

UiO : **Institutt for informatikk**

Det matematisk-naturvitenskapelige fakultet

# INF2260

Prosjektrapport høst 2017

Kristian Dahl - [kd@student.matnat.no](mailto:kd@student.matnat.no)

Ingvild Alette Jensen - [ingviaj@student.matnat.no](mailto:ingviaj@student.matnat.no)

Hans Jonathan Esperum Esbjug - [hansjes@student.matnat.no](mailto:hansjes@student.matnat.no)

Pio Nizar Finn Rasch-Halvorsen - [pnraschh@student.matnat.no](mailto:pnraschh@student.matnat.no)



<b>1</b>	<b>INNLEDNING</b> .....	<b>3</b>
1.1	DOMENE.....	3
1.2	UTGANGSPUNKT FOR PROSJEKTET .....	3
1.3	RAPPORTENS STRUKTUR .....	4
<b>2</b>	<b>METODOLOGI</b> .....	<b>4</b>
<b>3</b>	<b>METODE</b> .....	<b>7</b>
3.1	METODER FOR INNSIKT .....	7
3.2	METODER FOR UTVIKLING.....	8
3.3	METODER FOR EVALUERING .....	9
3.4	UTFORDRINGER .....	9
<b>4</b>	<b>DESIGNPROSESSEN</b> .....	<b>10</b>
4.1	INNSIKT OG INSPIRASJON.....	10
4.2	UTVIKLING AV LAVOPPLØSELIG PROTOTYPE .....	12
4.3	UTVIKLING AV HØYOPPLØSELIG PROTOTYPE.....	14
<b>5</b>	<b>SUMMATIV EVALUERING</b> .....	<b>15</b>
<b>6</b>	<b>KONKLUDERENDE DISKUSJON</b> .....	<b>17</b>
<b>7</b>	<b>SITERTE VERK</b> .....	<b>19</b>

# 1 Innledning

Dette prosjektet tar utgangspunkt i det pågående forskningsprosjektet Next Generation Decision Support Tools (NextGenDST) som gjennomføres i regi av SINTEF. Prosjektet tar for seg hvordan en avansert algoritme, ofte referert til som en kunstig intelligens, kan anvendes til å understøtte beslutningstaking i tidskritiske situasjoner. Hensikten er å gi mennesker et støtteverktøy som kan bidra til økt presisjon og sikkerhet innen komplekse domener som krisehåndtering og flygeledelse. Prosjektet ligger under SINTEF sin avdeling for Networked Systems and Services, og er et samarbeid mellom forskningsgruppene som jobber med automatisering og Human-Computer Interaction (HCI).

Klienten for prosjektet er SINTEF, og veileder for oppgaven er Amela Karahasanović.

## 1.1 Domene

Et av kravene til valg av domene var at det omfattet en tidskritisk og kompleks arbeidssituasjon hvor gode og hurtige avgjørelser er viktige. I denne sammenheng har vi valgt flygeledelse som domene for prosjektet.

For det første valgte vi flygeledelse fordi domenet passer godt til kravene beskrevet over. Flygeledere har ansvar for å dirigere flytrafikk; både på flyplasser og i luftrommet rundt flyplassene. Arbeidsoppgaver kan for eksempel være å dirigere fly inn mot rullebanene, og å sørge for at flyene kommer seg fra gate til rullebane i riktig rekkefølge. Det er en ansvarsfull jobb som krever presisjon, fokus og hurtige beslutninger. Flygeleiderne må være i stand til å ta inn store mengder med informasjon, og kunne fokusere på viktige detaljer samtidig som de beholder overblikk over situasjonen (Kleven, 2016, s. 20).

For det andre har SINTEF sitt prosjekt også hatt sterkt fokus på flygeledelse. Dette har medført at SINTEF innehar mye informasjon og kunnskap i tilknytning til domenet, noe vi ønsket å dra nytte av i prosjektarbeidet. Utover dette mener vi det hensiktsmessig at vårt prosjekt er nærliggende det SINTEF har gjort tidligere, fordi vi i større grad kan bidra til det overordnede forskningsprosjektet.

## 1.2 Utgangspunkt for prosjektet

Å utforske fremtidens beslutningsverktøy er en stor og kompleks oppgave. Derfor har vi valgt å snevre inn prosjektomfanget ved å fokusere på interaksjonen mellom flygeleder og beslutningsverktøyet.

Vi ønsker å oppfordre til diskusjon rundt hvordan man kan formidle intensjon, prioritering og rasjonale bak avgjørelser, både fra menneske til algoritme, og omvendt. Ut fra dette har vi formulert følgende diskusjonsspørsmål:

- a) Hvordan kan algoritmen formidle årsaken bak sine forslag til en flygeleder?
- b) Hvordan kan flygeleder formidle årsak for endringer som motstrider algoritmens forslag?

### 1.3 Rapportens Struktur

I denne rapporten vil vi beskrive hvordan vi har undersøkt disse diskusjonsspørsmålene. Dette vil vi gjøre ved å først forklare og begrunne valg av metodologi, for så å beskrive metodene vi har anvendt for datainnsamling, analyse, utvikling og evaluering. Videre vil vi ta for oss design- og prototypingsprosessen, samt evalueringene vi har gjort av prototypene. Avslutningsvis vil vi presentere og diskutere resultatene fra evalueringene, for så å avslutte med en konklusjon.

## 2 Metodologi

Som rammeverk for prosjektet har vi valgt metodologien **design fiction**. Sterling (2012) definerer design fiction som "The deliberate use of diegetic prototypes to suspend disbelief about change" (s. 1). En annen definisjon fra Lindley og Coulton (2015) beskriver design fiction som "(1) something that creates a story world, (2) has something being prototyped within that story world, (3) does so in order to create a discursive space." (s. 1).

Et fellestrekk i begge definisjonene er at prototypene ikke nødvendigvis er begripelige funksjonelle produkter, grensesnitt eller artefakter. I stedet er prototypene «diegetic»; de er narrative, de forteller en historie for å generere innsikt i en fiktiv situasjon, eller en medfølgende kultur (Bleecker, 2009, s. 76). I praksis innebærer design fiction å designe et framtidsscenario der en prototype eksisterer innen en gitt kontekst, for så å diskutere og undersøke hvilke muligheter som finnes her.

Valget av design fiction har vi gjort basert på begrensninger ved prosjektets bredde og omfang, samt ønsker fra klienten.

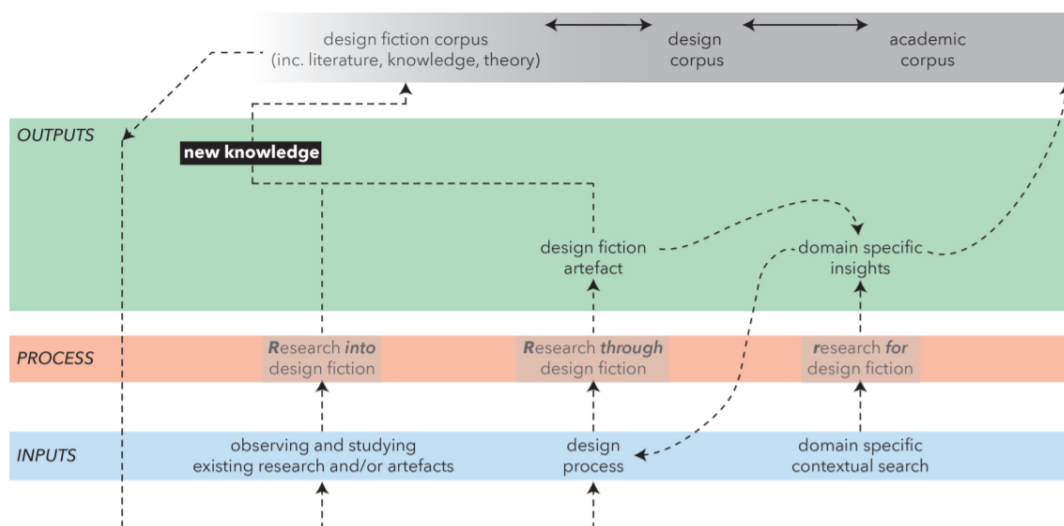
For det første er flygeledelse et komplekst domene med strenge sikkerhetskrav til implementasjon av ny teknologi (Ødegård, 2013, ss. 1-2). Dette kommer av at potensielle ulykker kan være livstruende for svært mange mennesker, noe som medfører at integrasjon av

ny teknologi i flygeledernes omgivelser skjer inkrementelt for å unngå feil (Hanseth & Lyytinen, 2010) (Ødegård, 2013). Siden målet med denne oppgaven har vært å utforske en radikal endring i flygeledernes arbeidskontekst, har vi sett det nødvendig å velge en metodologi som muliggjør en slik utforskning.

For det andre har flygelederne svært strenge arbeidsrelaterte krav til sikkerhet og rutiner, noe som har medført at vi ikke har hatt direkte tilgang til målgruppen. En forutsetning for prosjektet ble dermed at vi valgte en metodologi med metoder som muliggjør innsikt i problemområde og arbeidskontekst uten direkte kontakt med flygeledere.

For det tredje hadde SINTEF et sterkt ønske om at vi skulle utforske problemstillingen på en nytenkende og eksplorativ måte. De ønsket at vi skulle undersøke hvordan fremtidig teknologi potensielt kan anvendes i tidskritiske situasjoner. Vår rolle skulle være å bidra med konsepter og idéer i form av prototyper for å muliggjøre diskurs og generere innsikt som SINTEF kan anvende videre i sitt forskningsprosjekt.

Valget av metodologien har lagt føringer på hvilke metoder vi har brukt og hva slags data vi har samlet. Ifølge Lindley (2015) er design fiction i en kritisk utviklingsperiode, og det er uenighet knyttet til hva som er ideelle fremgangsmåter, metodevalg og prototypingsformer innen metodologien (ss. 2-3). På grunn av dette har vi valgt å følge Lindley sitt pragmatiske rammeverk for design fiction for å gi kontekst til designprosessen.



Figur 1: Forholdet mellom kategoriene i "A Pragmatics Framework for Design Fiction" (Lindley, 2015, s. 11)

Den første delen av rammeverket beskriver tre ulike prosesser:

Den første av disse er *research into design fiction*. Denne kategorien beskriver hvordan man får økt innsikt i design fiction som metodologi gjennom å undersøke teori om metodologien, eller eksempler på studier der design fiction er brukt. I dette prosjektet valgte vi å først undersøke design fiction som metodologi. Vi gjorde dette for å få innsikt i metodologien og hvordan vi skulle anvende den.

Den andre kategorien er *research for design fiction*. Denne kategorien kan beskrives som undersøkelser av kontekster som skal inngå i designprosessen ved å gi innsikt i et spesifikt domene. I denne sammenheng gjennomførte vi undersøkelser av konteksten rundt flygelederne i dag for å gi oss et forståelsesgrunnlag for designprosessen.

Den tredje kategorien er *research through design fiction*. Denne beskriver hvordan man produserer kunnskap ved å gjennomføre en designprosess. Lindley (2015, s. 4) argumenterer for at all design fiction er en form for forskning, i den forstand hensikten med designet er å produsere innsikt. Med prosjektarbeidet har vi forsøkt å gjennomføre en slik designprosess for å generere ny kunnskap om fremtiden til flygeledere og deres interaksjon med et beslutningsverktøy.

Den andre delen av rammeverket beskriver rammer for designprosessen. Lindley (2015) bemerker at det ikke er spesielle metoder som er foretrukne over andre, og at metodevalget avhenger av prosjektets kontekst og omfang (ss. 5-7). Når det gjelder hvordan man burde prototype innen design fiction bemerker Lindley (ibid.) at prototyper i form av filmer er passende verktøy for formidling av design fictions (s. 6). Alternativt foreslår Bleecker (2010) at man kan anvende fysiske props for å generere tanker og vekke diskusjon (s. 6). Uavhengig av hvilken form prototypen har, er det viktigst at den muliggjør evaluering og diskusjon (Lindley, 2015).

Det er viktig å påpeke hvordan vi forstår begrepet *bruker* innen design fiction. Siden hensikten med metodologien er å produsere innsikt og diskurs, er det nødvendig å gjøre et skille mellom brukerne som eksistere i *den narrative prototypen*, og de som skal diskutere fiksjonen og *dra nytte av designet*.

Brukerne i *prototypen* vil i dette tilfelle være fremtidens flygeledere. Dette er en brukergruppe som ikke eksisterer i dag, men som gir kontekst til prototypen ved å være jordet (grounded) i nåtidens flygeledelse. Brukerne som vi designer den *narrative prototypen* for er i dette prosjektets tilfelle forskerne og utviklerne som skal utvikle fremtidens decision support tools.

På grunn av den mulige forvirringen mellom disse to gruppene kalles sistnevnte ofte for publikum, eller tilskuere, innen design fiction (Kirby, 2010).

## **3 Metode**

Siden metodologien ikke legger klare føringer på metodevalgene har vi vurdert hvilke metoder som er hensiktsmessige å anvende ved de ulike fasene i designprosessen.

### **3.1 Metoder for innsikt**

#### **3.1.1 Uformelle intervju med domeneekspert**

Vi har valgt å gjennomføre ukentlige, uformelle intervju med vår kontaktperson og domeneekspert hos SINTEF. Vi valgte dette for å opprettholde et nært samarbeid med klienten, og fordi vi ønsket å ha en direkte kanal for kontinuerlig tilbakemelding i tilknytning til prosjektet, noe som er svært viktig innenfor HCI (Lazar, Feng, & Hochheiser, 2010, s. 187). Gjennom møtene fikk vi mulighet til å stille spørsmål tilknyttet domenet og beslutningsverktøyet. I tillegg fikk vi fortløpende tilbakemeldinger på idéer, konsepter og prototyper.

#### **3.1.2 Innholdsanalyse**

For å få innsikt i metodologi, domenet, bruk og brukerkontekst har vi gjennomført innholdsanalyse av relevant faglitteratur relatert til flygeledelse, design fiction, beslutningstaking og kunstig intelligens. I tillegg har vi gjennomført to multimedie-analyser av filmopptak av intervju med flygeledere. I denne sammenheng forstår vi innholdsanalyse som en systematisk og repliserbar teknikk for kategorisering av store innholdsmengder basert på koderegler (Sterling, 2001 i (Lazar, Feng, & Hochheiser, 2010, s. 285).

Måten vi har gjennomført analysene på har variert avhengig av innholdet og mediet vi har undersøkt. Ved undersøkelser av faglitteratur anvendte vi åpen koding for å danne oss kategorier for å skape en felles forståelse av sentrale teorier og begreper. Vi leste gjennom litteraturen hver for oss og noterte nøkkelfaktorer som vi videre sammenliknet, diskuterte og brukte til kategorisering. Dette gjorde vi for å sikre inter-coder reliabilitet (Lazar, Feng, & Hochheiser, 2010, ss. 297-299).

Ved undersøkelser av multimedier anvendte vi a priori koding basert på resultatene fra analysen av faglitteraturen. I denne sammenheng valgte vi ut kategorier vi vurderte som særlig relevante, som vi brukte til å forstå og organisere utsagn og bilder fra filmopptakene

ved forskjellige tidspunkt (Lazar, Feng, & Hochheiser, 2010, ss. 300-301). Her gjennomførte vi også separate analyser og tolkninger for å sikre inter-coder reliabilitet.

### 3.1.3 Brainwriting

For å generere idéer har vi anvendt metoden *brainwriting*. Brainwriting er en prosedyre for grupper der deltakerne først skriver flere idéer på egne ark, for så å diskutere de forskjellige idéene sammen (Paulus & Yang, 2000, s. 78). Metoden baserer seg på brainstorming (Dugosh, Paulus, Roland, & Yang, 2000, s. 722), men unngår fenomener som *sosial loffing* (social loafing), *evaluation apprehension* og illusjoner av produktivitet (Heslin, 2009, s. 131), noe vi vurderte som veldig fordelaktig ved arbeidet med ideasjon.

### 3.1.4 Empatisering

Fordi vi ikke har hatt muligheten til direkte kontakt med flygeledere, har vi valgt å anvende metoden *empatisering* (empathy) for å få økt forståelse av konteksten rundt brukerne og innsikt i brukernes hverdag. Empatisering er en metode som inngår i Design Thinking, der designeren forsøker å se en verden gjennom brukerens øyne, for å få innsikt i deres opplevelser, perspektiv og tanker. Som metode handler empati å oppnå en dypere forståelse av problemene og realiteten til de menneskene vi designer for (Dam & Siang, 2017).

## 3.2 Metoder for utvikling

### 3.2.1 Personas og AI Personas

For å bedre forstå og kommunisere trekk ved brukerne, benyttet vi metoden personas (Cooper, 2004, s. 142). Dette gjorde vi ved å lage relativt representative arketyper for flygeledere basert på innsamlet kunnskap om brukergruppen.

For å understøtte hvordan algoritmen inngår i samarbeid med mennesker, valgte vi å benytte oss av Artificial Intelligence (AI) personas (Karahasanović, Følstad, Schittekat, & Nordlander, 2017). En AI persona kan forstås som en representasjon og personifisering av algoritmens egenskaper. Her valgte vi å lage en persona som representerte algoritmene i støtteverktøyet, som vi ga personlighet, preferanser og behov.

### 3.2.2 Scenario

Fordi vi skulle designe en fremtidsfiksjon var det fordelaktig å anvende scenarier i denne prosessen. I følge Cooper *et al* er et scenario bruken av en spesifikk historie til å konstruere og illustrere designløsninger (Cooper, Reimann, Cronin, & Noessel, 2014, s. 104). I denne



sammenheng valgte vi å lage persona-based scenarios (ibid., s. 105) der våre personas og AI personas var sentrale aktører i scenariene.

### **3.2.3 Storyboards**

For å iscenesette scenariet har vi anvendt storyboards. Storyboards er en metode for å kommunisere et scenario gjennom sekvensielle tegninger av ledd i en interaksjon (Cooper, Reimann, Cronin, & Noessel, 2014, s. 128) Vi valgte dette fordi storyboards gjør det mulig å planlegge og strukturere et narrativ i fellesskap, og fordi de egner seg godt som grunnlag for prototyping via film.

## **3.3 Metoder for evaluering**

### **3.3.1 Workshop**

Vi valgte å gjennomføre en formativ evaluering av lavoppløselige prototyper ved å holde en workshop med 6 eksperter på flygeledelse. Vi valgte å gjøre dette fordi workshops gir mulighet for flere deltakere å delta i aktiviteter og diskusjoner rundt et spesifikt tema.

Formålet med workshopen var å avdekke behov og ulike synspunkter knyttet til prototypene for å videreutvikle dem og generere nye idéer.

### **3.3.2 Fokusgruppe**

For den summative evalueringen av prototypen valgte vi å holde en fokusgruppe med 5 deltakere som er eksperter på flygeledelse og som forsker på og utvikler decision support tools. Innen HCI brukes fokusgrupper for å bygge forståelse av behov, praksis, hensyn, preferanser og holdninger blant menneskene som interagerer med eksisterende eller fremtidige system (Lazar, Feng, & Hochheiser, 2010, ss. 187-89). Siden hensikten med prosjektet var å skape ny innsikt i fremtidens flygeledelse, mente vi at bruk av fokusgruppe egnet seg godt som evalueringsmetode fordi deltakerne kunne diskutere fiksjonen med hverandre og frembringe ulike perspektiver.

## **3.4 utfordringer**

Metodene vi har valgt har hatt medfølgende utfordringer, særlig når det gjelder analyse og tolking av kvalitativ data. I analysene er en sentral risiko at vi kan støte på tvetydig informasjon og komplekse situasjoner som er vanskelige å forstå. Dette kan føre med seg skjevheter ved tolkningen av innholdet (Lazar, Feng, & Hochheiser, 2010, s. 299). På grunn av dette har vi tatt høyde for at resultatene fra analysene er farget av våre egne erfaringer, perspektiver og tanker.

Utover dette må vi også nevne at deltakerne i evalueringene har hatt egeninteresse i prosjektet, noe som kan ha påvirket tilbakemeldingene vi har fått.

## 4 Designprosessen

Denne delen av rapporten handler om hvordan vi har jobbet med datainnsamling, og prosessen med utforming av lav- og høyoppløselige prototyper.

### 4.1 Innsikt og inspirasjon

Vi har valgt et domene som ingen i gruppen har kjennskap til fra før, og som er komplisert og lite transparent for dem som ikke er involvert. Store deler av starten av prosjektet ble derfor brukt på å sette seg inn i domenet. Til dette har vi brukt domeneeksperter fra SINTEF for å få innsikt i algoritmen og domenet vi skal ta utgangspunkt i. I tillegg har vi benyttet oss av innholdsanalyse av domenespesifikk litteratur og film, samt empatisering gjennom simulatorbruk.

Kjernespørsmålene i denne fasen har vært å (1) få kjennskap til flygeledernes arbeidsprosess, og hvordan DST kan inngå som en del av dette, samt å (2) finne mulige innfallsvinkler for å tenke ut hypotetiske løsninger og scenarier.

#### 4.1.1 Innholdsanalyse

Som tekstlig hovedkilde for innsikt i arbeidsoppgavene til flygeledere har vi benyttet oss av to masteroppgaver fra UiO, Ødegård (2013) og Kleven (2016), som begge er utarbeidet i forbindelse med et tidligere ledd av SINTEF sin forskning i dette feltet (SESAR-prosjektet). Disse oppgavene stiller begge med gode generaliserte fremstillinger av arbeidsoppgavene til flygeledere, og siden de begge dreier seg om visualisering og arbeidsflyt fokuserer de spesielt på deler som vi ønsket å utforske videre med prototypen vår.

Ødegaards oppgave (2013) fokuserer på bruk av visualisering som verktøy for egenevaluering, men gir også en indikasjon på at visualiseringsverktøy kan understøtte tolkningen av trafikkmønstre på flyplassen (s. 84). Kleven (2016) tar i sin oppgave for seg opplæring av flygeledere, og finner i sin forskning kognitiv kapasitet som en begrensende faktor i arbeidet som flygeleder. For å motvirke dette foreslås to tilnærminger: (1) visualiseringer fremstilt på en måte som senker kravet til mental prosessering, og (2) automatisere enkelte oppgaver slik at flygelederen får mindre å tenke på (ss. 89-90).

Det andre punktet foreslått av Kleven resonnerer godt med konseptet bak et Decision Support Tool. Selv om det ikke er snakk om fullstendig automatisering av noen aspekter, er hovedtanken å avlaste brukeren mentalt. Siden vår prototype ikke nødvendigvis trengte å være begrenset i sin grad av automatisering valgte vi å utforske dette nærmere med utgangspunkt i et rammeverk utviklet av Sheridan og Verplank for kategorisering av automatiseringsgrader fra fullstendig menneskestyrt til fullstendig automatisert (Sheridan & Verplank, 1978, ss. 168-70).

For å utdype forståelsen vi fikk av den tekstlige innholdsanalysen fant vi relevante videoer om flygeledelse, inkludert intervjuer med flygeledere og opptak av kommunikasjon mellom flygeledere og piloter. Disse kildene utførte vi en enkel, a priori koding av basert på tekstanalysen.

#### **4.1.2 Empatisering**

For å sette oss inn i den kognitive belastningen opplevd av flygeledere forsøkte medlemmene i gruppen seg på ATC-simulatorer på personlig datamaskin, både nettbaserte og mer avanserte varianter. Disse prøvene var mindre organiserte, og ingen form for formell koding ble forsøkt, men eksperimentet ga oss likevel verdifull innsikt. Det var fort merkbart hvor stor mental kapasitet som krevdes for å utføre oppgavene og holde styr på informasjon, og vi opplevde raskt task saturation i forsøkene. Det ble også åpenbart at det var stor forskjell på å styre luftrommet og å styre bakketrafikken—oppgaver som blir utført av forskjellige flygeledere (Kleven, 2016, s. 8).

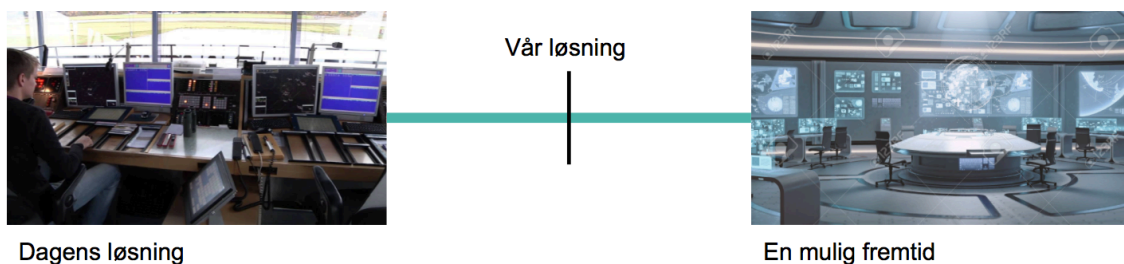
#### **4.1.3 Inspirasjon**

Som et ledd av den innledende utforskingen undersøkte vi også fremtidsrettede og fiktive måter for å representere informasjon. Vi studerte for eksempel det fiktive grensesnittet fra *Minority Report* (Spielberg, 2002), som også blir brukt som eksempel av Bleecker i hans presentasjon av Design Fiction (Bleecker, 2009, ss. 35-45). Ettersom både Kleven og Ødegård bruker elementer av gamification i sine prototyper, og vi hadde benyttet oss av spill i empatiseringen, hentet vi også inspirasjon fra spillgrensesnitt. Blant annet så vi på hvordan informasjon ble fremstilt i strategispill som *Sid Meier's Civilization*, der spilleren må holde oversikt over store mengder informasjon for å vinne spillet.

## 4.2 Utvikling av lavoppløselig prototype

Basert på innsikten vi hadde anskaffet oss gikk vi videre med utviklingen av en lavoppløselig prototype. Vi begynte med å diskutere graden av automatisering vi ville ha i scenariet vårt, da dette ville legge føringer for hvor futuristisk scenariet vårt måtte fremstilles for å være overbevisende.

Dette valgte vi å kontekstualisere ved å bruke teknikken jumping to the end (Jones, 2015). Vi startet med å se for oss en fremtid hvor alt er fullstendig automatisert, og menneskelige aktører ikke er nødvendig, for så å jobbe oss bakover mot en fremtid som er litt nærmere vår tid. Her kunne vi rotfeste (ground) prototypen for å gjøre den evaluerbar. Vi valgte derfor å ta utgangspunkt i nivå fem på skalaen til Sheridan og Verplank, der datamaskinen foreslår handlinger og utfører dem dersom brukeren godtar forslaget (Sheridan & Verplank, 1978, s. 169).



Figur 2: Vår løsning sett i forhold til dagens manuelle løsning og en fremtidig, helautomatisert løsning

### 4.2.1 Ideasjon

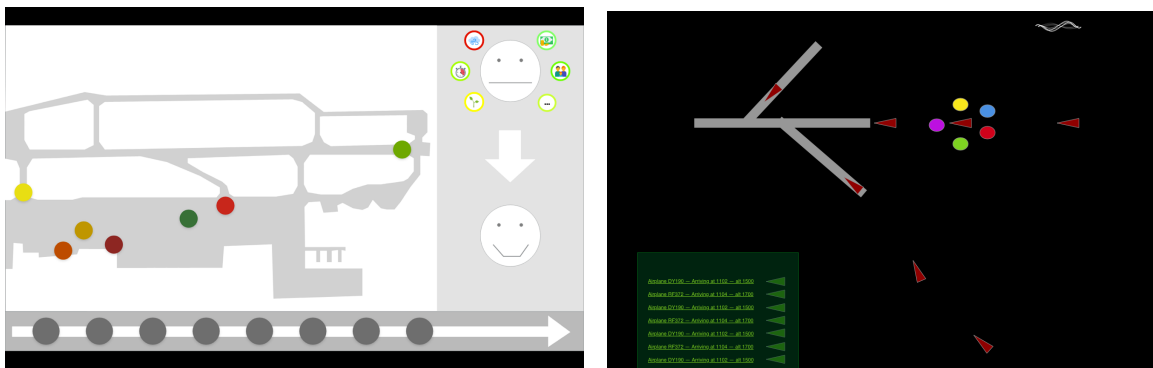
Ut fra dette utgangspunktet ønsket vi å komme frem til noen scenarier vi kunne jobbe videre med i en workshop med domeneekspertene hos SINTEF. I dette arbeidet hentet vi inspirasjon fra design thinking, der man har en divergent ideasjonsprosess som deretter snevres inn ved gjennomgang av idéene fra forrige steg (SITAT). Den divergente prosessen valgte vi å gjennomføre ved hjelp av brainwriting for å oppnå så mange videreførbare idéer som mulig (Heslin, 2009), og for å unngå å sette grenser på oss selv ut fra tanker om ”er dette gjennomførbart i dag?”

Etter å ha skrevet ned idéer i en gitt tidsperiode ble disse lest opp og diskutert, for deretter å kategoriseres ved hjelp av emergent koding. Gjennom denne diskusjonen kom vi frem til et mindre antall idéer som vi ønsket å jobbe videre med. Dette var både konseptuelle idéer og

mer konkrete forslag til visualiseringer. Av disse valgte vi ut tre konsepter vi ønsket å visualisere som lavoppløste prototyper som kunne presenteres på en workshop hos SINTEF. Disse ble utviklet ved hjelp av protoypingsverktøy som Adobe Experience Design og Sketch.

#### 4.2.2 Formativ evaluering og videreførte konsepter

Hovedkonseptene vi valgte å prototype dreide seg rundt visualisering av algoritmens løsning, og dens oppfattelse av situasjonen på flyplassen, samt tenkte grensesnitt for kommunikasjon av intensjon mellom flygeleder og algoritme.



Figur 3: To av de lavoppløste prototypene

Vi gjennomførte en formativ evaluering i form av en uformell workshop hos SINTEF, der domeneekspertene var med på å diskutere hvorvidt konseptene vi presenterte kunne representere et fremtidsscenario innen domenet. En slik ekspertevaluering håpet vi ville avdekke styrker og svakheter ved konseptene, samt ulike synspunkter mellom forskerne med ulik bakgrunn og kunnskap på feltet. Vi utførte innholdsanalyse av notater fra workshopen for å bestemme hvilke aspekter ved prototypene vi ønsket å jobbe videre med.

Mye av diskusjonen dreide seg om de grafiske elementene i grensesnittet. Det ble påpekt at grensesnittet opplevdes som lavoppløselig og rotete. Ekspertpanelet stilte seg spesielt positive til en tanke om humør som fremstilling av hvor godt situasjonen passer med algoritmens forslag, og fremstillingen av ønsket rekkefølge for avgang som en tidslinje. Det ble også påpekt at man kunne trenge en oversikt over hvordan de ulike flyenes “humør” samspilte, noe vi tok med oss videre. Det ble også ytret et ønske om bruk av Voice Control for interaksjon med algoritmen.

Ut fra dette valgte vi å integrere disse konseptene i én helhetlig prototype (Houde & Hill, 1997) der vi tok hensyn til ønsker om et “generelt humør”, og å fokusere på arbeidsoppgavene knyttet til styring av bakketrafikken på flyplassen. Vi kom frem til at kommunikasjon av

årsaken for foreslåtte endringer - både fra algoritmens og flygelederens side - kunne kommuniseres med kontekstuelle ikoner ved flyene. Vi la også mer vekt på det grafiske utseendet slik at dette ikke skulle være like forstyrrende under neste evaluering.

### 4.3 Utvikling av høyoppløselig prototype

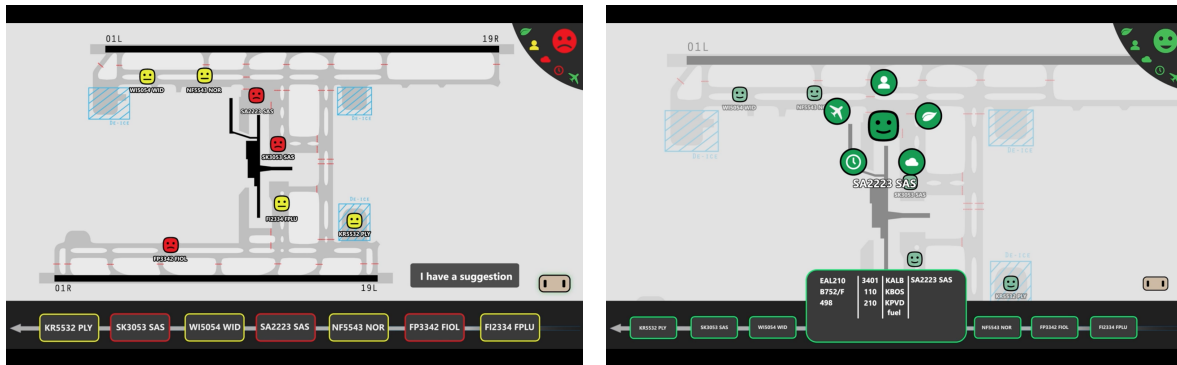
Etter ideasjonsprosessen med divergent utforskning av idérommet, og innsnevring og raffinering gjennom workshop, begynte vi arbeidet med å utforme en spesifikk fiksjon som utgangspunkt for den høyoppløste prototypen. Dette innebar å plassere den fiktive prototypen i tid, og å finne et narrativ som egnet seg til å demonstrere interaksjonene i den fiktive prototypen, det vil si et scenario.

For å lage scenariet valgte vi å undersøke ulike fremtidshypoteser ved å stille *what-if*-spørsmål. Dette beskriver Dunne og Raby (2013) som hensiktsmessig fordi det gir oss muligheten til å distansere oss fra umiddelbare begrensninger av hva som er gjennomførbart, og hvordan tilstander er i dag.

Det valgte scenariet ble en vinterdag i en relativt nær fremtid—år 2028—med dårlig vær og høy trafikk, og dermed mange trafikkavviklingsproblemer; en typisk situasjon der en flygeleder kan ønske å støtte seg på et støtteverktøy. Vi valgte å benytte oss av film som format for den høyoppløselige prototypen, da det er et format som egner seg til å formidle narrativer, og det gir oss gode muligheter til å simulere interaksjon. Vi benyttet oss av storyboards for å konkretisere narrativet i prototypen, slik at vi hadde et utgangspunkt for både selve filmingen, og for å lage det simulerte grensesnittet som benyttes i filmen.

#### 4.3.1 Det fiktive grensesnittet

For å illustrere bruken av den fiktive prototypen valgte vi å animere grensesnittet på en skjerm, slik at vi kunne simulere interaksjonen ved å ta på skjermen synkront med bevegelsene. For å finne frem til logiske animasjoner konstruerte vi deler av grensesnittet som en fungerende web-app. Selve grensesnittet som benyttet i den ferdige prototypen ble laget i *Adobe Illustrator*, og deretter animert i *Adobe After Effects*. Grensesnittet består av en hoveddel i form av et kart over flyplassen (OSL), samt en horisontal tidslinje som viser flyene under flygelederens kontroll. Flyene vises på kartet med smilefjes som indikerer ”humør,” og et tilsvarende ikon i større skala viser det ”generelle humøret,” som representerer hvor godt den totale løsningen passer med algoritmens forslag.



Figur 4: To skjermbilder fra den fiktive prototypen

### 4.3.2 Den narrative prototypen

Etter å ha fullført animeringen av det fiktive grensesnittet filmet vi en simulert interaksjon der animasjonene ble avspilt på en vanlig PC-skjerm, og hovedrollen i narrativen følger bevegelsene på skjermen som om han utfører handlingene. Filmen<sup>1</sup> ble redigert i *Adobe Premiere Pro* og *Adobe After Effects* og klippet sammen med *kontekstualiserende bilder* hentet fra nettkilder, som plasserer narrativen i kontekst av tid og sted.

## 5 Summativ Evaluering

For å validere den høyoppløselige prototypen gjennomførte vi en summativ evaluering med en fokusgruppe bestående av 5 domeneeksperter fra SINTEF. Samtlige deltakere var involvert i arbeidet med NextGenDST, og var også med i den formative evalueringen. Deltakerne var flerspråklige, noe som gjorde at vi ledet fokusgruppen på engelsk.

I forkant av evalueringen utviklet vi flere åpne spørsmål knyttet til spesifikke trekk ved prototypen. Vi valgte å gjøre spørsmålene så enkle som mulig for å skape et klart diskusjonsforum, og unngå forvirring (Lazar, Feng & Hochheiser, 2010, s. 194). For å sørge for at spørsmålene egnet seg til evalueringen, og for å utprøve logistikken ved evalueringen, gjennomførte vi en pilottest av evalueringen med interaksjonsdesign-studenter. Basert på vår vurdering av hvordan testen gikk, endret og presiserte vi spørsmålene og spørsmålsrekkefølgen til å passe bedre til sitt formål. Siden et av gruppe medlemmene er med i filmen, var han ikke med under evaluering for å unngå bias fra deltakerne.

Den summative evalueringen begynte med at vi spilte av den narrative-prototypen én gang foran deltakerne, mens de noterte ned sine tanker underveis. Etter dette holdt vi en

<sup>1</sup> Filmen kan sees på Youtube: <https://www.youtube.com/watch?v=wuElxbH-nbs>

spørsmålsrunde der vi stilte spørsmålene vi hadde forberedt, og lot hver enkelt deltaker svare. I tillegg la vi til rette for at de skulle gi kommentarer til hverandres svar, og diskutere høyt hva de tenkte. Vi avsluttet evaluering med en åpen runde, hvor alle fikk uttrykke generelle tanker om prototypen.

Vi valgte å ta lydopptak av sesjonen, som vi i etterkant transkriberte og analyserte. Vi utførte åpen koding, og kategoriserte nøkkelbegrep ut fra fra aspektene ved prototypen som vi laget spørsmål om.

Den summative evalueringen avdekket ulike synspunkter om hva som var av nytteverdi ved prototypen.

Når det gjaldt måten vi fremstilte algoritmens bakgrunn for valg, diskuterte deltakerne flere ulike aspekter ved prototypen.

- Samtlige deltakere var usikre på om smilefjes var en god måte å formidle status på. Dette fordi det kunne fremstå som barnslig, forstyrrende, eller demotiverende.
- Samtlige deltakere var svært positive til å bruke humør som en analogi for å formidle algoritmens status. En av deltakerne sa at dette “var noe de ønsket å bruke videre” i forskningsprosjektet ved utvikling av algoritmen. Her var det noe usikkerhet knyttet til hvor godt farger egnet seg for å representere humør: På den ene siden gjorde fargene det klart hvilken tilstand humøret var i. På den annen side kunne tilstedeværelsen av en negativt assosiert farge virke stressende på flygelederen.

Når det gjaldt framstilling av bevegelse påpekte en av deltakerne at det kunne være vanskelig å se hvilken retning flyene beveget seg i, og at skalering kunne bli et problem. Et alternativ kunne være å gi en retningsindikasjon til flyet i grensesnittet, eller plassere smilefjeset en annen plass på skjermen, og gjøre at det dukket opp når man velger det enkelte flyet.

Da fokusgruppen diskuterte anvendelsen av tidslinjen kom det frem flere ulike synspunkter.

- Deltakerne forholdt seg nøytrale til endringen fra en vertikal framstilling (flight strip) som anvendes i dag, til en horisontal tidslinje i scenariet. Deltakerne mente at opplæring kunne gjøre at denne overgangen ikke ville være vanskelig for flygelederne.
- Når det gjaldt innholdet i tidslinjen mente deltakerne at det var litt uklart hva som ble presentert. Deltakerne foreslo i dette tilfellet at det kunne være en tidslinje for hver



enkel rullebane. Et annet alternativ kunne være at hver flygeleder hadde sin egen tidslinje med både innkommende og utgående fly.

- Flere av ekspertene mente at det var positivt at tidslinjen økte muligheten for å planlegge fremover.

Videre diskuterte ekspertene symbolbruk i fiksjonen. Det kom frem at de ulike faktorene som påvirker flyenes individuelle og samlede humør var forståelige, men at de gjerne må deles opp i to kategorier. For eksempel er vær og flytype to faktorer som ikke varierer på en skala når det kommer til optimalisering, men heller operer som “true” eller “false”. Disse faktorene burde da skilles fra de andre og gjerne representeres på en annen måte.

Ved evaluering av hvordan man kan håndtere tilbakemeldinger til beslutningsverktøyet, oppstod en diskusjon om hvor objektive flygeledere er i etterkant av valgene sine. En av deltakerne påpekte at det å utsette tidspunktet for tilbakemelding kunne føre til skjevheter ved tilbakemeldingene. Ved å måtte utdype grunnlag for valg i øyeblikket kan man få en “sannere” forklaring, men dette tar også ekstra tid og kan være stressende for flygelederen. Det ble også påpekt at litt mer arbeid, selv i stressende situasjoner, kan utjevne arbeidsmengden gjennom dagen. Avslutningsvis ble det uttrykt et ønske om bruk av tale i kommunikasjon med algoritmen for å effektivisere prosessen.

## 6 Konkluderende diskusjon

Arbeidet med dette prosjektet har vist oss at det er utfordrende å jobbe med fremtidens løsninger i et komplekst domene. Vi opplevde design fiction som metodologi, med en narrativ prototype, som en god måte å utforske domenet på. Det krevde en del tilvenning fra vår side, da prosessen er veldig forskjellig fra de metodologiene og metodene vi har jobbet med tidligere: Både dualiteten ved brukerens rolle, og at målet med prototypen er diskusjon snarere enn en teknisk løsning, gjorde at vi stadig måtte styre forventningene og arbeidet vårt i en uvant retning.

Valideringen av den narrative prototypen vil først og fremst avhenge av om den skaper diskusjon rundt et ønsket tema. I denne sammenheng har de to diskusjonspunktene nevnt innledningsvis vært grunnlaget for både narrativet og den fiktive prototypen. De representerer aspektene ved fremtidsvisjonen vi ønsker å utforske nærmest, og dermed også skape diskusjon rundt.

Formidlingen av de bakenforliggende grunnene til algoritmens forslag var utfordrende å undersøke. Vi opplevde at humør som metafor ble hovedfokuset under evalueringene, og at det ble vanskelig å skape diskusjon rundt andre detaljer i fiksjonen. I lys av dette burde vi kanskje ha valgt andre metaforer, eller tilnærmet oss evalueringen på en annen måte. Samtidig viste humørmetaforen seg å være et konsept som oppfordret til diskusjon, og som tilsynelatende virket inspirerende for utvikling av nye løsninger. Dette kan tyde på at prototypen virket til sin hensikt.

Formidlingen av intensjon fra flygeleder til algoritme var enklere å skape diskusjon rundt. Dette kan være en følge av at narrativet viser denne interaksjonen mer tydelig gjennom de simulerte skjerminteraksjonene. Vi fikk sprikende tilbakemeldinger på hvor godt de presenterte konseptene for formidlingen fungerte. Dette førte til kreativ diskusjon og forslag til konkrete løsninger.

Gjennom den summative evalueringen opplevde vi at prototypen inspirerte til god diskusjon, der deltakerne deltok aktivt med forslag og tanker om fremtidens flygeledelse. I tillegg uttrykte deltakerne at de ønsket å bruke sentrale trekk ved fiksjonen videre i forskningsprosjektet. På grunn av dette, mener vi at prototypen var godt egnet til sitt formål, og at design fiction var en passende metodologi. Etter gjennomføringen av prosjektet sitter vi igjen med et ønske om å ha vært enda mer spekulative og kreative, for å utforske mulighetene i Design Fiction som metodologi i større grad.

Vi mener at arbeidet med prosjektet vårt viser at denne typen eksplorerende forskningsdesign kan være nyttig i arbeidet med å utvikle morgendagens teknologiske løsninger, spesielt innenfor felt der kompleksiteten begrenser mulighetene for innovasjon. I denne sammenheng mener vi at prosjektet har vært med på å utfordre, aktivere og inspirere klienten, og gitt ekspertene ved SINTEF et grunnlag for videre utforskning av interaksjonen mellom algoritmiske støtteverktøy og flygeledelse.

## 7 Siterte verk

- Bleecker, J. (2009). *Design Fiction: A short essay on design, science, fact and fiction*. Near Future Laboratory.
- Bleecker, J. (2010). *Design Fiction: From Props to Prototypes*. Hentet oktober 2017 fra Near Future Laboratory: <http://blog.nearfuturelaboratory.com/2010/09/21/design-fiction-from-props-to-prototypes-2010-01sj-biennial-catalog-essay/>
- Cooper, A. (2004). *The Inmates Are Running the Asylum*. Indianapolis: Sams Publishing.
- Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About Face: The Essentials of Interaction Design*. Indianapolis: Wiley & Sons.
- Dam, R., & Siang, T. (2017). *Design Thinking: Getting Started with Empathy*. Hentet november 2017 fra Interaction Design Foundation: <https://www.interaction-design.org/literature/article/design-thinking-getting-started-with-empathy>
- Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H.-C. (2000). Cognitive Stimulation in Brainstorming. *Journal of Personality and Social Psychology*, 79(5), ss. 722-735.
- Dunne, A., & Raby, F. (2013). *Speculative Everything: Design, Fiction, and Social Dreaming*. Cambridge: MIT Press.
- Hanseth, O., & Lyytinen, K. (2010). Design theory for dynamic complexity in information infrastructures: the case of building internet. *Journal of Information Technology*, 25(1), ss. 1-19.
- Heslin, P. A. (2009). Better than brainstorming? Potential contextual boundary conditions to brainwriting for idea generation in organizations. *Journal of Occupational and Organizational Psychology*(82), ss. 129-145.
- Houde, S., & Hill, C. (1997). What do Prototypes Prototype? I M. Helander, T. Landauer, & P. Prabhu (Red.), *Handbook of Human-Computer Interaction*. Amsterdam: Elsevier.
- Jones, M. (2015). *Jumping to the End - Practical Design Fiction*. Hentet fra Vimeo: <https://vimeo.com/121072011>
- Karahasanović, A., Følstad, A., Schittekat, P., & Nordlander, T. (2017). Putting a face on algorithms - personas for modelling artificial intelligence. *Utgift*. Oslo: SINTEF Digital.
- Kirby, D. (2010). The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development. *Social Studies of Science*, 40(1), ss. 41-70.
- Kleven, L. I. (2016). *Exploring Visualisation and Learning - Prototyping for Future Air Traffic Management Solutions*. Oslo: Universitetet i Oslo.
- Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. Glasgow: Riley Education.
- Lindley, J. (2015). A Pragmatics Framework for Design Fiction. *Proceedings of the 11th European Academy of Design Conference* (ss. 1-10). Paris: ADC.
- Lindley, J., & Coulton, P. (2015). Back to the Future: 10 Years of Design Fiction. *British HCI 2015* (ss. 210-211). Lincoln: ACM.
- Paulus, P. B., & Yang, H.-C. (2000). Idea Generation in Groups: A Basis for Creativity in Organizations. *Organizational Behavior and Human Decision Processes*, 82(1).
- Sheridan, T. B., & Verplank, W. L. (1978). *Human and Computer Control of Undersea Teleoperators*. Cambridge: MIT Man-Machine Systems Lab.
- Spielberg, S. (Regissør). (2002). *Minority Report* [Film].
- Sterling, B. (2012). *Bruce Sterling Explains the Intriguing New Concept of Design Fiction (Interview by Torie Bosch)*. Hentet september 2017 fra Slate: [http://www.slate.com/blogs/future\\_tense/2012/03/02/bruce\\_sterling\\_on\\_design\\_fictions\\_.html](http://www.slate.com/blogs/future_tense/2012/03/02/bruce_sterling_on_design_fictions_.html)
- Ødegård, S. S. (2013). *Exploring Visualization - Solutions for Air Traffic Control Workflow Productivity Improvement*. Oslo: Universitetet i Oslo.