener penger.

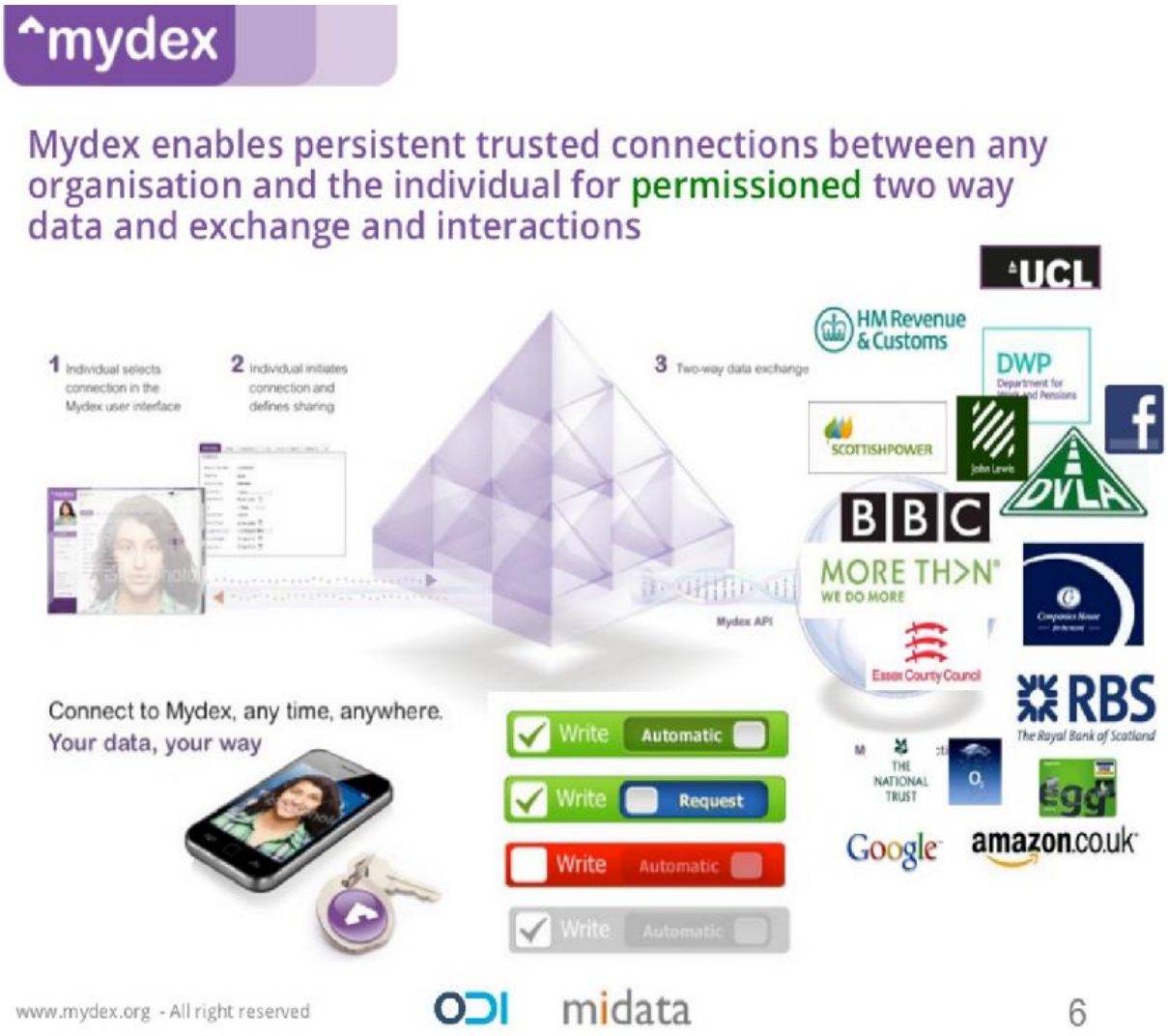
Arkitektur

Mydex bruker en skytjeneste hvor alle individer/brukere har sin egen PDS og det er bare brukerne som har administrator-tilgang til sin PDS. Et API sammen med standardiserte kontrakter blir brukt for å gi tjenestetilbydere tilgang til persondata, om brukeren (eieren av PDS'en) aksepterer det. Her er en tjenestetilbyder en av de som bruker Mydex sitt API, som foreløpig er 24 organisasjoner. I teorien kan alt fra Facebook til CandyCrush fungere som en samarbeidende tjenestetilbyder. For å illustrere arkitekturen ytterligere skal jeg se på to modeller, hvor den første er deres sikkerhetsmodell og den andre er mer en beskrivelse av informasjonsflyt.



Figur 1: Mydex sikkerhetsmodell

Figur 1 (dev.mydex.org): Her beskrives sikkerheten rundt lagring og tilgangsprosessen. Når en bruker kobler til en tjenestetilbyder, sendes data som tidligere har blitt samtykket om fra «File Storage Server» til «PDS Front End» og videre til tjenestetilbyderen. Først når brukeren logger inn/aktiverer Mydex kan «Front End» dekryptere data fra File Storage Server og sende det videre til tjenestetilbyderen her representert ved «Browser». Mydex bruker Eduserv som hosting tjeneste. Så med mindre man velger noe annet, blir data lagret i Storbritannia og drevet av «non-profit» bedrifter som har et godt rykte (Mydex CIC).



Figur 2: Mydex informasjonsflyt

Teksten tilhørende 1,2 og 3 i figur 2 er som følger; 1: Individet velger forbindelse i Mydex sitt brukergrensesnitt. 2: individet initierer forbindelse og definerer deling. 3: to-veis deling av data. Pyramiden representerer brukerens persondata. Tjenestetilbyderne til høyre integrerer seg med Mydex APIet som gir brukeren og tjenestetilbyderne mulighet til å utveksle informasjon. Brukeren har siste ordet og kontrollen på persondata-delingen, dette er representert de grønne og røde «sliderne» under pyramiden.

Trust framework

Mydex sitt rammeverk rundt tillit er definert av et sett med juridiske og tekniske regler som alle deltakende parter må akseptere (Mydex). Tosidig deling av informasjon mellom individet

og organisasjonen. Denne metoden reduserer risikoen for eksponering av flere brukernavn og passord.

All datadeling er gjennomført ved bruk av en standard for deling. Det er i delingskontrakten definert hvilke data som skal deles og hva data skal brukes til. Individet styrer prosessen for tilgang til persondata.

OpenPDS

OpenPDS er mer et PDS-økosystem enn bare en PDS. Dette er fordi de har en PDS-backend (hvor alt lagres i databaser) og en PDS-frontend som tar seg at kommunikasjonen mellom applikasjoner og backenden. Den er fortsatt i utviklingsstadiet og er konstruert av studenter fra MIT (Massachusetts Institute of Technology). Arkitekturen og funksjonene er enkle å forstå og oversiktlige.

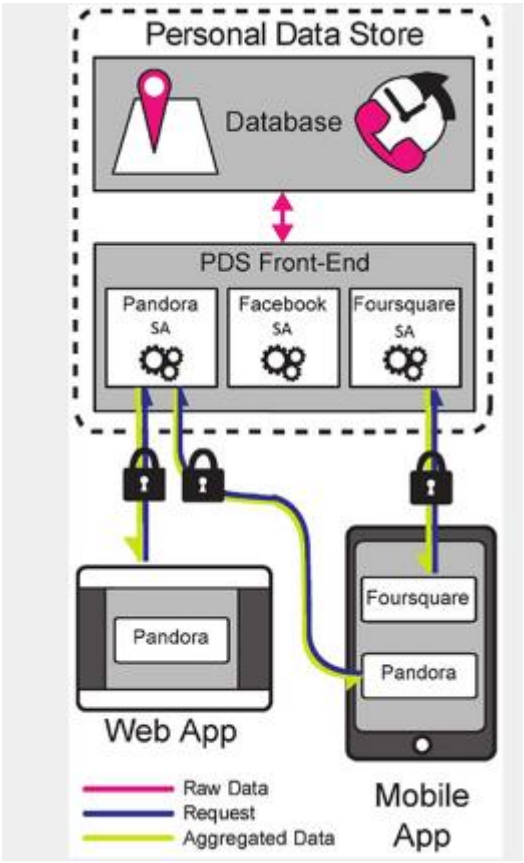
> *«We believe that a New Deal on data is needed. When it comes from data, "ownership" should to be thought of according to the old English common law. Data ownership would therefore be defined as the rights of possession, use, and disposal instead of a literal ownership.»* (OpenPDS)

De har fokusert på å sørge for at individet kan være helt anonym på nettet om de selv skulle ønske det.

Arkitektur

Her er et utdrag fra en forklaring av OpenPDS: LBSinc representerer en tilfeldig applikasjon som spør etter persondata.

> *"LBSinc web or mobile app sent a request to the user's openPDS. The request is passed on to the LBSinc SA module, which requests access to the database in order to retrieve the metadata needed to compute the answer. The SA module computes the answer, which is then validated by the PDS Front-End and send back to the web or the mobile app."* (OpenPDS, 2014)



Figur 3: OpenPDS arkitektur

Som bildet over viser er både databasen og frontenden en del av deres PDS terminologi. Pandora, Facebook og Foursquare er alle tjenestetilbydere som er ute etter data fra PDS-økosystemet. Når frontenden mottar en spørring fra en tjenestetilbyder, sendes rådata fra databasen til «frontend-processor» som her heter SafeAnswers, som igjen aggregerer et svar på spørringen eller benytter et ferdig prosessert svar. Dette svaret sendes tilbake til tjenestetilbyderen. I OpenPDS lagres rådata i databasen og det er bare frontend som har tilgang på denne data.

SafeAnswers

SA (SafeAnswers) moduler aksesserer backend databasen for å hente fram rådata som den kan bruke til beregningen av et svar. Den har bare tilgang til data som det ble samtykket om tidligere, da modulen ble installert. Etter at den har hentet data fra databasen vil den sende ut det beregnede/aggregerte svaret til frontend som i sin tur validerer svaret og sender det til applikasjonen som tilhører den aktuelle SA-modulen.

SA har også den egenskapen at den tilbyr «multiparty-computation». Dette kan for eksempel brukes for å skaffe seg data om en bestemt gruppe individer, men hvor det ikke vil være mulig å finne ut hvilket individ som hadde den bestemte data. Dette fungerer, om det er gitt tillatelse til det, på et overordnet lag. Slik at man kan stille en gruppe individer som passer den bestemte demografien, et spørsmål. Algoritmer kjører en form for mesh-kalkulering mellom individene og gir et svar når den er ferdig. Prosessen og kalkuleringen er gjort slik at det ikke vil bli mulig å spore tilbake informasjon til den enkelte.

Enigma

Enigma prosjektet fra MIT er det siste PDS-økosystemet vi skal se på. Det fortsatt i betafase og blir på nåværende tidspunkt testet. De kombinerer offentlig blockchain teknologi, som sikrer korrekthet og gjennomsiktighet på data og transaksjonene, med en privat del som skal sikre at bare du som eier/bruker har tilgang på din rådata. rådata blir splittet opp og sent til forskjellige steder og holdt styr på gjennom en DHT (Distribuert Hash-tabell). DHT lagrer referanser til data men ikke data selv. Tilgang på data er programmert inn i blockchainen, hvor Enigma bruker et API for å kontrollere dette.

> *"A peer-to-peer network, enabling different parties to jointly store and run computations on data while keeping the data completely private. Enigma's computational model is based on a highly optimized version of secure multi-party computation, guaranteed by a verifiable secret-sharing scheme."* (Zyskind, Nathan, & Pentland)

Blockchain teknologi

Nøkkelkonsept i blockchain teknologi er å kvitte seg med mellommannen. Arkitekturen består av et distribuert node-nettverk hvor alle noder har en log-kopi av alle tidligere transaksjoner. Alle historiske transaksjoner av enten bitcoin eller annen digital data lagres i en åpen hovedbok. Tilgang til den data er gitt til de som har den digitale nøkkelen. Halve nøkkelen er privat, bare eieren har denne delen. Den andre delen er offentlig og alle har tilgang på denne.

Mydex, OpenPDS og Enigma er tre forskjellige PDS-økosystemer som har valgt forskjellige retninger, men det de har til felles er at de er mer brukersentrerte enn dagens standard for

håndtering av persondata. Jeg vil avslutte bakgrunn-kapittelet med å se på noen prinsipper man bør ta hensyn til når man utvikler et IT-system som håndterer persondata. Dette vil være relevant i utviklingen av PDS-økosystem modellen.

Privacy By Design

«Privacy by design» beskriver syv fundamentale prinsipper om hvilke hensyn som bør tas når man utvikler et IT-system som blant annet håndterer persondata. Dette er relevant for oppgaven siden den skal utvikle en modell for et PDS-økosystem.

En kanadisk forsker ved navn Ann Cavoukian utviklet disse syv fundamentale prinsippene for «Privacy by Design» (Cavoukian, 2011)

1: Proaktiv ikke reaktiv; forebyggende fremfor avbøtende.

2: Personvern som standard; Graden av personvern bør være på maksimum som standard.

3: Personvern etablert i designet; og ikke som en tilleggsfunksjon.

4: Full funksjonalitet; både sikkerhet og personvern er mulig å få til. Ingen kompromisser skal inngås.

5: Ende til ende sikkerhet; hele livssyklusen til data skal være sikker, fra start til slutt.

6: Synlighet og gjennomsiktighet; alle komponenter og operasjoner skal være synlige for alle brukere og tilbydere.

7: Respekter brukerens personvern; sørg for å ha det brukersentrisk orientert.

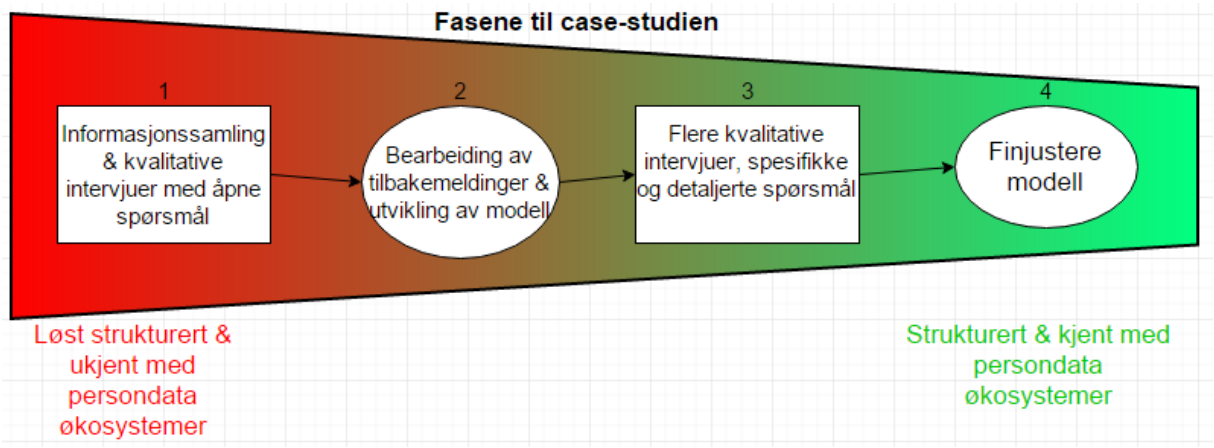
Ved at man utvikler et PDS-økosystem for håndtering, lagring og prosessering av persondata blir disse prinsippene bedre ivaretatt. Dette er fordi tjenestetilbydere fremfor alt skal utvikle en tjeneste, som gjør at sikring av persondata blir nedprioritert. Når man utvikler et PDS-økosystem står disse prinsippene høyeste prioritet fra første linje kode, fordi man nettopp er ute etter å sørge for en sikker håndtering av persondata.

3 Metode

I dette kapittelet beskrives oppgavens metodologiske fremgangsmåte i studien av håndtering, prosessering og lagring av persondata. Først kommer en kort oppsummering av hele prosessen hvor jeg kort beskriver de ulike stegene/fasene i studien. Dette etterfølges av en beskrivelse av hva en case-studie er og hvorfor jeg valgte en case-studie i denne oppgaven. Til slutt beskriver jeg i studiens fire faser og reflekterer kort over styrker og svakheter ved min tilnærming.

Case-studie delt opp i fire faser

Fokuset i denne case-studien er utviklingen av en modell for et PDS-økosystem. Denne modellen har et brukersentrisk perspektiv samtidig som den er interessant for tjenestetilbyderne. Brukeren har i denne modellen både rettslig og teknisk eierskap til sine persondata. Måten vi kom fram til denne modellen var ved først å avgrense området til fenomenet og finne aspekter som er sentrale for håndtering av private data. Senere ble en analyse av tre brukersentrerte PDS-økosystemer gjennomført, slik at vi fikk gå i dybden på modellenes (PDS-økosystemene) funksjoner. Her er «funksjoner» arkitektoniske løsninger for lagring og deling av data. Hvor vi til slutt inkorporerte de favoriserte funksjonene sammen med oppgavens modell.



Figur 4: Fasene til case-studien

Figur 4 illustrerer den overordnede prosessen i denne case-studien. De ulike fasene er nummerert 1 til 4, hvor de rektangulære boksene betegner faser hvor kvalitative intervjuer ble

utført. Gjennomgang av tilbakemeldinger og utvikling av modellen ble gjennomført i de runde/ovale fasene. Overgangen fra rødt til grønt viser til graden av ukjente aspekter og egenskaper ved PDS-økosystemer, hvor sterkere grønn farge betyr klarere visjon og konkretisert modell. Dette var en prosess hvor vi lette etter de optimale egenskapene (fra et brukersentrisk perspektiv), relevante funksjoner og avgrensninger (hva som ikke skulle inkluderes). En detaljert prosessbeskrivelse kommer senere i kapittelet. Under oppsummeres de fire fasene:

1: En eksplorerende tilnærming til feltet (håndtering, lagring og prosessering av persondata) hvor målet var å identifisere problemstillingen og avgrense fenomenet. Her ble det kartlagt hva som var sentralt i oppgaven og hva som falt utenfor. Detaljerte tekniske beskrivelser av protokoller og matematiske formler er eksempler på aspekter som falt utenfor kjernen til denne oppgaven. Eksempler på aspekter som oppgaven vektlegger er; Regulering, lagring- og tilgangs-modeller, brukersentrisme og tilpasning til markedet. Omentrent seks måneder ble lagt ned i denne fasen. Fasen ble avsluttet ved at jeg intervjuet tre sentrale personer over email eller Google docs samarbeid.

2: Fokus ble endret gradvis over fra eksplorerende til en mer deskriptiv utførelse. Her var det sentralt å beskrive fenomenet i den virkelige verden og gi en oversikt over årsakene til hvorfor fenomenet hadde utviklet seg til dagens standard for håndtering av private data. Her gikk jeg igjennom tilbakemeldingene fra fase en og fulgte opp på relevant faglitteratur. Denne fasen resulterte i utviklingen av en modell for håndtering, lagring og prosessering av private data. En modell for et persondata-økosystem/PDS-økosystem. Denne modellen er delvis en samling av funksjoner som er utviklet av andre og en del utviklet av oppgavens bidragsytere.

3: Målet i denne fasen var å få eksterne til å kritisk vurdere modellen og evalueringskriteriene: «gjennomsiktighet», «intensjon» og «enkelhet» som ble brukt til å analysere OpenPDS, Mydex CIC og Enigma. Sentralt var å finne de riktige intervjuobjektene og stille dem de riktige spørsmålene, slik at jeg i siste fase kunne bruke tilbakemeldingene for å finjustere modellen.

4: Siste fase gikk ut på ferdigstillelse av modellen i sin helhet, samt belyse dens styrker og svakheter. Konklusjonen viser i hvilken grad vi har klart å lage en sikker, gjennomsiktig og brukersentrisk modell for et PDS-økosystem.

Metodens rammeverk

Kort fortalt er fenomenet: håndtering, prosessering og lagring av persondata. Forskningsspørsmålet er: *hvordan kan et PDS-økosystem løse dagens utfordringer ved håndtering, prosessering og lagring av persondata?* Disse utfordringene og mulige løsninger på dem utforskes ved å se på forskjellige caser (Mydex, OpenPDS og Enigma). Jeg foreslår til slutt en generell modell for å løse de største utfordringene.

Case-studie

En case-studie er en studie om et fenomen i den virkelige verden, og hvor konteksten til fenomenet er sentral for å kunne forstå utviklingen til fenomenet. Her er fenomenet håndtering av persondata og konteksten er persondata-reguleringer, persondatainnsamlingskultur hos tjenestetilbyderne og økt fokus på brukersentrisme hos brukerne. Det finnes flere definisjoner og avgrensninger til hva som faktisk er en case-studie. Jeg valgt å bruke Yin (Yin, 2014) som referanse til denne oppgaven. Yin har en todelt definisjon av case-studie:

1: "A case study is an empirical inquiry that"

- "investigates a contemporary phenomenon (the case) in depth and within its real-world context, especially when"

- "the boundaries between phenomenon and context may not be clearly evident."

2: "A case study inquiry"

- "benefits from the prior development of theoretical propositions to guide data collection and analysis"

Punkt 1 forteller at case-studie som metode passer for et samtidsfenomen som man søker etter å undersøke i dybden og hvor man ser på konteksten til fenomenet. Denne oppgaven ser på lagring, håndtering og prosessering av private data/persondata (fenomenet), og hvordan dette har blitt normen (konteksten). Grensene mellom fenomenet og konteksten fikk jeg hjelp til å greie ut i den første intervjurunden. Grunnen til dette var at det ikke var klart hvilke elementer for var med på å forme dagens standard, samtidig som det var greit å få andres perspektiv på problemstillingen.

Til punkt 2 så har jeg i denne studien analysert tre modeller fra ulike PDS prosjekter, hvor jeg sammen med veilederne kombinerer funksjoner fra de ulike modellene til bruk i vår egen modell.

Yin beskriver videre hvordan en case-studie er en lineær, men iterativ prosess, i dette legger han at det finnes rom for redesigning av studien etter datainnsamling eller analyse. Dette passer med vår prosess, fordi vi ventet med å ferdigstille modellen etter tilbakemeldingene fra første intervjurunde, samt åpnet for redesigning etter andre intervjurunde. I tillegg til dette ble oppgaven skrevet om to ganger, grunnet ny informasjon ble gjort tilgjengelig.

Som et siste punkt i beskrivelsen av hvorfor dette er en case-studie, forteller Yin at en case-studie som metode bør vurderes om problemstillingen(e) starter med «hvordan» eller «hvorfor». Sentrale spørsmål i denne studien er; hvordan prosessering og håndtering av persondata blir utført på nåværende tidspunkt, samt hvorfor denne tilnærmingen ble normen. Og til slutt, hvordan vi kan forbedre PDS-økosystem modellene analysert i denne studien.

Eksplorerende og deskriptiv.

Metodikken brukt i denne oppgaven er ikke en «ren» eksplorerende eller deskriptive case-studie, men deler av prosessen preges av de forskjellige tilnærmingene.

En eksplorerende case-studie forsøker å finne de riktige forskningsspørsmålene, konkretisere problemstillingen og belyse sentrale aspekter ved et fenomen. Eksplorerende case-studie skal også finne den riktige fremgangsmåten for innsamling av data og metode design for fenomenet. Dette kan jeg relatere tilbake til starten av prosessen, hvor vi bare hadde en generell retning å forholde oss til. Vi visste at dagens håndtering av private data er langt fra optimal, spesielt med hensyn til brukernes kontroll over egne data. I tillegg hadde vi sett litt på PDS-konseptet. Gjennom en flere måneder lang prosess klarte vi å finne riktige forskningsspørsmål, framgangsmåte og datainnsamlingsmetode. Denne prosessen bestod av kvalitative intervjuer, diskusjoner med faglig sentrale personer og utforskning av litteratur.

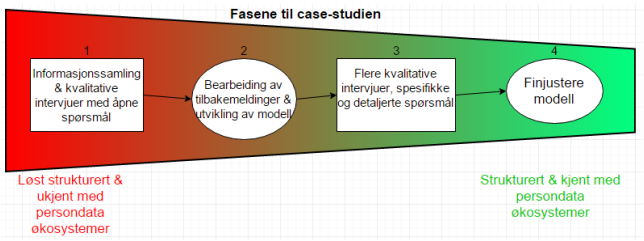
Målet med en deskriptiv fremgangsmåte er å beskrive fenomenet i sin kontekst, i den virkelige verden. For å kunne utvikle en modell som faktisk var relevant, ble vi nødt til å se på hvilke aspekter som ville gjøre den det. Det ble da viktig for oss i finne ut hvorfor dagens modell var normen og hvorfor den var en suksess. Dette gjorde vi ved å se på politiske

reguleringer, insentiver fra bedriftenes standpunkt og hvordan folkets økende forståelse av personvern setter det på dagsordenen.

Beskrivelse av selve prosessen

Fase én

↓



Prosessen begynte med PDS konseptet og hva slags funksjonalitet en «god» PDS bør tilby. Utbredt bruk av PDS vil løse mange utfordringer dagens IT-løsninger møter, hvor en av dem er muligheten til å identifisere samme bruker over forskjellige IT-systemer. Selv om de ulike systemene ikke har noen fellesnevnere. Denne praktiske problemstillingen møtte jeg med min startup-bedrift, og gjorde meg interessert i hvordan en PDS kan løse dette på en generell basis.

Som en naturlig fortsettelse så jeg på hvordan jeg kunne evaluere egenskapene til de ulike PDSene som fantes ute på markedet nå. Og hvilke ekstra funksjoner noen av de tilbyr. Dette strandet litt siden det bare fantes en PDS som hadde kommet forbi utviklingsfasen og inn på markedet. Så jeg bestemte meg for å se mer på hva som er basisen til PDS konseptet og hvilke egenskaper som er absolutt nødvendig for å sikre: gjennomsiktighet, teknisk eierskap, samtykke kontrakter og sikring av lang levedyktighet. Vår egen definisjon av PDS ble formulert for å støtte videre arbeid og på grunn av svak konsensus over definisjonen av konseptet hos gjeldene PDS-utviklere.

På et tidspunkt i første fase møtte vi Teknologirådet, et offentlig organ for rådgivning til Norske politikere. Under møtet fant vi ut at de var mer interessert i hva vi kunne vise til dem, enn omvendt. Jeg fikk inntrykk av at nye løsninger rundt personvernet og håndtering av persondata var på agendaen.

Noen veileder-møter senere bestemte vi oss for å ta kontakt med folk som vet mer om PDS konseptet og de ulike aspektene berørt av PDS terminologien. Vi ble enige om at en form for intervju ville vært praktisk. Derfor utviklet jeg et spørreskjema (se vedlegg 1) for å belyse sentrale og generelle aspekter rundt fenomenet, som hvorfor dagens standard er som den er og hvorfor går det så trått med utviklingen av PDS'er. Deretter ble to pilotintervjuer gjennomført, et over email og et ansikt til ansikt. Resultatet av dette var små justeringer av spørsmålsskjemaet, etterfulgt av flere sentrale personer ble kontaktet over mail og spurt om de vil være med på et intervju. Av de som sa ja inkluderte utviklerne av OpenPDS og Mydex CIC.

I første intervjurunde var strukturen i spørreskjemaet lagt opp slik at intervjuobjektene først fikk en kort beskrivelse av målet og meningen med studien, etterfulgt av et sett generelle spørsmål om dagens håndtering av private data. Avslutningsvis fikk intervjuobjektene et par spørsmål rettet mot deres fagfelt/ekspertise. Alle spørsmålene her var relativt åpne slik at intervjuobjektene fikk uttrykke seg fritt. Dette ble gjort for å få større variasjon i svarene, slik at jeg fikk en bedre oversikt over feltet.

En del av tilbakemeldingene fra pilotintervjuene var at noen av spørsmålene gikk litt over i hverandre og at PDS konseptet trengte bedre forklaring. Det var også noen kommentarer som påpekte at jeg burde utdype hvordan den nåværende geopolitiske situasjonen kan bremse reguleringen av private data. Her var terrorisme i et særskilt fokus.

Svaret fra rettsinformatikk, et fagområde som tar for seg juridiske problemstillinger innenfor informasjons teknologi, kom med tredje forsøket, og belyste problemstillingene rundt dataeierskap. Dette fikk meg til å innse enda mer hvordan dataeierskap, reguleringer og samtykke er et sentralt aspekt hos fenomenet.

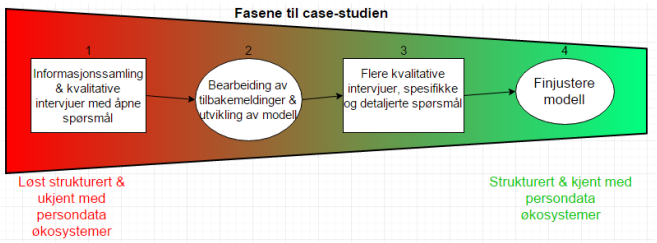
Tilbakemeldingen fra intervjuene med OpenPDS og Mydex ga meg generelt en bedre oversikt over de relevante aspektene til fenomenet håndtering, prosessering og lagring av persondata. OpenPDS fortalte at det kunne være lurt å skille mellom teknisk og juridisk eierskap til persondata og hvordan deres SafeAnswers-modul sikrer anonymiteten til brukeren. Mydex fokuserte på viktigheten av å kunne koble digital identitet til virkelig identitet, slik at alle parter i systemet kan stole på at de andre partene ikke er falske.

Mot slutten av første fase ble det klart at formålet med denne studien var å gi fremtidige PDS-utviklere en god grunnmur, både i form av teori og funksjonsbaserte løsninger. Slik at de vil

være bedre informert rundt problemstillingene fremstilt i denne studien. Det ville bli vårt bidrag til et brukersentrisk PDS-økosystem.

Fase to

↓



I denne fasen sluttførte jeg metodologien. Retningen var bestemt før denne fasen, men formalitetene kom på dette punktet. Andre hovedpunkter som skjedde i denne fasen var utformingen av modellen for persondata-økosystemet og evalueringen av OpenPDS, Mydex og Enigma. Jeg tok også beslutningen om at «insentiver», «gjennomsiktighet» og «design egenskaper/funksjoner» skulle være evalueringskriteriene. Disse kriteriene ble valgt av meg, på bakgrunn av tilbakemeldingene fra intervjuobjektene og gjennomgang av relevant litteratur.

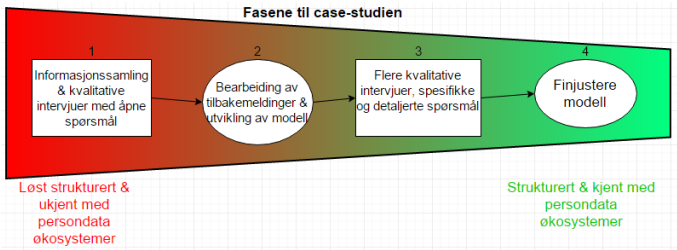
En detaljert beskrivelse av evalueringskriteriene står i analyse-kapittelet, men kort kan jeg si at «insentiver» var valgt på bakgrunn av de etiske og juridiske utfordringene ved dagens modell. «Gjennomsiktighet» ble valgt for å sikre en viss grad av brukersentrisme, slik at brukerne vet hva som skjer bak kulissene i PDS-økosystemet sitt. «Enkelhet» var et noe åpent kriteria, som muliggjorde ris og ros for forskjellige tekniske og arkitektoniske løsninger.

OpenPDS, Mydex og Enigma ble valgt fordi de var de mest eksponerte PDS-løsningene ute. Det fantes en del informasjon om dem, samt at jeg følte de representerte spekteret av PDS mangfoldet.

Utførelsen av modelldesignet kom etter flere diskusjoner med veiledere, samt gjennomgang av tilbakemeldinger fra intervjuobjektene.

Fase tre

↓



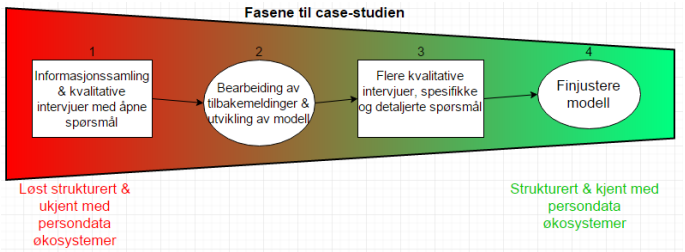
Her utviklet jeg spørreskjemaet for andre intervjurunde (se vedlegg 2). I dette skjemaet forklarte jeg begrepet PDS og spurte om insentiver, gjennomsiktighet og enkelhet var gode evalueringskriterier. Til slutt i skjemaet hadde jeg beskrevet modellen (PDS-økosystemet) og stilt et par spørsmål rundt den. Dette skjemaet hadde først spørsmål om hvorvidt man skjønte beskrivelsene mine. Disse ble litt ledende og lite nøytrale, så de ble fjernet. Jeg sendte da istedenfor skjemaet til 2 relativt tekniske bekjente som ga tilbakemelding på områder som var uklart. Resultatet av dette var at et konkret eksempel ble beskrevet i spørreskjemaet, samt andre mindre justeringer.

Vår modell og evalueringskriteriene var under fokus i andre intervjurunde. Her var spørsmålene mer på detaljnivå slik at vi fikk et bedre overblikk over modellen og om vi hadde oversett eller glemt viktige funksjonaliteter.

Skjemaet ble så sendt til noen utviklere. Tilbakemeldingene fra disse var at det var lettfattelig beskrevet og hadde ingenting å utsette på modellen. Et lengre intervju ble gjennomført med en professor ved UiO. Hans ekspertiseområde er persondata, noe som passet bra med kjernekunnskapen jeg var ute etter. Dette var et veldig ustrukturert intervju, hvor vi diskuterte ulike problemstillinger rundt personvern og persondata. Ett av de viktigste elementene jeg tok med meg fra dette intervjuet var at modellen må håndtere situasjoner som kan sammenlignes med «vouchere», eller «tusted authorities»

Fase fire

↓



Denne fasen er en avslutningsfase på studien. Her ble tilbakemeldingene fra intervjuene i fase 3 integrert i modellen. Modellen ble ferdiggjort og små justeringer i analysen av PDS-økosystemene gjorde evalueringskriteriene mer beskrivende. Forhåpentligvis vil denne studien hjelpe til med den videre utviklingen av PDS-økosystemer.

Tilbakemeldingene fra andre intervjurunde hadde en mindre påvirkende kraft på oppgaven enn første intervjurunde. Dette kan skyldes nok at i intervjuobjektene fra første runde var med på å belyse sentrale aspekter som er sentralt for utviklingen av et PDS-økosystem og håndtering av persondata. Muligheten for å bruke multiparty-computation og no-trust modeller, samt at det i dag er et skille mellom teknisk og juridisk eierskap til persondata er eksempler på sentrale tilbakemeldinger fra første intervjurunde. I den andre intervjurunden hadde intervjuobjektene mindre kjennskap til PDS-økosystemer enn de i den første runden, noe som kan være en kilde til svakhet i oppgavens metode.

Gjennom denne studien har jeg fått et mye bredere overblikk over relevante aspekter til håndtering av persondata og hvilke funksjoner som kan hjelpe til med å gjøre framtidens modeller mer brukersentrisk. Kort fortalt er det denne oppgaven jeg hadde trengt i starten av studien om jeg skulle utviklet et PDS-økosystem. Da hadde jeg fått et godt overblikk over; sentral litteratur, persondata-reguleringer og PDS-prosjekter for håndtering av persondata.

Neste side har en tabell som viser i hvilken fase de forskjellige intervjuobjektene ble intervjuet og hvilket bidrag de hadde. Pilotintervjuene er ikke tatt med.

Intervjuobjekt	Fase	Intervjumetode	Kompetanse	Resultat/ Svar	Anvendt resultat
Utvikler & kommunikasjonsansvarlig (Mydex)	1	Spørreskjema i Google docs	Software utviklere av PDS-økosystem	Omfattende beskrivelse Mydex og deres syn på PDSer	Ga meg oversikt over aspekter relatert til persondata håndtering, samt at alle parter i økosystemet må stole på at den andre parten eksisterer
Utvikler (OpenPDS)	1	Spørreskjema over email	Software utvikler av PDS-økosystem	Generelt overblikk over dagens persondata situasjon	Ga meg oversikt over aspekter relatert til persondata håndtering, samt multiparty-computation for anonymisering og det å holde rådata hemmelig
Rettsinformatiker	1	Spørreskjema over email	Rettsinformatiker Stipendiat - Institutt for privatrett	Generelt juridisk overblikk på persondata	Lenker til relevant litteratur og diskusjon. F.eks. Purtova
Universitetslektor (informatikk)	3	Ustrukturert intervju	Universitetslektor med interesse for personvern	Vouchere	Integrere vouchere i diskusjonen

Utvikler	3	Spørreskjema over email	IT-utvikler i Sannsyn AS.	Små kommentarer til modellen	Små justeringer av modell

Tilnærmingens styrker og svakheter

Den metodiske tilnærmingen var tilpasset tiden som var tilgjengelig og kunnskapen som fantes om PDS konseptet. Fenomenet (håndtering, prosessering og lagring av persondata) er et veldig vidt emne, med mange forskjellige og viktige aspekter som krysser hverandre på ulike måter. Valg måtte tas om hvilke aspekter som var relevante i denne oppgaven. En alternativ tilnærming kunne ha vært å gitt avkall på den generelle oversikten og heller sett på detaljløsninger, som valg og evaluering av protokoller eller andre mer tekniske og matematiske utbroderinger. Men jeg valgte altså å prioritere utviklingen av PDS-økosystem modellen og en generell oversikt over fenomenet.

Alternative metodiske tilnærminger

En annen aktuell metodologisk tilnærming jeg kunne tatt er en ren sammenligningsstudie, hvor jeg da hadde sammenlignet forskjellige PDS løsninger. For deretter å ha satt de opp mot hverandre ut i fra noen beskrivende kriterier. Fordelen med en slik studie er at jeg kunne ha gått mer i dybden på forskjellige PDS løsninger. Men da måtte jeg ha gitt avkall på tid jeg ville brukt på utvikling av modellen og diskusjon av sentrale aspekter rundt personvern og persondata.

Mulighet for ledende spørsmål og forklaringer

Siden spørsmålene krevde at intervjuobjektene hadde et minimum av forståelse om PDS konseptet, var det nødvendig for meg å forklare begrepet før jeg stilte spørsmål. Jeg følte at det ikke ville bli et stort hinder i første intervjurunde siden det ikke var nøyaktigheten av svaret jeg var ute etter, men heller øke vår forståelse av konseptet og ulike aspekter som påvirker det. Uansett er det er greit å være klar over at dette kan være en kilde til partiskhet.

Med partiskhet mener jeg her muligheten for at det brukersentriske perspektivet skal dreie spørsmålene i brukersentrismens favør.

I andre intervjurunde var detaljer mer i fokus, og hvordan jeg beskrev de ulike modellene vil nødvendigvis påvirke hvordan intervjuobjektet forstod fenomenet og konteksten. Her var jeg mer forsiktig i forhold til å unngå partiskhet. Perspektivet i denne studien er brukersentrisk, så total nøytralitet vil ikke være mulig.

4 Sentrale aspekter ved persondata

Dette kapittelet gjennomgår fire ulike aspekter ved håndtering, lagring og prosessering av private data og gir en gjennomgang på hvorfor de er sentrale i utviklingen av fremtidens persondataøkosystem. Meningen med dette kapittelet er å gi en forklaring på hvorfor tiden er moden for et PDS-økosystem slik det beskrives i kapittel syv. Aspektene som beskrives under er; General Data Protection Regulation fra EU, hvordan lagre persondata, sentralisering av identitet og anonymisering av persondata.

Siden denne oppgaven har en eksplorerende tilnærming til fenomenet er det sentralt å se på hvilke sentrale aspekter som gjør det mulig for en ny modell å slå igjennom på markedet.

GDPR

GDPR representerer den juridiske reguleringen av håndtering, lagring og prosessering av persondata. Reguleringen er ett av de mer sentrale aspektene som vil hindre at nye prosjekter ender opp som P3P. GDPR representerer den politiske viljen til å forsvare EU-innbyggernes rett til privatliv i den digitale sfæren. Uten GDPR ville det tatt mye lengre tid å få igjennom en seriøs persondata-reform av typen PDS. En av grunnen til dette; er at det er enklere å gjennomføre bestemmelsene i GDPR i et PDS-økosystem. I GDPR er det bestemt at brukeren har rett til å ha tilgang til alle persondata en tjenestetilbyder har lagret, i tillegg til at man har rett til å få dem slettet. Dette vil innebære at brukere for eksempel kan forlange å få tilgang til alle persondata lagret av Facebook, slik at de kan sette den inn i sin egen PDS, for så å kreve at Facebook sletter alle persondata. Brukerne kan så lage en ny Facebook-profil, hvor da Facebook blir nødt til å hente personopplysninger fra en PDS etter en felles bestemmelse ("consent form") mellom brukeren og Facebook. Dette krever at Facebook har utviklet funksjonaliteten for tilkoblingen mellom PDS-økosystemet og deres backend.

Hovedforskjellen fra dagens standard vil være at Facebook ikke lenger vil ha en grunn til å lagre dine personopplysninger for å kunne gi deg den beste brukeropplevelsen. Og når tjenestetilbydere ikke lagrer persondata, har man redusert muligheten for identitetstyveri og annet misbruk av personopplysninger/persondata.

GDPR sier også at det skal bli enklere å oppheve samtykket som tidligere har blitt gitt. I et tilfredsstillende PDS-økosystem vil det være enkelt å gjennomføre dette, siden man da har full oversikt over hvem som har tilgang på hvilke opplysninger. Dette vil være en enklere og sentralisert metode oppheving av samtykke. Om brukeren for eksempel har hørt negativ omtale om en bestemt tjenestetilbyder, kan brukeren enkelt gå inn i sitt PDS-økosystem som brukeren har kjennskap til og trekke samtykke tilbake. Å forholde seg til hver enkelt applikasjon/nettsides metode for å oppheve samtykket blir mer tungvint enn å kontrollere det i sitt eget PDS-økosystem.

En diskusjon som går nå (Kahn, 2016) er om GDPR er for streng på punktet om persondata samlet for et formål skal kunne brukes til noe annet. På den ene siden har man styresmakter som vil ha informasjon fra tjenestetilbydere for å kunne lokalisere kriminelle og redusere sannsynligheten for terrorisme og kriminalitet. På den andre siden må man skape tillit mellom tjenestetilbyder og bruker, samt vise hensyn til rettsvernet av persondata. Om man finner fram til et tilfredsstillende kompromiss i EU er det ikke sikkert at andre land utenfor EU vil akseptere løsningen.

Nylig ble en sjef for Facebook arrestert i Brasil etter en konflikt med Brasilianske myndigheter. Grunnen var at WhatsApp Inc. som eies av Facebook nektet å gi ut informasjon om spesifikke meldingsutvekslinger etter at Brasils domstol hadde beordret dette (Meyer, 2016). Det er vanskelig å se at denne interessekonflikten vil forsvinne med det første, om man fortsetter å prosessere persondata på samme måte. Nå var dette eksempelet i Brasil og ikke i EU, men den belyser fortsatt et viktig poeng. At styresmakter ikke vil stoppe med å forsøke å hente ut persondata fra tjenestetilbydere som lagrer og prosesserer mye av det.

Dette er en komplisert problemstilling som vi unngår i vårt PDS-økosystem. Siden data lagres hos brukeren vil det bli «umulig» å kreve den fra andre. rådata kan splittes opp og krypteres på en måte som gjør det særdeles vanskelig å få tak i tilstrekkelig mengde rådata for identifisering. Her vil kun brukeren sitte på oppskriften på hvordan rådata er fordelt og kryptert. Man vil selvfølgelig ikke fjerne problemet helt, siden noen gråsoner vil forekomme der man flytter digital informasjon mellom instanser. Eksempel på det siste er at internettilbyderen din har mulighet til å se datatrafikken.

Lagring av persondata

Hva har dukket opp de siste årene som kan tilby tilfredsstillende lagring av persondata?

Et av problemene med gratis eller billig lagring av persondata er at brukeren ofte selv blir produktet. Det vil si at bedriften som eier den fysiske lagringsplassen selger brukerens persondata videre for å finansiere lagringsplassen. Uansett om de selger persondata videre eller ikke er brukerne nødt til å stole på deres håndtering av persondata, enten blindt eller gjennom en trust modell. En trust modell er en modell som beskriver hvordan de sikrer brukernes persondata, ofte ved bruk at kryptering, der brukeren har eneste nøkkel.

«No-trust modell» er et begrep brukt om modeller som sikrer data på en slik måte at «part en» ikke trenger å stole på «part to» som tilbyr, i dette eksempelet, lagring. «Zero trust» eller «no-trust» blir brukt om hverandre og forskjellige definisjoner finnes, men det som ofte går igjen er «never trust, always verify» (Au, 2014).

Skal man ha en framtidsrettet modell for lagring av persondata, må man også ta hensyn til at verden utvikler seg. Lover, reguleringer og styresmakter i dag, vil ikke nødvendigvis være det samme om 20 år. Modeller rundt offentlig lagring og håndtering av persondata ble vurdert og forkastet i denne oppgaven. I stedet for å ha tillit til myndighetene, ville en no-trust modell sikre bedre konfidensialitet av persondata i årene fremover. I tillegg vil det være i styresmaktenes interesse å samle persondata for å hindre nye terrorangrep. Denne oppgaven har tatt valget å se brukersentrisk på interessekonflikten (personvern vs. terror) ved å gjøre det vanskelig å samle persondata uten brukerens samtykke. Dette kan gjennomføres med riktig no-trust modell.

En av de «nye» spennende teknologiene som har muligheten til å tilby en no-trust modell, er blockchain. Blockchain i seg selv er ikke privat. Dette kommer av den globale oppslagsboken som brukes til å verifisere noder's transaksjoner. Dette betyr at blockchain kan sikre systemets- og dataens integritet og tilgjengelighet, men ikke konfidensialitet. Kombinert med en eller flere tilleggsfunksjoner kan blockchain teknologien fungere mer konfidensielt.

«Storj» og «Enigma» er to nye modeller som gjør akkurat dette. De legger noen funksjoner til blockchainen slik at den kan sikre konfidensialitet rundt data. Det er utenfor denne studien å gå i detaljert forklaring rundt hvordan protokollene fungerer, men begge teknologiene bruker peer-to-peer protokoller sammen med blockchain for å oppnå ønsket grad av KIT (konfidensialitet, integritet og tilgjengelighet).

Mulighetene er mange med nye desentraliserte modeller som blockchain og nye protokoller for peer-to-peer. Om det ikke er blockchain (som begynner å møte sine begrensninger (Brock, 2016)) som fører privat lagring inn i fremtiden, er mulighetene store lignende modeller. For å oppsummere bør lagring av persondata gjennomføres med riktig no-trust modell, slik at kontroll over tilgang kun gis til bruker. Eksempler på lovende teknologier er kombinasjonen av blockchain og peer-to-peer protokoller.

Identifisere samme bruker på tvers av ulike systemer

De fleste kan nok kjenne seg igjen i følelsen av å måtte registrere seg nok en gang for å få tilgang til en tjeneste, enten det er mobilapplikasjon eller nettside. Man må overveie hvilket passord man skal velge denne gangen, eller velge det tryggeste alternativet, å lage ett nytt passord, som man har glemt neste gang man resetter nettleseren.

PDS-økosystemet kan erstatte registeringssiden med et visningsvindu som beskriver hvilke opplysninger denne tjenestetilbyderen ønsker persondata, og hvor brukeren samtykker eller avslår deling. Så blir det opp til tjenestetilbyder å håndtere situasjoner der brukere avslår ett eller flere samtykkepunkter.

Noen ganger trenger tjenestetilbydere å utveksle persondata med hverandre for å yte service til en bruker. Hvordan skal de få til det om de bruker forskjellige identifikasjonskreditter for deres brukere? Problemet med å identifisere brukere på tvers av IT-systemer kan et PDS-økosystem enkelt løse. Ved å sentralisere data rundt brukeren og legge til rette for en plattform som lar ulike applikasjoner (ulike systemer) snakke med hverandre, har man snudd hele problemstillingen på hodet og samtidig gjort problemet enkelt løsbart. Samtykket for at applikasjoner skal kunne dele informasjon med hverandre må selvfølgelig være gitt av brukeren. En av utfordringene ved å dele persondata mellom tjenestetilbydere i dagens

situasjon, er at man i noen omstendigheter er nødt til å spørre brukeren om lov (avhengig av sensitiviteten på dataen) og informere om delingen.

Ved bruk av PDS-økosystemer kan registrering/innlogging og deling av persondata mellom tjenestetilbydere gjøres mer strømlinjeformet. Ved å sentralisere identiteten hos brukernes PDS-økosystemer vil kontroll over samtykkekontrakter enkelt gi oversikt over hva brukeren deler med tjenestetilbydere og hva tjenestetilbydere deler seg imellom.

SafeAnswers o.l.

Håndtering og prosessering av persondata er like viktig som lagringen av den. PDS-økosystemet vil begrense ulike applikasjoners tilgang til å spørre om persondata, fordi data sendt er data kompromittert. Dette betyr at en metode for å anonymisere persondata vil gjøre at terskelen for å dele ut data blir lavere. PDS-økosystemet må også kunne skille mellom tjenestetilbydere som trenger tilgang på sensitive data for å yte tjenesten og tjenestetilbydere som ikke trenger sensitiv persondata. Banken er et eksempel på en tjenestetilbyder som er nødt til å få tilgang på brukerens mest sensitive persondata for å yte den servicen den skal. Dernest tjenestetilbydere som sosiale medier, spill og andre applikasjoner hvor link mellom virkelig identitet og digital identitet ikke er nødvendig, kan anonymiseringsmodeller benyttes.

Den anonymiseringsmetoden denne oppgaven har sett nærmere på er brukt av «SafeAnswers» (modul til OpenPDS) og «Enigma» og heter multiparty-computation (MPC). Denne metoden fungerer på følgende måte; en tjenestetilbyder vil ha svar på «hvor mange i Oslo bruker min applikasjon?» sendes spørring til et av brukernes PDS-økosystem (i dette eksempelet representerer et PDS-økosystem en node). PDS-økosystemet som har innebygd en modul for å gjenkjenne og benytte seg av MPC starter så prosessen med å sende spørringen videre til andre PDS-økosystemer. Til slutt blir samlingen av PDS-økosystemer enige om et svar og sender et «tall på hvor mange som benytter tjenesten i Oslo» tilbake til det PDS-økosystemet som mottok spørringen slik at det kan sendes tilbake til tjenestetilbyderen. Det er her ikke mulig å spore hvem som svarte hva når MPC benyttes slik som dette. MPC dekker kun spørringer om grupper og ikke spørringer direkte til en enkelt bruker. Spørringer direkte til en bestemt bruker, kan besvares gjennom ferdig aggregerte/kalkulerte svar. For eksempel, hvilken filmsjanger liker brukeren best. I stedet for å gi alle filmene brukeren har sett, gir SafeAnswers ut et ferdig kalkulert svar. Det sentrale her er at kalkuleringen av hvilken

filmsjanger brukeren foretrekker skjer i brukerens PDS-økosystem og ikke i tjenestetilbyderens «backoffice».

Aspektene drar i samme retning

Kombinasjonen av GDPR, brukersentrisk lagring av persondata, sentralisering av identitet og anonymisering av persondata gjør det aktuelt å se på et PDS-økosystem for håndtering, lagring og prosessering av persondata. Den nye reguleringen tvinger tjenestetilbyderne til å håndtere persondata på en mer ressurskrevende måte. Dette gjør det mulig å komme med en konkurrerende modell som ivaretar personvernet bedre. Samtidig er det sentralt å se på hvordan nye digitale økosystemer og plattformer gjør det mulig å kunne benytte seg av en applikasjon uten å gå igjennom en omfattende registreringsprosess.

Det her viktig å få med seg fremskrittene som gjøres rundt lagring og håndtering av persondata. Slik at vi kan benytte modeller som tilbyr sterkere anonymisering og bedre kontroll over egne persondata.

5 Analyse av tre utvalgte PDS-økosystemer

I dette kapittelet skal jeg analysere tre forskjellige PDS-økosystemer ut i fra kriteriene gjennomsiktighet, intensjon og enkelhet. Disse kriteriene ble valgt på bakgrunn av tilbakemeldingene fra intervjuene i fase en. Målet med dette kapittelet er å vise til hvor vi er i dag, og hvilke teknologier som finnes eller er under utvikling. To av de tre PDS-økosystemene benytter blockchain/peer-to-peer teknologi, noe som passer bra siden det eksisterer flere blockchain/peer-to-peer-løsninger enn uten. Det resterende markedet blir dekket gjennom Mydex sin løsning.

Det er verdt å merke seg at dette er en overordnet analyse av de tre modellene, hvor informasjonen er hentet mest fra andres forklaringer og beskrivelser av systemene. Det ble ikke gjort noen dyptgående matematiske analyser av protokoller osv. Med dette menes at oppgaven ikke har stress-testet protokollene eller modellene matematisk eller funnet ut hvor mange brukere som skal til for at systemet er selvopprettholdene. Noen av blockchain og peer-to-peer teknologiene trenger et visst antall brukere for at teknologien skal fungere optimalt. Dette faller utenfor oppgavens og analysens «scope».

Evalueringskriteriene

I denne oppgaven ble «gjennomsiktighet», «intensjon» og «enkelhet» valgt som evalueringskriterier og bli definert som følger:

Gjennomsiktighet defineres som «i hvor stor grad har bruker, tjenestetilbyder og utviklere mulighet til å observere persondataflyten og arkitekturen i PDS-økosystemet». Et eksempel kan være om brukeren har mulighet til å se alle som har en lese-rettighet til en bestemt del av rådata. Alle parter skal ikke ha gjennomsiktighet til alle deler av systemet, men de skal kunne se alt som er relevant for dem. Med dette menes at tjenestetilbydere ikke skal kunne få ubegrenset tilgang til en brukers rådata, men heller se mulighetene til PDS-økosystemet og være sikker på at ingen andre funksjoner i systemet tukler med deres dataflyt. Gjennomsiktighet ble valgt fordi den representerer brukersentrismen og «open-source». I denne oppgaven ble gjennomsiktighet målt etter hvor mye gjennomsiktighet og «open-

source» var i fokus hos utviklerne. Gjennomsiktighet gjør at brukeren og utviklere har tilstrekkelig oversikt for å se sammenhengen mellom dataflyt og samtykke. Dette gjør at designfeil enklere blir oppdaget, samt at man kan være sikker på at egen persondata ikke blir kompromittert.

Intensjon defineres som «på hvilket grunnlag er PDS-økosystemet utviklet?». Har for eksempel PDS-økosystem «x» blitt utviklet for å kapitalisere på brukernes bekostning, eller fordi utviklerne har et brennende ønske om å revolusjonere persondata håndteringen i en mer brukersentrisk retning. Her ses det på til hvilken grad PDS-økosystemet beskytter brukernes rettigheter mot insentivet til å tjene penger på PDS-økosystemet. Intensjon er et begrep som her går direkte på tilliten til PDS-økosystemet og hvorvidt brukerne kan stole på langsiktigheten til PDS-økosystemet. Er PDS-økosystemet bare en døgnflue som er laget for å samle inn raske penger eller har utviklerne en langsiktig plan? Et av de sentrale problemstillingene her var hvordan forretningsmodellen fungerte, mer spesifikt hvordan skal utviklerne ta seg betalt for utviklingsjobben og videre drift av systemet. Lagring av data krever hardware, og hardware krever strøm. Om intensjonen bare er å tjene penger, vil PDS-systemet få en lav score på dette kriteriet. Grunnen til dette er at da vil mest sannsynlig brukernes rettigheter settes til side for muligheten til å tjene mer penger. Her er det viktig med riktig prioritering. Å betale rundt 10 dollar i måneden for å kunne lagre og kontrollere sine egne persondata vil være realistisk, om systemet fungerer godt og er utbredt nok.

Fordelen med å ha en forretningsmodell der man tjener penger og som er profitabel når den skalerer er at den vil virke mer fristende for investorer. Dette gjør at man får utviklet et produkt/PDS raskere, i motsetning til mer altruistiske konkurrenter. Dette er ikke en enkel problemstilling, så man blir nødt til å ha en klar visjon og en god forretningsmodell.

Enkelhet er det siste evalueringskriteriet jeg skal beskrive og defineres som graden av ryddighet i sikkerhetsfunksjoner og brukervennlighet for brukere, tjenestetilbydere og utviklere. I dette begrepet legger oppgaven vekt på at sikkerhetsaspektet til PDS-økosystemet skal være enkelt å forstå for bruker, utvikler og tjenestetilbyder, samt at systemet skal være brukervennlig for alle parter. Eksempel på hva som blir sett på her er hvor mange sikkerhetsfunksjoner har PDS-økosystemet og er de fremtidsrettet nok. Til slutt er brukervennlighet, i form av hvor mye kompleks informasjon er det meningen at brukeren skal forholde seg til.

Jeg så på det som mer verdifullt å ha få, men beskrivende kriterier, enn flere detaljerte. Dette var fordi modellene varierer en del fra hverandre slik at et høyere abstraksjonsnivå vil treffe bedre. Tidsbegrensninger under studien var også en begrensende faktor.

OpenPDS

OpenPDS var det første prosjektet jeg møtte da jeg søkte etter moderne modeller for håndtering av persondata. OpenPDS er laget av en liten gruppe utviklere fra MIT (Massachusetts Institute of Technology) med visjonen; «… Data ownership would therefore be defined as the rights of possession, use, and disposal instead of a literal ownership.»

Gjennomsiktighet

En av hovedutviklerne fra OpenPDS, svarte på spørsmålet: hvordan kan OpenPDS sikre tilliten til brukerne? At «det er gjennomsiktig. Du har kontroll over rådata, du kan alltid oppheve tilgang til en app, du kan revidere hvilken app som spurte hva og hvilket svar som ble sendt tilbake.» (fritt oversatt fra engelsk)

Hele prosjektet ligger også ute på Github, samt at utvikleren uttrykket i et senere spørsmål at han er en stor tilhenger av open-source og vektlegger viktigheten av open-source i forhold til gjennomsiktighet.

Både arkitekturen og svarene fra utvikleren tyder i retning av høy gjennomsiktighet. Det virker som de har lagt til rette for at så mye som mulig skal være synlig for både bruker, tjenestetilbyder og utvikler.

Intensjon

Etter det jeg har forstått er de ute etter en total revolusjon innenfor lagring, håndtering og prosessering av persondata. De vil øke fokuset på brukeren og brukerens rettigheter. Utvikleren fra OpenPDS var den første som nevnte skillet mellom teknisk og juridisk eierskap og fortalte at brukere i dag bare har juridisk eierskap. Dette var noe OpenPDS ønsket å gjøre noe med. OpenPDS ønsker å anonymisere brukerne mest mulig der det lar seg gjøre. Til dette bruker de SafeAnswers multiparty-computation.

Det virker ikke som OpenPDS vil prøve å true inn en forretningsmodell inn i dette økosystemet. Intensjonen for å utvikle OpenPDS kommer fra en følelse av urettferdighet ovenfor brukeren i dagens håndtering av persondata. Dette gjør at OpenPDS får en høy score på dette kriteriet.

Enkelhet

En av de beste modellavgjørelsene til OpenPDS er å ikke gi ut rådata, men heller aggregerte svar på spørringer PDS-økosystemet mottar. Fordelen med å ikke gi ut rådata er at man reduserer kvaliteten på data. Om aggregert persondata lagres hos en tjenestetilbyder vil tjenestetilbyderen ha større vanskeligheter med å unikt identifisere brukeren, enn ved bruk av rådata. Aggregert persondata er mer anonymisert og generisk enn ubehandlet rådata. En annen fordel med aggregerte persondata er at den enklere og raskere kan benyttes, siden prosesseringen/kalkuleringen allerede er blitt gjort. Ferdig kalkulerte data innehar mer konsentrert informasjon og kan raskere benyttes.

Desentralisert teknologi kan benyttes for å anonymisere brukerne ytterligere. Dette gjør SafeAnswers mulitparty-computation metode. Det negative som OpenPDS nevner i intervjuet er «performance overhead», som per dags dato er signifikant når sikkerhet og personverns mekanismer er lagt til, samt at MC (multiparty-comp.) i seg selv genererer en del overhead.

Mydex

Mydex er en «ID-provider», PDS og CIC (Community Interest Company). Det var mye vanskeligere å forstå hvordan modellen til Mydex fungerte, når det kom til hva Mydex skulle være eller hva mydex ikke skulle være. For meg virket det som at de hadde mange ideer og hadde muligens endret kurs underveis opptil flere ganger. Det står beskrevet mye på deres hjemmeside, men en oversikt over helheten mangler.

Gjennomsiktighet

Flere punkter tyder på at de har et fokus på at gjennomsiktighet kreves for å gjøre systemet brukersentrisk. Noe de streber etter. Opptil flere ganger i spørreskjemaet fra fase en nevner utviklerne fra Mydex at problemene med dagens modell er nettopp mangel på

gjennomsiktighet, det at brukeren ikke ser hvilke data som går hvor. Det nevnes også på deres nettside (Mydex charter) at gjennomsiktighet er påtvunget applikasjoner som skal bruke Mydex plattformen.

Intensjon

Forretningsmodellen til Mydex er at organisasjoner som skal koble seg til denne plattformen skal betale for det, gjennom transaksjonskostnader, tilkoblingskostnader og årlige kostnader (Mydex tariff). Det er foreløpig gratis for brukere/individer, men det er ikke sikkert det alltid vil være slik (Mydex payment). Denne lite statiske og komplekse forretningsmodellen gjøre det hele litt uoversiktlig. De bør nok rydde litt opp i det, om de ikke vil bli utkonkurrert.

Enkelhet

Mydex bruker kjente teknologier hvor man kobler seg til plattformen, gjennom et API. Når man tilbyr et slikt API så er det viktig at utviklerne enkelt kan sette seg inn i sammenhengen mellom samtykke og persondata-henting. Samtidig som det tilbyr nok tilpasnings muligheter for organisasjonen som skal koble seg til.

Mydex beskriver det de kaller «trust framework», hvor alle brukerne er legitime identiteter og har muligheten til å kontrollere hvem de deler hva med. Dette blir sett på som en kjerneverdi i det å lage et PDS-økosystem, så det er et stort pluss. Organisasjoner kan stole på at brukerne er reelle og brukerne kan stole på at organisasjonene ikke misbruker persondata.

Enigma

Enigma er en teknologi fra MIT (Massachusetts Institute of Technology) som benytter seg av blockchain kombinert med en peer-to-peer protokoll, en modifisert utgave av kademlia, for sikring av konfidensialitet. Enigma ble oppdaget for sent i prosessen til at jeg fikk mulighet til å intervjue utviklerne.

Gjennomsiktighet

En av egenskapene til blockchain teknologien er at den er gjennomsiktig. Alle transaksjoner, her mellom applikasjoner og brukere vil bli logget i blockchain-nodenes hovedbok. Dette vil ifølge utviklerne fremme ærlig oppførsel siden man enkelt vil kunne oppdage uriktige transaksjoner i loggen.

Ellers eksisterer det bare et white paper som forklarer teorien rundt modellen. Om det blir gjennomsiktig for brukerne er en annen ting. Det nevnes ikke spesifikt, annet enn at blockchainen sikrer det.

Intensjon

Forretningsmodellen her var vanskeligere å oppdrive. Nettsiden Wired hadde skrevet noe tidligere som pekte på en form for betaling av å kunne hente ut multiparty-computation data: «Every time someone requests a computation from the Enigma network, he or she pays a bitcoin fee. A tiny part of that money is paid to a computer in the bitcoin network to record Enigma's metadata in the blockchain» (Greenberg, 2015).

Enkelhet

Blockchain teknologien begynner å møte motstand i form av ytelse. Om man ser på Bitcoin, kan verifisering av en transaksjon ta flere timer, eller dager om mange nok vil selge eller kjøpe samtidig. Grunnen til denne overheaden er at alle nodene skal ha en kopi av hovedboken som logger alle transaksjoner. Så skal alle nodene bli enige om at transaksjonen var riktig. Dette gjøres ved at 51% av nodene må ha den samme verdien på transaksjonen.

Alt i alt er det interessant at Enigma prøver å bruke blockchain til å lage et PDS-økosystem, og det skal de ha et pluss for.

Oppsummert ser de tre PDS-økosystemene slik ut:

PDS-økosystem	**Gjennomsiktighet:** I hvilken grad kan brukere, utviklere og tjenestetilbydere observere persondataflyt og arkitektur.	**Intensjon:** på hvilket grunnlag er PDS-økosystemet utviklet? (til hvilken grad vil de beskytte brukernes rettigheter fremfor å tjene penger?)	**Enkelhet:** graden av ryddighet i sikkerhetsfunksjoner og brukervennlighet for brukere, tjenestetilbydere og utviklere.
OpenPDS	Stor tilhenger av open-source og opptatt av «transparency» for alle parter.	Hovedsakelig brukersentrisk og ønsker et skifte til PDS fra dagens persondata håndtering. OpenPDS er fremdeles kun i prosjektfasen.	Veldig oversiktlig og enkel arkitektur for alle parter. Fokus på sikkerhet og konfidensialitet ovenfor bruker. Benytter «no-trust».
Mydex	Stor tilhenger av open-source og opptatt av «transparency» for alle parter, men litt rotete forklaringer og modeller.	Brukersentrisk, men uklar og kompleks betalingsmodell for tjenestetilbydere. Uklart hvordan de skal gi insentiver til tjenestetilbyderne.	Fokus på brukervennlighet ovenfor bruker, men ikke for utvikler og tjenestetilbyder. Forklaringer og illustrasjoner kan forbedres. Sikkerheten virker tilstrekkelig og de benytter «trust modell».
Enigma	Blockchainen skal sikre obseverbarhet	Bedrifter betaler for MPC data om	Det er ikke enkelt for brukere å sette

	av datatransaksjoner, men vanskelig å gjøre det oversiktlig for alle parter uten å lage mye overhead.	brukerne. Ellers lite beskrivelser.	seg inn i hvordan blockchain sikrer deres persondata. Benytter «no-trust». Brukervennlighet er ikke mulig å oppdrive på dette tidspunkt.

6 Diskusjon: krav til et PDS-økosystem

Her diskuterer jeg funnene i denne oppgaven basert på intervjuer, litteratur og diskusjon. Jeg baserer diskusjonen på PDS-økosystem aspektene; lagring, tilgang til data og samtykke, sikkerhet, identitetstilbyder og adopterbarhet for markedet. Håndtering, prosessering og lagring av persondata faller inn under flere av disse aspektene. Aspektene ble valgt på bakgrunn av «utviklingen et PDS-økosystem» og ikke fenomenet «håndtering, prosessering og lagring av persondata». Et PDS-økosystem skal tilby noe mer enn det dagens håndtering, prosessering og lagring av persondata gjør.

Lagring

Her diskuterer jeg spørsmålene hvor og hvordan persondata bør lagres og ser på mulige teknologier som kan tilfredsstille kravene for sikker og konfidensiell lagring. Jeg ser først på ulike teknologier og modeller, etterfulgt av en diskusjon om dette bør tilbys av det offentlige eller utvikles av det private markedet. Til slutt tar jeg for meg om lagringsteknologien er fremtidsrettet nok, eller hvilke grep som vil gjøre den bedre rustet for fremtidens utfordringer.

Cloud, peer-to-peer eller lokalt

Cloudlagring, peer-to-peer modeller eller lokalt hos bruker, er tre forskjellige modeller for lagring av persondata. Det er også rom for kombinasjoner av disse, enten som backup data eller splitting av data mellom forskjellige modeller. Splitting vil kreve en oppslagstabell (program) som holder oversikt over hvor data er spredt i forskjellige lagringsmedier.

Cloudlagring har den fordelen at den er tilgjengelig hele tiden og at data ikke forsvinner eller mister integriteten (forblir uforandret). Det negative med dagens cloudløsninger et at den sjelden er konfidensiell. Ofte har tilbyderen teknisk eierskap til data og har mulighet til å se hva du har lagret. Cloudtilbyderne kan si at de ikke vil lese data din, men de bestemmer hvordan din data blir kryptert og lagret (Single, 2011). Det skal ikke være nødvendig å stole

på en lagringstilbyder med så mye makt over dine data. Dette vil si at cloudlagring for persondata sikrer tilgjengelighet og integritet, men ikke konfidensialitet.

Peer-to-peer (p2p) er en teknologi som har eksistert lenge, og vært spesielt plagsom for musikk- og filmbransjen. p2p muliggjør enkel deling av data som kan bryte mot opphavsretten. p2p sammen med riktig algoritme kan fungere som en slags cloudløsning. «Medlemmene» av denne løsningen betaler lite eller ingenting for tjenesten, men vil i sin tur være en del av den, ved at man setter av en del av sin egen harddisk til disposisjon for systemet. Med riktig fordelings- og duplikasjons-algoritme, samt god kryptering kan denne teknologien tilby tilstrekkelig konfidensialitet, integritet og tilgjengelighet. Baksiden med p2p er at man kan få malware på lagringsplassen man selv tilbyr til tjenesten, samt at jo mer man deler opp data, desto mer trafikk får man over nettverket. Så her blir det påtvunget en tradeoff mellom konfidensialitet og tilgjengelighet. En utfordring utviklere må være oppmerksom på i p2p systemer, er sybil angrep og varianter av slike angrep. Sybil angrep er når er en ondsinnet bruker later som han er opptil flere noder i nettverket ved å forfalske identiteter. Dette vil ikke være fordelaktig for verken bruker eller tjenestetilbyder da dette tvinger tjenestetilbyderne å ha en strengere identifikasjonsmetode når brukere skal interagere med tjenesten. Dette vil gjøre det hele mer tungvidt og forhindre den sømløse interaksjonen mellom bruker og tjenestetilbyder, samt at brukere blir nødt til å gi fra seg flere identifikasjons-parametere for identifisering. Uten en sentralisert autoritet er det vanskelig å hindre sybil angrep (Piyawongwisal & Xia, 2011).

Ved å lagre data hos brukeren lokalt vil han/hun sikre konfidensialitet, integritet og tilgjengelighet. Hovedproblemet med denne løsningen er de praktiske utfordringene. Har man nok lagringsplass på alle enhetene sine? Har brukeren få enheter vil det være en stor risiko for å miste all persondata, grunnet få backup-medier. Dette vil også ta en god del av prosessorkraften til enheten og trekke mye batteri om aggregeringer og kalkuleringer gjøres ofte. Hva gjør brukeren om han/hun har fått ondsinnet programvare på en av enhetene når de skal synkroniseres? Dette vanskeliggjør denne type lagring og gjør at det mest sannsynlig blir for upraktisk å bruke denne modellen for lagring av persondata.

Et fjerde alternativ er en kombinasjon av to eller flere modeller for lagring. En modell for lagring som virker mer attraktiv enn andre er cloudlagring, hvor data er splittet opp og fordelt over flere cloudlagringstilbydere. Oppslagstabellen for hvor data er lagret eksisterer lokalt på enhetene til brukeren.

Det er også mulig å skille mellom persondata som kan kategoriseres som rådata og strukturerte/aggregerte data. Rådata er mer sensitiv informasjon og bør beskyttes bedre. Eksempler på rådata som kan brukes for å identifisere en bruker mot ens vilje er GPS-punkter med timestamps. Har tjenestetilbyderne nok av disse GPS/timestamps-punktene kan de ved hjelp av algoritmer identifisere en spesifikk bruker. Aggregerte persondata som for eksempel favorittfilmsjanger er mer anonymt og det stilles ikke samme krav til konfidensialitet.

Offentlig eller privat initiativ?

Er det offentlige eller private aktører som bør utvikle og drifte lagring av persondata? Jeg mener det er viktig å ha tatt opp den tematikken for å kunne vurdere forskjellige løsninger. Dette skal ikke være en debatt om hvilken grad av demokrati som trengs for at det offentlige ikke skal falle for fristelsen å misbruke persondata. Jeg ønsker heller å belyse at det er sterke insentiver for det offentlige å sniffe på persondata om de har muligheten og at det derfor ikke er noe vi kan se bort ifra.

Private aktører har sine insentiver som ofte er å tjene mest mulig penger innenfor lovens rammer. Et eksempel på dette er å se på dagens persondatamodell, her kan tjenestetilbydere selge brukerdata til tredjeparts organisasjoner om dette er tillatt i samtykkekontrakten mellom tjenestetilbyder og bruker. Om aktørene ikke taper brukere gjennom sine handlinger, samt opererer innenfor lovens rammer, er dette lukrativt. Dette kan by på utfordringer fra brukerens perspektiv i forhold til at det kan oppstå en interessekonflikt mellom beskyttelse av brukerens rettigheter og interessene til andre entiteter (f.eks: private aktører og styresmakter). Brukeren vil sikre at bare han/hun har tilgang til data, og full kontroll over dem. Vurderingen min er at både det private og offentlige ofte har sine egne insentiver, så om ikke utviklingen av et slikt PDS-økosystem gjøres ut i fra brukerens interesser bør det sikres gjennom regulering laget for beskyttelse av brukerens rettigheter.

Fremtidsrettet

Uansett hvilken modell man velger for lagring av persondata, er det viktig å ha en kryptering som tar hensyn til fremtidig prosessorkraft og kvantedatamaskiner. Det er ikke mulig å spå langt inn i framtiden, men utvikling av kvantedatamaskiner (Rich & Gellman, 2014) virker

som et relativt sikkert kort. Det finnes allerede flere krypteringsalgoritmer (PQCrypto) som tar hensyn til den nesten uendelige matematiske kraften til en kvantedatamaskin.

Politikk og lover er dynamiske. Siden regjeringer er ikke statiske, samt det økte fokuset på å finne «nye løsninger» for håndtering av terrortrusselen, gjør det vanskelig å gi offentlige instanser for mye makt over utvikling og drift av persondatalagringen.

For å oppsummere lagringsaspektet, ble det diskutert følgende modeller: cloudlagring, p2p og lokalt hos brukeren. Den modellen som oppgaven tar med til PDS-økosystem modellen er en kombinasjon av flere cloudtjenester med en lokalt lagret oppslagstabell for hvordan persondataen er splittet og fordelt mellom cloudtjenestene. Det ble også nevnt at kryptering som benyttes bør være av typen post-kvantedatamaskin, for å sikre rådata mot kvantedatamaskiners brute-force egenskaper.

Tilgang til data og samtykke

Her diskuteres samtykkekontrakter og hvordan tilgang til data kan styres for å hindre misbruk. Det er særdeles viktig at tilgang til data blir modellert riktig, da «hullete protokoller» vil muliggjøre uautorisert adgang og overdreven spørring etter data. Med «hullete protokoller» menes enkelt protokoller, eller samling av flere protokoller som ikke har tatt hensyn til alle omstendighetene i kommunikasjonsprosessen i et økosystem.

Antall spørringer av persondata bør minimeres slik at applikasjoner ikke spør etter mer data enn det som absolutt er nødvendig for gjennomføring av den aktuelle oppgaven. Grunnen til dette er at man vil minimere mengden med persondata som eksponeres og kompromitteres. Rådata bør ikke gis ut, da denne data er mer sensitiv og lettere kan brukes til å identifisere brukere. De fleste spørringer vil dessuten ikke ha bruk for rådata, men greie seg med aggregerte data. Et eksempel på dette er applikasjoner som vil gjøre markedsføring på applikasjoner mer relevant for bruker. Her kan da PDS-økosystemet selv kalkulere hvilket segment brukeren tilhører og sende det til applikasjonen, i stedet for å gi ut nettleserloggen («rådata»). Rådata bør håndteres og prosesseres på et trygt konfidensielt område i PDS-økosystemet, mens aggregerte data lettere kan gis ut til applikasjoner.

Eierskap til data

Sentralt i denne oppgaven er brukersentrismen og det juridiske aspektet som tydeliggjør hvem som eier persondata. Juridisk sett er det ikke tvil om hvem det er som eier persondata, det er brukeren. Om det ikke var brukeren hadde ikke selskapene trengt å låne den via «terms of service». Med dagens standard for håndtering, prosessering og lagring av data så er det nærmest umulig for brukeren å ha oversikt over hvem som eier hvilken data og hva den brukes til. Et sentralt skille her er mellom teknisk (kontroll over data og mulighet til å redigere/slette) og juridisk eierskap (hvem som etter loven har eierskap). Denne studien søker etter modeller som gjør det mulig for brukeren å ha både teknisk og juridisk eierskap til sine egne persondata.

Det er også viktig at samtykkekontraktene er oversiktlige, brukervennlige og redigerbare. Ved at de er redigerbare viser jeg til muligheten å kunne trekke tilbake samtykke tidligere gitt til en applikasjon for utlevering av persondata eller metadata. Det vil også være særdeles viktig at samtykkeprotokollen er designet universalt og er brukervennlig. Det bør være tilstrekkelig samtykkekategorier/persondatakategorier til at brukeren kan skille mellom det brukeren ønsker å dele og bare det. Med samtykkekategorier menes en gruppering/klassifisering av persondata. Om systemet velger å klassifisere persondata på bakgrunn av sensitivitet, vil for eksempel veldig sensitiv persondata grupperes i samme samtykkekategori, mens mindre sensitiv data grupperes i en annen samtykkekategori. Antall samtykkekategorier bør begrenses til et antall brukeren klarer å holde oversikt over. En bør unngå den strategien som applikasjoner bruker i dag, hvor brukeren blir møtt med så mye informasjon til at det blir slitsomt å holde seg oppdatert på hva brukeren faktisk deler av persondata.

Skille mellom tjenestetilbydere etter grad av tillit

Forskjellige applikasjoner bør behandles etter grad av tillit. Med dette mener jeg at applikasjoner fra Skatteetaten og banken stoler brukerne ofte mer på enn for eksempel Facebook og ymse fitness-applikasjoner på mobilen. Det bør således være mulig å skille disse fra hverandre på en oversiktlig måte. En teknologi som kan gjennomføre dette er å lage en form for plattform for å håndtere applikasjoner med lav grad av tillit. Det vil da fungere slik at man har tillit til plattformen men ikke applikasjonene på den. Måten dette vil fungere på er at plattformen kommuniserer med PDS-økosystemet og har en egen protokoll for å hente ut

persondata fra PDS-økosystemet. Plattformen vil tilby et API til andre tjenestetilbydere, slik at andre applikasjoner kan koble seg på denne plattformen som har bedre aggregerings- og anonymiserings-algoritmer enn selve PDS-økosystemet. Plattformen vil da fungere som en slags brannmur eller filter som sørger for å «ufarliggjøre» ulike tjenestetilbydere. Enten å anonymisere ved hjelp av MPC (multiparty-computation), eller aggregering/prosessering av data på en måte som gjør linken mellom persondata og brukeren svakere og mer generisk. Plattformen vil også gjøre det mer oversiktlig og sikkert for brukeren, ved at plattformen kan si ifra om uregelmessig adferd hos forskjellige applikasjoner/tjenestetilbydere, slik at brukeren ikke trenger å gå i sømmene til alle applikasjonene.

Spørringer bør logges slik at sikkerhetsapplikasjoner kan dra nytte av loggen ved å finne ut om applikasjoner oppfører seg ondsinnet eller mistenksomt. Med sikkerhetsapplikasjoner menes applikasjoner som kjører på PDS-økosystemet med hensikten å sikre økosystemet på en eller annen måte. Det blir da særdeles viktig at disse sikkerhetsapplikasjonene har gode skussmål for ærlig oppførsel. Strenge krav bør dermed settes for sikkerhetsapplikasjonene. Ved at alle datatransaksjoner logges, kan systemet også gå i sømmene for å kontrollere at samtykkeovertramp rapporteres og gjøres synlig for brukeren. Selv om det er mer optimalt at slike overtramp ikke blir teknisk mulig.

Siden det kan være vanskelig og tidkrevende å etterse alle applikasjoners tilgang, bør en se på muligheten til å begrense transaksjoner over hele systemet med en hovedbryter eller flere hovedbrytere. Et eksempel på en hovedbryter kan være at ingen applikasjoner skal kunne få ut personnummer om denne er skrudd på. Dette vil gjøre at PDS-økosystemet får en begrensning over dataflyten på et generelt grunnlag. Dette kan også innebære at brukeren alltid skal dobbeltsjekke transaksjoner med kredittkort informasjon, eller aldri gi ut adresse-informasjon om man skulle ønske det. Mens vi bruker terminologien med hovedbryter kan jeg også introdusere «dimmeren» (kap. 7 kalt «level of paranoia»). Denne skal øke eller minke datastrømmen til flere ferdig utfylte kategorier. Meningen er at personvern-eksperter skal fylle ut hvilke datakategorier som er mer sensitive enn andre, slik at brukeren ikke trenger å gå dypt inn i problemstillingen selv. Problemet med en slik dimmer er at det vil ganske sikkert påvirke et ukjent antall applikasjoner på en utilsiktet måte. Da applikasjonene ikke får tilstrekkelig med persondata til å gjennomføre oppgavene sine optimalt.

Fremtidsrettet

For at samtykke og tilgang skal være framtidsrettet nok, bør en ta hensyn til GDPR, samt sikre at teknisk eierskap også tilhører brukeren. Dette vil sikre en mer helhetlig beskyttelse av data, og gjøre det vanskeligere for tjenestetilbydere å misbruke persondata. Om teknisk eierskap ikke blir gitt til brukeren vil GDPR sine retningslinjer uansett sørge for bedre sikring og håndtering av data. Så håndtering av persondata går en lysere fremtid i møte selv om det tar lengre tid å flytte persondata fra tjenestetilbydere til brukere.

Et annet punkt her er muligheten til å la AI (Artificial Inteligence) og adaptive algoritmer sørge for at alle komponentene (i PDS-økosystemet) spiller etter reglene. Mer spesifikt at AI'en lærer seg de nye mulige exploitene til systemet og stopper applikasjoner som prøver å utnytte de. Når en AI har lært det fra et PDS-økosystem, vil den fortelle om metoden til andre AI'er som overser andre brukeres PDS-økosystem.

Plattformer kan brukes for å skille mellom tjenestetilbydere brukeren stoler på og ikke, ved at de tjenestetilbydere brukeren har lav tillit til kjører på en plattform som brukeren har tillit til. Andre sentrale punkter i denne seksjonen er sikkerhetsapplikasjoner som går igjennom transaksjonslogger for å finne avvik og brytere som brukeren enkelt kan bruke for å strupe informasjonsflyten til tjenestetilbydere.

Sikkerhet

Sikkerhet er en av de viktigste pilarene til et PDS-økosystem. Brukerne ønsker å få beskyttet persondata fra ondsinnede aktører og unngå identitetstyveri. Sikkerhetselementer har blitt nevnt i diskusjonen over, fordi det kommer innpå andre områder som samtykkekontrakter og lagringskryptering. Det er også viktig med en skikkelig gjennomgang av sikkerhet i sin helhet, noe jeg skal se nærmere på her.

Generell sikkerhet og open-source

Forskjellige teknologier krever forskjellige sikkerhetsmekanismer. Det gir for eksempel ingen mening å snakke om et sybil angrep i et system uten noder, da sybil angrep per definisjon forfalsker noder i et nettverk. Nå er både blockchain og p2p modeller nevnt tidligere og da er det viktig å være klar over utfordringene til slike distribuerte/desentraliserte modeller. En av

hovedutfordringene til distribuerte/desentraliserte modeller er falske noder i nettverket. Som oppgaven var inne på under diskusjonen av p2p modeller, kan en løsning være at alle har offentlige nøkler, hvor så en sentralisert autoritet kvalitetssikrer brukerne/nodene opp mot et offentlig nasjonalt register.

På et generelt grunnlag er det viktig at krypteringen for lagring og transportering av data er tilstrekkelig gjennomført. Dette er for å sikre at data er uleselig for alle andre parter enn de aktuelle. I tillegg til kryptering er det viktig med solide protokoller som tidligere har blitt stresstestet. Strenge rutiner for deling og lagring av persondata i protokollen vil gjøre det enklere å debugge eventuelle systemfeil og «exploits», ved at antallet variabler begrenses.

Ved bruk at open-source kan utviklere velge å eksponere sikkerhetsaspektene til et system slik at utviklere med gode og ondsinnede hensikter kan evaluere styrken på metodene og protokollene. Dette er en kjent metode for evaluering av sikkerhetsmekanismer. Det faller utenfor området til denne studien å diskutere om hvorvidt dette bringer mer positivt enn negativt til sikkerheten av et system, da dette er en kjent teknikk (Gallivan, 2001)

Det er også viktig å ta hensyn til det praktiske. Et scenario der mobiltelefonen blir stjålet, så kan ikke PDS-økosystemet la data bli kompromittert. En løsning her kan være bruk av OTP (One Time Passwords) i form av biometri. For eksempel et krav om fingeravtrykk identifisering hver gang brukeren prøver å administrere applikasjoner i økosystemet. Oppgaven favoriserer teknologi som sentraliserer kontroll hos brukeren, gjerne gjennom en hovednøkkel som bare brukeren har tilgang til. Denne nøkkelen kan heller ikke bli stjålet om den kun kan aktiveres av brukerens biometri (fingeravtrykk, iris, ansiktsgjenkjenning eller lignende). Systemet må være designet slik at å stjele verdiene til de biometriske egenskapene ikke muliggjør autentisering. Det skal også ikke være mulig å kunne stjele disse verdiene fra systemet, da dette bør gjøres i en lukket del av systemet. Denne lukkede delen skal håndtere den biometriske autentiseringen, og dermed linken mellom digital identitet og virkelig identitet. Denne delen bør iverksettes hver gang nye samtykkekontraker skal godkjennes, samt når andre tillatelser skal godkjennes.

Egne sikkerhetsapplikasjoner utviklet gjennom «open source» burde ha mulighet til å kjøre i økosystemet. Dette vil kunne øke sikkerheten betraktelig ved at mange forskjellige utviklere får testet deres nye ideer for økt sikkerhet, konfidensialitet og integritet. For at tredjeparts-sikkerhetsapplikasjoner skal kunne fungere optimalt i et slikt økosystem, må man ha et

poengsystem der gode applikasjoner blir vurdert høyt og dårlige vurdert lavt. For at dette poengsystemet ikke skal bli misbrukt, med for eksempel gi masse poeng til farlige eller dårlige applikasjoner. Kan man sørge for at alle bare kan stemme en gang gjennom PDS-økosystemet identifikasjon.

Fremtidsrettet

Om man vektlegger Ann's «Privacy by design» punkter, bør man konstant vurdere sikkerheten når en utvikler et nytt system. Det er mye enklere å få til god sikkerhet om man implementerer det tidlig i en utviklingsfase, enn å legge det til som en funksjon senere.

For å oppsummere trenger node-modeller som blockchain og p2p en sentralisert autoritet for å hindre varianter av sybil angrep. Oppgaven fremmer også bruken av open-source i utviklingen av sikkerhetsapplikasjoner som igjen er linket til et rankingsystem. Dette vil gjøre at gode sikkerhetsapplikasjoner blir utviklet og benyttet av flest mulig PDS-økosystemer. Til slutt trenger brukeren en hovednøkkel som ikke kan bli stjålet når han eller hun skal godkjenne nye tillatelser eller samtykkekontrakter. Dette kan gjennomføres bra ved bruk at biometri og gjerne kombinasjoner av flere typer som fingeravtrykk og stemmegjenkjenning.

Identitetstilbyder

Et sentralt diskusjonstema er om PDS-økosystemet skal være koblet til den virkelige identiteten, eller om brukere kan lage så mange digitale identiteter de ønsker på flere PDS-økosystemer. På den ene siden virker det åpenbart å koble virkelig identitet til digital identitet, både for å hindre svindel og node forfalskninger. Derimot det er ikke sikkert alle vil blottlegge sin virkelige identitet av ulike grunner. Politisk forfølgning er et klart eksempel. Selv om systemet ikke skal være mulig å kompromittere, så er det alltids en risiko at maktpersoner får tak i persondata til bestemte individer. Styresmakter har ofte mye ressurser tilgjengelig til slike formål. Brukere kan også løse det enkelt om de vil være «helt» anonym så skrur de av PDS-økosystemet når de surfer på nett eller anvender applikasjoner. Mydex bruker sikring av identitet som en del av deres trust framework. Slik at både tjenestetilbyder og bruker kan stole på at den andre parten ikke er falsk.

Vouchere

En voucher defineres her som en entitet/tredjepart som skal sikre betalingsevnen og identiteten til en bruker ovenfor en tjenestetilbyder, og hindre at tjenestetilbydere lurer brukere. Voucheren skal løse problemet med at brukere ikke vil ha kredittopplysningene spredd over nettet, samt hindre «liten skrift» svindel. Med dette mener jeg situasjoner der man bestiller en vare til en billig penge, men hvor det står med liten skrift at man også blir trukket et tilleggsbeløp hver måned.

Voucheren fungerer slik at brukeren kun legger tillit til voucheren og dermed slipper å ha tillit til at tjenestetilbyderne skal behandle betalingskredittene til brukeren sikkert og konfidensielt. Vanligvis når brukeren betaler på en nettside/tjenestetilbyder, lagres de opplysningene, samt at nettsiden spør banken om denne brukeren har tilstrekkelig betalingsevne. Når brukere benytter en voucher sendes betalingsopplysningene til voucheren, hvor voucheren så går god for brukeren, ved at de selv sjekker med banken og bare sender en «godkjent» identifikator til nettsiden/tjenestetilbyderen. Det er sentralt her at både brukeren og tjenestetilbyder stoler på voucheren. Voucheren selger på en måte tillit som et produkt.

Konseptet med vouchere trenger ikke integreres i PDS-økosystemet. Grunnen til dette er at tjenestetilbydere kan utveksle informasjon direkte med bank-modulen som kjører på PDS-økosystemet. Dette er forklart under gjennom «identifisering mellom tjenestetilbydere».

Identifisering mellom tjenestetilbydere

En kjent utfordring for IT-systemer og bedrifter er å identifisere brukere mellom systemer og databaser. Her antas det at tjenestetilbydere trenger en sikker identifikasjonsmetode, slik at de kan benytte seg av banktransaksjoner. Det antas også at tjenestetilbyderne ikke har hatt noe å gjøre med hverandre fra før. De har også forskjellige systemer og identifikatorer på personene i databasen. Det finnes da ingen generell metode eller program som enkelt kan løse denne problemstillingen. Eneste kjente og sikre metoden er manuell tilpassing uten at tjenestetilbyderne involverer brukerne (personene i databasen). Det kan også hende at tjenestetilbyderne må involvere brukerne, grunnet for svak korrelasjon mellom identifikatorene brukt i brukerdatabasene.

Denne problemstillingen kan løses i et PDS-økosystem ved at brukeren godkjenner en samtykkekontrakt for deling av informasjon mellom ulike tjenestetilbydere (hvor begge har en modul i PDS-økosystemet). Dette er en av fordelene med å sentralisere persondata hos brukeren og ikke hos alle forskjellige tjenestetilbydere/applikasjoner.

En sentralisert autoritet kan sikre at et PDS-økosystem tilhører en bestemt bruker, vil dette være nyttig for tjenestetilbydere da de kan være sikre på at digital identitet tilhører en virkelig identitet. Dette sammen med muligheten for at tjenestetilbydere kan utveksle informasjon om en bruker over PDS-økosystemet gjør det mulig å tilby det som i dag gjøres av vouchere. Her vil da tillit være bundet til handlingsmønstre i protokollen, mot tillit til en organisasjon som det er i dag.

Gjør PDS-økosystemet adopterbart for markedet

Det er viktig å ikke glemme hva markedet ser etter i en PDS-løsning. Om PDS-utviklere ikke får med seg tjenestetilbyderne vil introduksjonen av PDS-løsningen i beste fall ta lang tid, eller ikke bli benyttet i det hele tatt. I denne seksjonen skal vi se på hvilke insentiver PDS-økosystemet kan lokke med til det private markedet.

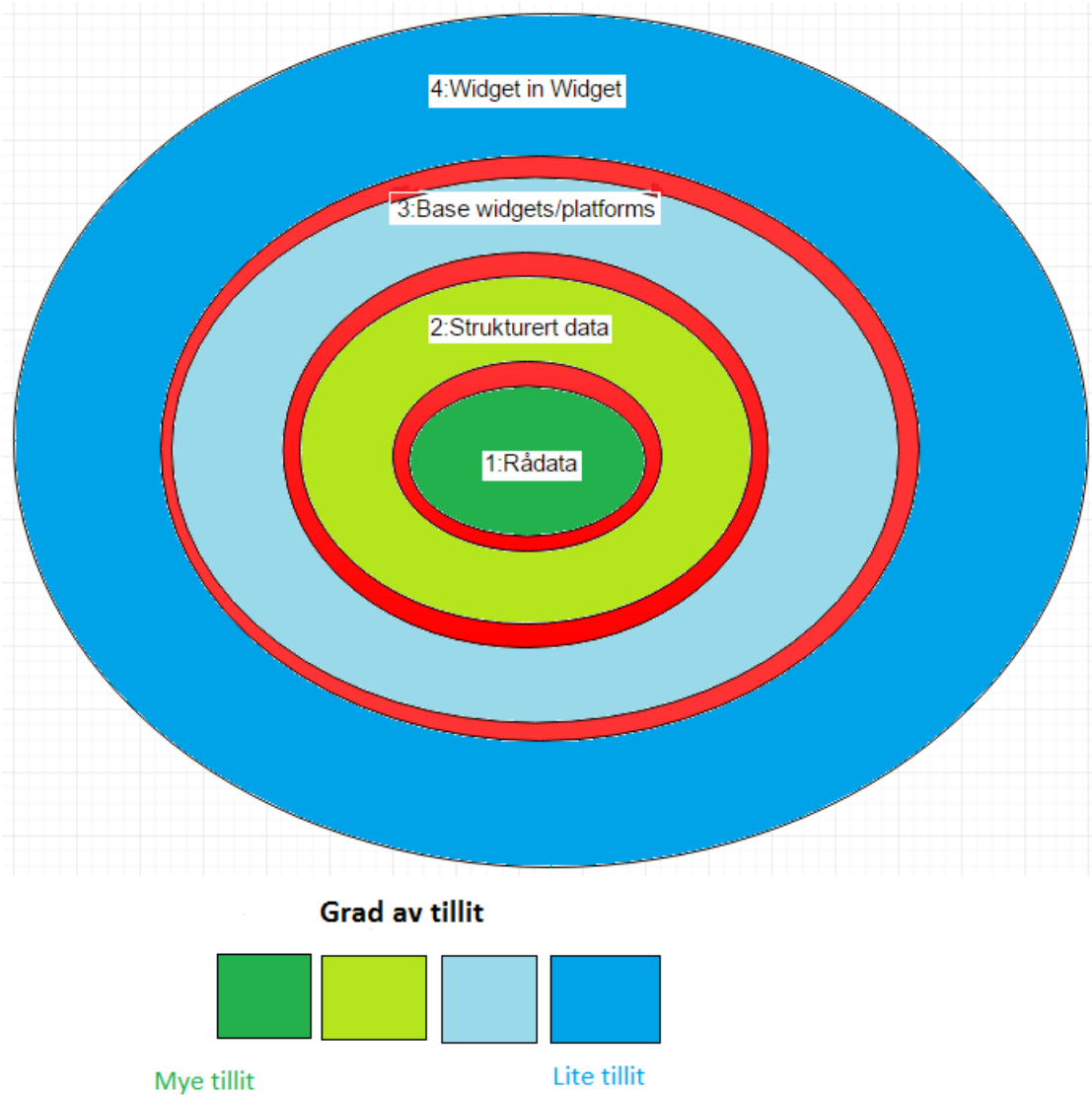
Denne oppgaven ser for seg at en widget in widget (WiW) løsning, hvor en tjenestetilbyder mange har tillit til utvikler en plattform andre tjenestetilbydere kan benytte seg av. På denne måten kan utviklingskostnadene forbli små, ved å enkelt integrere sin egen applikasjon via en API'et til plattformen. WiW løsning vil gjøre utvikling av en plattform mer attraktivt for store selskaper som Facebook og Google. De vil kunne lytte på trafikken mellom widget'ene og PDS-kjernen (der persondata er lagret). De vil også ha mulighet til å sMPC (Secure Multiparty-comp.) mellom forskjellige PDS-økosystemer for aggregerte/anonymiserte gruppe spørringer. Dette er mindre brukersentrisk enn det jeg ser etter i studien, men det vil sannsynligvis få utviklingen mot PDS-økosystemer til å gå raskere. Her blir det en slags trade off mellom brukersentrisme og hurtig utviklingstempo av PDS-økosystemet.

WiW vil gjøre det mulig for mindre bedrifter/applikasjoner å ha en lettere backend, siden man slipper å holde styr på store bruker-databaser. Registrering og innloggings skjemaer vil også forsvinne, og bli byttet ut med API integrering til en plattform. Base-widget'en eller plattformen vil som nevnt tidligere kunne gjøre mye «Big data» ved hjelp av sMPC. Dette insentivet er sentralt for å få med seg de store aktørene/tjenestetilbyderne.

Om en slik modell forbedrer opplevelsen og flyten til brukerne, har ikke markedet råd til å ikke være med på utviklingen. Brukerne vil oppleve bedre flyt ved bruk av applikasjoner, samt at de ikke trenger å logge seg inn eller registrere seg hos nye tjenestetilbydere. Til gjengjeld kan samtykke for deling av persondata gjøres på en oversiktlig og brukersentrisk måte. Oppgaven ser på WiW teknologi som det viktigste insentivet for å få med de store tjenestetilbyderne med på persondata-byttet fra dagens situasjon til PDS-økosystemer.

7 Modellen

I dette kapittelet beskriver jeg en modell for et PDS-økosystem. Dette er hovedsakelig en abstrakt modell som betyr at den ikke definerer tekniske detaljer (som valg av protokoller), men heller funksjonalitet og muligheter. Denne modellen kommer fra en prosess der jeg sammen med veilederne har diskutert hvilke funksjoner et PDS-økosystem bør tilby. Dette ble diskutert etter å ha gått igjennom tilbakemeldingene fra de tidligere intervjuene. Denne modellen beskriver en samling av funksjoner og strukturer som tilsammen skal representere et PDS-økosystem.



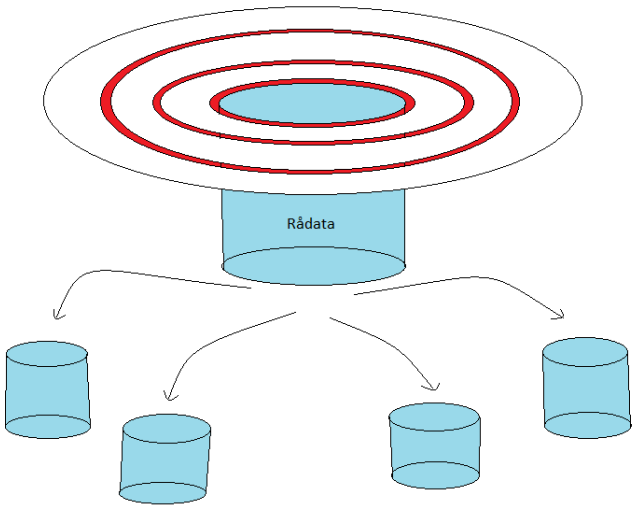
Figur 5: PDS-økosystem modellen, overordnet

Modellen har en tjenesteorientert arkitektur der tjenestetilbydere plassert i ring 3 og 4 kobler seg på en interaksjonsprotokoll slik at de kan utveksle informasjon mellom seg og ring 1 eller 2. Persondata som lagres i ring 1 og 2 kan utveksler til tjenestetilbyder gjennom et API om brukeren (eieren av PDS-økosystemet) godkjenner samtykkekontrakten.

Spørringer fra tjenestetilbydere gjøres fra de ytre lagene og innover, mens svar på spørringene går fra indre lag og utover. De røde ringene representerer grader av tillit som tjenestetilbydere må passere for uthenting av persondata. Dette vil si at jo mer brukeren stoler på tjenestetilbyderen, desto mer sannsynlig er det at tjenestetilbyderen får kommunisere med ringen under (nærmere sentrum). Tillitssirkler representerer her en abstrakt egenskap for å vise til at modellen lagrer mer sensitiv persondata i sentrum av «løken». Under står de forskjellige ringene beskrevet i kronologisk rekkefølge. Etter det kommer et konkret eksempel på kommunikasjon mellom applikasjon/tjenestetilbyder og bruker.

1: Rådata: Lagring av rådata er representert ved et område der persondata med metadata er lagret i ubehandlet form. Data kan være alt fra rom-tid punkter (GPS med timestamps) til nettleser loggen. Innhenting av rådata kan skje ved at brukeren krever sin persondata fra for eksempel Facebook eller Google. Denne informasjonen kan brukerne kreve gjennom de nye bestemmelsene i GDPR (General Data Protection Regulation, EU). Brukerne kan også ha widgets (applikasjoner som kjører på PDS-økosystemet) som samler inn rådata kontinuerlig. En kombinasjon av forskjellige måter å hente inn rådata på vil nok fungere best da det muliggjør at brukeren får mye rådata til å starte med og at widgeten sørger for at rådata er oppdatert til enhver tid.

Når det gjelder den fysiske lagringen av rådata, vil denne modellen fremme «no-trust» konseptet. Dette innebærer funksjoner og protokoller som sikrer at ingen andre enn brukeren selv har tilgang til data. Dette kan for eksempel gjennomføres ved at data deles opp og spres på en måte som gjør at tilgang til det ene lagringsstedet ikke kompromitterer data. Data deles opp og krypteres med en tilstrekkelig algoritme.



Figur 6: PDS-økosystemets rådata håndtering/lagring

Om man velger DHT (Distributed Hash-Table) for å slå opp data spredt utover de forskjellige serverne er det viktig å være klar over mulige angrep (Sit & Morris).

2: Strukturert data: Formålet med dette laget er å sørge for at spørringer som skjer ofte ikke trenger å ha tilgang til rådata. Denne ringen representerer en lokal buffer, hvor data det spørres ofte etter kan lagres. Navn, epost og favorittfilmer er typiske ting som vil ta opp plass i dette bufferet, samt aggregerte data. Mesteparten av data som finnes i dette bufferet skal ha blitt prosessert på en eller annen måte for å hindre at rådata kommer på avveie.

3: Base widget: Også kalt «plattformer» har plass i denne ringen. Her kan man sette inn en eller flere plattformer som håndterer datatrafikken mellom de to innerste ringene og applikasjoner/tjenestetilbydere. Man kan se på dette som et sub-økosystem i økosystemet, som fortrinnsvis er open-source. En av de viktige oppgavene til plattformen er å sørge for sikker transport av data, samt analysere de ulike applikasjonenes oppførsel og rapportere overtramp. Samtykkekontrakter kan håndteres av plattformens brukergrensesnitt, men det er også en samtykkekontakt mellom PDS-økosystemet og plattformen.

Brukeren må ikke bare ha plattformer i ring 3. Organisasjoner og applikasjoner man har mer tillit til, som for eksempel brukerens bank, kan ha sin egen modul i ring 3.

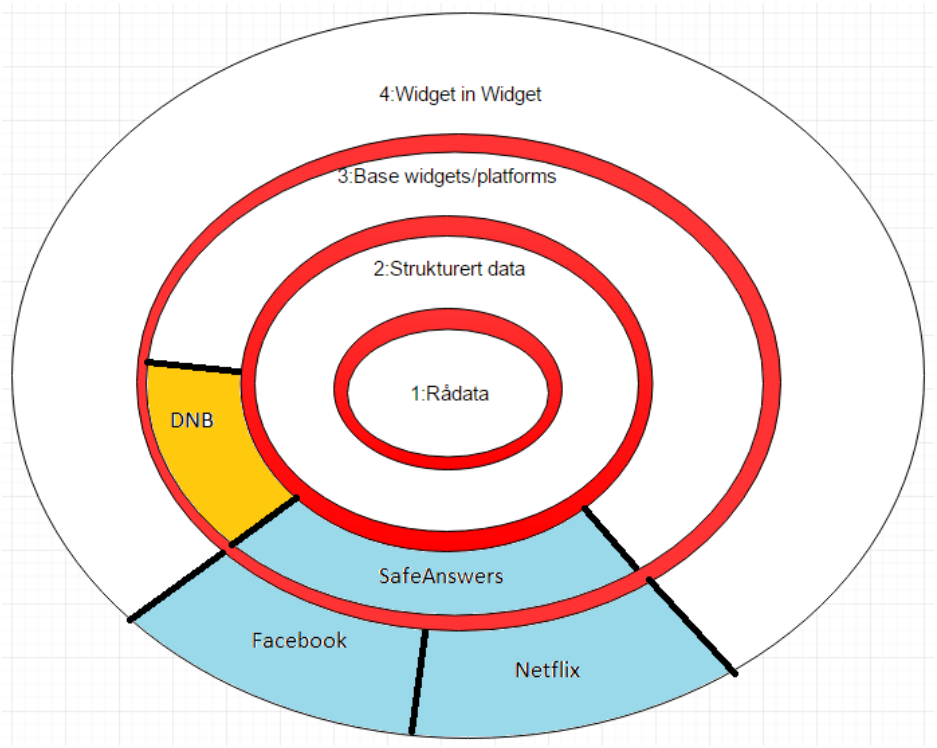
SafeAnswers («SA») brukes som et eksempel på en slik plattform. SA bruker ferdig aggregerte svar og multiparty-computation for å sikre konfidensialiteten til rådata og anonymiteten til brukeren.

4: Widget in widget: Dette er topplag hvor applikasjons-moduler kan kjøre på spesifikke plattformer. Tanken her er at widget'ene benytter et enkelt API (generert av plattformen) for å hente de data applikasjonen trenger. En nettside kan være ett eksempel på en topplag-applikasjon. I stedet for å logge seg inn på nettsiden (og ha en bruker der) spør siden om bestemt informasjon om deg. Du kan så velge å godta eller avslå samtykkekontrakten. Om du velger å godta, vil det bli gjennomført en datatransaksjon, utført av plattformen, mellom ring 1 og 2 til ring 4 (nettsiden).

De røde ringene representerer **tilgangs logg** og en begrensning kalt «**Level of paranoia**». Tilgangs logg er en enkel egenskap i økosystemet som skal gjøre det enklere å sikre konfidensialiteten og sikkerheten til data og modellen. Hver eneste dataforflytting mellom ringene/lagene blir logget med nok informasjon til at man er helt sikker på at ingen utnytter systemet eller gjør transaksjoner som det ikke er gitt tillatelse til. Det er også meningen at egendefinerte sikkerhets-widgets skal kunne utnytte denne loggen.

«Level of paranoida» illustrerer en kontrollmekanisme som gir brukeren en enkel mulighet til å begrense forflytting av persondata mellom ringene/lagene. Det kan også diskuteres om man skal skille mellom persondata og sensitiv persondata, samt hvilke data som faller under hvilket begrep. Om man setter «level of paranoia» til maksimum, vil widget'ene bare kunne få tak i en minimal mengde med persondata fra økosystemet. Lavere grad av paranoia betyr at applikasjonene/widget'ene får tilgang til mer persondata. Denne egenskapen vil sette restriksjoner på dataforflyttingen i en generell forstand, muligens kanskje med 3-5 forskjellige nivåer.

Eksempel: på figuren under ser man at Den Norske Bank (DNB) og SafeAnswers, en applikasjon og en plattform man stoler på ligger i «tillitsring» 3 mens, Facebook og Netflix ligger i «tillitsring» 4, plassert på SafeAnswers plattformen.



Figur 7: PDS-økosystemets plassering av applikasjoner/plattformer

Konkret eksempel mellom bruker og Netflix: Når en bruker åpner applikasjonen Netflix, spør Netflix brukeren om han/hun har installert SafeAnswers (SA) plattformen eller lignende plattformer. Netflix har da allerede rutiner for integrering mot SA-APIet, slik at de kan kommunisere med SafeAnswer-plattformen på brukerens PDS-økosystem. Første gangen brukeren åpner applikasjonen vil brukeren bli møtt med en enkel samtykkekontrakt, litt som den man takker ja til når man installerer nye applikasjoner på Android-telefoner. Bortsett fra at brukeren har større rettigheter til å velge hva han/hun skal godta eller avslå. Det må her skilles mellom hva Netflix MÅ ha for å kunne tilby tjenesten (identitet) og hva Netflix vil ha for å kunne tilby en optimal tjeneste (finne favoritt sjanger osv.).

Informasjonsflyten blir i dette eksempelet som følgende: Netflix spør etter avtalt informasjon fra SA (SafeAnswers), hvor SA spør «strukturert data» (ring 2) eller «rådata» (ring 1) om informasjonen det spørres om. SA aggregerer et svar fra ring 1 eller finner det i ring 2 og sender det til Netflix.

8 Studiens begrensninger og framtidig arbeid

I dette kapittelet skal jeg ta opp studiens utfordringer og begrensninger. Med hovedfokus på tid, dybde, GDPR og utfordringer rundt rettsinformatikk som disiplin. Etter det vil jeg si noen ord om framtidig arbeid og videreutvikling av modellen.

Tid og dybde: mesteparten av denne studien ble gjennomført mot slutten av den tiden som var tilgjengelig. Mye tid i starten ble brukt til å lete etter riktig problemstilling og litteratur. Det var utfordrende å finne ut hva jeg skulle søke etter og hvordan jeg skulle gå fram for å finne mer utfyllende informasjon. Nå skal det sies at det var usikkert om studien skulle benytte seg av PDS-løsninger. Helt i starten ble mange tankeeksperimenter rundt variasjoner av dagens løsning diskutert og modellert. De seks siste månedene ble jeg mer sikker på den videre utviklingen av studien og mye skulle falle på plass på den lille tiden. To runder med intervjuer skulle testes og gjennomføres, der mye modellutvikling skulle ta plass mellom dem. Den første intervjurunden ble forsinket, grunnet at flere av deltakerne ikke hadde tilstrekkelig med tid (rett før og i juleferien). Det var ikke slik at jeg kunne bytte intervjuobjekter, siden de jeg skulle spørre var utviklerne av OpenPDS og Mydex. Mye stod «på vent» i den første intervjurunden da jeg skulle bruke tilbakemeldingen fra denne runden til å evaluere problemstillingen og perspektivet i studien. Jeg skulle også bruke tilbakemeldingen som veiledning til å finne mer litteratur. Det viktigste å få fram her er at jeg ønsket å gå nærmere inn i detaljene, når det gjelder protokoller og funksjoner, om jeg hadde hatt mer tid mot slutten.

GDPR: det er sentralt å få med seg at denne oppgaven ble ferdigsluttet våren 2016, hvor det fortsatt er god tid for tjenestetilbydere å finne smutthull og frie tolkninger i GDPR. Som tidligere nevnt har alle som prosesserer, håndterer og lagrer persondata ca. 2 år på seg til å gjennomføre nødvendige endringer i praksisen slik at de overholder GDPR-reguleringene. Det kan også forekomme forandringer eller tekniske finurligheter som gjør at bestemmelsene ikke blir implementert på en tilsiktet måte. Dette vil da erodere vekk noe av grunnpilarene til denne studien, siden den tar GDPR som den står i dette tidspunkt som utgangspunkt for diskusjonene og modellen.

Området som omfatter **personvern** og **persondata**, ligger mellom informasjons teknologi og rettsvitenskap. Selve reguleringen av personvernet gjøres av rettssystemet, mens modelleringen og implementeringen/utviklingen av protokollene utføres at tekniske eksperter. At de som definerer persondata-reguleringen ofte er juridiske eksperter og ikke systemarkitekter, kan være en utfordring når teori skal bli praksis. Dette gjelder også andre veien, ved at de som designer systemer ikke er juridiske eksperter. I tillegg til denne utfordringen av å ligge mellom to disipliner, er den moderne personvern-problemstillingen relativt fersk, historisk sett. Det finnes også veldig store verdier i persondata, noe som har hatt en innvirkning på utviklingen.

Siden lover, reguleringer og teknologisk utvikling er dynamiske faktorer, vil de forandre seg med tiden. Tiden vil også forandre samfunnet og konteksten disse faktorene relateres til. For å sette det litt på spissen, vil det å håndtere persondata-problemstillingen perfekt, bli litt som: å skyte to leirduer (regulering og tekniske muligheter) med ett skudd (implementeringen). Da er det veldig viktig med riktig ståsted (innfallsvinkel).

Framtidig arbeid

Meningen med denne studien var å gi en oversiktlig fremstilling av PDS-konseptets plass og muligheter i fenomenet: lagring, prosessering og håndtering av persondata. Her er de viktigste aspektene ved fenomenet belyst og diskutert, og oppgaven vil være relevant for utviklere og andre som ønsker overblikk over PDS-økosystemer. Forhåpentligvis vil denne oppgaven være behjelpelig med å sette fremtidige interessenter raskere inn i fenomenets utfordringer og muligheter.

Et økosystem-bytte fra dagens modell til PDS ville også jevnet ut konkurransen mellom store etablerte bedrifter som eier digre datasentre fylt med persondata og nyetablerte startup-bedrifter. I hvert fall i forhold til brukerpreferanser og muligheten til å skreddersy opplevelsen til hver enkelt bruker.

Det er også selvfølgelig ønskelig at noen videreutvikler modellen eller bruker den til å finne løsninger på lignende problemstillinger. En av de tingene modellen trenger er en spesifisering av protokoller og matematiske algoritmer med tilhørende logikk.

Litteraturliste

Bibliografi

Au, D. (2014). Steps to Implementing a Zero Trust Network. *url: http://www.securityweek.com/steps-implementing-zero-trust-network.*

Boitnott, J. (2014). Why We (Mostly) Don't Care What Social Media Sites Do With Our Data. *url: http://www.inc.com/john-boitnott/why-we-mostly-don-t-care-what-social-media-sites-do-with-our-data.html*.

Brock, A. (2016). Perspectives on Blockchains and Cryptocurrencies. *url: http://artbrock.com/blog/perspectives-blockchains-and-cryptocurrencies*.

Cavoukian, A. (2011). Privacy by Design. *url: https://www.ipc.on.ca/images/resources/7foundationalprinciples.pdf*.

Cranor, L. (2012). P3P is dead, long live P3P! *url: http://lorrie.cranor.org/blog/2012/12/03/p3p-is-dead-long-live-p3p/*.

Davies, H., & Marks, S. (2015). Revealed: how Google enlisted members of US Congress it bankrolled to fight $6bn EU antitrust case. *url: http://www.theguardian.com/world/2015/dec/17/google-lobbyists-congress-antitrust-brussels-eu*.

EU. (2012). Data protection reform: Frequently asked questions. *url: http://europa.eu/rapid/press-release_MEMO-12-41_en.htm?locale=en*.

EU. (2015). Financial Services: Commission requests Portugal to ensure proper application of EU rules for investment funds. *url: http://europa.eu/rapid/press-release_IP-10-1246_en.htm?locale=en*.

EU. (2015). GDPR. *url: http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P7-TA-2014-0212&language=EN*.

EU. (u.d.). Transferring your personal data outside the EU. *url: http://ec.europa.eu/justice/data-protection/data-collection/data-transfer/index_en.htm*.

Gallivan, M. J. (2001). Striking a Balance Between Trust and Control in a Virtual Organization: A Content Analysis of Open Source Software Case Studies.

Greenberg, A. (2015). MIT's Bitcoin-Inspired 'Enigma' Lets Computers Mine Encrypted Data. *url: http://www.wired.com/2015/06/mits-bitcoin-inspired-enigma-lets-computers-mine-encrypted-data/*.

Grothaus, M. (u.d.). I Asked a Privacy Lawyer What Facebook's New Terms and Conditions Will Mean for You. *url: http://www.vice.com/read/i-asked-a-lawyer-how-facebooks-new-terms-will-affect-my-online-life-183*.

Hansen, K. M., & Manikas, K. (2012). Software ecosystems - A systematic literature review.

Hardekopf, B. (2015). The Big Data Breaches of 2014. *url: http://www.forbes.com/sites/moneybuilder/2015/01/13/the-big-data-breaches-of-2014/#488227de3a48*.

Heimes, R. (2016). Top 10 operational impacts of the GDPR: Part 1 – data security and breach notification. *url: https://iapp.org/news/a/top-10-operational-impacts-of-the-gdpr-part-1-data-security-and-breach-notification/*.

Hunton & Williams. (2015). The EU General Data Protection Regulation. *url: https://www.huntonprivacyblog.com/2015/12/17/the-eu-general-data-protection-regulation/*.

Kahn, J. (2016). UK Government Told to Revise Spy Law. *url: http://www.bloomberg.com/news/articles/2016-02-01/u-k-should-revise-proposed-spying-law-parliament-group-says*.

Meyer, D. (2016). Brazil Arrests Senior Facebook Exec Over WhatsApp Aid In Drug Case. *url: http://fortune.com/2016/03/01/brazil-facebook-arrest/*.

Mydex charter. (u.d.). url: https://pds.mydex.org/mydex-charter.

Mydex CIC. (u.d.). Mydex Security Model. *url: https://dev.mydex.org/fyi/security-model.html*.

Mydex payment. (u.d.). url: https://community.mydex.org/question/why-dont-individuals-pay-mydex-services#node-807.

Mydex tariff. (u.d.). url: https://pds.mydex.org/tariff-table.

Mydex. (u.d.). Trust framework. *url: http://openidentityexchange.org/ resources/oix-trust-frameworks-2/mydex-trust-framework/*.

OpenPDS. (u.d.). Homepage. *url: http://openpds.media.mit.edu/*.

OpenPDS. (u.d.). Protecting the Privacy of Metadata through SafeAnswers. *url: http://journals.plos.org/plosone/article?id=10.1371/ journal.pone.0098790#pone.0098790-Schwab1*.

P3P. (2006). Platform for Privacy Preferences. *url: https://www.w3.org/P3P/*.

Piyawongwisal, P., & Xia, P. (2011). Sybil Attack and Defense in P2P Networks. *url: http://pbg.cs.illinois.edu/courses/cs538fa11/lectures/19-Henry-Pratch.pdf*.

PQCrypto. (u.d.). url: https://pqcrypto.org/.

Purtova, N. (2015). The Illusion of Personal Data as No One's Property.

Rich, S., & Gellman, B. (2014). NSA seeks to build quantum computer that could crack most types of encryption. *url: https://www.washingtonpost.com/world/national-security/nsa-seeks-to-build-quantum-computer-that-could-crack-most-types-of-encryption/2014/01/02/8fff297e-7195-11e3-8def-a33011492df2_story.html*.

Sandland, V. (2016). Økning i ID-tyverier. *url: https://norsis.no/2016/02/okning-i-id-tyverier/*.

Single, R. (2011). Dropbox Lied to Users About Data Security, Complaint to FTC Alleges. *url: http://www.wired.com/2011/05/dropbox-ftc/*.

Sit, E., & Morris, R. (u.d.). Security Considerations for Peer-to-Peer Distributed Hash Tables.

Stone Business Law, P.C. (u.d.). Why aren't contracts written so that ordinary people can read them? *url: http://www.stonebusinesslaw.com/resource/general-business-law/why-aren%E2%80%99t-contracts-written-so-ordinary-people-can-read-them*.

Thylmann, O. (2011). Why did p3p fail? *url: https://www.quora.com/Why-did-p3p-fail*.

UK Government. (2014). Review of the midata voluntary programme. *url: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/327845/bis-14-941-review-of-the-midata-voluntary-programme-revision-1.pdf*.

Yin, R. K. (2014). *Case Study Research.* SAGE.

Zyskind, G., Nathan, O., & Pentland, A. '. (u.d.). Enigma: Decentralized Computation Platform with. *url: http://enigma.media.mit.edu/enigma_full.pdf*.

Vedlegg 1

Intervjuguide for første runde intervjuer. Spørsmålene til OpenPDS og Mydex var like, sett bort i fra at der det står OpenPDS stod det Mydex i Mydex sitt spørreskjema. Spørreskjemaet til rettsinformatikk varierer noe fra det OpenPDS og Mydex fikk og er vedlagt under.

Questions to OpenPDS.

Purpose: Develop a framework and select its core focus/facilities for sharing personal data between a PDS (Personal Data Storage) and Apps/commercial actors.

Current plan of action: Two sessions where in the first one I will ask a few questions about current standard and how OpenPDS may be part of the solution. The key point in the first exploratory round is to identify and map the various dimensions. Further, it will help to draw the boundaries of the thesis. After this first round I will set up a framework, which will be the focus in the last question-session.

Foundation: First we will look at aspects of privacy and data ownership, including legal and trust dimensions of the current standard. Here we are interested in finding out where the shoe pinches.

Questions concerning foundation:

Introductory questions:
x) Currently the online service providers have ownership and control over the data gathered from users.

Q1: Why is this the standard?

Q2: What are the strengths and weaknesses to this model?

Q3: Does it exist any alternatives? What's their strengths and weaknesses?

Q4: How should one go about when introducing a new model? Obstacles and/or incentives?

5: What do you think is the main pitfalls on creating a system for sharing data between a personal PDS and applications?

6: Do you think it is possible to "save" the personal data privacy in the long run, and do you have any thoughts how we can achieve this?

7: How does OpenPDS ensure the trust of the users?

8: Is the OpenPDS still in development?

9: Any thoughts on preparing the system (OpenPDS) for the future? And the amount of resources this will require?

10: From what I can understand OpenPDS is pro open source. Is there any aspect of the OpenPDS system that is not open source?

Questions for IT/Law

Purpose: Develop a framework and select its core focus/facilities for sharing personal data between a PDS (Personal Data Storage) and Apps/commercial actors.

Current plan of action: Two sessions where in the first one I will ask a few questions about current standard (The accumulation of user-data stored and used at organizations and businesses). The key point in the first exploratory round is to identify and map the various dimensions. Further, it will help to draw the boundaries of the thesis.
After this first round I will set up a framework, which will be the focus in the last question-session.

Foundation: First we will look at aspects of privacy and data ownership, including legal and trust dimensions of the current standard. Here we are interested in finding out where the shoe pinches.

Definition of PDS:

PDS

PDS is short for personal data store, and is a term describing a system that stores all data and/or meta-data about the user. Think of it as a cloud service working for you, logging every action, click, GPS-coordinates of your very life. This is all done today by Apps and websites owned by organizations, companies, and other digital entities monitoring your life.

Just the basics (standard contract for sharing data) not analytics tools running on the data.

Alternate terms used to describe the same system; Personal data vault, personal data storage and personal data lockers.

Main features:
1. (level 1) Secure storage (encryption),
2. (level 1) Store identity credentials and be able to view who they have shared the data with.
3. (level 1) Standard contract (with time-tokens) for use of personal credentials.
4. (level 2) Store meta data on user behaviour.
5. (level 2) All data processing and/or analytics happens inside the PDS.
6. (level ?) Identity manager/provider.

Level 1: Is must have to be considered a PDS. Level 2 and up is considered additional features but can also be included in the PDS' range of operation.

Definition of Online service provider:

Organizations, individuals, or entities offering a digital service, often in the form of an application running on your phone or pc/mac.

Questions

Introductory questions:
x) Currently the online service providers have ownership and control over the data gathered from users.

Q1: Why is this the standard?

Q2: What are the strengths and weaknesses to this model?

Q3: Does it exist any alternatives? What's their strengths and weaknesses?

Q4: How should one go about when introducing a new model? Obstacles and/or incentives?

5: What do you think is the main pitfalls on creating a system for sharing data between a personal PDS and applications?

6: Do you think it is possible to "save" the personal data privacy in the long run, and do you have any thoughts how we can achieve this?

7: Any thoughts on preparing the system for the future? And the amount of resources this will require?

8: Can you think of any legal implications on letting the users own their data in a global distributed system?

9: Is there anything major/central I have missed or some issue I haven't addressed?

Vedlegg 2

Intervjuguide for andre runde intervjuer.

Spørreskjema om PDS

I dette skjemaet skal jeg forklare konseptet «PDS» og stille tre spørsmål rundt det. Etter det kommer en modell for håndtering, prosessering og lagring av persondata, med to tilhørende spørsmål.

Hva er en PDS?

PDS står for Personal Data Store, og beskriver et system som lagrer all data og metadata om en bruker. Det finnes mange forskjellige definisjoner av PDS. Noen gir begrepet flere egenskaper, andre færre. Her bruker vi begrepet i en ganske grunnleggende form hvor vi utvider begrepet til «PDS-økosystem» for å legge til egenskaper som vil gjøre modellen attraktiv. Andre termer som beskriver personal data store er: personal data vault og personal data locker.

En grunnleggende PDS innehar en enkel protokoll for deling av data, samt innhenting av den. Flere abstraksjonslag som analyseverktøy og mer komplekse handlinger vil gjøre at konseptet faller inn under PDS-økosystem.

Kort liste over funksjoner hvor grad 1 tilhører bare PDS og grad 2 tilhører PDS-økosystem.

Grad 1: - Sikker lagring av data.

- Lagring av ID-kreditter
- Se hvem du deler hvilke data med, med mulighet for å endring
- Standard kontrakter for samtykke, for bruk at dataen.
- Muligheten til enkel prosessering og innhenting av dataen.

Grad 2: - Analyse og mer avansert prosessering av dataen.

- Komplekse strukturer for deling av dataen.
- Multi-party-computation. Muliggjør avansert form for anonymisering.

Hva kan et PDS-økosystem brukes til?

Mulighetene til et PDS-økosystem er veldig mange. I stedet for at dataen din lagres hos ulike aktører som Facebook, Netflix, Goolge etc. så sentraliseres den hos brukeren/deg. Dataen

sentraliseres hos brukeren i form av at det kun er brukeren som har tilgang til rådataen, og brukeren bestemmer hvem andre som skal ha hvilken tilgang. Dette vil si sin tur føre til at man ikke trenger å registrere seg på hver eneste nettside man besøker, man gir bare nettsiden tilgang til den dataen den trenger. Eiendeler kan enkelt kobles opp til den digitale identiteten, og identiteten vil interagere lettere med «Internet of Things» løsninger.

Sannsynligheten av identitetstyveri og misbruk av dataen vil minkes kraftig, siden all dataen samles under brukerens kontroll.

Evalueringskriterier

I denne studien blir har vi valgt å se på OpenPDS, Mydex og Enigma. Disse tre prosjektene har forskjellige angrepsstrategier for å introdusere PDS til massene. Vi skal ikke gå innpå hvordan de fungerer her, men jeg er interessert i om jeg har valgt de riktige evalueringskriteriene.

Jeg har valgt å bruke egenskapene: gjennomsiktighet, intensjon og modelldesign/funksjoner for å evaluere hvorvidt det bestemte PDS-løsningen holder vann.

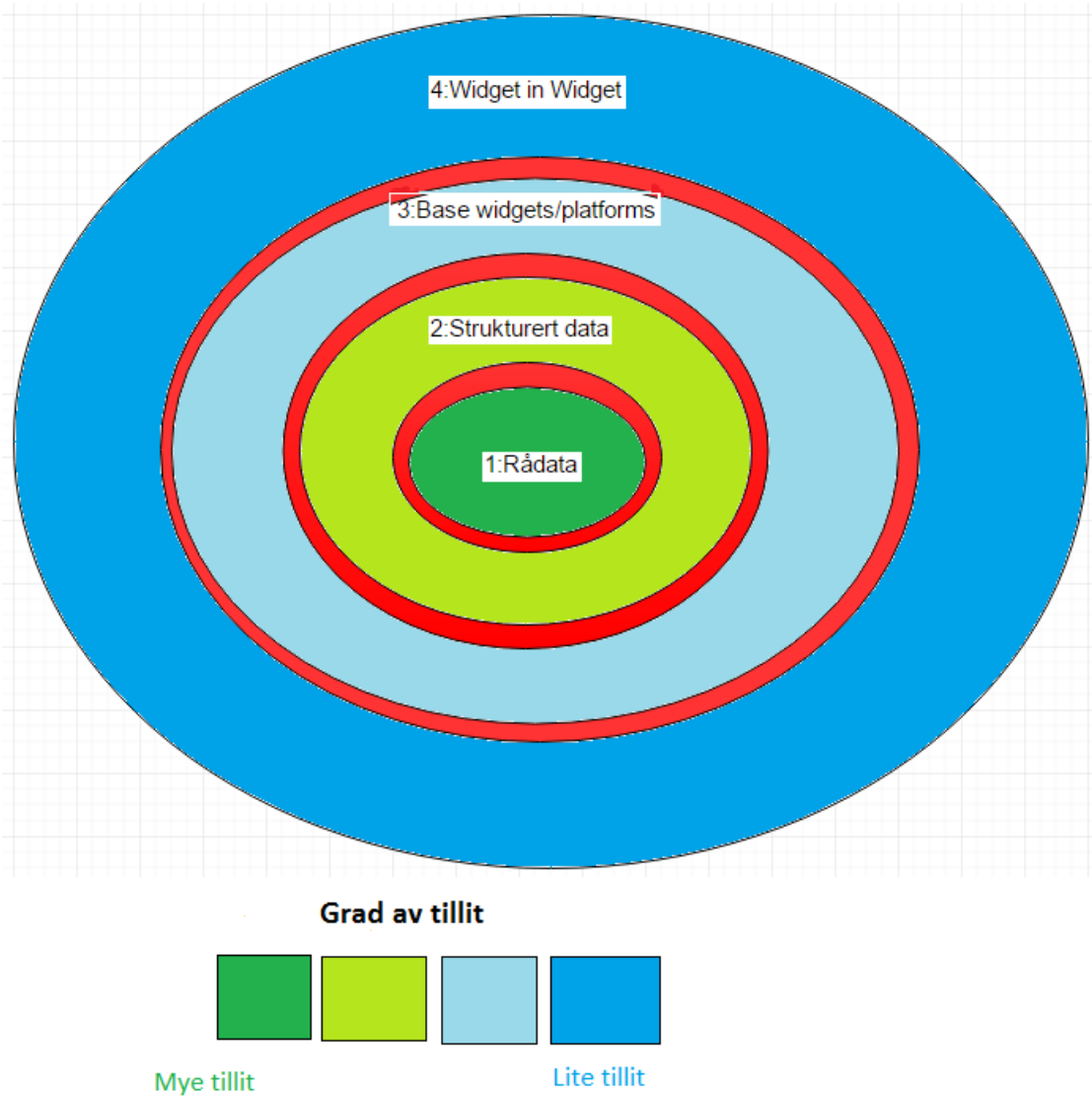
Gjennomsiktighet: muligheten til å se all dataflyt sammen med hva slags samtykke som er gitt.
Intensjon: hvilke intensjoner har bedriften for å tilby en PDS-løsning. Her vil vi se på forretningsmodellen til bedriften.
Modelldesign/funksjoner: arkitektoniske og tekniske løsninger på ulike utfordringer. Nøkkelbegreper her er sikkerhet, fremtidsrettet og brukervennlighet.

Er kriteriene dekkende nok?

Ville du brukt flere eller færre kriterier?

Om du synes dette var få kriterier, hvilke andre kriterier ville du ha inkludert?

Modellen

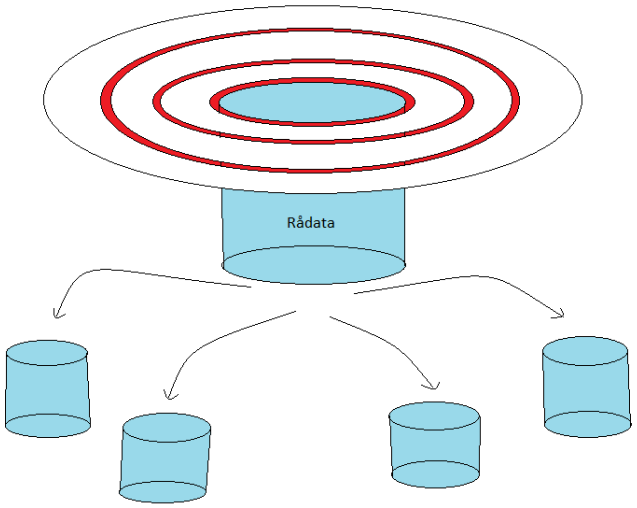


Spørringer gjøres fra de ytre lagene og innover, mens svar på spørringene går fra indre lag og utover. De røde ringene representerer grader av tillit som man må passere, dette utdypes under. Tillitssirkler representerer her bare en abstrakt egenskap for å vise til at mer sensitiv data lagres i sentrum av «løken». Under står de forskjellige ringene beskrevet i kronologisk rekkefølge. Nederst er det også et konkret eksempel på kommunikasjon mellom applikasjon og bruker.

1: Rådata: Lagring av rådata er representert ved et område der persondata med metadata er lagret i ubehandlet form. Denne dataen kan være alt fra «spatio-temporal» punkter (GPS med timestamps) til nettleser loggen. Samling av dataen kan skje gjennom en stor chunk av persondata, fra for eksempel Facebook eller Google. Denne informasjonen kan man kreve gjennom de nye paragrafene i GDPR (General Data Protection Regulation, EU). Man kan også ha widgets som samler inn denne dataen kontinuerlig. En kombinasjon av begge vil nok

fungere best, med at man får mye data til å starte med og at den vil være oppdatert til enhver tid.

Når det gjelder den fysiske lagringen av rådataen, vil denne modellen fremme «no-trust» konseptet. Dette innebærer funksjoner og protokoller som sikrer at ingen andre enn brukeren selv har tilgang til dataen. Dette kan for eksempel gjennomføres ved at dataen deles opp og spres på en måte som gjør at tilgang til det ene lagringsstedet ikke kompromitterer dataen. Dataen deles opp og krypteres med en tilstrekkelig algoritme, og den eneste som har oversikt over hvordan dataen er spredt er brukeren.



Om man velger DHT (Distributed Hash-Table) for å slå opp dataen spredt utover de forskjellige serverne er det viktig å være klar over mulige angrep [https://web.eecs.umich.edu/~zmao/eecs589/papers/p2pSec.pdf].

2: Strukturert data: Formålet med dette laget er å sørge for at frekvente spørringer ikke trenger å fiske opp rådata. Denne ringen representerer en lokal buffer, hvor data det spørres ofte etter kan lagres. Navn, epost og favoritt filmer er typiske ting som vil ta opp plass i dette bufferet, samt aggregerte og kumulerte data. Mesteparten av dataen som finnes i dette bufferet skal ha blitt prosessert på en eller annen måte for å hindre at rådata plukkes opp av utilsiktede entiteter.

3: Base widget: Eller plattformer har plass i denne ringen. Her kan man sette inn en eller flere plattformer som håndterer datatrafikken mellom de to innerste ringene og applikasjoner/nettsider. Man kan se på dette som et sub-økosystem i økosystemet, som

fortrinnsvis er open-source. En av de viktige oppgavene til plattformen er å sørge for sikker transportering av data, samt analysere de ulike applikasjonenes oppførsel og rapportere overtramp. Samtykkekontrakter kan håndteres av plattformens brukergrensesnitt, men det er også en samtykkekontakt mellom PDS-økosystemet og plattformen.

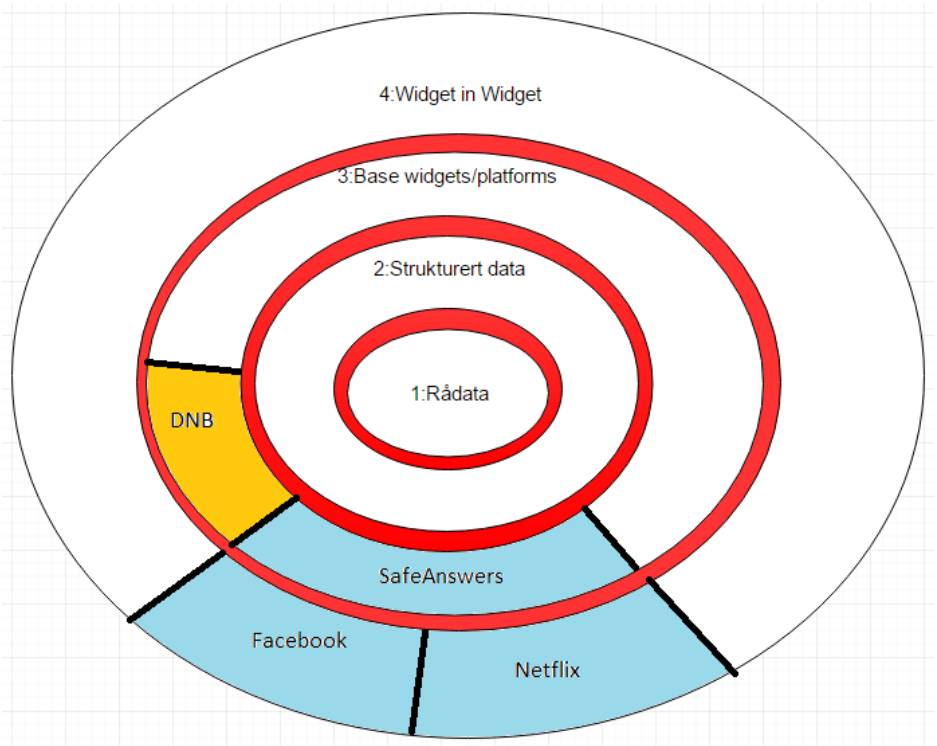
> Man trenger ikke bruke disse plattformene for å dele persondata mellom brukeren og applikasjoner. Organisasjoner og applikasjoner man har mer tillit til, som for eksempel brukerens bank, kan ha sin egen modul i ring 3.
>
> SafeAnswers brukes som et eksempel på en slik plattform. SA bruker ferdig aggregerte svar og multiparty-computation for å sikre konfidensialiteten til rådataen og anonymiteten til brukeren.

4: Widget in widget: Dette er topplag hvor applikasjoner kan kjører på spesifikke plattformer. Tanken her er at widget'ene benytter et simplistisk API (generert av plattformen) for å hente den dataen applikasjonen trenger. En nettside kan være ett eksempel på en topplag-applikasjon. I stedet for å logge seg inn på nettsiden (og ha en bruker der) spør siden om bestemt informasjon om deg. Du kan så velge å godta eller avslå samtykkekontrakten. Om du velger å godta, vil det bli gjennomført en datatransaksjon, utført av plattformen, mellom ring 1 og 2 til ring 4 (nettsiden).

De røde ringene representerer **tilgangs logg** og en begrensning kalt «**Level of paranoia**». Tilgangs logg er en enkel egenskap i økosystemet som skal gjøre det enklere å sikre konfidensialiteten og sikkerheten til dataen og modellen. Hver eneste dataforflytting mellom ringene/lagene blir logget med nok informasjon til at man er helt sikker på at ingen utnytter systemet eller gjør transaksjoner som det ikke er gitt tillatelse til. Det er også meningen at egendefinerte sikkerhets-widgets skal kunne utnytte denne loggen.

«Level of paranoida» illustrerer en kontrollmekanisme som gir brukeren en enkel mulighet til å begrense dataforflyttingen av persondata og sensitiv data mellom ringene/lagene. Det kan også diskuteres om man skal skille mellom persondata og sensitiv data, samt hvilke data som faller under hvilket begrep. Om man setter «level of paranoia» til maksimum, vil widget'ene bare kunne få tak i en minimal mengde med data fra økosystemet. Lavere grad av paranoia betyr at applikasjonene/widget'ene får tilgang til mer data. Denne egenskapen vil sette reskriksjoner på dataforflyttingen i en generell forstand, muligens kanskje med 3-5 «levels».

Eksempel: på figuren under ser man at DNB og SafeAnswers, en applikasjon og en plattform man stoler på ligger i «tillitsring» 3 mens, Facebook og Netflix ligger i «tillitsring» 4, plassert på SafeAnswers plattformen.



Konkret eksempel mellom bruker og Netflix: Når en bruker åpner applikasjonen Netflix, spør Netflix brukeren om han/hun har installert SafeAnswers plattformen eller lignende plattformer. Netflix har da allerede rutiner og API som kan kommunisere med SafeAnswer-plattformen til brukeren. Første gangen brukeren åpner applikasjonen vil brukeren bli møtt med enkel samtykkekontrakt, litt som den man takker ja til når man installerer nye applikasjoner på Android-telefoner. Bortsett fra at brukeren har større rettigheter til å velge hva han/hun skal godta eller avslå. Det må her skilles mellom hva Netflix MÅ ha for å kunne tilby tjenesten (identitet) og hva Netflix vil ha for å kunne tilby en optimal tjeneste (finne favoritt sjanger osv.).

Informasjonsflyten blir i dette eksempelet som følgende: Netflix spør etter avtalt informasjon fra SA (SafeAnswers), hvor SA spør «strukturert data» (ring 2) eller «rådata» (ring 1) om informasjonen det spørres om. SA aggregerer et svar fra ring 1 eller finner det i ring 2 og sender det til Netflix.

Ser du noen umiddelbare juridiske, tekniske eller arkitektoniske utfordringer?

Hva skal til for at utviklere ønsker å omfavne en slik modell?