

UiO • **Institutt for informatikk**

Det matematisk-naturvitenskapelige fakultet

INF2260

Prosjektrapport høsten 2016

ivarsjo@uio.no snhagen@uio.no
ehfugler@uio.no sodaniel@uio.no



1 Innledning	3
1.1 Domene	3
1.2 Rapportens struktur	3
2 Design fiction	4
3 Metode	5
3.1 Metoder for innsikt	5
3.2 Metoder for evaluering	7
3.3 Utdfordringer	7
4 Designprosessen	8
4.1 Innsiktsfase	8
4.1.1 Litteraturstudie	9
4.1.2 Intervju med domeneekspert	10
4.1.3 Fremtidens flygeledelse	10
4.2 Utvikling av et fremtidsscenario	11
4.3 Future Workshop	11
4.3.1 Fremtidige teknologier	12
4.4 Utvikling og evaluering av storyboards	13
4.4.1 Implikasjoner fra evaluering	13
4.5 Høyoppløselig prototype	14
5 Evaluering	15
5.1 Ekspertevaluering	15
5.2 Spørreundersøkelse	17
6 Konkluderende diskusjon	18
Litteratur	20

1 Innledning

Denne prosjektrapporten inngår i SINTEFs pågående forskningsprosjekt Next Generation Decision Support Tools (NextGenDST). Målet med prosjektet er å utvikle beslutningsunderstøttende systemer som kan bidra til økt prestasjon i tidskritiske og komplekse domener – i dette tilfellet kontroll av flytrafikk. Vår rolle i dette forskningsprosjektet har bestått i å utforske hvordan mennesker som operatører kan kommunisere med algoritmer og automatiserte systemer, med utgangspunkt i følgende spørsmål:

- Hvordan kan *virtual reality* som grensesnitt understøtte bruk av automatiserte systemer for å fatte beslutninger?

Formålet med oppgaven har vært å fremstille og evaluere løsninger i en fremtidskontekst, ved å ta i bruk *design fiction* som overordnet metodologisk tilnærming. Vi legger i denne rapporten frem designimplikasjoner for fremtidens løsninger, basert på evaluering av en mulig fremtidsløsning, fremstilt i form av en kontekstuell, narrativ videoprototype (se lenke i litteraturliste).

1.1 Domene

SINTEF NextGenDST er et forskningsprosjekt som befinner seg i krysningen mellom HCI og automatisering. Prosjektet retter seg mot tidskritiske og komplekse domener hvor dårlige beslutninger kan ha alvorlige konsekvenser, eksempelvis i flygeledelse og beredskapsledelse. SINTEFs mål er å introdusere et neste generasjons støtteverktøy for å gjøre disse valgene mer effektive og sikre, samtidig som man ikke forringer kompetansen til den menneskelige operatøren. Slik SINTEF legger det frem krever dette at man tenker på menneske-maskin-samarbeid på en helt ny måte, der graden av automatisering burde være dynamisk og fleksibel, tilpasset operatørens, oppgavens og situasjonens behov. Prosjektet ligger under SINTEF sin avdeling for Networked Systems and Services, og er et samarbeid mellom forskningsgruppene som jobber med automatisering og HCI.

1.2 Rapportens struktur

I kapittel 2 tar vi for oss de viktigste trekkene ved, og hvordan vi har forholdt oss til, prosjektarbeidets overordnede metodologiske tilnærming. I kapittel 3 beskrives de viktigste

metodene for datainnsamling og innsikt, før vi i kapittel 4 presenterer resultatene fra analysen og fremstillingen av vår endelige prototype. I kapittel 5 presenterer vi sentrale funn i vår evaluering. Til slutt oppsummerer og drøfter vi de viktigste resultatene i kapittel 6.

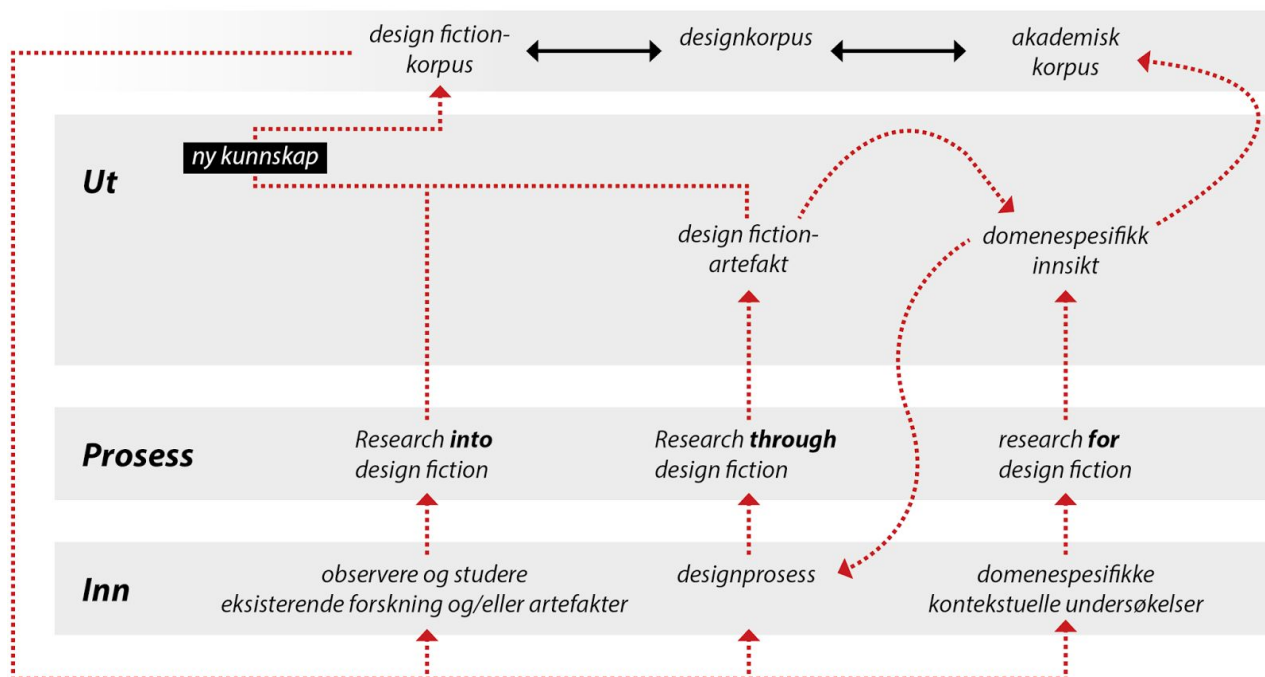
2 Design fiction

Domenet *flygeledelse* er stort og komplekst, noe som medfører at endring og fornying oftest foregår i små inkrementer. Målet med denne oppgaven har imidlertid vært å undersøke hvilke implikasjoner en radikal endring av dagens arbeidsmetoder vil medføre. Streng sikkerhetsregler og rutiner blant flygeledere har i tillegg medført at vi ikke har hatt tilgang til hverken den aktuelle målgruppen eller konteksten de opererer i. På grunn av dette, og klare ønsker fra SINTEF både om å utforske teknologier som er tidkrevende å prototype, samt å tilnærme oss problemstillingen på en eksperimentell måte, har vi valgt å tilnærme oss problemstillingen ved hjelp av metodologien *design fiction*.

Begrepet design fiction er relativt nytt innen Human-Computer Interaction (HCI), og krediteres ofte forfatteren Bruce Sterling, som i 2005 beskrev metodologien som «the deliberate use of diegetic prototypes to suspend disbelief about change». Ordet *diegetic* leses best på norsk som *narrativ*. Design Fiction benytter seg således av narrative prototyper i en fremtidskontekst, altså konteksten disse prototypene eksisterer i, for å undersøke og kommunisere mulige fremtidige teknologier (Tanenbaum, 2014).

Vi støtter oss i denne oppgaven på rammeverket Joseph Lindley (2015) beskriver som *research through design fiction*. Med dette antar vi at det er mulig å akkumulere kunnskap ved å gjennomføre design fiction i praksis. Målet med denne prosessen er slik sett ikke å fremstille et materielt sluttprodukt, men å produsere domenespesifikk innsikt i problemområdet vi undersøker. I dette inngår også *research for design fiction*: kontekstuelle undersøkelser utført for å understøtte designet av en fremtidsfiksjon (Lindley, 2015). Resultatene fra vår analyse vil kunne indikere hvorvidt design fiction egner seg til å løse denne typen problemstillinger, og vi mener derfor dette prosjektet også kan tolkes som *research into design fiction*. Ved å ta i bruk design fiction som overordnet metodologi for vår designprosess, har vi blant annet unngått å låse oss til én bestemt teknologi. Dette har igjen gjort arbeidet med å fremstille testbare prototyper enklere. Samtidig har vi ikke vært

bundet til de muligheter dagens teknologi kan understøtte. Vi ser dette som relevant da både *virtual reality* og flygeledelse er domener i stadig endring.



Figur 1: Joseph Lindleys pragmatiske rammeverk for design fiction.

3 Metode

Virtual reality som grensesnitt for visualiseringsløsninger i flygeledelse er lite utforsket. Vi anser derfor dette som en eksplorativ studie. På bakgrunn av dette, og at målet med prosjektet var å fremstille en prototype plassert i en spesifikk kontekst, så vi en kvalitativ tilnærming til problemstillingen som nødvendig, spesielt da vi selv har begrenset kjennskap til omfanget og kompleksiteten i domenet flytrafikk og flygeledelse. Data presentert i denne oppgaven er samlet inn gjennom litteraturstudier, et forskningsintervju, workshops med domeneeksperter samt en spørreundersøkelse. Intervjuobjektene er et strategisk utvalg, da informantene har kvalifikasjoner som er relevante med tanke på problemstillingen.

3.1 Metoder for innsikt

For å gjennomføre søk etter domenespesifikk litteratur baserte vi oss på nøkkelord presentert for oss i SINTEFs oppgavebeskrivelse. Noen eksempler på dette er *Air Traffic Controller*, *Future Air Traffic Control* og *Automation*. Søkeresultater i Google Scholar ble vurdert basert på antall siteringer, og relevans for oppgaven. Enkelte søkeord ga treff i et bredt spekter av artikler, og ble av

den grunn kombinert med andre søkeord for å oppnå mer presise resultater. Av de utvalgte artiklene ble flere forkastet da de viste seg ikke å være relevante for vår oppgave, mens andre igjen ga verdifull informasjon.

Forskningsintervjuet ble gjennomført i semi-strukturert, eksplorativ form, med intensjon om å gi økt forståelse av flygelederes arbeidsprosesser, og det komplekse domenet disse inngår i. Intervjuet tok utgangspunkt i overordnede temaer og forhåndsdefinerte spørsmål. Med en opportunistisk tilnærming til oppfølgingsspørsmål og intervjuobjektets kommentarer kunne vi i tillegg forvente å få svar på spørsmål vi ikke ville vært i stand til å formulere på forhånd (Lazar et al., 2010, s. 189).

Videre gjennomførte vi en *future workshop*, med intensjon om å generere innsikt i teknologiske og kontekstuelle muligheter, samt mulige utfordringer, i nær fremtid. Vi tok i dette prosjektet utgangspunkt i Åbo universitets ACTVOD-modell, et format typisk brukt til å utforske praktiske spørsmål knyttet til et domenes – eksempelvis én spesifikk industri – fremtid. En slik fremgangsmåte kan bidra til å belyse komplekse problemer ved å avdekke mange, ofte motstridende, synspunkter (Lautamäkki, 2014). I tillegg til å samle inn og generere informasjon kunne vi forvente å få bedre innsikt i hvilke sosiale implikasjoner en potensiell fremtidsløsning vil ha. Deltakerne besto av domeneeksperter og en student som i dette tilfellet deltok som representanter for nåværende og fremtidige passasjerer – en gruppe som direkte eller indirekte kan påvirkes av fremtidige løsninger. Den praktiske gjennomføringen er beskrevet i detalj i kapittel 4.5.

For å kommunisere ideer både innad i gruppen og overfor oppdragsgiver har vi benyttet oss av scenarier og *storyboards*. Hensikten med dette var å visualisere menneskelige aktiviteter og oppgaver på en måte som ga informantene rom til å diskutere og utforske ikke bare våre løsningsforslag, men også konteksten de eksisterer i (Preece et al., 2015). I tillegg tok vi internt i gruppen i bruk tekstlige beskrivelse av *use cases* for å bryte konseptuelle ideer ned i mindre komponenter. Dette så vi som spesielt viktig da et sentralt aspekt ved vår oppgave var å identifisere hvilke aktører forskjellige ansvarsområder potensielt kan tilfalle i fremtiden.

Underveis i arbeidet med fremtidsfiksjonen har vi korrigert og endret designet basert på gjennomgang med veileder og domeneekspert Amela Karahasanovic. Ved å korrigere designet basert på tilbakemeldinger fra en domeneekspert kunne vi forvente å unngå åpenbare feil eller inkonsistens i løsningen. Hadde vi ikke rettet disse feilene kunne vi risikert å trekke fokuset bort fra

selve løsningen, og satt informantene ute av stand til å identifisere de aspektene vi ønsket å undersøke i vår sluttevaluering (Lazar et al., 2010, s. 256-257).

3.2 Metoder for evaluering

Den endelige evalueringen av vår fremtidsfiksjon ble gjennomført i to faser. Et sentralt aspekt ved design fiction er å analysere et publikums – en målgruppes – reaksjon på den aktuelle fremtidsvisjonen. I vårt tilfelle med en intensjon om å genere og innhente synspunkter som kan bidra til designimplikasjoner for fremtidige løsninger. I denne oppgaven benyttet vi oss av forskere ved SINTEF som proxybrukere for den aktuelle målgruppen, flygeledere, i form av en fokusgruppe. Ved å gjennomføre evalueringen på denne måten kunne vi forvente å avdekke både fellestrekk og motstridende synspunkter (Lazar et al., 2010, s. 192). Vi ser dette som en styrke, da det minsket sjansen for at subjektive meninger skulle bli for fremtredende i vår analyse. I tillegg anser vi enighet hos informantene som en viktig indikator på hvorvidt aspekter ved vår løsning er relevante med tanke på å generere designimplikasjoner for fremtidige løsninger.

Parallelt med ekspertevalueringen gjennomførte vi en nettbasert spørreundersøkelse, med intensjon om å måle subjektiv tilfredshet – i dette tilfellet med fokus på tillit. Målgruppen for undersøkelsen var alle norske reisende som har tatt minst én reise fra en norsk flyplass, mens utvalgsrammen for undersøkelsen var vårt eget kontaktnettverk – distribuert gjennom sosiale medier. Respondentene ble bedt om å se den endelige prototypen, for så å ta stilling til påstander som ble besvart i form av likertskalaer. Likertskalaer ble brukt da det var ønskelig for oss å undersøke hvorvidt vår løsning ble oppfattet som utopisk eller dystopisk, samt eventuelle politiske eller sosiale implikasjoner ved løsningen. Med tanke på at likertskalaer er en bipolar skaleringsmetode, kunne vi forvente å få et inntrykk av om respondentene stilte seg negativt, positivt eller nøytralt til vår løsning (Allen og Seaman, 2007).

3.3 utfordringer

Metoden har vært preget av fleksibilitet, og har således gitt oss mulighet til å tilpasse oss erfaringer underveis. I all kvalitativ forskning må vi anta at det oppstår et påvirkningsforhold mellom forsker og informanter, og at vår egen forforståelse vil påvirke hvordan vi tolker data (Thaagard, 2009). Ved å gjennomføre gjentatte intervjuer og workshops har vi hatt mulighet til å korrigere vår egen forforståelse, samtidig som vi har fått nye data som fortolkes.

Som en umiddelbar utfordring ved å involvere domeneeksperter bør det påpekes at deres kunnskap om flygeledelse er basert på egne studier, noe som medfører at våre data i stor grad er basert på sekundærkilder. Bruken av workshops med praktisk oppgaveløsning og gruppesamtaler kan imidlertid bidra til å hindre at subjektive oppfatninger blir for dominerende, blant annet ved å avdekke fellestrekk blant informantene, samt legge godt til rette for individuell oppfølging og gruppediskusjon (Lazar et al., 2010, s. 192). En annen utfordring ved dette var imidlertid at ekspertene vi har forholdt oss til har vært ansatt ved SINTEF, og at de som vår oppdragsgiver i dette prosjektet kan ha personlige interesser i vår studie.

Om bruken av eget kontaktnettverk som utvalgsramme for vår spørreundersøkelse vil vi nevne at dette kan utgjøre et visst utvalgsskjevhet. Vi så det imidlertid som en kostnadseffektiv måte å oppnå kontakt med tilstrekkelig antall respondenter på. Vi vil også argumentere for at denne måten å rekruttere deltakere på kan sies å være selvvalgt undersøkelse, noe som kan være hensiktsmessig med tanke på at vi undersøker et nytt bruksfenomen (Lazar et al., 2010, s. 109). En ulempe kan på den annen side være at vi oppnår større respons fra de som er spesielt interesserte i fly og teknologi, til tross for at løsningen potensielt vil påvirke reisene til alle passasjerer. Med denne utvalgsrammen kan vi heller ikke anta at estimatene fra undersøkelsen vil dekke alle aldersgrupper, fordi vårt nettverk i stor grad består av personer på vår egen alder. Det er grunn til å tro at personer på vår egen alder i større grad er oppdatert på og har tillit til digitale løsninger – noe vi måtte ta hensyn til i analysen. Det er derimot grunn til å tro at meninger om sosial og politisk påvirkning av løsningen til en viss grad vil være jevnt fordelt i flere aldersgrupper.¹

4 Designprosessen

I dette kapittelet vil vi gjøre rede for gjennomførte datainnsamlinger, samt diskutere hvordan vi har analysert, og hvilken innsikt dette har gitt oss.

4.1 Innsiktsfase

Flygeledelse er et komplekst domene som ikke er allment kjent, noe som førte til at vi brukte lang tid på innledende innsiktsfase av studiet. Vi hadde heller ikke tilgang på flygeledere og måtte derfor

¹ Resultater fra siste stortingsvalg og skolevalg viser forholdsmessig lik oppslutning om stortingspartiene på nasjonal basis (Norsk Rikskringkasting, Samfunnsveven, 2013).

basere oss på domenespesifikk litteratur og domeneeksperter. Under denne fasen i prosjektet formulerte vi to hovedmål som skulle være veiledende for datainnsamlingen.

1. Få innsikt i flygeleders arbeidsprosesser og hvordan dagens systemer støtter opp under disse.
2. Avdekke noen sannsynlige fremtidsscenarier og sentrale temaer i diskusjonen om automatisering innenfor flygeledelse.

4.1.1 Litteraturstudie

Dette kapittelet tar i korte trekk for seg tidligere arbeid og forskning relatert til vårt prosjekt, knyttet til temaene informasjonsvisualisering, effektivisering og automatisering. Av relevant litteratur har vi lagt vekt på hvorvidt visualiseringsverktøy og automatisering vil medføre økt effektivitet blant flygeledere.

Ødegaard (2013) legger i sin masteroppgave til grunn at man ved å visualisere arbeidsflyt for flygeledere kan identifisere flaskehalser i trafikkscenarier, og konkluderer med at en løsning som visualiserer trafikkflyt legger til rette for vurdering av trafikkmønstre. I tillegg viser studien at visualiseringsløsninger kan brukes av flygeledere for å vurdere eget arbeid.

I diskusjonen om bruk av automatisering av flygelederes oppgaver støtter vi oss på Panel on Human Factors in Air Traffic Controls *The Future of Air Traffic Control: Human Operators and Automation* (1998). Panelet anbefaler at tiltak for automatisering burde understøtte flygelederes beslutningstaking i alle systemfunksjoner, og burde fokusere på tilegnelse, integrasjon og presentasjon av informasjon. I diskusjonen om graden av automatisering argumenteres det for en høy grad av automatisering av oppgaver som involverer liten risiko og usikkerhet. Der det motsatte gjelder burde automatiseringen kun foreslå foretrukne alternativer. Automatisering over dette nivået burde designes for å forhindre tap av aktsomhet, tap av situasjonsbevissthet, forringelse av ferdigheter, samarbeid og kommunikasjon. Det vil være spesielt viktig at automatisering gir beslutningshjelp relatert til å håndtere konflikter i lufta og avstand mellom fly, spesielt i forstadiet av landing.

I forlengelsen av dette trekker vi med oss prinsipper presentert i Don Normans *How Might Humans Interact With Robots?* (2004). Norman advarer mot å automatisere oppgaver i en slik grad at menneskelige operatører blir redusert til å monitorere systemer, og utelukkende ta kontroll dersom

noe går galt. Videre påpeker Norman viktigheten av fullstendig kommunikasjon mellom datasystemer og mennesker, for på den måten å unngå at operatører må bruke unødvendig lang tid på å sette seg inn i situasjonen, dersom systemet skulle bryte sammen.

4.1.2 Intervju med domeneekspert

Studiets første datainnsamling var et intervju med en domeneekspert, som er prosjektkoordinator for NextGenDST, og vår veileder. Hun har i tillegg gjennomført et kurs for flygeledere via EUROCONTROL. Opptak fra intervjuet ble transkribert, og videre gjennomførte vi en a-priori koding ut i fra noen overordnede kategorier. Disse var basert på etablerte ideer vi identifiserte fra domenespesifikk litteratur, i tillegg til litteratur om konseptualisering av bruk. Eksempler på dette er oppgaver, kontekst, verktøy, ansvarsfordeling og sikkerhet. I tillegg innførte vi kategorien *fremtidig kontekst*, på bakgrunn av undersøkelsens formål. For å motvirke skjevhet, som er iboende i kvalitativ analyse, kodet alle i gruppa de samme dataene (Lazar et al., 2010, s. 296). Her så vi etter mening og sammenhenger mellom forekomster i de ulike kategoriene, slik som kontekst og sikkerhet. Deretter så vi dette i sammenheng med innsikt fra litteraturstudien.

4.1.3 Fremtidens flygeledelse

Innsikt fra intervjuet ga oss et overordnet blikk over nåværende arbeidsprosesser for flygeledere. Eksempelvis hvordan arbeidsprosessene understøttes av samspillet mellom pilot, flygeledere på ulike nivåer, fysiske verktøy og systemet. Angående fremtidig utvikling innenfor domenet var det viktigste vi tok med oss fra intervjuet mulighetene for dynamiske flysoner, flygeledelse på anbud, sentralisering av flygeledelse, samt automatisering. Vi antok derfor at endrede relasjoner og ansvarsfordeling mellom mennesker, og mellom menneske og maskin, er et sentralt tema i diskursen om fremtidens flygeledelse.

Denne antagelsen stemmer overens med funn fra litteraturstudien. Et gjentagende argument i diskursen var viktigheten av fullstendig kommunikasjon mellom menneske og maskin. Her ble det presisert at mennesket alltid skal ha kontroll over kritiske avgjørelser og maskinen kun foreslå foretrukne alternativer. Slik unngår man situasjoner der mennesker er blitt mindre oppmerksomme samt forringelse av operatørens kunnskap. Ved innføring av nye systemer i tidskritiske domener trekkes spesielt tillit og aksept fra interessenter frem, både i intervjuet og litteraturen, som helt sentralt for at dette skal være gjennomførbart.

Wickens et al. (1998) argumenterer for at en stor grad av automatisering vil føre til en mer komplisert og tett flytrafikk, som vil gå utover operatørens situasjonsbevissthet.

Dette kan ses i lys av innsikt vi fikk fra intervjuet om hvordan kontekstuelle faktorer legger til rette for operatørens situasjonsbevissthet, som er avgjørende for gjennomføring av operatørens arbeidsoppgaver.

Vi antar at et fremtidsscenario som ikke tar hensyn til implikasjoner beskrevet ovenfor vil medføre en høyere risiko ved håndtering av krisesituasjoner enn i dagens situasjon. Det ble derfor viktig å ta med denne innsikten i videre undersøkelser.

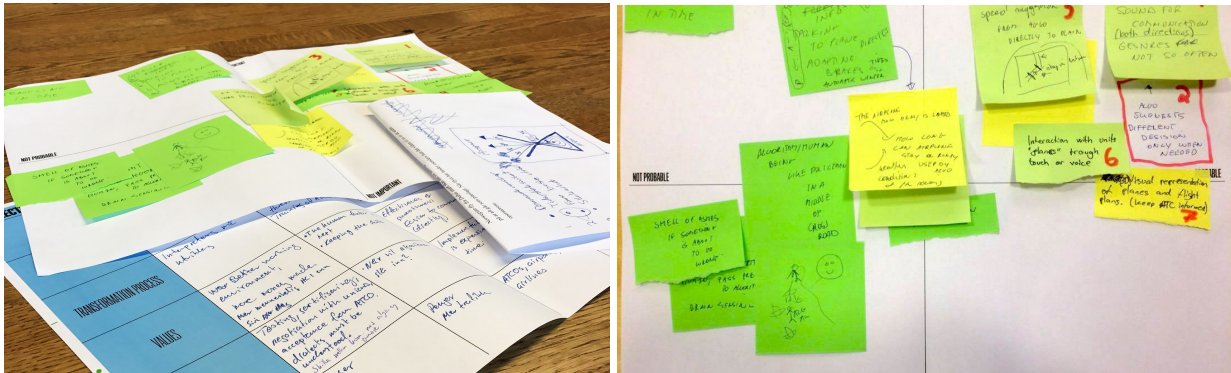
4.2 Utvikling av et fremtidsscenario

Basert på innsikten beskrevet i foregående avsnitt laget vi et scenario som beskrev en fremtidig kontekst for kontroll av flytrafikk, satt til 2025. Intensjonelt valgte vi å utelate en del detaljer rundt interaksjon med algoritmen og konkret teknologi, fordi dette var det vi ønsket videre innsikt i.

4.3 *Future Workshop*

For å få bedre innsikt i hvordan mennesker kan interagere med maskiner i fremtiden, arrangerte vi en *future workshop* i SINTEFs lokaler i Oslo. Deltakerne på workshopen var tre ansatte på SINTEFs prosjekt NextGenDST, en masterstudent og tre av gruppens medlemmer. To av deltakerne fra SINTEF har forsket på DST for flygeledelse over flere år, noe som gjør dem til domeneeksperter. Ved å inkludere en student fikk vi også med et passasjerperspektiv. Den tredje deltakeren fra SINTEF ble inkludert på bakgrunn av hans kunnskap om emergente teknologier, spesielt *virtual reality* og *mixed reality*.

På starten av workshopen presenterte vi scenarioet, i tillegg til en 3D-prototype av en mulig løsning, for å sette i gang fantasien. Vi opplevde at scenarioet fungerte godt til å stimulere til diskusjon, og skjemaene ga deltakerne spesifikke knagger å basere diskusjonen videre på. For gjennomføringen av workshopen delte vi deltakerne i to grupper, med så jevn balanse av deltakernes bakgrunn som mulig. Vi brukte et *futures wheel* til brainstorming og deltakerne brukte sin ekspertise til å kategorisere mulige løsninger og utfordringer i fremtiden basert på hvor viktige og sannsynlige de er. Videre ble disse ideene ført inn i en *futures table*, for å gå i dybden på hver av løsningene, blant annet ved å identifisere aktører, verdier og utfordringer.



Figur 2: Futures table (venstre) og futures wheel (høyre) benyttet i future workshop.

4.3.1 Fremtidige teknologier

Vi benyttet oss av en standardisert prosedyre for koding utviklet av Åbo universitet. Her ble alle begrep etablert i futures table vurdert opp mot hverandre, for å se hvilke som kan eksistere sammen i mulige fremtidsscenarioer (Lautamäkki, 2014). Dette ga oss også forhåndsdefinerte kategorier basert på deres rammeverk for å undersøke dommer i en fremtidig kontekst, og vi så klart hvilke trender som gikk igjen. Talebasert interaksjon ble av mange trukket frem som en god løsning. Det samme ble visuell representasjon – altså at informasjon ikke bare skal være i form av tekst og tall, men at man skal kunne se og interagere med objektene. Under workshopen ble også sammenhenger mellom forekomster under de ulike kategoriene identifisert. Der vi oppdaget mest konflikter i henhold til fremtidige løsninger, var hvordan ansvar skal fordeles mellom menneske og algoritmene. Her ble det diskutert hvilke beslutninger som kan tas av algoritmene uten å holde mennesket informert, og hvorvidt mennesket skal foreslå alternativer til løsninger eller omvendt. I tillegg var verdiene sikkerhet, tillit og effektivitet gjennomgående i alle løsningene som ble diskutert.

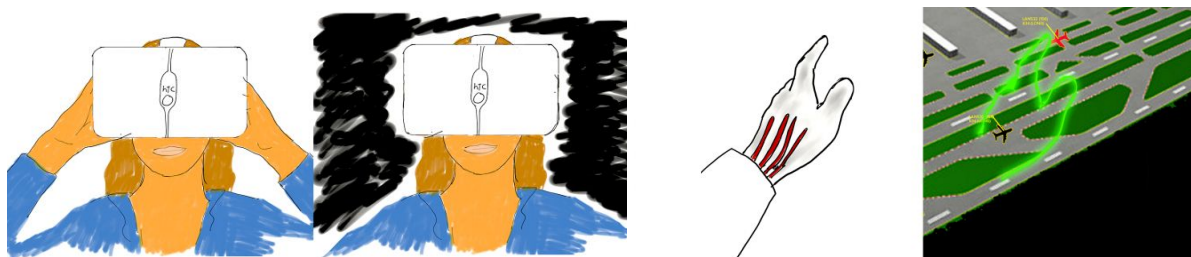
Dataene fra workshopen bekreftet det vi fant i tidligere datainnsamling på flere områder. Spesielt ble diskusjonen om fordeling av ansvar mellom menneske og algoritmen trukket frem begge steder. Under workshopen ble ulike slike fordelinger foreslått som mulig, men det var enighet om at sikkerhet var det viktigste å overholde. Derfor må mennesket fortsatt sitte med ansvar for alle kritiske beslutninger, noe som samsvarte med det vi fant i tidligere datainnsamling. Et annet interessant bidrag til denne diskusjonen fra workshopen, var muligheten til å overstyre systemet.

Innsikt fra litteraturstudiet og workshopen samsvarte også i oppfatningen om at fremtidige systemer må legge til rette for økt situasjonsbevissthet og ikke føre til for stor kognitiv belastning hos

operatøren. Her fant vi i litteraturstudien hvordan for stor grad av automatisering kan føre til tetthet og kompleksitet i flytrafikk, mens i workshopen ble det lagt spesiell vekt på teknologier som talebasert grensesnitt visuelle representasjoner og *tangible* interaksjon.

4.4 Utvikling og evaluering av *storyboards*

Med utgangspunkt i funn fra workshopen videreutviklet vi vårt opprinnelige scenario til to ulike. Dette ga oss mulighet til å diskutere hvordan vi kunne implementere det som ble vektlagt av ekspertene og hvordan dette kunne fungere i en sammenheng. Basert på dette laget vi også to tilhørende *use cases* som tydeligere illustrerte oppgaver, interaksjon og feedback. For å gi en mer visuell representasjon av kontekst og interaksjon laget vi i tillegg *storyboards*. Det som hovedsakelig skiller de to versjonene er bruk av *virtual-* eller *mixed reality* som grensesnitt, grad av automatisering og former for interaksjon. *Storyboards* ble grunnlaget for videre undersøkelser.



Figur 3: Utklipp fra lavoppløselig prototype av interaksjon med automatisert system for flygeledelse.

Vi utførte en formativ evaluering med de lavoppløselige prototypene med en ekspert på emergente teknologier fra SINTEF. Formålet var å ta avgjørelser om hvilke aspekter vi skulle gå videre med fra de to løsningene. Sammen med eksperten gikk vi gjennom *storyboardet* og diskuterte det med fokus på brukbarhet og teknisk gjennomførbarhet i nær fremtid.

4.4.1 Implikasjoner fra evaluering

Ut fra denne evalueringen fant vi at *mixed* eller *augmented reality* er et mer realistisk grensesnitt med tanke på domenet, siden operatøren trolig vil foretrekke å beholde eksisterende verktøy. Oppdragsgiver så imidlertid *virtual reality* som foretrukket grensesnitt i den høyoppløselige prototypen, på bakgrunn av at dette er et mindre utforsket designrom (eng.: *design space*) i vårt domene. *Virtual reality* gir også mulighet til å utforske om man kan plassere flygelederen i et hvilket som helst kontor, uten å behøve andre skjermer eller vinduer å se gjennom. Her ville spesielt sikkerhet være et sentralt tema i videre undersøkelser.

Vurderingen av gjennomførbarhet av foreslåtte løsninger ble gjort ut i fra ekspertens antagelser om hvorvidt dagens teknologiske utvikling vil legge til rette for dette i nærmeste fremtid. Med tanke på den siste tidens utvikling av talegjenkjenning ble det antatt at det vil være godt nok selv til kritiske oppgaver i fremtiden. Evalueringen førte også til at vi gikk bort fra en interaktiv hanske, og heller baserte videre arbeid på teknologi tilsvarende dagens *Leap Motion* (bevegelsesgjenkjenning) for å interagere med grensesnittet.

Eksperten antok at sikkerheten kan øke dersom systemet tar over kommunikasjonen med piloter. Her må man vurdere hvorvidt dette skaper et behagelig miljø for operatøren siden man mister det menneskelige aspektet. Videre ble det foreslått at det også burde settes krav til omgivelsene med hensyn til lys og farger. Dårlige design kan gå utover villigheten til å akseptere systemet av brukerne, også fordi tilliten man har til andre mennesker ofte er sterkere enn den man har til en maskin. Det er derfor viktig at systemet presenterer forslagene på en måte som gjør at mennesket forstår det rasjonelle bak forslagene. Da kan operatøren reagere og gjøre sine eventuelle endringer med forståelse for konsekvensene. At operatøren skal akseptere forslag fra en algoritme er også avhengig av en generell tillit til datasystemer. Det er antatt at ikke alle vil ha stor nok tillit i dag, men etter hvert som slike systemer innføres i hverdagslige gjøremål, eksempelvis gjennom Apples *Siri*, kan denne tilliten øke.

4.5 Høyoppløselig prototype

Med utgangspunkt i tilbakemeldinger på våre lavoppløselige *storyboards*, valgte vi å gå videre med bruk av *virtual reality* som grensesnitt. I tillegg la vi vekt på et prinsipp om at algoritmen skal foreslå løsninger som mennesket skal reagere på, for at den menneskelige operatøren fremdeles skal ha oversikt over situasjonen til enhver tid. Videre valgte vi, på bakgrunn av innsikten vi fikk i det innledende ekspertintervjuet, å fremstille prototypen i en fremtidskontekst der flygeledelse i stor grad er sentralisert og basert på dynamiske flysoner. I dette ligger at flygelederens ansvarsområde ikke bestemmes av geografisk tilhørighet, men ut i fra trafikkmengde og behov i den aktuelle arbeidsperioden.

Dagens *virtual reality*-teknologi er et tidkrevende og kunnskapsintensivt verktøy for å utvikle prototyper, og strekker seg langt utover rammene for vårt prosjekt. Bruk av video ble derfor et mer naturlig valg for vår endelige prototype. En annen avgjørende faktor for dette valget var prosjektets formål om å se lengre enn hva dagens teknologi tillater. Slike prototyper kan i tillegg enklere spres

til mange ulike interessenter for en mer inkluderende diskusjon rundt en mulig fremtidig designløsning.



måten å koordinere arbeidet på vil også virke godt med en av måtene det utforskes å styre flytrafikk på i dag – at én operatør styrer et fly på hele reisen, i stedet for dagens system, der man opererer med en Air Traffic Controller for hver av sonene et fly ferdes i. Samtidig er dette en koordinering av ansvar og interaksjon som ble beskrevet som fremtidsrettet, og tilpasset en fremtid der man potensielt vil kontrollere førerløse fly. Med utviklingen av og interessen som er for førerløse kjøretøy i dag er dette noe som bør prioriteres ved utvikling av et nytt system for flygeledelse.

Det å bruke bevegelsesgjenkjenning, som *Leap Motion*, er en effektiv og enkel måte å interagere med noe man ser på i *virtual reality*, men bør vurderes opp mot HMS og faren for at operatører blir for slitne eller får belastningsskader. Samtidig er det alltid en mulighet for at talegjenkjenning kan misforstå en kommando, uansett hvor bra tolkeren er. Det må derfor være sikkerhetsmekanismer som hindrer avgjørelser systemet vurderer som for risikabelt. Dette kan igjen hindre operatøren å ta avgjørelser som av erfaring kan regnes som trygg.

Ekspertene som evaluerte videoen mente at muligheten til å velge parametere som skal prioriteres, og fremstillingen av konsekvenser dette medfører, var bra, men at det bør diskuteres om disse prioriteringene skal defineres av ledelsen på hver flyplass. Av innspill kom det også at man bør vurdere å prioritere fly etter konsekvenser av forsinkelse på hvert av dem, og å vise dette i grensesnittet til operatøren. Det er eksempelvis langt større – både praktiske og økonomiske – konsekvenser for passasjerer og flyselskaper om et fly med passasjerer som skal mellomlande før et videre fly til et annet kontinent blir forsinket, enn om et lite fly med få passasjerer på en lokalrute i Norge blir forsinket.

Ekspertene fra SINTEF var usikre på om man vil, eller burde, bruke flightstrips i fremtiden. Metoden er gammel, og baserer seg på at informasjon om innkommende og utgående fly struktureres på papirlapper i rekkefølge. Det kom frem i evalueringen at det ikke nødvendigvis er best å digitalisere flightstrips i 3D-grensesnittet, men at man forsøker å utvikle et nytt system for å strukturere informasjonen. Det ble også uttrykt ønske om mulighet for å simulere valg man tar, og å få visualisert dem. For eksempel ved å prøve ut en omprioritering av fly, og 3D-simulere hvordan trafikk på flyplassen vil se ut da. Da vil flygeledere få et større oversiktsbilde av hvordan trafikk på rullebanene går parallelt, og det blir lettere å få med seg situasjoner hvor det er travelt og potensielle risikoer. Om man i fremtiden har en flygeleder som styrer et fly helt fra avgang fra gate og til

landing på destinasjonen vil det være lettere å gå bort fra flightstrips. Da kan man heller vise et bilde av flyene i deres faktiske posisjon gjennom hele ruten.

5.2 Spørreundersøkelse

En omfattende grad av automatisering av flygeledelse endrer ikke bare arbeidsplassen til flygeledere, men påvirker også passasjerene. Vi gjennomførte derfor en undersøkelse med vanlige passasjerer, for å få deres meninger om et scenario hvor algoritmer i stor grad tar over jobben fra mennesker, og hvor dårlige avgjørelser i verste utfall kan få fatale konsekvenser. Dette fordi vi, som nevnt, anser aksept fra flere interessenter som en viktig forutsetning for gode løsninger. På spørreundersøkelsen hadde vi 65 respondenter, og alle spørsmål var obligatorisk å svare på. Respondentene svarte på en likertskala fra 1 til 7, hvor 1 er helt uenig og 7 er helt enig.

	s	mo	me	v	o	st
Dette datasystemet fremstår som trygt	4,4	5	5	2,3	6	1,5
Jeg har tillit til at mennesker kan utvikle kritiske datasystemer uten feil	4,2	6	5	3,6	6	1,9
Jeg har tillit til at datamaskiner kontrollerer flytrafikk på en trygg og forsvarlig måte	5,1	5	5	2,1	5	1,5
Jeg tror datamaskinen vil gi flygelederen upartisk informasjon	5	6	5	3,1	6	1,8
Jeg tror datamaskinen vil føre til mindre forsinkelser	5,2	6	6	2,9	6	1,7
Jeg har tillit til at flygeledere vil oppdage eventuelle feil datamaskinen gjør	4,2	4	4	2,5	6	1,6
Jeg tror dette datasystemet vil gjøre flyplasser mer effektive	5,3	6	5	2	6	1,4
Om en datamaskin kontrollerer flytrafikken vil det være mindre kø på flyplassene	4,6	7	5	3,4	6	1,8
Jeg synes det er greit at datamaskiner tar over arbeidsoppgaver fra flygeledere	4,7	7	5	3,8	6	2
Flytrafikk bør styres fra ett felles kontrolltårn	4	4	4	2,6	6	1,6

Figur 5: Resultater fra spørreundersøkelse rundet av til nærmeste tittel.

s = snitt, mo = modus, me = median, v = varians, o = område, st = standardavvik

Svarene på spørsmålene spriker over hele skalaen, men spørsmålene som omhandler tekniske aspekter har en overvekt av svar på den positive siden. Aspektet ved fiksjonen som omhandler å samle kontroll av flytrafikk på et felles sted, stiller passasjerene seg i snitt nøytrale til. En mulig kilde til dette kan være at domenet flygeledelse er komplekst, og at de færreste har satt seg inn i denne problematikken. Utover dette har vi ikke grunnlag til å si hvorfor passasjerene stiller seg nøytrale til å sentralisere flytårn, men ser vi til ekspertevalueringen ble det påpekt at det medfører høy risiko å samle kritiske systemer på ett sted.

Vi ser også at passasjerene jevnt over har lavest tillit til menneskene i og bak systemet, selv om svarene i snitt er på positiv side. Om vi ser på sammenhenger mellom respondenters besvarelser, ser vi at de med svar i området 1-3 på spørsmålet om de har tillit til at mennesker kan utvikle sikre systemer (altså liten tillit), også stiller seg tvilende til at flygeledere vil oppdage kritiske feil (snitt: 3,5). Denne gruppen er også kritiske til hvorvidt datasystemet i fiksjonen fremstår som trygt (snitt: 3,5). Ser vi til passasjerene med svar i området 5-7 på spørsmålet om de har tillit til at mennesker kan utvikle sikre systemer, er de på sin side langt mer positive til at datasystemet i fiksjonen fremstår som trygt (snitt: 4,9), samt til at flygeledere vil oppdage kritiske feil (snitt: 4,8).

6 Konkluderende diskusjon

I diskusjonen om fremtidens flygeledelse, og designimplikasjoner for fremtidige løsninger, vil vi fokusere på tre sentrale aspekter: Interaksjonstyper, visualisering og ansvarsfordeling.

Basert på resultatene fra evalueringen vil vi argumentere for at kombinasjonen av tale og bevegelse er en god måte å interagere på. All interaksjon bør ikke foregå med bevegelse fordi flygelederne blir for sliten, og ikke all interaksjon er like lett å lage korte og effektive talekommandoer på. Brukt i kombinasjon ser vi at det gjør jobben til flygeledere mer effektiv. Vi mener også at interaksjonen mellom flygeleder og pilot bør bruke datasystemet som kommunikasjonsledd. Da vil verken flygeleder eller pilot trenge å rette all oppmerksomhet mot hverandre for en samtale, og flygelederen kan gjøre flere oppgaver parallelt.

Med de grafiske mulighetene bruk av *virtual reality* gir burde et system for flygeledelse også kunne kjøre simuleringer flygelederen ber om. Å få visualisert konsekvenser av avgjørelser er en lettere måte å vurdere på enn en liste av tall og tekst, som mennesker må legge mer ressurser i å tolke. For å gjøre simuleringer til et best mulig verktøy bør flygelederen kunne interagere med simuleringen og gjøre endringer i den, etter hvert som man ser potensielle problemer eller farer dukke opp.

Et nytt system bør også være rustet for at flygeledere kan kontrollere både dynamiske flysoner, eller ett enkelt fly gjennom hele reisen. Hvordan flygeledelse vil foregå i fremtiden er uvisst, men dette synes å være de to store trendene som forskes på. Uavhengig av hva man velger av dynamiske flysoner eller å styre flyet gjennom hele reisen bør man gå vekk fra flightstrips. Vi mener

teknologien som er tilgjengelig i dag muliggjør utvikling av helt nye måter å strukturere informasjonen på, som er bedre tilpasset fremtidens arbeidsmåter. På et sanntidskart kan man for eksempel tegne inn ruter og hvor flyet burde være på ethvert tidspunkt for å nå planlagt ankomsttid. Dette gir sanntidsinformasjon på en langt bedre måte enn det dagens teknologi støtter.

Spesielt viser evaluering med eksperter at våre funn rundt automatisering av flygeledelse er konsistente med tendenser som går igjen i denne rapportens relevante litteratur på interaksjon med roboter, hvor særlig fullstendig kommunikasjon og oppmerksomhet vektlegges. Slik tidligere publikasjoner på interaksjon med roboter har foreslått, må alle kritiske avgjørelser også innen flygeledelse tas av mennesker. Etter dette bør systemer for flygeledelse være verktøy som understøtter flygelederen til å ta de beste avgjørelsene, fremfor å ta over oppgaver fullstendig. På denne måten unngår man også å forringe kompetansen til flygelederen, samtidig som man hjelper dem å jobbe effektivt.

Skulle vi endret noe ser vi at det kunne vært en fordel å eksponere respondentene på spørreundersøkelsen for dagens arbeidsverktøy og -metoder innen flygeledelse. Dette kunne gitt respondentene bedre grunnlag til eksempelvis å svare på spørsmål om hvorvidt vårt fremtidssystem vil være mer effektivt enn dagens løsninger. Videre ville det vært interessant å fremstille systemet i vår fiksjon som en testbar, fungerende prototype. Om vi på den måten kunne gjennomført eksperimenter med aktive flygeledere må vi anta at vi i større grad kunne utarbeidet konkrete designimplikasjoner for fremtidige løsninger.

Sett i lys av rammeverket vi har fulgt for design fiction vil vi argumentere for at dette som metodologi med hell kan benyttes for å fremprovosere innsikt i et spesifikt domene. En annen effekt, ikke diskutert i detalj i denne oppgaven, er hvordan vår fremtidsfiksjon ble ansett av SINTEF som et hjelpemiddel i deres videre arbeid med NextGen DST, i den forstand at den konkretiserer og visualiserer ideer det kan være krevende å diskutere rent konseptuelt.

Vi vil påpeke at undersøkelser i denne rapporten må sees som eksplorative. Både det faktum at bruk av *virtual reality* innen flygeledelse er et uutforsket område, og at dette prosjektarbeidet ikke har involvert aktive flygeledere, taler for at videre forskning må til før løsninger som tar i bruk denne teknologien kan realiseres.

Litteratur

Allen, Elaine og Christopher Seaman (2007): Likert Scales and Data Analyses. Hentet fra:

<http://asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html>

Lazar, J. et al. (2010): *Research Methods in Human-Computer Interaction*. Glasgow: Wiley Publications.

Lautamäkki, Villa (2014): *Practical Guide for Facilitating a Futures Workshop*. Finland: Åbo Universitet

Norman, Don (2004): *How Might Humans Interact With Robots?*. Hentet fra:

http://www.jnd.org/dn.mss/how_might_humans_int.html

NRK (2013, 7. oktober): Valgresultat 2013. Hentet fra:

<http://www.nrk.no/valg2013/valgresultat/>

Preece, Jenny, Yvonne Rogers og Helen Sharp (2015): *Interaction design. Beyond human-computer interaction* (4. utgave). Glasgow: Wiley Publications.

Samfunnsveven (2013, 3. september): Resultatrapport 2013. Hentet fra:

<http://samfunnsveven.no/skolevalg/pdf/skolevalgresultater2013.pdf>

Thaagard, Tove (2009): *Systematikk og innlevelse*. Bergen: Fagbokforlaget

Wickens, C., Mavor, A., Parasuraman, R. og McGee, J. (1998): *The future of air traffic control: Human operators and automation*. Washington, DC: National Academies Press.

Ødegård, S. S. (2013): *Exploring Visualization - Solutions for Air Traffic Control Workflow Productivity Improvement* (Masteroppgave). Universitetet i Oslo.

Lenke til endelig prototype: <https://vimeo.com/191373067>