

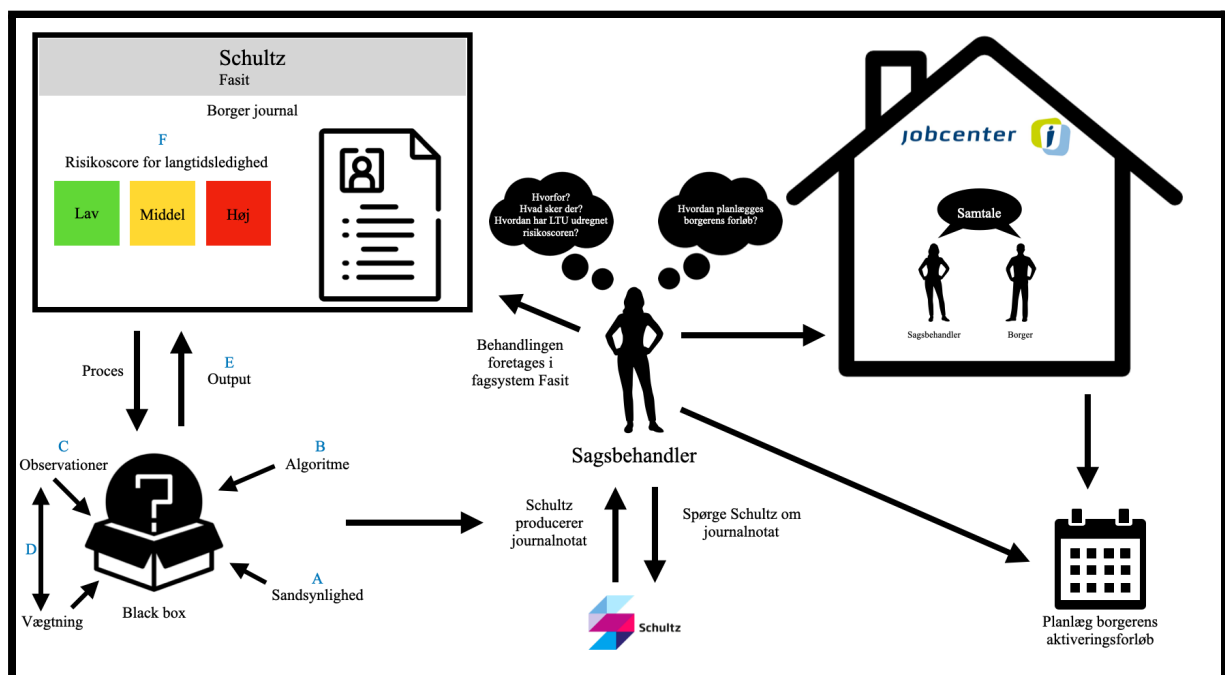
# Transparens i en virksomheds data-softwaressystem

Fagmodul i Informatik F2022 (25336) • Roskilde Universitet • HumTek-Forår 2022

Vejleder: Lars Bogetoft

Afleveringsdato: 01-06-2022

Anslag: 125972



Gruppe: S2225230497

Christine Van Wulffeld (71998)

Laura Sofie Juel Nielsen (72005)

## **Abstract**

Denmark is a frontrunner when it comes to digitalisation, which is especially seen within the public sector. Caseworkers in job centers use artificial intelligence (AI) tools as part of their work process. Concurrently with the rising use of AI in the public sector, there is an increased demand for transparency. AI used in casework can weaken the protection of citizens' rights that ensures openness and insight into public case work. Many companies try to create transparency by making the code available for their customers, but transparency fails as the code is conveyed in technical language, which is why transparency needs to be communicated in other ways. This project will investigate how to create greater transparency of an AI tool used in case work so that it complies with the law as well as the caseworkers' questions about the AI tool. The project will create and convey transparency of the AI tools through models so that it is understood by the users.

# Indholdsfortegnelse

<b>Abstract</b>	<b>2</b>
<b>Indholdsfortegnelse</b>	<b>3</b>
<b>Begrebsafklaring</b>	<b>5</b>
<b>1. Indledning</b>	<b>7</b>
1.1 Problemfelt	8
1.2 Problemformulering	9
1.3 Afgrænsning	9
<b>2. Metode</b>	<b>11</b>
2.1 Empiri	11
2.2 Interview	11
2.2.1 Ustruktureret interview	11
2.2.2 Semistruktureret interview	12
2.3 Modellering	13
2.3.1 Modelleringsteknikker	13
<b>3. Teori</b>	<b>15</b>
3.1 Transparens	15
3.1.1 Begrebet transparens i IT	15
3.1.2 Elementer i transparens	16
3.1.3 Niveau af transparens	16
3.2 UML Diagrammer	17
3.2.1 Domænemodel	17
3.2.2 Use case diagram	20
3.2.3 Sekvensdiagram	22
3.3 Rich picture	24
3.4 Rationelle beslutningsmodeller	24
3.4.1 Economic man	24
3.4.2 Administrative man	25
3.5 Mintzbergs organisationsformer	25
3.5.1 Koordineringsmekanismer	25
3.5.2 Basale dele af en organisation	26
3.5.3 Fagbureaukratiet og maskinbureaukratiet	26
<b>4. Empiri</b>	<b>28</b>
4.1 Sagsbehandlerens arbejdspraksis og it-systemer	28
4.1.1 Lovgivning	28
4.1.2 Digital sagsbehandling	30
4.1.3 Sagsbehandlerens arbejdsgang	30

4.1.4 Sagsbehandlerens holdning til LTU	31
4.2 LTU	32
4.2.1 Input	33
4.2.2 Proces (Matematiske beregninger)	34
4.2.3 Output	35
4.3 Journalnotat	36
4.4 Delkonklusion	38
<b>5. Analyse</b>	<b>40</b>
5.1 Begreber	40
5.2 Domænemodel	41
5.3 Use case diagram	43
5.4 Sekvensdiagram	44
5.5 Rich Picture og egne udarbejdede modeller	46
5.5.1 Rich Picture	46
5.5.2 A: Sandsynlighed	47
5.5.3 B: Algoritme	49
5.5.4 C: Observationer	50
5.5.5 D: Fortegnets effekt på risikoscoren	52
5.5.6 E: Output	54
5.5.7 F: Risikoscore	55
5.6 Delkonklusion	56
<b>6. Diskussion</b>	<b>58</b>
6.1 Skaber modellerne transparens?	58
6.2 Kan modellerne stå alene eller kræves der en beskrivelse?	60
6.3 Modellering og transparens	61
6.4 Delkonklusion	62
<b>7. Perspektivering</b>	<b>63</b>
7.1 Fagbureaukrati og maskinbureaukrati	63
7.2 Economic man og administrative man	64
<b>8. Konklusion</b>	<b>65</b>
<b>Litteraturliste</b>	<b>67</b>

# Begrebsafklaring

## Algoritme

Algoritmer er en samling af step-by-step instruktioner, som bliver udført af en computer.

## Arbejdspraksis

Samlet begreb for alt hvad der vedrører brugeren arbejdet, her skal man således anse arbejdet som funktion, opgave og proces

## Black Box

Betegner teknologier som er tilstrækkelige komplekse således at teknologien indeholde elementer som ikke ligetil at fortolke af mennesker.

## Features

Featurenavne er et begreb som anvendes i journalnotatet udarbejdet af Schultz. Featurenavne repræsenterer variabelernes navne, som anvendes i LTU algoritmen.

## Journalnotat

Tekstdokument som indeholder en vurdering af en potentiel borgers risiko for langtidsledighed. Indeholder oversigt over den pågældende borgers data (features) samt features beskrivelse, vægtning, spredning og gennemsnit.

## Kunstig Intelligens

Begreb for teknologier som har evnen til at have menneskelige egenskaber såsom læring, ræsonnement, beslutninger mm. Der eksistere forskellige former af kunstig intelligens heribland machine learning. Machine learning indebærer algoritme der kan analysere data hvoraf analysen anvendes til at forudsige eller gætte et udfald.

## LTU

LTU (Long term unemployment) er en algoritme udarbejdet af Schultz, hvor outputtet er en risikoscore for en borgers risiko for langtidsledighed.

### **Observationer**

Observationer er et begreb som anvendes i journalnotatet udarbejdet af Schultz, der repræsenterer variabelens værdi.

### **STU**

STU (Short term unemployment) er en algoritme udarbejdet af Schultz, hvor outputtet er en risikoscore for en borgers risiko for kortidsledighed.

### **Transparens**

Begreb for forståelig gennemsigtighed af et objekt, system, proces mm.

### **Variabler**

Variabler er de parametre som anvendes i LTU algoritmen og som ligger til grund for beregningen af risikoscoren. I journalnotatet fremgår det, at en variabel indeholder parameterens navn (featurenavn), parameterens beskrivelse, parameterens observation samt parameterens enhed.

### **Værdi**

Observationerne i en variabel indeholder en værdi, som kommer til udtryk ved enten at være et numerisk tal, sandt eller falsk sandhedsværdi, eller en kategori.

# 1. Indledning

Danmark er frontløber for anvendelse af it og digitale løsninger, hvilket blandt andet ses ved at Danmark i 2021 indtog førstepladsen i digitalisering i EU (Digitaliseringsstyrelsen, 2021). Denne digitalisering ses blandt andet indenfor den offentlige sektor, hvor man i stigende grad anvender teknologier, til at optimere den offentlige sektor samt løse eksisterende udfordringer og problemstillinger (Djøf & Tech DK Kommissionen, 2022).

Digitaliseringen i den offentlige sektor omfatter en række forskellige teknologier herunder kunstig intelligens. Kunstig intelligens er en fællesbetegnelse for teknologier, som har evnen til at have menneskelige egenskaber såsom læring, ræsonnement, beslutninger mm.

Rambøll og Dansk IT udgav i 2021 en rapport, som blandt andet viser at 41% af offentlige organisationer anvender kunstig intelligens. Derudover forudser Rambøll og Dansk IT at foregående procentdel vil stige til 91% inden 2024. Den offentlige sektor er derved igang med at indtræde i en ny epoke, hvor man i højere grad ønsker at anvende nye teknologier (Ginman, 2021).

Denne stigende udvikling afspejles også i Kommunernes Landsforenings seneste digitaliseringspolitik. Denne digitaliseringspolitik blev offentliggjort i maj 2022, og fokuserer på at den danske velfærd skal gentænkes, hvor digitalisering er central for denne gentænkning. Det påpeges at eksisterende og nye teknologier skal digitalisere kommunernes services (Kommunernes Landsforening, 2022).

Denne digitalisering ses også indenfor beskæftigelsesområdet. Beskæftigelsesområdet har til formål at få ledige borgere i henholdsvis arbejde eller uddannelse. Kunstig intelligens indgår i offentlige beslutningsprocesser, og har til formål at optimere sagsbehandlingen af borgere ved at supplere sagsbehandlerens arbejde. Eksisterende kunstig intelligens værktøjer indebærer eksempelvis databaseret match mellem ledig borger og virksomhed, vurdering af risiko for langtids- og kortidsledighed mm. (Schultz, n.d.-b). I forlængelse af den danske digitalisering udvikles nye kunstig intelligens værktøjer jævnlige. Odense Kommune er igang med at udvikle et kunstig intelligens værktøj som skal anbefale beskæftigelsesindsatser som bedst muligt bringer den ledige i arbejde (Borgmesterforvaltningen, 2021).

Den stigende anvendelse af kunstig intelligens i den offentlige sektor, tilføjer behovet for at sagsbehandlere kan formidle, hvorfor den kunstige intelligens giver de anbefalinger den gør. Dette behov kan sagsbehandlerne på nuværende tidspunkt ikke opfylde, da der generelt ikke

forekommer transparens af kunstig intelligens. Dette betyder, at sagsbehandlerne ikke har et grundlag for at kunne forstå kunstig intelligens.

## **1.1 Problemfelt**

Den øget anvendelse af kunstig intelligens i beskæftigelsesområdet og den fortsat manglende forståelse for denne, fremhæves af Djøf som en problematik. Kunstig intelligens i sagsbehandling kan svække beskyttelsen af borgernes rettigheder, der sikre åbenhed og indsigt i offentlig sagsbehandling. Djøf anbefaler dertil at give borgeren en gennemsigtighedsgaranti, som beror på borgernes ret til indsigt i sagsbehandling og dens afgørelse. Dette omfatter en skriftlig forklaring på, hvordan kunstig intelligens er indgået i denne behandling. Hertil fremhæver Djøf at den skriftlige forklaring kan understøttes af modeller (Djøf & Tech DK Kommissionen, 2022).

Den manglende forståelse for kunstig intelligens, kan ses i forlængelse med manglende transparens. Transparens henviser til formidling og gennemskueliggørelse af teknologier, herunder kunstig intelligens. Kunstig intelligens refereres ofte til som en black box, hvor der er hele eller dele af teknologien som ikke kan forklares. I sådan tilfælde anvendes brugbar transparens, hvor det er den tilgængelige information som skal formidles. Transparens kan således anvendes til at skabe forståelse for den pågældende teknologi.

Problematikken som Djøf fremhæver ses også i Lyngby-Taarbæk kommune som anvender kunstig intelligens værktøjet, LTU (Long Term Unemployment). LTU er en algoritmen som forudser en borgers risiko for langtidsledighed. Sagsbehandlerne oplever ligeledes manglende forståelse for LTU, og kan dermed ikke redegøre for dens forudsigelser. Denne rapport har til formål at åbne dele af LTU's black box, ved at undersøge og modellere algoritmen.

Modellernes formål er at skabe større transparens af LTU og dermed større forståelse hos sagsbehandlerne. Transparensen og den øgede forståelse hos sagsbehandlerne imødekommer Djøf's anbefaling af en gennemsigtighedsgaranti.

Projektgruppen har i denne sammenhæng indgået et samarbejde med virksomheden Schultz, som er leverandør af LTU. Dette samarbejde er indgået med det formål, at få adgang til ikke offentlige oplysninger samt opnå en dybdegående viden om LTU. Ligeledes har dette samarbejde muliggjort besvarelsen af problemformuleringen.



## 1.2 Problemformulering

Hvordan kan modelleringsværktøjer bidrage til at skabe og formidle transparens af algoritmen LTU, således at LTU bliver forståelig for brugeren?

### Arbejdsspørgsmål:

- Hvad indebærer transparens?
- Hvad er modelleringsværktøjer?
- Hvordan er sagsbehandlerens arbejdspraksis, hvori LTU anvendes?
- Hvad er sagsbehandlerens holdning til den nuværende transparens af LTU?
- Hvordan er algoritmen LTU's input, proces og output?
- Hvordan kan modeller skabe og formidle transparens af LTU?
- Hvorvidt har modellerne skabt transparens?

## 1.3 Afgrænsning

I forbindelse med udarbejdelse af rapporten, har projektgruppen foretaget fravalg af en række emner og områder. Dette skyldes både rapportens omfang samt relevansen i henhold til problemstillingen.

Vi har valgt at anvende metoden modellering til at skabe transparens af en teknologi. Metoden modellering og teorien om transparens fokuserer udelukkende på teknologi, hvor projektgruppen afgrænser fra andre aspekter.

I rapporten anvendes 3 forskellige UML diagrammer til modellering. Disse diagrammer er anvendt med henblik på, at beskrive relevante begreber, brugerens interaktion med systemet samt meddelesudveksling til og fra forskellige objekter i et pågældende scenarie.

Diagrammerne modellerer forskellige aspekter af LTU. I forlængelse af dette, forekommer der afgrænsning af andre UML diagrammer. Yderligere forekommer der afgrænsning af teorien om de tre UML diagrammer, hvor teorien inddrager de elementer, som indgår i analysen.

Teori om beslutningsmodeller og Mintzberg anvendes i rapportens perspektivering, hvor teorien fokuserer på rationelle beslutningsmodeller samt Mintzbergs organisationsformer; fagbureaukrati og maskinbureaukrati. Rapporten afgrænser dermed fra andre beslutningsmodeller samt andre teorier fra Mintzberg. Denne afgrænsning er foretaget, da

perspektivering kun inddrager de rationelle beslutningsmodeller og de førnævnte organisationsformer.

Rapportens empiri som beskriver sagsbehandlerens arbejdspraksis og deres holdning til LTU, tager udgangspunkt i Lyngby-Taarbæk kommune. Denne afgræsning forekommer, da projektgruppen ikke har haft mulighed for at interviewe et jobcenter som anvender LTU. Yderligere har den tilgængelig information taget udgangspunkt i sagsbehandlere, og rapporten afgrænser dermed fra at inddrage jobcenteret som organisation. Sagsbehandlere skal overholde en række forskellige love, hvor empirien beskriver de love som direkte influerer sagsbehandlere, ved anvendelse af LTU.

Rapporten tager udgangspunkt i teknologien LTU, og afgrænser dermed fra andre teknologier som sagsbehandlere kan anvende. Denne afgræsning indebærer teknologier såsom STU og STAR algoritmen.

Empirien om LTU inddrager overordnede beskrivelser af algoritmens fremtrædende matematiske modeller, og afgrænser fra dybdegående matematiske beskrivelser. Yderligere fremgår det i empirien, at variabler kan være både numeriske og kategoriske, hvor analysens modeller inddrager numeriske variabler. De kategoriske variabler er afgrænset i analysens modeller grundet rapportens omfang.

Afslutningsvis, har projektgruppen afgrænset til at teste modellerne på almene borgere fremfor den tiltænkte brugergruppe, sagsbehandlere. Dette skyldes, at det ikke har været muligt at få kontakt med relevante jobcentre. Årsagerne til dette uddybes i afsnit 6. Diskussionen.

## **2. Metode**

Følgende afsnit indeholder en beskrivelse af rapportens metodiske tilgang til indsamling af empiri. Dernæst redegøres for metoden interview samt hvordan projektgruppen har anvendt metoden. Afslutningsvis beskrives modellering hvor der redegøres for spændet mellem kompleksitet og abstraktion. Derudover inddrages modelleringsteknikker som er anvendt i rapporten.

### **2.1 Empiri**

Empirien i denne rapport beror på både primært og sekundært materiale til besvarelse af problemformuleringen.

Rapportens primære materiale er indsamlet ved metoden interview, hvor vi har afholdt interviews med ansatte fra Schultz. Formålet var at indsamle viden om algoritmen LTU på et generelt og teknisk niveau. Derudover har vi afholdt interviews med en række forskellige personer, som repræsentere almene borgere i samfundet. Disse interviews havde til formål at teste hvorvidt modellerne i analysen skaber større transparens. Denne viden anvendes i udarbejdelsen af modellerne, samt i rapportens diskussion.

Rapportens sekundær materialer omfatter Schultz' interne dokumenter samt korrespondance med ansatte i Schultz. Materialerne indeholder powerpoints, databasediagram, journalnotat og opfølgende spørgsmål. Yderligere anvendes forskningsprojekter, bøger og internetkilder. Projektgruppen har underskrevet fortrolighedskontrakt med Schultz hvilket betyder at projektgruppen har foretaget anonymisering af materialerne modtaget fra Schultz. Derudover frabad alle interviewpersonerne at blive optaget, og projektgruppen har derfor foretaget noter under de forskellige interviews. Noterne er udarbejdet i stikordsformat, som fremgår i bilagslisten.

### **2.2 Interview**

Metoden interview er anvendt i projektet, for at indsamle viden fra Schultz og til at teste rapportens modeller på almene borgere. Projektgruppen har anvendt interviewformerne; ustruktureret og semistruktureret interviews.

#### **2.2.1 Ustruktureret interview**

Det ustruktureret interviews overordnede mål er, at indsamle dybdegående viden og erfaringer fra interessenten. Dette udføres i en samtaleform, hvor der er planlagt emner og

dagsorden for samtalen, men uden at planlægge konkrete forudbestemte spørgsmål. Samtalen bliver derfor vejledt af de overordnede projektmål, emner som skal indgå i projektet og emner som interessen og interviewereren finder relevante. Dermed er det både interviewereren og interessen som kan påvirke det ustruktureret interviews retning, hvor dagsordenen for interviewet kan ændres undervejs i samtalen. Dette kræver at interviewereren forholder sig åben for uforudsete emner. Derudover skal samtalen forholde sig relevant til projektet, og derved sørge for at der ikke forekommer gentagelser. Intervieweren anvender en "interviewguide" som indeholder en dagsorden. Denne interviewguide anvendes til at vejlede interviewet således, at de ønskede emner afdækkes i samtalen (Wilson, 2014).

Projektgruppen har anvendt ustruktureret interview, ved det første interview med Schultz, hvor formålet med samtalen var at skabe en bred viden om emnefeltet. Ved planlægningen af det ustruktureret interview, indeholdt interviewguiden en dagsorden med emner såsom Schultz' systemer, transparens i systemerne og efterspørgsel fra deres eksisterende kunder.

### **2.2.2 Semistruktureret interview**

Et semistruktureret interview er en kombination med foruddefinerede spørgsmål og en åben tilgang til interviewet. Det giver mulighed for at stille uddybende spørgsmål undervejs, eller at samtals retning ændres ved udfoldelsen af interviewet. Derfor er et semistruktureret interview en kombination af det struktureret og ustruktureret interview. Det semistruktureret interviews mål er, systematisk at indsamle viden om bestemte emner, hvoraf det er muligt at udforske eventuelle nye problemstillinger der kan forekomme undervejs. Denne interviewform anvendes, når interviewereren allerede indehaver en vis viden om emnet eller problemstillingen, men stadig har behov for mere viden og flere detaljer. Interviewet kan inddrage både åbne og lukkede spørgsmål, hvilket kan resultere i både kvantitative og kvalitative data. Ved denne type interview, indeholder "interviewguiden" blandt andet formålet med interviewet, emnerne som skal indgå samt spørgsmål til emnerne (Wilson, 2014).

Projektgruppen har anvendt denne interviewform til størstedelen af de udførte interviews. Anvendelsen af det semistruktureret interview, gav mulighed for at have en interviewguide med foruddefinerede spørgsmål til hvert emne, hvor en åben tilgang gav plads til at forme nye spørgsmål undervejs i interviewet.

## 2.3 Modellering

Modeller anvendes indenfor en række fagområder, og er særlig effektive til at skabe overblik og forenkler kompleksitet. Dette resulterer i at modeller skaber en forenklet repræsentation af virkeligheden. Præmissen for modellens effektivitet er, at mennesket fundamentalt har begrænset kapacitet til at absorbere informationer, og derved forstå komplekse systemer. Anvendelse af modeller forenkler kompleksiteten og begrænser mængden af informationer. Modellering udføres ved at kompleksiteten analyseres og transformeres til modeller, som repræsenterer en abstrakt og forenklet version af systemet. Det fremhæves i teksten, "The Principle of Limited Reduction", at formindskning af kompleksitet øger usikkerheden, hvortil det i forsøget på at imødekomme usikkerheden øger kompleksiteten (Mathiassen & Stage, 1992).

Abstraktion anvendes til at formindske kompleksiteten, hvilket medfører minimering af detaljer. Abstraktion resulterer dermed i, at modellerne er en simplificeret repræsentation af virkeligheden. Modellerne er derfor ikke identiske med virkeligheden eller det pågældende system. Relevante overvejelser er hvorvidt abstraktionen har minimeret relevante detaljer, samt om repræsentationen er pålidelig i henhold til det reelle system. Det er dermed vigtigt at være opmærksom på forholdet mellem abstraktion, forenkling og kompleksitet, samt de abstraktions usikkerheder dette medfører (Mathiassen & Stage, 1992).

Indenfor IT anvendes modeller som et metodisk værktøj i udvikling og design af nye systemer. Derudover anvendes modeller til at dokumentere og skabe forståelse for eksisterende systemer (Mathiassen & Stage, 1992). Modellering er særlig fordelagtigt til at nedbryde komplekse teknologier, hvilket opnås ved at analysere teknologien fra forskellige aspekter. Særligt indenfor IT er det fordelagtigt at anvende UML diagrammer, da disse fokuserer på forskellige aspekter af et system. Diagrammerne anvendes til at nedbryde kompleksiteten samt til formidling af systemet, hvor diagrammerne er forenkledede og visuelle repræsentationer af systemet. Dette medfører at brugeren opnår større forståelse for teknologien. Modellering i rapporten anvendes til at skabe forståelse for algoritmen, LTU samt at formidle denne forståelse.

### 2.3.1 Modelleringsteknikker

Projektgruppen anvender modelleringsteknikken UML, hvor use case diagram, sekvensdiagram og domænemodel indgår i rapporten. Diagrammerne bliver anvendt til at dokumentere systemets funktionaliteter, nedbryde kompleksiteten i LTU og til at analysere

den kontekst LTU anvendes i. I forlængelse af dette, bliver der skabt abstrakte repræsentation af forskellige aspekter.

Projektgruppen har udarbejdet domænemodellen ud fra klasse diagrammets notationer. Modellen anvendes til at identificerer konceptuelle begreber i og omkring LTU, hvortil der bliver skabt et visuelt overblik af begrebernes relationer. Use case diagrammet anvendes til at illustrerer sagsbehandlerens interaktion med fagsystemet Fasit, herunder LTU. Denne model fokuserer på, at identificerer brugerens interaktion med det pågældende system. Yderligere anvendes sekvensdiagram til at illustrerer interaktion mellem objekter i et system. Der forekommer to sekvensdiagrammer i rapporten, hvoraf det ene uddyber en use case fra use case diagrammet og det andet illustrerer LTU's input, proces og output.

Udover UML anvendes modelleringsteknikken rich picture samt egne udarbejdet modeller. Rich picture anvendes som et supplement til domænemodellen, hvormed rich picture tilføjer flere aspekter og detaljer, såsom sagsbehandlerens arbejdspraksis i henhold til anvendelsen af LTU. Yderligere indeholder rapporten egne udarbejdet modeller, som hver fokuserer på forskellige aspekter af LTU.

## 3. Teori

I følgende afsnit vil vi redegøre for udvalgte teorier. Først vil vi beskrive begrebet transparens og dets elementer i henhold til teknologi. Dernæst vil vi redegøre for udvalgte UML diagrammer, herunder de notationer som anvendes i analysen. Afslutningsvis beskrives rationelle beslutningsmodeller og Mintzbergs koordineringsmekanismer, grundmodellen og organisationsformerne; fagbureaukratiet og maskinbureaukratiet. Dette beskrives med henblik på at belyse måder hvorpå mennesker træffer beslutninger samt hvordan en organisation ideelt opdeler og koordinerer arbejdet.

### 3.1 Transparens

Den danske ordbog definerer begrebet transparens som følgende "*åbenhed og gennemskuelse i en sag eller sammenhæng*" (Den Danske Ordbog, n.d.). Dog forekommer den danske ordbogs definition ikke fyldestgørende, da der indenfor IT er udarbejdet yderligere definitioner. Følgende afsnit vil redegøre for begrebet transparens i henhold til teknologi med fokus på kunstig intelligens. Redegørelsen af begrebet tager udgangspunkt i kapitel 10 "Transparency" skrevet af Nicholas Diakopoulos, i bogen "The Oxford Handbook of Ethics in AI" og Artikel 29 "Guidelines on transparency under Regulation 2016/679" udarbejdet af European Commission.

#### 3.1.1 Begrebet transparens i IT

Diakopoulos og GDPR (General Data Protection Regulation) beskriver transparens i en kontekst hvor teknologi behandler data med det formål at anvende resultatet som beslutningsstøtte. I disse tilfælde kan resultatet være en rangeringsscore, klassifikation og/eller associationer mellem data (European Commission, 2018).

I "Transparency" kapitlet definerer Diakopoulos transparens som følgende "*is about information, related both to outcomes and procedures used by an actor, and it is relational, involving the exchange of information between actors*" (Diakopoulos, 2020, s. 198). I dette citat refererer "actor" til brugeren af den pågældende teknologi. Transparens har til formål at formidle og gennemskeleliggøre teknologien, hvor input, proces, output og aktører er relevante elementer at beskrive.

Yderligere indgår begrebet transparens i GDPR forordningen, hvor det har til formål at sikre formidling og kommunikation om behandling af persondata "*requires that any information and communication relating to the processing of those personal data be easily accessible and*

*easy to understand*” (European Commission, 2018, 2016, s. 6). GDPR påkræver derved, at brugeren skal oplyses om deres behandling af data, og denne oplysning skal forekomme let tilgængeligt samt forståeligt for brugeren.

Diakopoulos påpeger at begrebet transparens hidtil har skelnet mellem transparens af systemets resultatet og transparens af processen i systemet (Diakopoulos, 2020). Hvormed Diakopoulos og GDPR definere transparens således, at det indebærer formidling af både processen i systemet, dets resultatet samt hvordan resultatet indgår i yderligere behandling.

### **3.1.2 Elementer i transparens**

Overordnet ses det at begrebet transparens indeholder forskellige elementer, som kan kategoriseres under følgende: input, proces, output, forståelighed og kommunikation.

I henhold til kunstig intelligens forekommer input i form af data, processen som matematiske modeller og outputtet er resultatet af foregående. Hvortil antagelser og normer er en del af den overordnede proces (Diakopoulos, 2020; European Commission, 2018).

Yderligere fremtræder forståelighed og formidling som centrale elementer, hvor forståelighed henviser til hvorvidt brugeren forstår formidlingen af systemets transparens.

Diakopoulos argumenterer følgende *“this information must be presented to recipients and stakeholders in ways that they can actually make sense of and connect to their specific goals”* (Diakopoulos, 2020, s. 204). Han pointerer dermed, at forståelighed og formidling skal anskues i kontekst til brugeren og deres mål. Dertil fremhæves det, at transparens kræver overvejelser om, hvilke informationer der skal inddrages samt hvordan dette skal formidles til brugeren (Diakopoulos, 2020).

I GDPR påpeges det, at formidlingen af transparensen skal henvende sig til en almindelige bruger af den tilsigtede brugergruppe (European Commission, 2018). Formidlingen forekommer i forlængelse af forståelighed, hvor formidlingen ifølge GDPR bør være *“easy to understand, and that clear and plain language be used”* (European Commission, 2018, s. 6). Diakopoulos og GDPR fremhæver derved, at transparens opnås når den almene brugere forstår teknologien, og at transparensen er formidlet i et forståeligt sprog.

### **3.1.3 Niveau af transparens**

Kunstig intelligens, herunder algoritmer omtales ofte som “black boxes”. Black box fremhæver problematikken omkring manglende forståelse for teknologiers processer. Det kan forekomme at man har kendskab til teknologiens overordnede proces(ser), men at dele af



systemet er uforståelig (Merriam-Webster, n.d.). Fænomenet black box er iblandt en af årsagerne til at man som udvikler skal vurdere niveauet af transparens, da det ikke altid er muligt at skabe fuld transparens af kunstig intelligens. I relation til kunstig intelligens henviser Diakopoulos til “usable transparency”. Han fremhæver, at på trods af black boxes, indeholder den pågældende teknologi stadig potentielle, relevante og forståelige informationer (Diakopoulos, 2020). Usable transparency henviser dermed til et niveau af transparens, som formidler de oplysninger der er tilgængelige. Derudover vejleder European Commission at formidlingen af informationer bør tage udgangspunkt i effektiv og kortfattet beskrivelse, således at man undgår informationstæthed. Dette kan opnås ved at undlade informationer, og dermed skabe forskellige niveauer af transparens. European Commission argumentere for anvendelse af usable transparency. (European Commission, 2018).

Transparens kan forekomme i tekstformater, dokumenter og forskellige visuelle modeller, som hver henviser til forskellige niveauer af transparens (Diakopoulos, 2020).

## **3.2 UML Diagrammer**

UML (Unified Modeling Language) beskrives som “*a visual language for specifying, constructing and documenting the artifacts of systems*” (Larman, 2005, s. 11), hvor modeller vises grafisk i form af diagrammer. UML tilbyder 14 forskellige diagrammer som opdeles i henholdsvis struktur- og adfærdsdiagrammer (Seidl et al., 2015) .

Struktur diagrammerne anvendes til modelleringen af strukturen i et system, hvor den dynamiske opførelse af elementerne ikke er inkluderet. Ændringerne over tid bliver derfor ikke taget i betragtning. Adfærdsdiagrammerne viser konsekvenser af mindst ét objekts handling, altså påvirkningen af hvordan objekters tilstande bliver ændret med tiden.

Kort fortalt, beskriver modellerne en del af virkeligheden, som diagrammerne giver et overblik over. Konkret udtrykker UML diagrammerne hvilke brugere der benytter hvilke funktionaliteter og hvordan systemerne er opbygget, uden at specificere en decideret implementering (Seidl et al., 2015).

### **3.2.1 Domænemodel**

En domænemodel referer også som en konceptuel model, da det er en visuel repræsentation af konceptuelle objekter eller klasser som er skabt på baggrund af virkelige situationer.

Domænemodellens fokus beskrives som “*focusing on explaining ‘things’ and products important to a business domain*”(Larman, 2005, s. 134). Dette betyder, at udvælgelsen af

begreberne er en vigtig faktor for forståelsen af den konceptuelle model. Begreberne i modellen skal afspejle det tiltænkte domæne.

Domænemodellen anvender UML klassediagrammets notationer. Ved at anvende klassediagrammets notationer, vil domænemodellen indeholde konceptuelle klasser, associationer og attributter, hvortil metoder udelades. Forskellen på klassediagrammet og domænemodellen ses ved at klassediagrammet, kan indeholde software og konceptuelle perspektiver, hvor domænemodellen kun tilbyder konceptuelle perspektiver. Det konceptuelle perspektiv illustreres ved konceptuelle klasser eller objekter og attributter af en konceptuel klasse samt associationer mellem disse (Larman, 2005).

### **3.2.1.1 Konceptuel klasse**

En konceptuel klasse beskrives uformelt som en idé, ting eller objekt. Mere formelt, skal en konceptuel klasse tage dens symbol, intention og forlængelse i betragtning. Symbol henviser til ord eller billeder der repræsenterer en konceptuel klasse. Intention henviser til definitionen af en konceptuel klasse og forlængelsen henviser til det sæt eksempler, som den konceptuelle klasse er gældende for. Et eksempel kan være en købstransaktion, hvor symbolet henviser til den konceptuelle klasse med navnet "Salg". Intentionen er at salg repræsenterer en købstransaktion, hvor der fremgår en dato og et tidspunkt. Forlængelsen af salg er alle eksempler på salg, altså sættet af alle salgsforekomster (Larman, 2005).

For at finde og navngive konceptuelle klasser, kan 3 strategier anvendes. Den første strategi er at genanvende og ændre i eksisterende modeller. Den anden strategi er at anvende en kategoriliste, hvorfra man kan udarbejde en liste af mulige konceptuelle klasser. Tredje strategi, henviser til at identificere navneord som kan anvendes til mulige konceptuelle klasser (Larman, 2005).

### **3.2.1.2 Associationer**

En association er forholdet mellem klasserne, som indikerer en forbindelse som er meningsfuld og interessant. I UML beskrives en association som "*the semantic relationship between two or more classifiers that involve connections among their instances*" (Larman, 2005, s. 150).

Associationer i domænemodellen inddrages til at fremhæve relationer mellem de konceptuelle klasser eller objekter. Associationen bliver vist i form af en linje mellem klasserne, hvor et associationsnavn kan indgå. Associationsnavnet formidler klassernes relation. Derudover kan man anvende en "læseretningspil", som angiver læseretningen for

associationen. Hvis denne pil undlades, er normen at læse fra venstre mod højre, eller fra top til bund (Larman, 2005).

Associationlinjernes ender kaldes en “rolle” og kan indeholde multiplictetsværdier, som angiver det numeriske forhold mellem klasserne (Larman, 2005). Multiplicitetsværdierne ses i nedenstående figur 1: Multiplicitetsværdier.

Multiplicitetsværdier	Betydning
*	nul eller flere; “mange”
1..*	En eller flere
1..40	En til 40
3	Præcis 3
2,5,7	Præcis 2,5 eller 7

Figur 1: Multiplicitetsværdier (Larman, 2005)

### 3.2.1.3 Attributter

En attribut defineres som en logisk dataværdi for et objekt, hvor den i UML har syntaksen:

“*visibility name: type multiplicity = default {property-string}*” (Larman, 2005, s. 159).

Visibility specificerer, hvem der har, og hvem der ikke har tilladelse til at få adgang til attributten (Seidl et al., 2015). De forskellige typer af visibilities, kan ses i nedenstående figur 2: Visibilities

Navn	Symbol	Betydning
Public	+	Attributter der er markeret som public, har alle klasser adgang til.
Private	-	Attributter der er markeret som private, beskytter mod uautoriseret adgang, hvor kun objektet selv kender til værdierne af attributterne.
Protected	#	Attributter der er markeret som protected, har kun klassen selv og dens underklasser adgang til.
Package	~	Attributter som er markeret med package, har klasser der er i samme pakke som denne klasse, adgang til attributten.

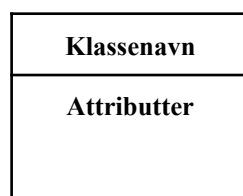
Figur 2: Visibilities (Seidl et al., 2015)

Hvis attributten ikke har angivet et visibility tegn, bestemmer UML ingen visibility tegn som standard (Seidl et al., 2015). Dog som en konvention, vil de fleste udviklere antage, at attributter har en private visibility (-), hvis der ikke er angivet andet (Larman, 2005). En attribut har som minimum et navn, hvorefter den kan have en type. De mulige attributtyper

kan omfatte primitive datatyper eksempelvis heltal (Integer) eller streng (String). Ved at angive navn og en datatype som f.eks. "Attributnavn: String", defineres Attributnavn med typen String.

Attributtens multiplicity angiver, hvor mange værdier attributten kan indeholde. Hvis attributten ikke har en multiplicitetsværdi angivet, antages værdien som standard at være 1, hvilket henviser til en attribut med en enkelt værdi. Multiplicitetsværdien angives i firkantede parenteser "[ ]", hvor syntaksen for multiplicity er "[minimum .. maksimum]". Minimum og maksimum er tal, der henviser til den nedre og øvre grænse, hvor minimumsværdien skal være mindre end eller lig med værdien af maksimum. Hvis der ikke er en grænse for maksimum, indikeres dette med en stjerne (\*). Hvis minimum og maksimum er identiske, kan man nøjes med at skrive tallet i de firkantede parenteser. Yderligere kan attributten have en standardværdi, som kan defineres ved at anvende "= Default". Default er en værdi eller et udtryk som er brugerdefineret. Til sidst i attributtens fulde syntaks, indeholder den "properties", som kan skrives i krøllede parenteser "{}". Properties angiver egenskaber, som f.eks. egenskaben {readOnly}, der henviser til at attributværdien ikke kan ændres, når først den er blevet initialiseret (Seidl et al., 2015).

Attributterne placeres i en anden boks end klasse navnet (Larman, 2005). Dette illustreres i nedenstående figur 3: Konceptuel klasse boks.



Figur 3: Konceptuel klasse boks

### 3.2.2 Use case diagram

Et use case diagram anvendes til at illustrere use case navne, aktører og forholdene mellem dem i et system (Larman, 2005). Ligeledes udtrykker use case diagrammet, hvad et system skal gøre, uden at specificere en implementering, datastruktur, algoritmer osv. Use case diagram anvendes til udvikling af et nyt system, samt til dokumentation af et eksisterende. Dokumentationen indebærer systemets funktionaliteter, og registrere hvilke brugere der har tilladelser til de forskellige funktionaliteter i systemet (Seidl et al., 2015).

Systemet i use case diagrammet, illustreres som en rektangel, som indeholder use cases. Use cases repræsenterer funktioner i et system, som bliver forbundet med aktørerne ved brug af

association. Aktørerne er placeret udenfor systemet (Seidl et al., 2015). Aktører, associationer og use cases bliver uddybet i de følgende underafsnit.

### **3.2.2.1 Aktør**

Aktører interagerer med systemet i sammenhæng med deres use cases som de er associeret med. Aktører illustreres i form af tændstikfigurer, valgfrit definerbart symbol eller rektangler hvor «aktør» indgår. Derudover repræsenterer en aktør ikke en specifik bruger, men en rolle som en bruger kan anvende. Aktøren er altid placeret udenfor systemet, hvor der gennem associationer illustreres hvilke use cases de kan anvende (Seidl et al., 2015).

I use case diagrammet findes tre typer aktører; primær aktør, sekundære aktør og offstage aktør. Primær aktør placeres til venstre for systemet og de resterende placeres til højre for systemet. En primær aktør er en bruger, som anvender systemets funktionaliteter, hvor den sekundære aktør leverer en service til systemet. Den tredje type aktør, offstage aktør, har en interesse i systemet, men er hverken primær eller sekundær aktør (Larman, 2005).

En aktør kan være menneskelig eller ikke-menneskelig, og kan enten have en aktiv eller passiv rolle. Den aktive rolle henviser til at en aktør interagerer med systemet, og dermed igangsætter udførelsen af en use case. Den passive rolle henviser til, at aktøren bliver brugt af systemet og yderligere giver funktionalitet til udførelse af use cases (Seidl et al., 2015).

### **3.2.2.2 Associationer**

En association forbinder en aktør og en use case, og illustreres i form af en solid linje. Ved at forbinde en use case med en aktør kan aktøren både kommunikere med systemet og anvende en bestemt funktion i systemet. En aktør skal som minimum have en association til mindst én use case, ellers interagerer aktøren ikke med systemet. Ligeledes skal en use case altid have en association til en aktør, da den ellers ville være irrelevant, i og med at den aldrig bliver anvendt. Associationer fremgår som binær, hvor den er specificeret mellem én aktør og én use case. Ved associationens ender, kan man inddrage multiplicitetsværdier, hvor den samme UML syntaks fremgår som skrevet i figur 1: Multiplicitetsværdier i afsnit 3.2.1.2

Associationer. I tilfælde af, associationens ende ved aktøren har en multiplicitetsværdi som er større end 1, henviser den til at der er flere end én instans af en aktør, som er inddraget i udførelsen af use casen. Hvis der ikke fremgår en multiplicitetsværdi for aktøren, anvendes 1 som standardværdien (Seidl et al., 2015).

### **3.2.2.3 Use case**

En use case kan blive illustreret i form af en ellipse, hvor use case navnet fremgår i eller under ellipsen. Use cases kan beskrive de funktionaliteter, som kunden forventer at der bliver udviklet i et fremtidigt system. Derudover kan use cases anvendes til at dokumentere den eksisterende funktionalitet, som et system har. Dog beskriver use case diagrammet ikke den interne struktur og en eventuel implementering af use casene. Use casene befinder sig på et højt abstraktionsniveau (Seidl et al., 2015).

Use cases kan have relationer til hinanden, herunder “include” og “extend” forhold. Ved anvendelsen af include og extend, benyttes en stipletpil fra en use case til en anden use case, hvor teksten “«include»” eller “«extend»” indsættes ved pilen.

Include anvendes når en use case inkludere en anden use case. Den første use case omtales som “base use case”, hvor den use case der bliver inkluderet omtales som en “included use case”. Hver gang “base use case” udføres, skal den “included use case” også udføres. I et extend forhold, anvendes begrebet “base use case” ligeledes ved den første use case, hvor den anden use case bliver omtalt som en “extending use case”. Extend anvendes når en use case har mulighed for, at anvende adfærden fra en anden use case, uden at det er obligatorisk. Dermed kan begge use cases udføres uafhængigt af hinanden (Seidl et al., 2015).

### **3.2.3 Sekvensdiagram**

Et sekvensdiagram er et interaktionsdiagram, der illustrerer interaktionerne mellem objekterne i et system, og hvor samspelet mellem interaktionspartnere omfatter en sekvens af meddelelser. Sekvensdiagrammet illustreres som et todimensionelt diagram. På den horisontale akse fremvises interaktionspartnere, hvor den lodrette akse fremviser den kronologiske rækkefølge af interaktionen (Seidl et al., 2015). Interaktionsdiagrammet, interaktionspartnere og meddelelser bliver uddybet nærmere i følgende underafsnit.

#### **2.2.3.1 Interaktionsdiagram**

UML tilbyder fire forskellige interaktionsdiagrammer, hvor sekvensdiagrammet oftest anvendes. Interaktionsdiagrammerne anvendes til at specificere interaktioner, hvor man modellerer efter et konkret scenarie. Sekvensdiagrammer fremhæver udveksling af meddelelser mellem forskellige interaktionspartnere. Ved at modellere et konkret scenarie, foregår meddelesudvekslingen indenfor en specifik kontekst, som har til formål at løse en bestemt opgave. Derudover er det generelt kun en bestemt del af scenariet, som interaktionerne beskriver. Interaktionsdiagrammer kan anvendes i forskellige situationer, da

interaktionernes meddelelser kan beskrives i henhold til modtagerens kompetencer. De kan f.eks beskrives til udviklere eller slutbrugere (Seidl et al., 2015).

### **2.2.3.2 Interaktionspartnere**

Interaktionspartnere i sekvensdiagrammet henvises som “livslinjer”, hvor en livslinje bliver illustreret ved en lodret stiplede linje. Øverste ende af livslinjen kaldes hovedet og illustreres gennem symboler som f.eks en tændstikfigur eller en rektangel, hvor navnet på objektet indgår. Hvert objekt bliver tildelt en separat livslinje, hvor hver livslinje illustrerer objektets levetid (Seidl et al., 2015). På en livslinje kan der forekomme en til flere aktiveringsbarer, som indikerer hvor og hvornår et objekt udfører en proces. Aktiveringsbaren illustreres ved en rektangel placeret på livslinjen (Larman, 2005; Seidl et al., 2015).

### **2.2.3.3 Meddelelse**

Sekvensdiagrammet anvender meddelelser fra afsender til modtager, som bliver illustreret i form af en pil. Typen af pilen henviser til typen af kommunikationen. Ved anvendelse af en synkron meddelelse, er pilen illustreret ved at linjen er kontinuerlig og hvor hovedet af pilespiden er en udfyldt trekant. Asynkrone meddelelser illustreres med en pil med ubrudt linje og hvor pilespiden er åben. Ved anvendelsen af den synkrone meddelelse, afventer afsenderen på at fortsætte, indtil afsenderen har modtaget en svarmeddelelse.

Svarmeddelelsen illustreres af en pil med en stiplede linje, hvor pilespiden er åben. Hvis afsenderen fortsætter direkte efter en sendt meddelelse, anvendes den asynkrone meddelelse. En meddelelse identificeres med et navn, hvor valgfrie specifikationer af parametre og returværdi kan indgå (Seidl et al., 2015).

I sekvensdiagrammet kan man anvende “found message”, som illustreres med en sort cirkel. Found message symboliserer, at en afsender af en meddelelse enten er ukendt eller ikke relevant. Den sorte cirkel kan dermed anvendes fremfor en interaktionspartner. Derudover kan man anvende “lost message”, som ligeledes illustreres med en sort cirkel. Denne type meddelelse symboliserer, at modtageren er enten ukendt eller ikke relevant, og dermed anvendes en sort cirkel som modtager (Seidl et al., 2015).

### **2.2.3.4 Alternativ interaktion**

I sekvensdiagrammet kan man anvende et alt-fragment til at illustrere alternative sekvenser. Et alt-fragment har minimum 2 forløb, hvor hvert forløb viser en alternativ vej. Hvert forløb

har også en “guard”, som er et boolsk udtryk der illustreres i firkantede parenteser. Guards anvendes i et sekvensdiagram til at udvælge det forløb, der skal udføres (Seidl et al., 2015).

### **3.3 Rich picture**

Rich picture er et modelleringsværktøj som skaber en visuel repræsentation af det undersøgende felt. Formålet med rich picture er at kortlægge og overskueliggøre forskellige elementer i feltet. Grundlæggende indeholder værktøjet følgende tre komponenter; struktur, proces og bekymringer. Disse tre komponenter identificerer interessenter, deres bekymringer og spørgsmål, samt de underliggende strukturer i arbejdspraksissen. Udarbejdelsen af rich picture tager udgangspunkt i den primære bruger. Derefter identificeres strukturen omkring brugeren, brugerens bekymringer samt relationer mellem brugeren og interessenter. Strukturen repræsenterer den primære brugers strukturelle forhold, såsom organisatorisk kontekst, fysisk placering mm. Processen skal identificere og kortlægge interessenter og relationerne imellem. Derudover fremhæves den primære brugers bekymringer, hvilket udformes i spørgsmål, udfordringer og kommentarer (Monk & Howard, 1998).

### **3.4 Rationelle beslutningsmodeller**

Rationelle beslutningsmodeller henviser til, at beslutningstagerne handler fornuftigt og dermed rationelt. Der forekommer 2 typer af rationelle modeller, som er economic man og administrative man (Christiansen et al., 2001).

#### **3.4.1 Economic man**

Economic man henvises som den ideale rationelle beslutningsmodel, hvor man søger den “bedste” løsning. Denne type, er den fuldt ud rationelle beslutningstager, som vælger den bedste løsning ud fra flere forskellige muligheder (Christiansen et al., 2001).

Economic man forudsætter, at man har fuldstændig information og overblik over alle mulighederne, samt detaljeret viden om konsekvenserne. Dette resulterer i at economic mans beslutningsgrundlag er en række muligheder. Economic man kan rangere de forskellige alternative muligheder efter en valgt værdiskala, hvorefter man kan vælge det alternativ der er højeste på værdiskalaen. Ligeledes kan værdiskalaen overføres til fremtidige beslutningssituationer af samme type, således at hvis situationen opstår igen, er valget det samme (Christiansen et al., 2001; Simonsen, 1994).



### **3.4.2 Administrative man**

Den administrative man søger den “bedst mulige” løsning, hvoraf løsningen blot skal være tilfredsstillende. Dermed vælger denne type, den bedst mulige løsning ved at sammenligne med et begrænset antal tilfredsstillende alternativer. Derved, stræber den administrative man efter “det bedst mulige alternativ” kontra “den bedste løsning”. Den administrative mans udgangspunkt er, at rationaliteten afhænger af hvem, der skal løse problemstillingen. Dette betyder at rationaliteten er tilknyttet de beslutningstager, der skal træffe beslutningen. Ved at rationalet er personafhængig, bliver resultatet en forenklet model af virkeligheden, som stammer fra den individuelle beslutningstagers situationsforståelse (Christiansen et al., 2001; Simonsen, 1994).

## **3.5 Mintzbergs organisationsformer**

Mintzberg har blandt andet opstillet en organisatorisk grundmodel, som danner udgangspunkt for hans organisationsformer. Mintzbergs fem organisatoriske former er idealtyper, hvor han henviser til at disse er forestillingsbilleder. Derfor er organisationer i virkeligheden ikke altid struktureret identisk med idealtyperne (Christiansen et al., 2001). Derudover har han udarbejdet 5 koordineringsmekanismer, som uddybes i følgende afsnit.

### **3.5.1 Koordineringsmekanismer**

Mintzberg har udarbejdet 5 forskellige koordineringsmekanismer, hvor disse mekanismer håndtere kompleksiteten i arbejdsopgaverne forskelligt.

De 5 koordineringsmekanismer er:

1. Direkte overvågning: henviser til at en leder giver specifikke ordre og koordinerer dermed deres arbejde.
2. Standardisering af arbejdsprocesser: henviser til at arbejdet koordineres ved indførelse af standarder til udførelse af arbejdet, hvilket forekommer ved regler, fastlagte procedure mm.
3. Standardisering af output: henviser til at arbejdet koordineres efter standardperformance samt specifikationer for arbejdets output.
4. Standardisering af færdigheder: henviser til at arbejdet koordineres gennem individets standard egenskaber og viden. Dette opnås ofte inden individet tiltræder stillingen.

5. Gensidig tilpasning: henviser til at individet koordinere eget arbejde ved at kommunikere med andre ansatte.

(Mintzberg, 1980).

### **3.5.2 Basale dele af en organisation**

Mintzbergs grundmodel danner udgangspunktet for Mintzbergs fem organisatoriske former. Grundmodellen inddrager topledelsen, mellemlidelsen, produktionskernen, teknostrukturen og støttestaben. De fem aspekter indgår forskelligt i de fem organisatoriske former, alt efter hvilket omfang områderne fylder i organisationen (Christiansen et al., 2001).

Topledelsen består af bestyrelse og direktion, hvor der varetages strategiske opgaver.

Koordineringsmekanismerne i topledelsen vil oftest være direkte overvågning gennem ordregivning, konfliktløsning, personalepolitikker osv. Mellemlidelsen består af funktionschefer, værkførere, gruppeledere, formænd osv. og varetager den daglige drift.

Mellemlidelsen anvender koordineringsmekanismerne direkte overvågning af produktionskernen. Produktionskernen består af medarbejdere der udfører selve arbejdet, og dermed kan produktionskernen bestå af specialarbejdere, maskinarbejdere, værkførere, konsulenter osv. Virksomhedens produkter eller tjenesteydelser er virksomhedens output, som bliver fremstillet af produktionskernen. Teknostrukturen består af planlæggere, økonomer, metodeteknikker osv. hvor deres opgaver omfatter at styre virksomhedens taktiske, strategiske og operative niveau. De kan anvende forskellige styringsredskaber som f.eks. planlægningssystemer og økonomistyringssystemer. I teknostrukturen sker koordinering gennem standardisering af arbejdsprocesser, arbejdsresultat og faglig viden. Resterende enhed er støttestaben, som kan bestå af støttefunktioner som f.eks. en reception, lønningskontor eller en kantine. Formålet med støttestaben er, at støtte de forskellige dele af organisationen (Christiansen et al., 2001).

### **3.5.3 Fagbureaukratiet og maskinbureaukratiet**

Fagbureaukratiet karakteriseres ved at der findes mange højtuddannede ansatte i produktionskernen. Dette betyder at højtuddannede ansatte udfører arbejdet hvor organisationens produkter eller tjenesteydelser fremstilles. Koordineringen i fagbureaukratiet sker gennem standardisering af færdigheder. Dermed er produktionskernen det centrale aspekt i idealtypen. Denne idealtypen har en produktion som beror på viden, hvoraf der træffes afgørelse på baggrund af deres færdigheder. Disse færdigheder er opnået ved eksempelvis uddannelse (Christiansen et al., 2001).

Maskinbureaukratiet karakteriseres ved at teknostrukturen er det centrale aspekt i idealtypen. Dette betyder at koordinering af arbejdet tager udgangspunkt i standardisering af arbejdsprocesser, hvor produktionskernen følger regler og systemer. Disse regler og systemer er udarbejdet af planlæggere og økonomer i teknostrukturen (Christiansen et al., 2001).

## 4. Empiri

Anvendelse af LTU er situeret i sagsbehandlernes arbejdspraksis, hvor følgende afsnit vil beskrive forskellige aspekter af denne arbejdspraksis, og dernæst beskrive teknologien LTU. Afsnittet vil først belyse udvalgte lovgivninger, dernæst beskrive sagsbehandlernes IT-system samt deres arbejdsgang hvori LTU indgår. Til forståelsen af sagsbehandlernes arbejdsgang med LTU, tager beskrivelsen udgangspunkt i Lyngby-Taarbæk kommune, da de anvender LTU til at vurdere de nyledige dagpengemodtagere. Ligeledes fremhæves sagsbehandlernes holdninger og bekymringer til LTU's transparens.

Derudover beskrives teknologien LTU med fokus på input, proces og output, hvoraf denne beskrivelse tager udgangspunkt i viden indsamlet gennem interview med Schultz, samt supplerende materiale. Beskrivelsen har til formål at give en overordnet forståelse for LTU. Afsltningsvis beskrives Schultz' journalnotat, hvilket inddrages med henblik på at forstå LTU's nuværende transparensniveau.

### 4.1 Sagsbehandlerens arbejdspraksis og it-systemer

Sagsbehandlernes arbejdspraksis i jobcentre er situeret i beskæftigelsesområdet, som overordnet drejer sig om arbejdsmarkedet i Danmark. Arbejdsmarkedet er styret af politik, lovgivning og forskellige indsatser. Arbejdspraksis er et samlet begreb for alt hvad der vedrører brugerens arbejde, her skal man således anse arbejdet som funktion, opgave og proces (Bødker et al., 2008). Beskæftigelsesindsats har hovedsageligt til formål, at få ledige i arbejde. Disse indsatser kan indeholde aktiveringsforløb som virksomhedsrettede tilbud, uddannelse, opkvalificering mm. (Beskæftigelsesministeriet, n.d.).

Beskæftigelsesindsatserne udføres på henholdsvis nationalt, regionalt og kommunalt niveau. Det nationale niveau varetager det overordnede ansvar for beskæftigelsespolitikken, hvor det regionale niveau understøtter politiske reformer og sikre implementering af disse i jobcentrene. Lokalt udføres beskæftigelsesindsatserne i kommunernes jobcentre, hvor deres primære opgave er at få ledige i henholdsvis arbejde eller uddannelse (Styrelsen for Arbejdsmarked og Rekruttering, n.d.).

#### 4.1.1 Lovgivning

Sagsbehandlerens arbejde er underlagt en række komplekse lovgivninger. Centralt i sagsbehandling, hvori kunstig intelligens indgår, er forvaltningsloven, offentlighedsloven

samt databeskyttelsesloven. I forvaltningsloven indgår blandt andet officialprincippet og begrundelseskravet.

Officialprincippet bestemmer at myndigheden har ansvar for at oplyse om en sag, inden der foretages en afgørelse. Dette sikre og understøtter at afgørelser foretages på baggrund af lovlige beslutninger. I en sag kræves det, at retsgrundlaget og relevante oplysninger indgår. I tilfælde af at teknologi anvendes til hel eller delvis automatiserede afgørelser, skal sagen suppleres med kriterierne for teknologien, samt hvilke juridiske konsekvenser denne har. (Folketingets Ombudsmand, n.d.-c).

Derudover indgår begrundelsekravet i forvaltningsloven, hvor borgeren har ret til en begrundelse af en afgørelse. Begrundelsen skal indeholde en forklaring på afgørelsen samt angive de hensyn og retsregler som afgørelsen har lagt vægt på. Det er dog ikke altid at retsreglerne kan anvendes, hvor man alternativ kan anvende skøn. Skøn henviser til at lovregler kan forekommer ufuldstændige, hvilket resulterer i at forvaltningen vurdere hvilke kriterier der skal være til grund for afgørelsen (Christensen, 2009). Derudover skal begrundelsen indeholde en redegørelse, af oplysninger om sagens omstændigheder. Ifølge forvaltningsskik anbefales det yderligere, at inddrage væsentlige elementer, som har haft betydning for sagens afgørelse. Det er dog ikke ved lov angivet at elementer såsom procesledende beslutninger skal indgå i begrundelsen. Hel eller delvis automatisering kan kategoriseres som procesledende beslutninger, og kan dermed undlades i begrundelsen (Folketingets Ombudsmand, n.d.-a; Folketingets Ombudsmand, n.d.-b).

Under offentlighedsloven forekommer retten til aktindsigt, hvor man som borgere har ret til at blive gjort bekendt med dokumenter, der indgår i den offentlige sagsbehandling.

Dokumenter er både skriftlige, digitale, fotografier mm. (Justitsministeriet, n.d.).

Yderligere er en offentlig sagsbehandling underlagt databeskyttelsesloven, herunder GDPR. GDPR artiklerne 13 og 14 henviser til, at databehandleren har oplysningspligt ved indsamling af persondata (Den Europæiske Unions Tidende, 2016). Disse artikler indgår i sagsbehandlerens arbejde, hvor de skal oplyse borgeren om en eventuel indsamling af data samt formålet med denne. Derudover er artikel 22 “Automatiske individuelle afgørelser, herunder profilering” relevant for sagsbehandlerens arbejde, når dele eller hele arbejdsprocesser er automatiseret. Artikel 22 indebærer retten til ikke at blive underlagt automatisk behandling af personoplysninger, hvor behandlingen har til formål at analysere eller forudsige forhold vedrørende den registrerede. Artiklen er gældende, hvis den automatiserede behandling har en retsvirkning eller betydelig påvirkning på den registrerede.

Den registrerede har dermed altid krav på en menneskelig vurdering (Den Europæiske Unions Tidende, 2016; Datatilsynet, 2018).

#### **4.1.2 Digital sagsbehandling**

Over halvdelen af de danske jobcentre anvender IT-plattformen Fasit, til sagsbehandling. Overordnet er Fasit et fagsystem, som understøtter “samtlige arbejdsgange i jobcentret”, herunder koordinering af borgerens forløb (Schultz, n.d.-a). Førstkommende indeholder Fasit en samlet oversigt over den specifikke borger, herunder kontaktoplysninger, CV, samtaleforløb mm. Fasit er det grundlæggende sagssystem, hvortil Schultz tilbyder en række tillægsapplikationer, herunder applikationen ASTA. ASTA tilbyder et kunstig intelligens værktøj, som skaber overblik over de vigtigste informationer om borgeren, matcher borgere med jobs samt vurderer risikoen for langtidslidighed og korttidslidighed. Til vurdering af risikoen for langtidslidighed, har Schultz udviklet egen algoritme “LTU”, hvor LTU er en forkortelse for Long Term Unemployment. Yderligere har Schultz udviklet algoritmen “STU”, som er en forkortelse for Short Term Unemployment. STU vurderer risikoen for korttidslidighed (Bilag 1: Interview (LTU); Bilag 2: LTU PowerPoint). Derudover faciliterer Fasit muligheder for at vedhæfte journalnotat og bilag i den specifikke sag. Dette resulterer i, at sagsbehandlerens digitale arbejdsproces overholder officialprincippet, begrundelse og GDPR. Ligeledes understøtter Fasit borgernes ret til aktindsigt, da det i Fasit er muligt at generere en rapport over sagen, som sendes til borgeren. (Bilag 3: Borgere i Schultz Fasit).

#### **4.1.3 Sagsbehandlerens arbejdsgang**

Implementeringen af LTU har betydet, at sagsbehandlerne som udgangspunkt skal vurdere de nyledige dagpengemodtagere ud fra LTU’s anbefaling. Dermed er LTU’s vurdering blevet en del af deres arbejdsgang. Dog er det frivilligt om sagsbehandlerne følger vurderingen, hvor deres egen vurdering vægter højere end LTU’s resultat. Sagsbehandlerens egen vurdering foretages ved første møde med borgeren. Hvis tilfældet er, at sagsbehandleren vurderer borgeren anderledes end LTU’s vurdering, skal sagsbehandleren dokumentere dette på borgerens sag, og kan tilknytte borgeren de aktiveringer som sagsbehandleren finder relevante. Dermed kan sagsbehandleren afvige fra LTU’s anbefalede aktiveringsforløb (Flügge et al., 2022).

Ved at sagsbehandleren følger LTU's anbefalinger, kan det påvirke følgende to forhold; tidspunktet for borgerens indkaldelse til samtaler samt den aktivering som den ledige bliver tilbudt af jobcenteret. Det første møde borgeren har som "ledig" med jobcenteret, henvises som en "fælles samtale". Hvornår fælles samtalen afholdes, afhænger af LTU's risikovurdering. Borgere der vurderes til at være i middel eller høj risiko, indkaldes til fælles samtalen efter 4-6 uger, og borgere i lav risiko indkaldes efter 6-8 uger. Hvis borgeren befinder sig i lav risiko, skal borgeren indgå i et online aktiveringsforløb på Candeno, som er en digital platform der hjælper borgeren med at optimere deres cv og jobansøgninger. Middel eller høj risiko vurderede borger skal derimod have et aktiveringsforløb, hvor de tilknyttes en personlig jobformidler og ligeledes tilknyttes Candeno. Fælles samtalen og LTU's vurdering, er udgangspunktet for helhedsvurdering af den ledige borger, som sagsbehandleren skal udarbejde. Kategorierne lav, middel og høj refereres ofte i Lyngby-Taarbæk kommune til farverne grøn, gul eller rød, hvor grøn repræsenterer lav, gul repræsenterer middel og rød repræsenterer høj risiko for langtidsledighed (Flügge et al., 2022).

#### **4.1.4 Sagsbehandlerens holdning til LTU**

Som beskrevet i afsnit 4.1.3 Sagsbehandlerens arbejdsgang, anvendes LTU som beslutningsstøtte til vurdering af borgerens risiko for langtidsledighed. Derfor er det relevant at indsamle empiri om sagsbehandlerens holdning til LTU. Af flere årsager har det ikke været muligt at etablere interviews med relevante jobcentre. Ved første møde med Schultz blev 3 relevante jobcentre identificeret; Lyngby-Taarbæk, Skive og Frederiksberg Kommune. Lyngby-Taarbæk kommune var ikke en mulighed, da Schultz frabad projektgruppen at kontakte dem. Dette skyldes interne konflikter forårsaget af Københavns Universitets forundersøgelse "Er du grøn? Algoritmer til beslutningsstøtte i det offentlige". Derudover ønskede Skive kommune ikke at deltage, da de hel eller delvis var stoppet med at anvende LTU. Ligeledes anvender Frederiksberg Kommune ikke LTU men STU. Vi har derfor valgt at empiri om kundernes holdning til LTU, tager udgangspunkt i den førnævnte forundersøgelsen fra Københavns Universitet. Forundersøgelsen undersøger sagsbehandlerens arbejdspraksis, og deres anvendelse af LTU til vurdering af nylediges risiko for langtidsledighed i Lyngby-Taarbæk Kommune (Bilag 6: Afslag på jobcentre interviews).

Ifølge sagsbehandlerne er et gennemgående problem deres manglende forståelse for algoritmens kategoriseringen af borgerne. Dette fremhæves ved at sagsbehandlerne oplever "at der er sager, hvor de ikke kan forklare, hvorfor borgerne bliver risikovurderet, som de

gør” (Flügge et al., 2022, s. 3). Sagsbehandlerne påpeger at LTU ikke forekommer transparens, da de oplever *“at der er nogle af de 50 variable som ASTA vurderer ud fra, der ikke er gennemsigtige”* (Flügge et al., 2022, s. 3). Variablerne og dets betydning fremstår dermed uklare for sagsbehandlerne. Schultz har forsøgt at imødekomme uklarhederne ved, at udlevere journalnotat til sagsbehandlerne. Journalnotatet har dog foreløbigt ikke resulteret i større transparens for sagsbehandlerne, hvilket fremhæves ved følgende citat; *“SI, at de ikke har fået noget at vide om, hvordan de skal bruge hver enkelt variabel.”* (Flügge et al., 2022, s. 14). Sagsbehandlerne ved dermed ikke hvordan de skal anvende oplysningerne om variablerne i deres arbejdspraksis. Derudover udtrykkes det af henholdsvis sagsbehandlerne og ledelsen, at der ikke forekommer dokumentation eller forklaring på hvor præcis LTU er, hvilket udtrykkes i følgende citat; *“Den kommer ikke med nogen forklaring af, hvor god den er til at ramme rigtigt”* (Flügge et al., 2022, s. 17).

Sagsbehandlerne udtrykker derved, at de først og fremmest ikke forstår hvorfor borgerne kategoriseres som de gør. Derudover forekommer algoritmens variabler uklare, hvortil variabernes relation til hinanden og vægtningen ikke er transparens. Aflsuttetnignisvis påpeges det, at der ikke foreligger en beskrivelse af hvor god algoritmen er.

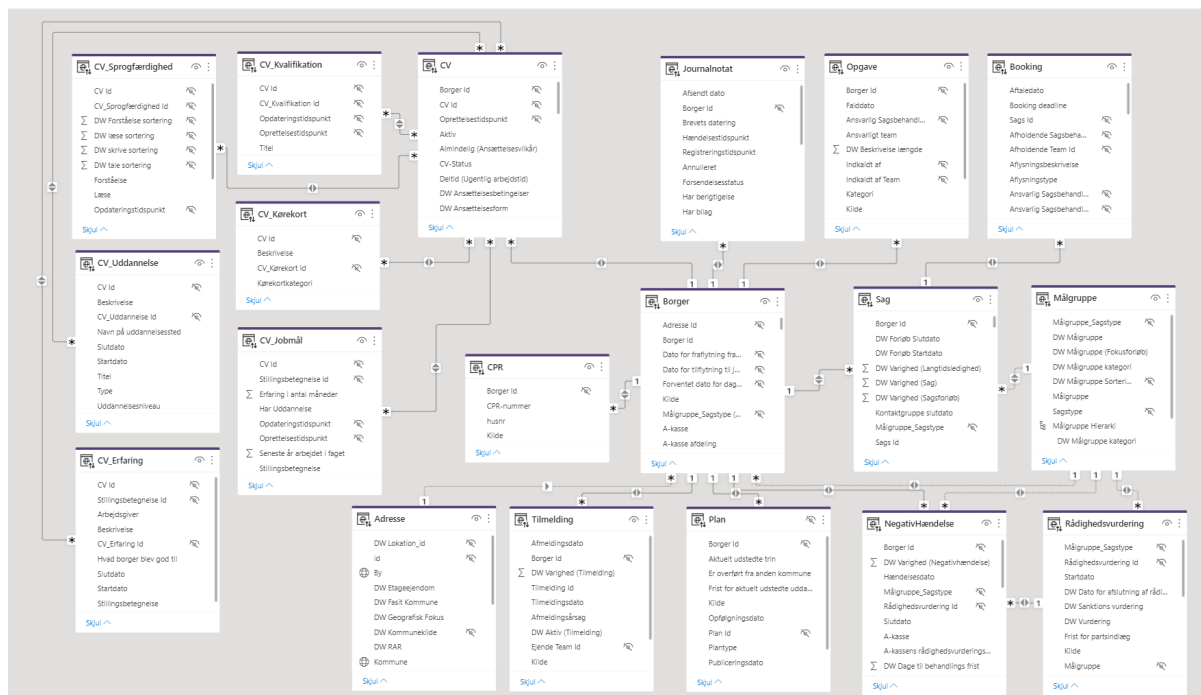
## 4.2 LTU

Projektet om at skabe LTU startede i 2017, og udspringer af en idé om at kunne identificere borgere som er i risiko for langtidsledighed, og dermed gøre sagsbehandlingen nemmere. I 2019 blev første version af LTU (1.0) færdig, men fremstod ikke transparens, hvor hverken Schultz eller sagsbehandlere kunne redegøre for dens resultater. Dette betød, at sagsbehandleren udelukkende fik en risikoscore som var enten lav, middel eller høj, uden en begrundelse for, hvordan dette var udregnet. Sagsbehandlerne efterspurgte en begrundelse for borgerens kategorisering, hvilket Schultz ikke kunne redegøre for. Dette betød at både sagsbehandlere og Schultz ønskede større transparens af LTU, og resulterede i at Schultz udarbejdede en ny algoritme i 2022. Den nye algoritme (LTU 2.0) indebærer at sagsbehandleren kan modtage et journalnotat, hvor variablerne som ligger til grund for resultatet indgår. Ligeledes indeholder journalnotatet en oversigt over variablenes vægtning (Bilag 1: Interview (LTU); Bilag 2: LTU PowerPoint; Bilag 4: Journalnotat).



## 4.2.1 Input

LTU algoritmens input er data om borgeren. Dataen hentes fra Schultz' egen database, hvoraf dataen indsamles fra Fasit og andre offentlige tilgængelige datakilder (Bilag 2: LTU PowerPoint). Nedenfor ses Schultz' database repræsenteret i et databasediagram.



Figur 4: Schultz' database diagram

Databasediagrammet viser at LTU har adgang til en række forskellig data om borgeren hvilket eksempelvis er deres CV, tidligere sager, uddannelse, joblog, demografi mm. (Bilag 5: Mail).

LTU algoritmen anvender ikke al data der forekommer om en borger, men kan derimod have maksimalt 50 variabler som input. Variabler henviser til datapunkter om borgeren, eksempelvis borgerens alder. Antallet af variabler bestemmes efter hver gentræning, hvilket betyder at antallet af variabler, som input til algoritmen, varierer. Interviewpersonen fremhæver at de 50 variabler som algoritmen maksimalt kan inddrage, er fast udvalgte datapunkter. Hver gentræningen af algoritmen udvælger de variabler som indgår i algoritmen ud fra disse 50 fastlagte variabler. De udvalgte 50 variabler er overført en-til-en fra den gamle algoritme (LTU 1.0), hvortil Schultz ikke har foretaget yderligere vurdering om variablernes relevans. Vurdering af hvilke variabler der er relevante at inddrage, foretages ud fra statistiske beregninger, hvor modellen fravælger variabler som ikke er signifikante (Bilag 2: LTU PowerPoint; Bilag 4: Journalnotat).

Statistisk signifikant kan anvendes til at beregne sandsynligheden for sammenhæng mellem to eller flere variabler. Hvis variablen er signifikant, er der statistisk belæg for sammenhæng mellem variablerne. Det er dermed kun signifikante variabler der indgår i vurdering af risiko for langtidsledighed (Torgersen & Kvittingen, 2019; Bilag 5: Mail)

Den seneste gentræning har vurderet, at ca. 30 ud af de 50 variabler er statistisk signifikante og dermed relevante at inddrage i LTU. Schultz anvender og fravælger variabler udelukkende fra et statistisk perspektiv og forholder sig dermed ikke til etiske aspekter. (Bilag 1: Interview (LTU)).

Variablerne som anvendes i LTU ses i journalnotatet, hvor der ved seneste gentræning eksempelvis inddrages følgende data; borgerens alder, statsborgerskab, adresse, antal tidligere forløb og antal tidligere jobs (Bilag 4: Journalnotat).

#### 4.2.2 Proces (Matematiske beregninger)

LTU algoritmen tager udgangspunkt i følgende matematiske områder; sandsynlighed og statistik, hvor LTU anvender den matematiske metode, logistisk regression.

Regressionsanalyse er en matematisk formel, for at bestemme forholdet mellem afhængige og uafhængige variabler. En uafhængig variabel, er den variabel som påvirker en anden variabel, altså den afhængige variabel. Regressionsanalyse anvendes til forudsigelser og prognoser, hvor logistiske regression konverterer en eller flere variabler til en sandsynlighed (Javatpoint, n.d.). Algoritmen LTU har til formål, at analysere og vurdere forudsigelser ud fra flere variabler som er numeriske, kategoriske eller begge dele. Numeriske variabler har en talværdi, hvilket eksempelvis kan være antal år. Kategoriske variabler indebærer to eller flere kategorier, hvilket eksempelvis kan være køn hvor flere kategorier indgår (Aarhus Universitet, n.d.; Bilag 1: Interview (LTU)).

Sandsynligheden for LTU er givet ved funktionen:

$P = f(w_0 + w_1 * x_1 + w_2 * x_2 + \dots + w_p * x_p)$	
P	er ltu-sandsynlighed
F	er sigmoid-funktionen
w_0	er intercepted
w_1, ..., w_p	er feature vægte

$x_1, \dots, x_p$	er normaliserede feature værdier
-------------------	----------------------------------

Figur 5: Formel for LTU algoritme

I formlen for LTU indgår sigmoid funktionen, som er blevet populær indenfor kunstig intelligens. Sigmoid funktionen anvendes i sammenhæng med logistisk regression, hvor sigmoid funktionen sikre, at resultatet for sandsynligheden er mellem 0-1 (Wood, n.d.; Saeed, 2021).

#### 4.2.2.1 Normalisering

Normalisering anvendes til forskellige formål, og har derved forskellige betydninger. Indenfor statistik og sandsynlighed er normalisering transformation af variabelens værdi, hvor man omformer værdien til en ønskelige numerisk form. Dette muliggør at man kan sammenligne forskellige måleenheder, da man omformer værdien til samme numeriske form. I LTU algoritmen, anvendes gennemsnit og spredning til at normalisere variabelens observation ved følgende formel:  $\frac{\text{Observation} - \text{Gennemsnit}}{\text{Spredning}}$  (Bilag 1; Interview (LTU); Bilag 5: Mail).

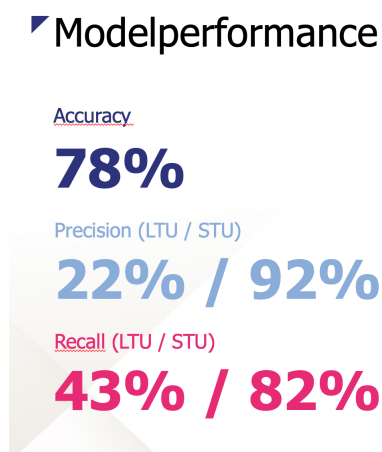
Observationer normaliseres således, at variabler med forskellige måleenheder kan sammenlignes f.eks variablerne alder og køn. Variabler som ikke er numeriske skal forinden normaliseringen omformes til en numerisk værdi. Dette udføres med metoden “one hot encoding” (Bilag 5: Mail). One hot encoding anvendes til at transformere ikke numerisk data, til numerisk data. Dette ses eksempelvis ved variabelen køn, som kan indeholde 2 til flere kategorier, og disse kategorier får hver tildelt en numerisk værdi. Metoden er særlig fordelagtigt i kunstig intelligens, da den numeriske værdi resulterer i, at algoritmen kan foretage mere nuancerede forudsigelser (Vasudev, 2017; Educative, 2021; Bilag 5: Mail).

#### 4.2.3 Output

Resultatet af LTU algoritmen er en vurdering af borgerens risiko for langtidsledighed, og vises i form af en risikoscore som er et tal mellem 0 og 1. Schultz har kategoriseret risikoscoren i henholdsvis lav, middel og høj. Outputtet af algoritmen er lav mellem 0,00-0,06, middel mellem 0,06-0,54 og høj mellem 0,54-1,00. Resultatet for LTU algoritmen indgår i journalnotatet, som Schultz producerer, sammen med de variabler der ligger til grund for beregningen, og variabernes tilhørende vægte (Bilag 4: Journalnotat).

### 4.2.3.1 Accuracy, precision og recall

Schultz anvender standardmålene accuracy, precision og recall til at måle kvaliteten af LTU og STU til klassifikation (Bilag 2: LTU PowerPoint). Resultatet af standardmålene illustreres i nedenstående figur.



Figur 6: Modelperformance (Bilag 2: LTU PowerPoint)

Accuracy henviser til hvor mange gange modellen samlet set var korrekt. Den fortæller dermed LTU og STU's samlede procentvis korrekte klassificerede elementer.

Accuracy er 78%, hvilket betyder at ud af f.eks. 100 forudsigelser, har modellen korrekt klassificerede 78 gange. Precision for LTU er 22% hvor STU er 92%. Precision henviser til hvor mange af borgerne som LTU og STU forudsagde til at være i risikokategorierne lav, middel og høj, rent faktisk er i disse kategorier. Recall for LTU er 43% hvor STU er 82%. Recall henviser til forholdet mellem det korrekt identificerede og det faktuelle tal. Dermed hvor god modellen er til at finde borgere i kategorierne lav, middel og høj ift. det faktuelle tal (Goodfellow et al., 2016; H. Christiansen , 2021; Bilag 2: LTU Powerpoint).

### 4.3 Journalnotat

På baggrund af sagsbehandlerens efterspørgsel på større transparens af LTU, har Schultz produceret et journalnotat. Dette journalnotat kan fremsendes på efterspørgsel af sagsbehandleren, hvor journalnotatet tager udgangspunkt i den specifikke borger det vedrører. I journalnotatet indgår borgerens risikoscore for langtidsledighed, datoen for selve udregningen, grundlaget for vurderingen og beregningen. Grundlaget for vurderingen og beregningen er inddelt i to separate tabeller (Bilag 1: Interview (LTU); Bilag 4: Journalnotat).

Tabellen der beskriver grundlaget for vurderingen i journalnotatet, kan ses i nedenstående figur 7: Journalnotat - Grundlag for vurdering.

Grundlag for vurderingen.

Nedenstående parametre ligger til grund for beregningen:

featurenavn	beskrivelse	observation	enhed
AllowEmailNotification	Tillader borger, at jobcenteret sender dem emails?	False	Boolsk
CancelledInterviewsCurrentCount	Hvor mange samtaler har borger aflyst?	0	Antal
PreviousSLUCasesCount	Hvor mange langtidsledige forløb har borger tidligere haft?	1	Antal
PlacementJobQualificationCurrentCount	Hvor mange vejlednings- og opkvalificeringsplaceringer har borger haft	0	Antal
InterviewTotalCurrentCount	Hvor mange samtaler er borger blevet indkaldt til i alt?	1	Antal
AgeAtCaseStart	Hvad var borgers alder, da borgers forløb startede?	33	År
AllowTextNotification	Tillader borger, at jobcenteret sender dem sms'er?	False	Boolsk
IsNordicCitizen	Er borger en nordisk borger?	True	Boolsk
HealthIssuesCurrentCount	Hvor mange helbredsudfordringer har borger p.t. registreret i FASIT?	0	Antal
PreviousCasesDistinctMunicipalityCount	I hvor mange kommuner har borger tidligere haft sager?	1	Antal
Gender	Borgers køn.	Mand	Kategorisk
PreviousCasesCount	Hvor mange tidligere sager har borger haft?	1	Antal
Longitude	Hvad er længdegrad for borgers bopæl?	12.53327199	Grader
LivesInApartment	Bor borger i lejlighed?	True	Boolsk
PermanentJobPosition	Er borger åben over for permanent ansættelse?	False	Boolsk
PartTimeEmployment	Er borger åben over for deltidsansættelse?	True	Kategorisk
TemporaryJobPosition	Er borger åben over for midlertidig ansættelse?	True	Boolsk
StartAsAgreedOn	Er borger villig til at starte efter aftale?	True	Boolsk
StartAsSoonAsPossible	Er borger villig til at starte hurtigst muligt?	True	Boolsk
JobGoalHasEducation	Har borger en uddannelse, der matcher mindst ét jobmål?	False	Boolsk
AvgScheduledInterviewStartTime	Hvilket tidspunkt er borgers samtaler gennemsnitligt sat til?	.Andet	Kategorisk
DaysBeforeFirstInterview	Hvor mange dage gik der, før borger var til første samtale med jobcenteret?	5	Dage
OtherExperienceCount	Hvor mange andre ikke-job erfaringer har borger?	0	Antal
PreviousEnrolmentClosure	Hvis borger har været tilmeldt før, hvorfor blev de så senest afmeldt?	.Andet	Kategorisk
FirstInterviewCategory	Hvilken slags samtale blev borger først indkaldt til?	CV samtale	Kategorisk
LatestSustenanceDomain	Hvis borger har været på offentlig forsørgelse, hvad er det seneste domene før borgers nuværende forløb?	Aktivering, kontanthjælp	Kategorisk
LanguagesSpokenCount	Hvor mange sprog taler borger?	3	Antal
TotalExperienceCount	Hvor mange job erfaringer har borger?	9	Antal
JobgoalsCount	Hvor mange jobmål har borger?	1	Antal
JoblogUpdatesCount	Hvor mange gange har borger opdateret sin joblog?	0	Antal
JobSearchGeoRegionsCount	Hvor mange geografiske regioner har borger søgt jobs indenfor?	0	Antal
JobSearchSourceHeadhuntCount	Hvor mange jobmuligheder er borger blevet headhunnet til?	0	Antal

Figur 7: Journalnotat - Grundlag for vurdering

På figur 7: Journalnotat - Grundlag for vurdering fremgår 32 forskellige featurenavne, tilhørende beskrivelser, observationer og enheder. Featurenavne repræsenterer variabelens navn. Et eksempel fra journalnotatet er featurenavnet “AgeAtCaseStart”, med beskrivelsen “Hvad var borgers alder, da borgers forløb startede”, observationen “33” og enheden “år”. Journalnotatets beregning hvori vægtninger, gennemsnit og spredning indgår kan ses i nedenstående figur 8: Journalnotat - Beregning.

Til beregning af sandsynlighed er benyttet \*logistisk regression\* med følgende vægte:

featurenavn	kategori	vægt	gennemsnit	spredning
Intercept		-10.402	nan	nan
AgeAtCaseStart		0.141	39.054	12.468
AllowEmailNotification		0.013	nan	nan
AllowTextNotification		0.317	nan	nan
AvgScheduledInterviewStartTime	Eftermiddag	0.552	nan	nan
AvgScheduledInterviewStartTime	Formiddag	0.648	nan	nan
CancelledInterviewsCurrentCount		-0.265	0.292	0.538
DaysBeforeFirstInterview		0.382	5.716	8.94
FirstInterviewCategory	Anden samtale	-0.443	nan	nan
FirstInterviewCategory	CV samtale	-3.003	nan	nan
FirstInterviewCategory	Fælles Jobsamtale med dagpengemodtager	-0.551	nan	nan
FirstInterviewCategory	Informationsmøde	-0.539	nan	nan
FirstInterviewCategory	Jobsamtale	-0.38	nan	nan
FirstInterviewCategory	Rådighedssamtale	-2.468	nan	nan
Gender	Mand	0.008	nan	nan
HealthIssuesCurrentCount		0.092	0.03	0.176
InterviewTotalCurrentCount		-0.179	0.82	1.33
IsNordicCitizen		-0.36	nan	nan
JobSearchGeoRegionsCount		0.124	0.278	0.737
JobSearchSourceHeadhuntCount		0.064	0.174	1.861
JobgoalHasEducation		0.008	nan	nan
JobgoalsCount		0.164	2.789	2.092
JoblogUpdatesCount		-0.107	7.533	25.542
LanguagesSpokenCount		0.075	2.823	1.213
LatestSustenanceDomain	A-dagpenge mv.	2.568	nan	nan
LatestSustenanceDomain	Aktivering, kontanthjælp	-0.777	nan	nan
LatestSustenanceDomain	Aktivering, sygedagpenge m.fl.	-1.161	nan	nan
LatestSustenanceDomain	Fleksjob og øvrige ordninger	-0.834	nan	nan
LatestSustenanceDomain	Kontanthjælp	-0.122	nan	nan
LatestSustenanceDomain	Ressourceforløbsydelse	-1.176	nan	nan
LatestSustenanceDomain	Revalidering, ledighedsydelse mv.	-0.587	nan	nan
LatestSustenanceDomain	Syge og barselsdagpenge	-1.302	nan	nan
LatestSustenanceDomain	Uddannelseshjælp	-1.322	nan	nan
LivesInApartment		0.125	nan	nan
Longitude		0.134	10.95	1.382
OtherExperienceCount		0.149	0.351	0.9
PartTimeEmployment		2.462	nan	nan
PermanentJobPosition		-0.292	nan	nan
PlacementOpkvalificeringCurrentCount		0.104	0.024	0.157
PreviousCaseDistinctMunicipalityCount		-0.25	0.712	0.496
PreviousCasesCount		-0.493	2.304	3.71
PreviousEnrollmentClosure	Afmeldes grundet ikke jobklar/jobparat/åbenlyst udd. parat	1.901	nan	nan
PreviousEnrollmentClosure	E 303 (Jobsøgning i udlandet)	0.556	nan	nan
PreviousEnrollmentClosure	Sygdom	0.103	nan	nan
PreviousEnrollmentClosure	Ændret ydelseskategori	0.302	nan	nan
PreviousLtuCasesCount		0.274	0.709	1.105
StartAsAgreedOn		1.061	nan	nan
StartAsSoonAsPossible		3.568	nan	nan
TemporaryJobPosition		1.272	nan	nan
TotalExperienceCount		-0.035	4.897	4.092

hvor vægt er parametervægtene i regressionsmodellen, og gennemsnit samt spredning bruges til at normalisere observationerne.

Figur 8: Journalnotat - Beregning

Figur 8: Journalnotat - Beregning, viser 50 featurenavne. For hver af disse featurenavne vises en eventuel kategori, featurenavnets vægt, gennemsnit og spredning. Et eksempel er featurenavn "AgeAtCaseStart" som ikke har en kategori. Denne variabel har en vægt på "0,141", hvor gennemsnittet er "39,054" og spredningen er "12.468".

Under gennemsnit og spredning forekommer der "nan" som er en forkortelse for "Not a Number", hvilket betyder at observationen for den pågældende featurenavn har en kategorisk værdi. Det er i disse tilfælde at Schultz anvender one hot encoding til at normalisere observationen (Bilag 5: Mail).

#### 4.4 Delkonklusion

Ved anvendelse af hel eller delvis automatiseret sagsbehandling, skal sagsbehandleren være særlig opmærksom på de lovgivninger de er underlagt. Loven bestemmer, at sagsbehandleren skal kunne begrunde og redegøre for afgørelsen. Anvendelse af LTU tilføjer problematikken, hvori sagsbehandleren ikke kan redegøre for dets proces og resultatet. Dertil har borgeren,

ifølge GDPR, ret til ikke at blive underlagt hel eller delvis automatiserede behandling, i tilfælde af at behandlingen har en retsvirkning eller betydelig påvirkning på borgeren. Derudover bestemmer LTU som udgangspunkt borgerens forløb i Lyngby-Taarbæk Kommune. Sagsbehandleren kan ved første møde med borgeren tilsidesætte dette forløb, hvis sagsbehandleren vurderer borgeren anderledes end LTU. Dermed vægter sagsbehandlerens vurdering højere end LTU's vurdering. Yderligere udtrykker sagsbehandlerne manglende forståelse for LTU's input, proces og output.

Algoritmen LTU har til formål at vurdere borgers risiko for langtidsledighed. Algoritmens input er data om borgeren, hvor den kan behandle maksimalt 50 variabler. Dataen er hovedsageligt indhentet fra fagsystemet Fasit og borgernes CV'er. Dernæst foretager algoritmen forskellige matematisk beregninger, hvori variabler, vægtning, gennemsnit og spredning indgår. Beregningen har til formål at forudsige borgerens langtidsledighed, hvortil outputtet er et numerisk tal mellem 0 og 1, der repræsenterer en risikoscore. Risikoscoren kategoriseres som lav, middel eller høj, hvor hver kategori repræsenterer et interval mellem 0 og 1.

Schultz har udarbejdet journalnotat, som redegørelse for LTU's vurdering af borgerens risiko for langtidsledighed. Journalnotatet er skabt med det formål, at skabe større transparens af LTU, hvortil journalnotatet sendes på forespørgsel fra sagsbehandleren. Sagsbehandleren udtrykker, at journalnotatet ikke skaber større forståelse for LTU.

## 5. Analyse

Følgende analyse har til formål at skabe større transparens af LTU ved modellering, og beror på præmissen om at transparens skal øge brugernes forståelse for den pågældende teknologi. Schultz har øget niveauet af transparens for LTU, ved at optimere algoritmen således, at Schultz kan redegøre for dens input, proces og output. Ligeledes har Schultz udviklet et journalnotat med henblik på at skabe transparens af LTU til sagsbehandlerne. Dog ses det at sagsbehandlerne i Lyngby-Taarbæk kommune fortsat mangler forståelse for LTU, på trods af journalnotatet. De udtrykker, at de ikke ved hvordan journalnotatet konstruktivt skal indgå i deres arbejde, og at de ikke ud fra journalnotatet kan forklare hvorfor borgerne bliver risikovurderet som de gør. Vi kan ud fra empirien om LTU udlede, at det kan være fordelagtigt for sagsbehandlerne at se hvilke variabler LTU anvender, og hvordan risikoscoren for disse beregnes. For at skabe transparens af beregningen inddrages forholdet mellem vægtning og observation.

I analysen anvendes en domænemodel som er udarbejdet med notationer fra UML klassediagram, med det formål at afdække relevante begreber og deres relation til hinanden. Derudover anvendes UML diagrammerne; use case diagram og sekvensdiagram. Use case diagrammet dokumenterer funktioner sagsbehandleren kan anvende i systemet, og sekvensdiagrammet dokumenterer interaktioner mellem objekter.

Efterfølgende anvendes modelleringsværktøjet rich picture, som har til formål at skabe et visuelt overblik af sagsbehandlernes arbejdspraksis, hvori LTU indgår. Dette suppleres med egne udarbejdede modeller, hvor hver model uddyber et element i rich picture modellen. LTU er en black box, hvor analysens modeller har til formål at åbne denne black box, og dermed skabe “usable transparency”. Modellerne fokuserer på at skabe transparens af LTU’s input, proces og output til den almene borger. Dermed anvendes modellerne ligeledes til formidling af transparens.

### 5.1 Begreber

Ved udarbejdelsen af denne rapport, er det tydeligt at LTU er et komplekst system, hvor der anvendes en række forskellige begreber, særligt it-faglige begreber. Dette kombineres med sagsbehandlernes fagområde, hvor der også forekommer faglige begreber. Vi har derfor fundet det relevant at udarbejde en model som beskriver centrale begreber. Gennem begrebsafklaring sikres, at brugerne forstår de begreber som anvendes i modellerne.



Ligeledes medfører det en bedre forudsætning for at skabe transparens, da brugerne kender til begreberne og deres betydning inden modellerne fortolkes.

Begreberne og tilhørende forklaringer ses i nedenstående figur.

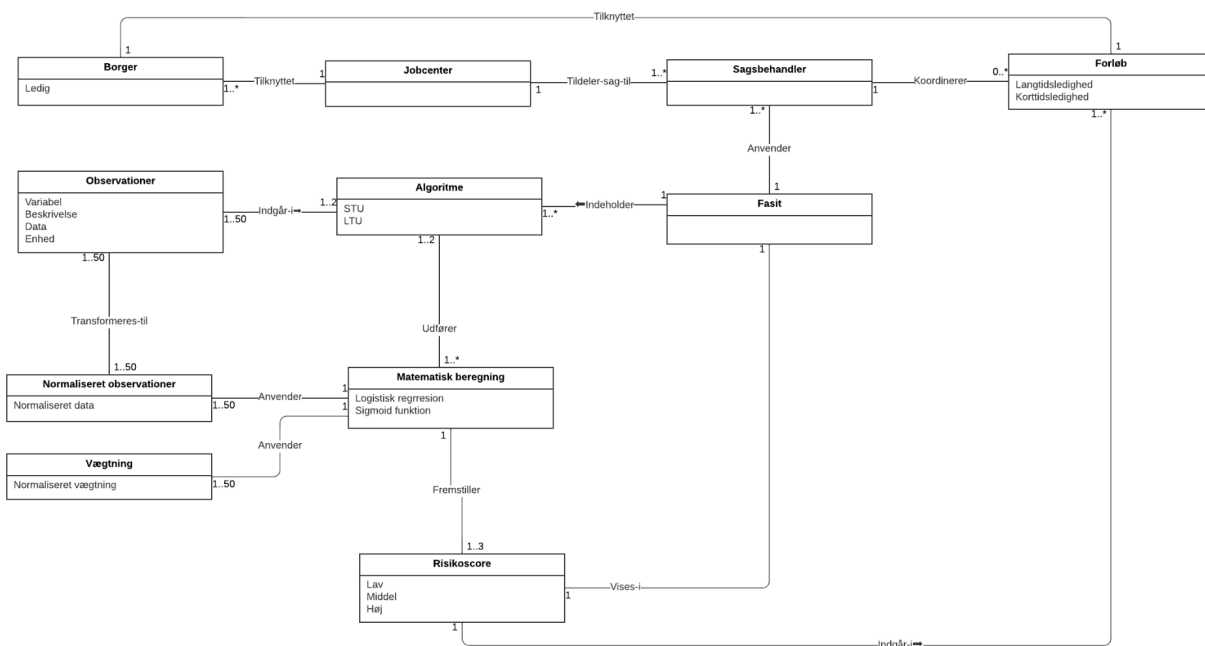
Begreber	Forklaring
Variabel	Navnet på de datapunkter som ligger til grund for beregningen. Variabler er eksempelvis AgeAtCaseStart (alder) og HealthIssuesCurrentCount (helbredsudfordringer)
Observation	Variablens indhold. Variablen AgeAtCaseStart (alder) indeholder antal år eksempelvis 40 år.
Vægtning	Hvor meget variablen betyder for risikoscoren. Vægten er statistisk beregnet ud fra borgernes CV og data fra Fasit
Normaliser	Omregner forskellige måleenheder til den samme, således at observationerne kan sammenlignes
Sigmoid funktion	Matematisk formel som er særlig fordelagtig til at beregne forudsigelser. Giver en risikoscore mellem 0 og 1
Rød pil op ↑	Repræsenterer at observationens risikoscore påvirker den samlede risikoscore negativt. Pilen har retning op, da det pågældende tals fortegn er plus og derved forøger den samlede risikoscore.
Grøn pil ned ↓	Repræsenterer at observationens risikoscore påvirker den samlede risikoscore positivt. Pilen har retning ned, da det pågældende tals fortegn er minus og derved formindsker den samlede risikoscore.

Figur 9: Begrebsafklaring

Figurens første kolonne indeholder følgende begreber; variabel, observation, vægtning, normaliser og sigmoid funktion. Anden kolonne indeholder korte forklaringer, med formålet om at skabe en overordnet forståelse.

## 5.2 Domænemodel

Domænemodellen er udarbejdet med henblik på at skabe et overblik over konceptuelle objekter i domænet. Domænet henviser til sagsbehandlernes arbejdsproces hvori LTU indgår.



Figur 10: Domænemodel

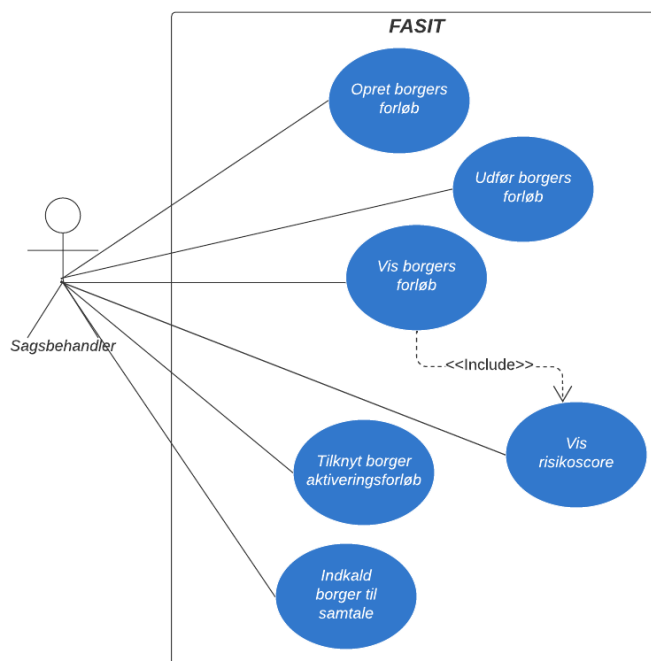
Ovenstående figur 10: Domænemodel illustrerer, konceptuelle objekter som repræsenteres i boksene. Relationen mellem de konceptuelle objekter forbindes med associationer hvor hver association har et associationsnavn. Derudover indeholder modellen multiplicitetsværdier som angiver det numeriske forhold mellem de konceptuelle objekter. Modellen læses fra venstre mod højre eller fra top til bund. Hvis dette ikke er tilfældet illustreres læseretning med en læseretningspil.

Domænemodellen illustrerer at en ledig borger tilknyttes et jobcenter, hvortil et jobcenter kan have en eller flere borgere tilknyttet, men en borger kun kan være tilknyttet et jobcenter. I jobcenteret arbejder sagsbehandlere, som får tildelt borgernes sag. Den ledige borger får tilknyttet et forløb, som sagsbehandleren koordinerer. Som en central del af borgerens forløb, indgår en vurdering af borgerens langtidsledighed og korttidsledighed. Sagsbehandleren anvender fagssystemet Fasit til koordinering og planlægning af borgerens forløb, hvor algoritmen LTU og STU forekommer. Algoritmen indeholder observationer som input, en matematisk beregning som process samt en risikoscore som output. I observationerne indgår et variabelnavn, en beskrivelse, dataen og enheden. Observationerne transformeres til normaliserede observationer, hvor disse normaliserede observationer og vægtningen indgår i den matematiske beregning. Den matematiske beregning kan indeholde op til 50 variabler. Den matematiske beregning fremstiller et output som er en risikoscore. Risikoscoren kan være enten lav, middel eller høj, og vises i Fasit. Ved at risikoscoren vises i Fasit, indgår den dermed i sagsbehandlerens koordinering af borgerens forløb.

Domænemodellen anvendes til at identificere relevante elementer, samt forholdet mellem disse. Modellen har et højt abstraktionsniveau, hvilket resulterer i at modellen skaber en overordnet forståelse for emnet. I forlængelse af højt abstraktionsniveau indeholder modellen få detaljer, med det formål at formindske kompleksiteten. Således at brugeren skaber en overordnet forståelse for sagsbehandlerens arbejdspraksis hvori LTU indgår.

### 5.3 Use case diagram

Use case diagrammet dokumenterer de funktioner en aktør kan anvende i systemet. Use case diagrammet illustrerer en sagsbehandlerens interaktion med Fasit. Fasit indeholder flere funktioner end illustreret, men disse er udeladt da fokus er sagsbehandlerens koordinering af en ledig borgers forløb.



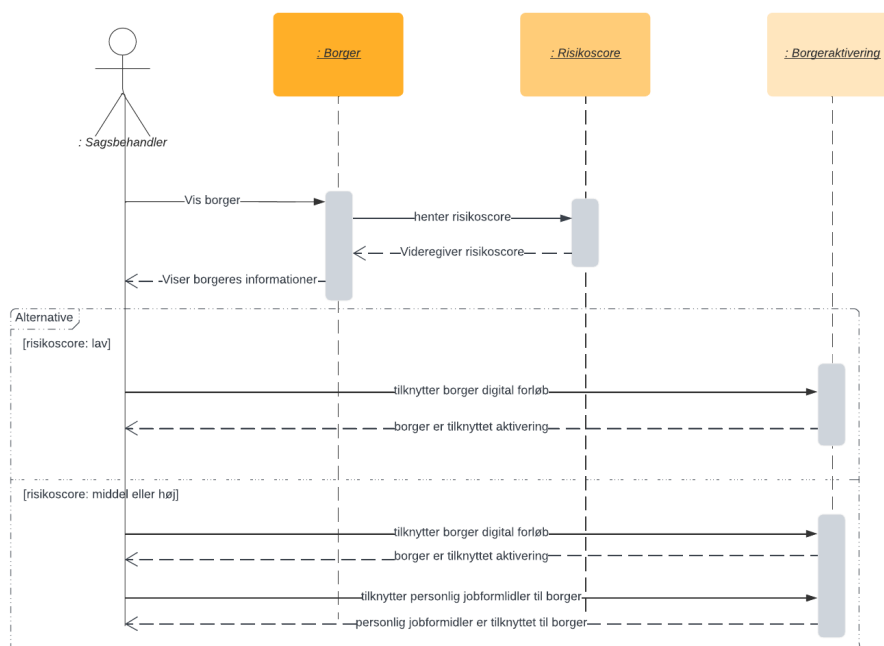
Figur 11: Use case diagram

Figur 11: Use case diagram indeholder aktøren sagsbehandleren, og 6 forskellige use cases i systemet Fasit. Use cases repræsenterer funktioner som kan anvendes i systemet. Use casene forbindes med aktøren gennem associationer. Use casene anvendes uafhængigt af hinanden medmindre andet er angivet. I modellen fremgår et include forhold, mellem use casene “Vis borgers forløb” og “Vis risikoscore”. “Vis borgers forløb” fremgår som base use case og “Vis risikoscore” er included use case. Disse to use cases har et include forhold, da risikoscoren automatisk vises i borgerens forløb.

Use case diagrammet illustrerer en forenklet repræsentation af aktørens interaktion med Fasisit. Forenklingen er et resultat af, at diagrammet udelukkende fokuserer på sagsbehandlerens koordinering af en ledig borgers forløb, hvortil andre funktioner i Fasisit udelades. Yderligere, indeholder modellen udelukkende detaljer, som er nødvendige for at illustrerer fokusområdet. Modellen er dermed en abstrakt repræsentation af aktørens interaktion med Fasisit.

## 5.4 Sekvensdiagram

Projektgruppen har udarbejdet to forskellige sekvensdiagrammer. Det første sekvensdiagram illustrerer use casen “Tilknyt borger aktiveringsforløb” fra ovenstående use case diagram. Det andet sekvensdiagram illustrerer LTU’s input, proces og output. Nedenstående figur illustrerer det første sekvens diagram.



Figur 12: Sekvensdiagram “Tilknyt borger aktiveringsforløb”

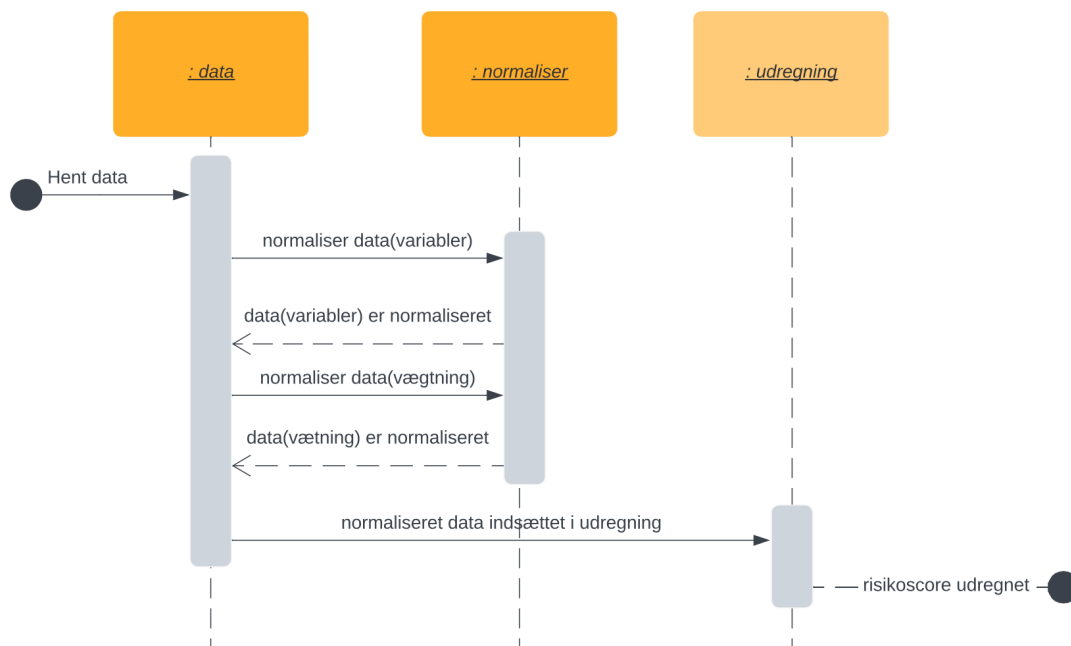
Ovenstående figur 12: Sekvensdiagram “Tilknyt borger aktiveringsforløb” illustrerer interaktionen mellem de forskellige interaktionspartnere, hvor interaktionen fremgår som meddelelser.

Interaktionspartnerne er følgende; sagsbehandler, borger, risikoscore og borgeraktivering. Sagsbehandleren repræsenterer et menneskeligt objekt, som starter diagrammets sekvens af meddelelser. Den menneskelige interaktionspartner illustreres med symbolet tændstikfigur, hvor de resterende interaktionspartnere er ikke-menneskelige, og illustreres med firkantede

bokse. Pilene i diagrammet repræsenterer meddelelser mellem interaktionspartnerne og forekommer i kronologisk rækkefølge.

I empirien fremgår det, at en borger tilknyttes forskellige aktiveringsforløb, afhængig af borgerens LTU risikoscore. Dette illustreres ved anvendelse af alt-fragment, hvori to guards indgår. Dette viser at der kan forekomme to forskellige sekvenser af meddelelser. I alt-fragment boksen illustrerer den første guard, at hvis borgeren har en lav risikoscore, tilknyttes borgeren et digitalt forløb. Den anden guard illustrerer, at hvis borgeren kategoriseres i henholdsvis middel eller høj risiko, tilknyttes borgeren både digital forløb og en personlig jobformidler.

Nedenstående figur 13: Sekvensdiagram LTU illustrerer LTU's input, proces og output.



Figur 13: Sekvensdiagram LTU

Ovenstående figur 13: Sekvensdiagram LTU indeholder følgende interaktionspartnere; data, normaliser og udregning. Sekvensen af meddelelser starter og afsluttes med en found og lost message, som repræsenteres ved sorte cirkler. Dette betyder at afsenderen eller modtageren er ukendt eller vurderes ikke relevant. I denne model henviser begrebet data til både variabler og dets vægtning. LTU's proces starter ved at udvalgt data hentes om den specifikke borger. Derefter normaliseres variablernes værdi, hvilket sendes tilbage til data objektet. Hvorefter vægtning normaliseres, hvilket ligeledes sendes tilbage. Derefter sendes den samlede

normaliserede data til objektet udregning, hvilket afsluttes i at risikoscoren udregnes og sendes til en ukendt eller ikke relevant modtager.

Sekvensdiagrammerne fokuserer overordnet på sekvenser af meddelelser. De to sekvensdiagrammer tilføjer detaljer til to specifikke perspektiver. Første sekvensdiagram tager udgangspunkt i en abstrakt use case, hvor sekvensdiagrammet øger detaljegraden og dermed kompleksiteten. Det andet sekvensdiagram forsøger gennem modellen, at åbne black boxen ved trinvis at beskrive LTU's proces. Da fokus er LTU's proces, anvendes lost og found message. I denne repræsentation anvendes found message da afsenderen af dataen er irrelevant, ligeledes anvende lost message da modtageren af risikoscoren også forekommer irrelevant.

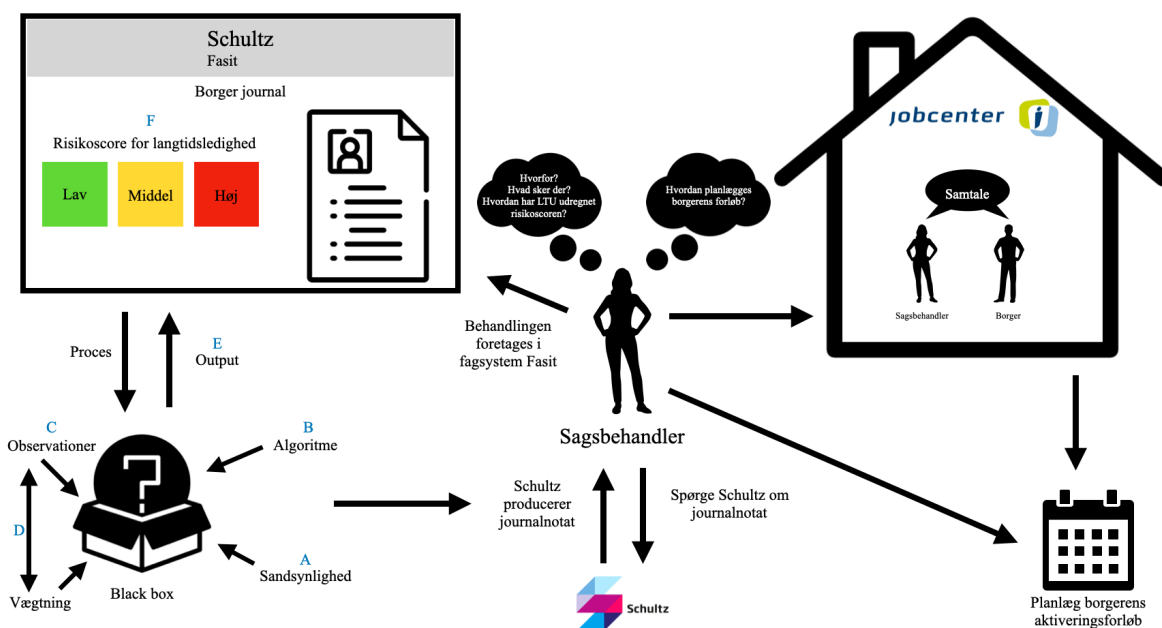
Begge sekvensdiagrammer nedbryder det komplekse system, og er to forskellige repræsentationer af LTU. Detaljerne i modellen relaterer sig direkte til det pågældende aspekt i sekvensdiagrammet, hvilket betyder at der er undladt unødvendige detaljer. Modellen er dermed en abstraktion af det virkelige system.

## **5.5 Rich Picture og egne udarbejdede modeller**

Til at skabe transparens af LTU, har vi udover UML anvendt modelleringsteknikken Rich Picture og egne udarbejdede modeller, som har til formål at uddybe elementer i rich picture. Disse modeller er udviklet til at øge forståelsen for LTU's kompleksitet, ved at nedbryde det og anskue det fra forskellige aspekter.

### **5.5.1 Rich Picture**

Rich Picture modellen er en visuel repræsentation som identificerer interessenter, bekymringer samt strukturer i henhold til sagsbehandlerens arbejdspraksis, hvori LTU anvendes. Yderligere er rich picture et supplement til domænemodellen, hvor rich picture tilføjer menneskelige faktorer og øger detaljegraden. I modellen forekommer der bogstaverne A-F, som hver repræsenterer et aspekt som repræsenterer med modeller. Disse modeller uddybes i analysens del afsnit: 5.5.2-5.5.7.



Figur 14: Rich Picture

Figur 14: Rich Picture tager udgangspunkt i den primære aktører, sagsbehandleren. Modellen er struktureret således, at sagsbehandleren er placeret i midten. Til højre for sagsbehandleren illustreres arbejdsprocessen, hvor sagsbehandleren koordinerer en ledig borgers forløb. Til venstre for sagsbehandleren illustreres sagsbehandlerens interaktion med LTU i systemet Fasit. Yderligere illustreres den manglende forståelse af LTU, ved symbolet black box. Omkring black boxen fremhæves LTU's forskellige elementer, som har tilknyttet et bogstav. Bogstaverne henviser til en til flere modeller, som uddyber det pågældende element. Derudover er sagsbehandlerens bekymringer og spørgsmål centralt for rich picture, hvilket repræsenteres ved tænke- og talebokse hvori almindelig tekst indgår.

### 5.5.2 A: Sandsynlighed

Centralt for sagsbehandlerens bekymringer, er deres manglende forståelse for hvordan risikoscoren beregnes. I empirien er det beskrevet at der forekommer forskellige beregninger i algoritmen. Algoritmens overordnet formel illustreres i to modeller. Begge modeller er en repræsentation den rigtige formel, hvoraf graden af simplificeringen varierer.

Nedenstående model illustrerer en høj grad simplificering af formlen.

$$\text{Variabel 1} \times \text{Vægt 1} + \text{Variabel 2} \times \text{Vægt 2} + \dots + \text{Variabel x} \times \text{Vægt x} = \text{Riskoscore}$$

Figur 15: Formel for beregning af LTU (høj grad simplificering)

Figur 15: Formel for beregning af LTU (høj grad simplificering) fremhæver i en forenklet version formlens centrale elementer. Den illustrerer at den samlede risikoscore beregnes ved at gange variablerne og deres tilhørende vægt, hvor disse beregninger lægges sammen.

Formlen for beregningen af LTU illustreres med lavere grad af simplificering ved nedenstående model.

$$P = f(w_0 + w_1 * x_1 + w_2 * x_2 + \dots + w_p * x_p)$$

$P$  = LTU sandsynligheden  
 $F$  = Sigmoid funktion  
 $w_0$  = Sandsynligheden for LTU når alle andre værdier er 0  
 $w_p$  = Variablens vægt  
 $x_p$  = Normaliseret værdi af variabel

Figur 16: Formel for beregning af LTU (lav grad simplificering)

Figur 16: Formel for beregning af LTU (lav grad af simplificering) formidler LTU's matematiske formel. Den matematiske formel illustreres øverst i modellen, hvortil betydningen af hvert element forklares i nederste del af modellen. Elementet  $w_p$  og  $x_p$  er et udtryk for en potentiel normaliseret variabel og dens tilhørende vægt.  $w_1$  og  $x_1$  fremgår som den første normaliseret variabel og dens tilhørende vægt,  $w_2$  og  $x_2$  er et udtryk for den anden normaliserede variabel og dens tilhørende vægt. Dette fortsætter indtil algoritmens input ikke har flere variabler. Forklaringen af elementerne tilknyttes formlen ved streger. Derudover visualiseres denne tilknytning ved at elementet i formlen har samme farve for tilhørende forklaring.

Figur 15 og 16 repræsenterer hver en abstraktion af LTU's formel. Figur 15 er modelleret således, at den inddrager matematiske elementer, uden at illustrerer den konkrete formel. Modellen har dermed et højt abstraktionsniveau og få detaljer. Det høje abstraktionsniveau medfører usikkerheder i modellen. Disse usikkerheder afdækkes i figur 16, ved at øge detaljerne og kompleksiteten. Figur 16 illustrerer den konkrete formel, hvortil de matematiske elementer kort forklares.

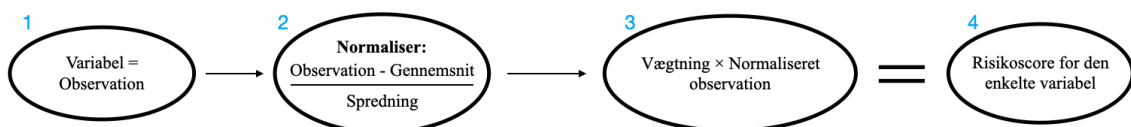


Disse modeller bevæger sig i spændet mellem kompleksitet, abstraktion og usikkerhed. Kompleksiteten af formelen nedbrydes, hvorefter der skabes abstraktioner. Disse abstraktioner medfører færre detaljer og større usikkerhed. Usikkerheden afdækkes ved at øge detaljegraden og kompleksiteten. Disse modeller supplerer dermed hinanden, og henvender sig til brugernes forskellige kompetencer.

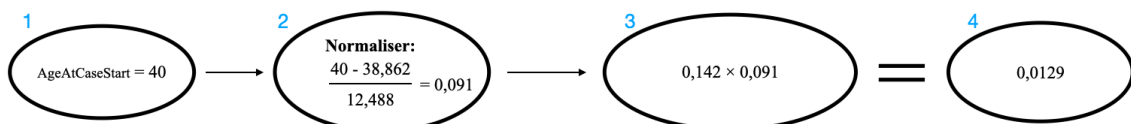
### 5.5.3 B: Algoritme

Udover manglende forståelse for den overordnede beregning, udtrykte sagsbehandlerne ligeledes manglende forståelse for, hvordan de enkelte variabler relaterede sig til risikoscoren. Beregningen er kompleks, og det er derfor nødvendigt at nedbryde den. For at kunne forstå hvordan variabler relateres til risikoscoren, skal man først og fremmest forstå hvordan enkelte variables risikoscore beregnes. Nedenstående figur illustrerer beregningen af en enkelt variabels risikoscore.

#### Processen for beregning af sandsynligheden for den enkelte variabel



#### Eksempel beregning af variabel: AgeAtCaseStart



Figur 17: Risikoscore af enkelt variabel

Figur 17: Risikoscore af en enkelt variabel illustrerer formlerne for beregningen af risikoscoren for den enkelte variabel, samt et konkret eksempel på disse beregninger. I cirkel 1 illustreres den enkelte variabel som har en observation. Denne observation anvendes i cirkel 2, hvor den skal normaliseres således, at alle observationer har samme måleenhed. Observationen normaliseres ved at minusse observation med gennemsnittet for den enkelte variabel, hvilket derefter divideres med spredningen. Resultatet af cirkel 2 anvendes i cirkel 3, hvor den ganges med vægtningen for tilhørende variabel. Dette resulterer i risikoscoren for den enkelte variabel som illustreres i cirkel 4.

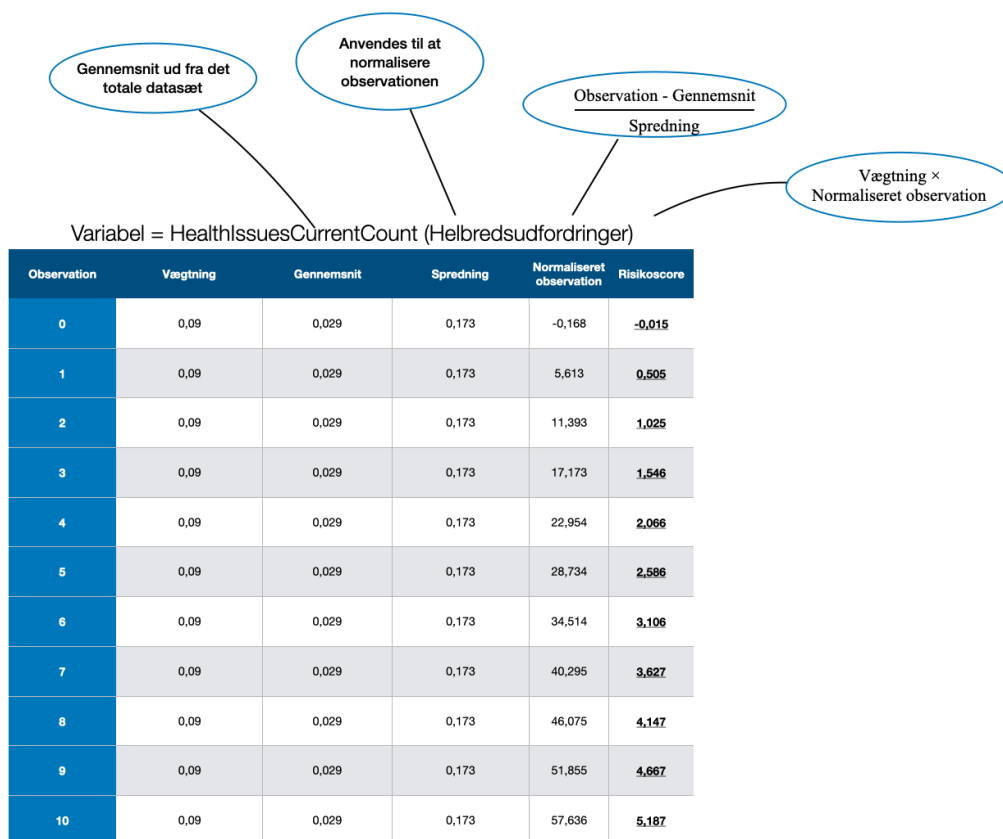
Eksemplet i den ovenstående model tager udgangspunkt i den enkelte variabel: AgeAtCaseStart, hvor observationen for denne variabel har en værdi på 40. Værdien 40 normaliseres ved at minusses med gennemsnittet 38,862, og derefter divideres med spredning, 12,488 hvor resultatet er 0,091. Resultatet ganges dernæst med vægten som er 0,142. Risikoscoren for AgeAtCaseStart med observationen 40 er dermed 0,0129.

Formlerne og eksemplet i modellen skal anvendes i forlængelse af hinanden, da formlerne beskriver den overordnede beregning, hvortil eksemplet konkretiserer formlerne og sætter det i relation til en aktuel variabel. Formlerne er en abstraktion af beregningen, hvor eksemplet imødekommer de usikkerheder abstraktionen medfører.

Derudover er indholdet i cirkel 3, en simplificeret repræsentation af følgende element:  $w_p * x_p$  i den overordnede formel, som ses i figur 16. Denne model skaber derved en anden repræsentation samt øger detaljegraden for figur 15 og 16.

#### **5.5.4 C: Observationer**

Nedenstående figur er udarbejdet med det formål, at illustrerer en enkelt variabels mulige forskellige observationer og dets relation til risikoscoren. Tabellen inddrager yderligere formlerne beskrevet i ovenstående figur 17. Tabellen kan ses som et supplement til journalnotatets beregning ift. en enkelt variabel.



Figur 18: Helbredsudfordringer

Figur 18: Helbredsudfordringer tager udgangspunkt i variabelen HealthIssuesCurrentCount. I journalnotat fremgår det, at variabelen beskriver antallet af helbredsudfordringer borgeren p.t. har registreret i Fasit. Tabellen indeholder 6 kolonner; observation, vægtning, gennemsnit, spredning, normaliseret observation og risikoscore. Den første kolonne “Observation”, repræsenterer mulige antal helbredsudfordringer. I denne tabel fremgår værdierne fra 0 til 10. Kolonne 2 er “vægtning”, som bestemmes for hver gentræning af LTU. Vægtningen bliver ikke påvirket af antallet af observationer, og fremgår dermed som det samme tal for hver observation i tabellen. Kolonne 3 og 4 er “gennemsnit” og “spredning”, som repræsenterer gennemsnittet og spredningen i det pågældende datasæt. Gennemsnittet og spredningen er udregnet fra Schultz’ datasæt om helbredsudfordringer, hvor disse bliver anvendt til at normalisere værdien.

Kolonne 5 er “normaliseret observationer”, som illustrerer værdien for den normaliseret observation, hvor følgende matematiske formel anvendes:  $\frac{Observation - Gennemsnit}{Spredning}$ . Den normaliserede observation er relevant at inddrage i tabellen, da den skaber mere transparens, for hvilke tal der ligger til grund for observationernes risikoscore, og skaber dermed mere forståelse for processen. Kolonne 6 er “Risikoscore”, som repræsenterer risikoscoren for den enkelte variabels observation, udregnet ved at gange vægten og den normaliserede

observation. Risikoscoren kan have et fortegn der er minus eller plus, afhængig af den foregående beregning. Betydningen af dette beskrives i afsnit 5.5.5 D: Fortegnets effekt på risikoscoren.

### 5.5.5 D: Fortegnets effekt på risikoscoren

Nedenstående figur 19: Enkelt variabels effekt på risikoscoren, indeholder samme struktur som figur 18. Figur 19 tager udgangspunktet i variabelen AgeAtCaseStart, og illustrerer hvorvidt den enkelte variables risikoscore har en positiv eller negativ effekt på den endelige risikoscore.

AgeAtCaseStart					
Observation	Vægtning	Gennemsnit	Spredning	Normaliseret observation	Risikoscore
30	0,142	38,862	12,488	-0,710	-0,101
32	0,142	38,862	12,488	-0,549	-0,078
33	0,142	38,862	12,488	-0,469	-0,067
34	0,142	38,862	12,488	-0,389	-0,055
35	0,142	38,862	12,488	-0,309	-0,044
36	0,142	38,862	12,488	-0,229	-0,033
37	0,142	38,862	12,488	-0,149	-0,021
38	0,142	38,862	12,488	-0,069	↓ -0,010
39	0,142	38,862	12,488	0,011	↑ 0,002
40	0,142	38,862	12,488	0,091	0,013
41	0,142	38,862	12,488	0,171	0,024
42	0,142	38,862	12,488	0,251	0,036
43	0,142	38,862	12,488	0,331	0,047
44	0,142	38,862	12,488	0,411	0,058
45	0,142	38,862	12,488	0,492	0,070
46	0,142	38,862	12,488	0,572	0,081
47	0,142	38,862	12,488	0,652	0,093

Figur 19: Enkelt variabels effekt på risikoscoren

Figur 19: Enkelt variabels effekt på risikoscoren illustrerer hvornår variabelens riskoscore skifter fra negativt fortegn til positivt fortegn. Dette skift illustreres med en blå rektangel, som omkranser to rækker i tabellen. Det kan udledes, at observationer til og med 38 har en risikoscore med et minus fortegn, hvilket har en positiv effekt på den endelige risikoscore. Effekten er positiv, da en risikoscore med minus fortegn, bringer den endelige risikoscore ned. Modsat ses det at risikoscore fra og med observation 39 har et plus fortegn, hvilket har en negativ effekt på den endelige risikoscore. Dette skyldes at den enkelte variabels risikoscore, hæver den endelige risikoscore. Det kan dermed udledes, at den endelige

risikoscore bliver lavere, hvis risikoscoren for den enkelte variabel er minus. Hvorimod den endelige risikoscore bliver højere, hvis risikoscoren for den enkelte variabel er plus. Dette illustreres ligeledes med pile. En positiv effekt illustreres med en grøn pil, som har retningen ned, da den overordnede risikoscore minimeres. En negative effekt illustreres med en rød pil, som har retning op, da den endelige risikoscore hæves.

Modellen er en repræsentation af hvordan den enkelte risikoscore påvirker den endelige risikoscore. Modellen har høj detaljegråd i henhold til den enkelte variabel, da modellen sammenligner potentielle observationer for den samme variabel. Dette betyder ligeledes, at der skabes et grundlag for at kunne sammenligne hvorvidt en observations risikoscore er høj eller lav. Denne sammenligning foretages med andre mulige observationer for den specifikke variabel.

Modellens fokus resulterer i, at andre variabler udelades, og er dermed kun en repræsentation for en enkelt variabel. Modellen muliggør dermed ikke at brugeren kan sammenligne med andre variabler.

Vi kan i forlængelse af ovenstående model udlede at fortegnet for vægtning og den normaliseret observation, har en effekt på den endelige risikoscore. Dette er sammenfattet i nedenstående figur, som illustrerer det matematiske fortegns betydning for den endelige risikoscore.

### Fortegn betydning

Vægtning		Normaliseret observation	=	Enkelt variabels risikoscore	Påvirkning på endelige risikoscore
—	×	—	=	+	↑
—	×	+	=	—	↓
+	×	—	=	—	↓
+	×	+	=	+	↑

Figur 20: Fortegn betydning

Figur 20: Fortegn betydning indeholder 5 kolonner: vægtning, gangetegn, normaliseret observation, lighedstegn, enkelt variabels risikoscore og påvirkning på endelig risikoscore. Rækkerne repræsenterer forskellige kombinationer af fortegn og dets påvirkning på den endelige risikoscore. Vægtningen ganges med den normaliseret observation.

Pilene i sidste kolonne har retningerne op og ned, som viser hvorvidt den endelige risikoscore stiger eller falder. Dertil er pilene henholdsvis røde eller grønne, som viser hvorvidt det har en positiv (grøn) eller negativ (rød) effekt på den endelige risikoscore.


Første række illustrerer at hvis vægtning og normaliseret observation begge har minus fortegn, bliver den enkelte variabels risikoscore et positiv tal. Dette medfører at den endelige risikoscore stiger, og dermed har en negativ effekt. Række 2 og 3 illustrerer, at hvis vægtning og normaliseret observation har forskellige fortegn, vil den enkelte variabels risikoscore have et minus fortegn. Dette bringer den endelige risikoscore ned, og har dermed en positiv påvirkning på den endelige risikoscore. Den sidste række, række 4, illustrerer at hvis vægtningen og den normaliseret observation begge har plus fortegn, bliver fortegnet for den enkelte variabels risikoscore ligeledes plus. Dette påvirker den endelige risikoscore negativt.

Figur 20 indebærer et højt abstraktionsniveau, og er en simplificeret repræsentation. Modellen illustrerer kun fortegnets betydning, og inddrager ikke den enkelte variabels risikoscore. Usikkerheden ligger i, at den enkelte variabels risikoscore uafhængig af fortegnet, og i sig selv, påvirker den endelige risikoscore. Dette betyder, at man ikke udelukkende kan anvende modellen til at vurdere effekten på den endelige risikoscore.

### **5.5.6 E: Output**

Tidligere modeller har fokuseret på at skabe transparens af, hvordan risikoscoren udregnes med fokus på enkelte variabler. De enkelte variabler kan anvendes til at fortolke hvilken data LTU inddrager, samt hvor meget de forskellige variabler vægter. Dog pointerede interviewpersonen fra Schultz, at sagsbehandlere ikke udelukkende skal fortolke enkelte variabler, da den endelige risikoscore er et udtryk for alle variabler. Dette stemmer dog ikke overens med sagsbehandlernes arbejdsgang og behov, da de ikke kan redegøre for hvorfor borgeren bliver risikovurderet som de gør. Vi har derfor udarbejdet en model som inddrager den endelige risikoscore og risikoscore for de enkelte variabler, som anvendes i beregningen af den endelige risikoscore.

Variabelnavn	Risikoscore
Variabel	Variabels risikoscore
Variabel	Variabels risikoscore
Variabel	Variabels risikoscore
Variabel	Variabels risikoscore
...	...
...	...
I alt	Den endelige risikoscore



Variabelnavn	Risikoscore
AgeAtCaseStart	0,0129
HealthIssuesCurrentCount	0,505
PreviousCasesCount	-0,39
DaysBeforeFirstInterview	0,005
...	...
...	...
I alt	X

Figur 21: Den endelige og de enkelte variabelers risikoscore

Figur 21 indeholder to tabeller, som har henholdsvis to kolonner og adskillige rækker. Kolonne 1 repræsenterer variabelnavn, og kolonne 2 repræsenterer risikoscoren for den enkelte variabel. Rækkerne repræsenterer det antal variabler som indgår i LTU's beregning, hvor den sidste række repræsenterer summen af de enkelte variabelers risikoscore, og dermed den endelige risikoscore. Disse to tabeller inddrager blot et udsnit af de variabler som indgår i LTU's beregning.

Tabel 1 illustrerer en grundmodel, der kan anvendes til konkrete risikovurderinger. Grundmodellen anvender begreberne; variabel og risikoscore, hvor udvalget af variableerne varierer ved hver gentræning. Derudover indeholder tabellen to rækker, som har indholdet "...", dette illustrerer at der kan tilføjes yderligere variabler og deres respektive risikoscore. Tabel 2 eksemplificerer grundmodellen, og inddrager konkrete variabler og potentielle risikoscore til disse. Figur 21 fremhæver de enkelte variabler, og deres respektive risikoscore og viser det i henhold til den endelige risikoscore.

Da modellen blot skal give et overblik over den endelige og de enkelte variabelers risikoscore, har modellen højt abstraktionsniveau med få detaljer, og med dette formindskes kompleksiteten. Modellen indeholder ikke beskrivelse af variabelens observation eller udregning af resultatet. Dette er undladt for at undgå informationstæthed, og dermed udelukkende formidle en oversigt over variabelernes risikoscore, samt den endelige risikoscore.

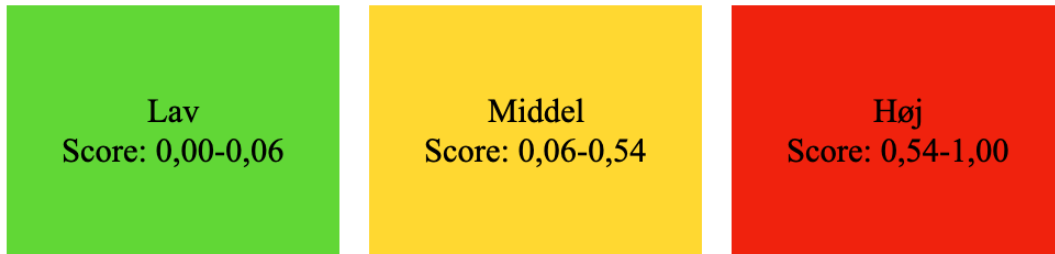
### 5.5.7 F: Risikoscore

Nedenstående figur 22: Risikoscore, illustrerer i hvilke intervaller risikoscoren bliver inddelt i kategorierne lav, middel og høj. Disse repræsenteres med farverne grøn, gul og rød.

Yderligere indgår risikoscoren og intervallerne for kategoriseringerne også i journalnotatet.

# Risikoscore

=



Figur 22: Risikoscore

Figur 22 illustrerer, at LTU's risikovurdering af borgerens sandsynlighed for langtidsledighed beregnes således, at outputtet er et numerisk tal mellem 0 og 1. Borgere med en risikoscore i intervallet 0,00-0,06 kategoriseres som lav risiko, hvor borgere med en risikoscore i intervallet 0,06-0,54 kategoriseres som middel risiko. Sidste interval 0,54-1,00 kategoriserer borgerne til at være i høj risiko. Da Lyngby-Taarbæk kommune ofte omtaler borgerne i farverne grøn, gul og rød i henhold til de tre kategorier, illustreres dette ligeledes i figur 22. Denne model fokuserer på at formidle outputtet visuelt. Hvor modellen inddrager både numeriske intervaller, kategoriseringen i lav, middel og høj, samt deres tilhørende farver.

## 5.6 Delkonklusion

Metoden modellering bevæger sig i spændet mellem abstraktion, kompleksitet og usikkerheder. Kompleksitet formindskes og repræsenteres i modeller som er abstrakte repræsentationer af virkeligheden. Formindskning af kompleksitet øger usikkerheden, hvortil det i forsøget på at imødekomme usikkerheden, øger kompleksiteten. Dette spænd er gennemgående i analysens modeller.

Analysens modeller anvendes til at nedbryde LTU's kompleksitet, hvor modellerne belyser forskellige aspekter samt supplerer hinanden. Modellerne er dermed en abstrakt repræsentation af LTU. Modellerne supplerer hinanden således, at analysen afdækker de usikkerheder, som forekommer i forlængelse af modellernes forskellige niveauer af abstraktion.

Transparens skal formidles således, at den almene bruger og/eller almene borger, opnår forståelse for teknologien. Modellerne har til formål at skabe transparens af LTU's input, proces og output, ved at nedbryde LTU's kompleksitet. Disse modeller skaber dermed



abstrakte repræsentationer af LTU, hvortil det teoretisk er nemmere for den almene borger at forstå LTU. Modellerne anvendes dermed til at skabe transparens.

## 6. Diskussion

I denne diskussion vil vi diskutere, hvorvidt rapportens modeller skaber større transparens af LTU, samt hvordan modellerne skal formidles til den almene borgeren for at opnå størst forståelse. Transparensen er LTU er henvendt til sagsbehandlere, og de udgøre den tiltænkte brugergruppe. Ideelt bør modellerne testes på sagsbehandlere, for at kunne vurdere hvorvidt modellerne skaber transparens. Vi har dog ikke haft mulighed for at teste på sagsbehandlere af flere årsager. På første møde med Schultz, blev projektgruppen anbefalet 3 kommuner som anvender LTU, og fremstod derfor som særlig fordelagtige at teste modellerne på. Disse kommuner var Lyngby-Taarbæk, Skive og Frederiksberg. Projektgruppen måtte dog senere erkende, at vi på trods af anbefalingen ikke havde mulighed for at teste modellerne i disse kommuner. Frederiksberg kommune var igang med at implementere STU, som forudsiger risikoen for kortidsledighed. Kommunen blev vurderet irrelevant, da rapportens modeller tager udgangspunkt i LTU. I forlængelsen af deres implementering af STU, havde de endnu ingen spørgsmål vedrørende transparens af STU. Skive kommune ønskede ikke at deltage, da de ikke længere anvendte LTU i deres arbejdspraksis. I henhold til Lyngby-Taarbæk kommune frabad Schultz, at vi kontaktede dem, hvilket skyldtes en konflikt mellem Schultz og kommunen. Denne konflikt opstod i forlængelse af forundersøgelsen “Er du grøn? Algoritmer til beslutningsstøtte i det offentlige”. Projektgruppen har derfor alternativt valgt at teste modellerne på almene borgere. Denne løsning understøtter GDPR’s regler om transparens, hvori det fremhæves at transparens skal kunne forstås af et almindeligt individ. Derudover kræves det, i borgernes ret til begrundelse, at der skal indgå en forklaring på en afgørelse. Hvis der er anvendt hel eller delvis automatiseret behandling, skal dette også fremgå i begrundelsen. Transparensen som opnås gennem modellerne, skal derfor også kunne forstås af den almene borger.

### 6.1 Skaber modellerne transparens?

I hovedtræk var modellerne forståelige for interviewpersonerne, dog forekom der uklarheder hvilket fremhæves i dette afsnit.

Fælles for interviewene var, at domænemodellen blev læst som tiltænkt, hvilket udledes af at interviewpersonerne fulgte domænemodellens flow. Sekvensdiagrammet var derimod sværere at forstå, hvilket udledes af at én af interviewpersonerne, slet ikke forstod diagrammet, hvorimod de resterende kunne ræsonnere sig frem til at det repræsenterede en proces (Bilag 7: Interview (Test af modeller)). Særligt fremtrædende var elementet interaktionspartnere,

hvor forståelsen først blev opnået efter en beskrivelse. Der forekom yderligere tvivl om meddelelserne og aktiveringsbarene i sekvensdiagrammet. Meddelelserne forekom uklare, da interviewpersonerne udtrykte tvivl om hvorvidt det symboliserede funktioner, handlinger eller andet. Ligeledes forstod interviewpersonerne ikke hvad aktiveringsbarene repræsenterede (Bilag 7: Interview (Test af modeller)).

Use case diagrammet blev overordnet forstået som tiltænkt, altså at en sagsbehandler kan anvende funktioner i Fasit. Dog udtrykte interviewperson 1, tvivl om hvordan use casene skulle læses. Denne tvivl handlede om hvorvidt use casene skulle læses oppefra og ned, om de var sammenhængende eller om de skulle forstås som separate aktiviteter (Bilag 7: Interview (Test af modeller)).

Yderligere forekom rich picture modellen forståelig for alle interviewpersonerne uden supplerende forklaring, modsat forekom der uforstående elementer i de uddybende punkter A-F, som indeholder figur 15-22.

Interviewpersonerne var ikke enige i hvorvidt figur 16: Formel for beregning af LTU (lav grad simplificering) var forståelig og brugbar til at skabe transparens. Interviewperson 1 forstod slet ikke modellen, interviewperson 3 fandt den irrelevant, hvorimod interviewperson 4, fandt det relevant med en mere matematisk repræsentation af beregningen. Da interviewpersonerne udtrykte forskellige holdninger til hvorvidt den matematiske formel var relevant, udarbejdede projektgruppen en yderligere repræsentation af denne formel (figur 15). Denne repræsentation differentierer fra figur 16, ved at have højere grad af simplificering. Interviewpersonerne fandt figur 17: Risikoscore af enkelt variabel forstående, da modellen både havde en overordnet beskrivelse af processen samt et konkret eksempel.

Interviewperson 4 fremhævede, at kombinationen af en overordnet beskrivelse og et konkret eksempel øgede hans forståelse, da han *“kører lidt imellem dem når man læser den”* (Bilag 7: Interview (Test af modeller), s. 67). Derudover pointerede han, at den overordnede formel blev konkretiseret. Derimod forekom figur 18: Helbredsudfordringer sværere at forstå, hvor interviewperson 3 og 4 udtrykte tvivl om modellens fokus. Denne tvivl blev afdækket med en beskrivelse af modellen, hvilket medførte at den blev forståelig for interviewpersonerne.

Beskrivelsen af figur 18 resulterede i, at interviewpersonerne oplevede at figur 19 var nemmere at forstå, da modellerne indeholder samme tabelstruktur. Yderligere forstod interviewpersonerne at modellen fremhævede en risikoscore der var minus og plus.

Interviewperson 4 udtrykte modellens formål, ved at pointere at en ledig borger er dårligere stillet som 39 årig end som 38 årig.

Ved figur 20: Fortegn betydning, er det særligt relevant at fremhæve, at interviewperson 1 fokuserede på hvorvidt fortegnet skulle ganges eller divideres. Da dette ikke er modellens fokus, valgte projektgruppen at inddrage et gangetegn, således at den almene borger fokuserede på fortegnetets betydning, fremfor hvorvidt de skulle ganges eller divideres (Bilag 7: Interview (Test af modeller)).

Derudover udtrykte interviewpersonerne at figur 21: Den endelige og de enkelte variables risikoscore, fremstod som forstående og relevant. Interviewpersonerne fremhævede, at figur 21 opsummerede den viden, de havde opnået gennem tidligere fremlagte modeller. Endvidere pointerede interviewperson 4, at figuren gav et godt overblik. Afslutningsvis forekom figur 22: Risikoscore nem at forstå, da modellen er enkel, både i henhold til dens farver og tekst (Bilag 7: Interview (Test af modeller)).

## **6.2 Kan modellerne stå alene eller kræves der en beskrivelse?**

Det kan udledes at interviewpersonerne overordnet forstod modellerne, men at der i deres beskrivelser af modellerne forekom mange gæt som både var korrekte og forkerte. Ved uforstående elementer, oplevede projektgruppen, at en beskrivelse øgede deres forståelse af modellen. Yderligere efterspurgt interviewpersonerne en begrebsafklaring. De pointerede, at modellerne ville være nemmere at forstå, hvis man forinden var bekendt med begrebets betydning. Derudover pointerede interviewperson 4, at han var bekendt med begrebet “normaliser” i henhold til database, og denne forståelse ville han inddrage i fortolkningen af modellerne. Begrebet normaliser har en anden betydning i rapportens modeller, end det blev fremlagt af interviewperson 4. Dette argumenterer for at en begrebsafklaring sikre, at den almene borger opnår samme forståelse for begrebets betydning, som de anvendes i modellerne. Et af interviewspørgsmålene var “*Hvad er din holdning til at modellerne suppleres med en beskrivelse*” (Bilag 7: Interview (Test af modeller), s. 64), hvor alle interviewpersonerne udtrykte, at modellerne med fordel kunne suppleres med en beskrivelse.

Uklarheder kan blive en usikker faktor for borgeren, hvilket kan medføre at modellerne forekommer upålidelige og/eller at borgeren opgiver at forstå modellerne. Denne usikkerhed kan imødekommes ved, at modellerne formidles med en skriftlig eller mundtlig beskrivelse. Det kan dermed argumenteres, at rapportens modeller uden skriftlig eller mundtlig formidling, ikke nødvendigvis øger forståelse af LTU for den almene borger. Dette betyder også, at der ikke nødvendigvis skabes større transparens af LTU.

Hvis modellerne derimod formidles med tekst eller tale, opnås større forståelse og dermed større transparens. Derudover ses det at modellerne, modsat almindelig tekst, simplificerer og visualiserer LTU's kompleksitet. De visuelle modeller tilføjer en abstraktion, som kan være nemmere for den almene borger at forstå.

For at skabe større transparens, ses det at modellerne ikke kan stå foruden beskrivelser, men at transparensen ligeledes ikke kan skabes udelukkende med beskrivelser. Da modellerne ikke kan formidles alene, kan de anses som formidlingsstøtte, og dermed skal både modeller og beskrivelser inddrages for at skabe transparens.

### **6.3 Modellering og transparens**

Modellerne bevæger sig på forskellige niveauer af transparens, hvilket kommer til udtryk ved at modellerne repræsenterer forskellige aspekter af LTU, hvor detaljegraden i repræsentationerne varierer. LTU er fortsat en black box, hvor rapportens modeller har til formål at åbne dele af denne black box. Det argumenteres derved, at denne rapport ikke arbejder med fuld transparens, men derimod med "usable transparency".

Projektgruppen har udarbejdet modellerne med udgangspunkt i sagsbehandlerne spørgsmål til LTU, og de oplysninger projektgruppen har haft tilgængelige. Modellerne skal skabe transparens af LTU til sagsbehandlerne og almene borgere. Dette betyder, at modellerne skal formidles til en gruppe, som ikke nødvendigvis har tekniske kompetencer til at forstå LTU's kompleksitet. Dette påvirker niveauet af transparens, da modellerne skal transformere tekniske faglige begreber til et alment sprog. Mængden af information og detaljer har dermed været centrale overvejelser i udarbejdelsen af modellerne. Dette fremhæves ligeledes i interviewene hvor modellerne testes. I dette interview, var interviewperson 3 og 4 uenige om relevansen af figur 16. Interviewperson 3 fandt ikke figuren relevant, da han ikke forstod den. Hvorimod interviewperson 4 forstod formlen, og pointerede at den gav en mere teknisk vinkel på LTU. Interviewpersonernes forskellige holdninger fremhæver, at modellerne formidles til borgere med forskellige tekniske kompetencer, og dermed har modellerne ligeledes forskellige niveauer af transparens. Modellering supplerer dermed transparens, da modellering muliggør at kompleksitet kan nedbrydes, hvor denne modellering formidles således, at aspekter af den komplekse teknologi forstås af den almene borger. Det kan dermed argumenteres, at modellering er fordelagtig til at skabe transparens.

## 6.4 Delkonklusion

I hovedtræk forekommer modellerne forståelige for den almene borger, under forudsætning af at de forstår begreberne som anvendes i modellerne.

Modellerne beskriver forskellige aspekter af LTU, i forsøget på at åbne black boxen. Da det ikke altid er muligt at skabe fuld transparens af kunstig intelligens, skaber modellerne usable transparency, ved at formidle tilgængelige informationer.

Yderligere argumenteres det, at modeller der har til formål at skabe transparens, ikke kan formidles foruden beskrivelser, og at transparens ligeledes ikke kan skabes udelukkende med beskrivelser. Da modellerne ikke kan formidles alene, kan de anses som formidlingsstøtte, og dermed skal både modeller og beskrivelser inddrages for at skabe transparens.

Modellering supplerer dermed transparens, da modellering muliggør at kompleksitet kan nedbrydes, hvor denne modellering formidles således, at aspekter af den komplekse teknologi forstås af den almene borger. Det kan dermed argumenteres, at modellering er fordelagtig til at skabe transparens.

## 7. Perspektivering

LTU automatisere delvis sagsbehandlerens arbejdsproces, som planlægger og koordinere borgerens ledighedsforløb. Anvendelse af kunstig intelligens til at automatiserer arbejdsprocesser, perspektiveres til Mintzbergs organisationsformer; fagbureaukrati og maskinbureaukrati. Afslutningsvis inddrager perspektivering hvorvidt et jobcenter eller den individuelle sagsbehandler anvender LTU, hvilket anskues ud fra beslutningsmodellerne economic man og administrative man.

### 7.1 Fagbureaukrati og maskinbureaukrati

Anvendelse af kunstig intelligens i organisationer tillader hel og/eller delvis automatiseret arbejdsprocesser. Disse arbejdsprocesser kan være situeret i en arbejdspraksis, hvor arbejdsprocessen har direkte indflydelse på en borger, såsom offentlig sagsbehandling. Offentlig sagsbehandling er i høj grad baseret på faglig viden, hvor det ses at denne faglige viden suppleres med automatiserede processer, bl.a. i form af kunstig intelligens. Arbejdspraksisser baseret på faglig viden og automatiseret processer, anskues ud fra Mintzberg' organisationsformer; fagbuerakati og maskinbureaukrati.

Fagbureaukrati handler om at organisationens arbejde i høj grad beror på ansattes faglige viden, hvortil de ansatte udfører og tilrettelægger arbejdet efter de deres faglige kompetencer. Disse kompetencer opnås ofte gennem uddannelse. Modsat handler maskinbureaukrati om at organisationens arbejde beror på standardiseret processer, hvor der ikke kræves en højere faglig viden til at udføre og tilrettelægge arbejdet.

Kunstig intelligens anvendes ofte til at standardisere og optimere arbejdsprocesser, hvortil det ses, at kunstig intelligens standardisere arbejdsprocesser, som hidtil har været betinget af faglig viden. Kunstig intelligens tillader netop at man omformer faglig uddannelse til automatiseret arbejdsprocesser, hvor faglig viden gennem algoritmen standardiseres. Dette ses netop ved jobcentre, som anvender LTU. Førhen blev borgernes risiko for langtidsledighed udelukkende vurderet ud fra sagsbehandlerne faglige viden, hvor anvendelsen af LTU har ændret sagsbehandlerne arbejdspraksis. Denne ændring har medført, at borgerne som udgangspunkt vurderes ud fra LTU's anbefaling. Det kan derved argumenteres at fagbuerakati såsom jobcentre, bevæger sig mod maskinbureaukrati. Anvendelse af kunstig intelligens i sagsbehandling berører dermed spørgsmålet om, hvorvidt

faglige kompetencer og uddannelse kan standardiseres således, at arbejdet kan udføres automatisk.

Yderligere argumenteres det, at kunstig intelligens ikke fuldkommen kan automatiserer sagsbehandlerens arbejdspraksis. Dette skyldes, at denne arbejdspraksis er underlagt love, som bestemmer at en borger har ret til at blive menneskelig vurderet. Derudover fremhæves det, at kunstig intelligens ikke er i stand til at identificere menneskelige faktorer. En menneskelig faktorer kan eksempelvis være borgerens hygiejne, hvilket en sagsbehandler først kan vurdere ved menneskelig kontakt. Afslutningsvis påpeges det, at en automatiseret behandling kan medføre, at sagsbehandleren tilsidesætte deres faglige vurdering og skaber uhensigtsmæssige konsekvenser for borgeren (Frost, 2022; Flügge et al., 2022).

## **7.2 Economic man og administrative man**

Hvorvidt et jobcenter eller den individuelle sagsbehandler anvender kunstig intelligens, eksempelvis LTU, kan anskues ud fra hvilken beslutningsmodel de befinder sig i. Fælles for beslutningsmodellerne economic man og administrative man er, at der træffes beslutninger ud fra et rationelt grundlag.

Hvis jobcentret eller sagsbehandleren stiller sig tilfreds med LTU som den bedst mulige løsning, og dermed anvender den i deres arbejdspraksis, befinder jobcentre og sagsbehandleren sig under modellen administrative man. Herunder vil detaljerne om LTU ikke være centrale, såfremt at det forekommer som det bedst mulige alternativ. Hvis dette er tilfældet, vil en administrativ man anvende LTU som beslutningsstøtte. Hvis jobcentret eller sagsbehandleren befinder sig i beslutningsmodellen economic man, vil det være centralt at vide hvordan LTU fungerer, før den anvendes som beslutningsstøtte, da den ellers ikke ville kunne fremstå som den bedste løsning. Jobcentret eller sagsbehandleren har ikke fuldkommen information og forståelse for LTU, da de ikke kan redegøre for hvordan LTU beregner den ledige borgeres risikoscore. LTU's transparens vil dermed ikke være tilfredsstillende, hvilket medfører at den ikke fremkommer som den bedste løsning. Den bedste løsning for economic man, vil derfor være sagsbehandlerens faglige vurdering, da de derved har fuld information og kan redegøre for vurderingen.



## 8. Konklusion

Denne rapport har til formål at skabe transparens af kunstig intelligens algoritmen LTU (Long Term Unemployment), med fokus på dens input, proces og output. Da det ikke altid er muligt at skabe fuld transparens af kunstig intelligens, fokuserer rapporten på at skabe usable transparency ved at formidle tilgængelige informationer.

Algoritmen LTU har til formål at vurdere borgers risiko for langtidsledighed i beskæftigelsesområdet. Sagsbehandlere er underlagt forvaltningsloven, offentlighedsloven og databeskyttelsen loven, hvortil de skal redegøre og begrunde en sagsbehandlings afgørelse. Sagsbehandlere i Lyngby-Taarbæk Kommune anvender LTU som beslutningsstøtte, hvor LTU som udgangspunkt bestemmer borgernes ledighedsforløb. De udtrykker manglende forståelse for LTU's input, proces og output, hvilket tilføjer problematikken, at de ikke kan begrunde LTU's risikovurdering af borgeren. Schultz har udarbejdet journalnotat med det formål at skabe transparens og forståelse for LTU, dog udtrykker sagsbehandlerne at de fortsat ikke har opnået større forståelse af LTU.

Algoritmens input er data om borgeren, hvor dataen anvendes i algoritmens matematiske beregninger. Centralt for disse beregninger er logistisk regression, hvori variabler, vægtning, gennemsnit og spredning indgår, og har til formål at forudsige borgerens langtidsledighed. Algoritmens output er et numerisk tal mellem 0 og 1, der repræsenterer en risikoscore. Risikoscoren kategoriseres som lav, middel eller høj, hvor hver kategori repræsenterer et interval mellem 0 og 1.

Analysens modeller beror på metoden modellering, som bevæger sig i spændet mellem abstraktion, kompleksitet og usikkerheder. Kompleksitet formindskes og repræsenteres i modeller som er abstrakte repræsentationer af virkeligheden. Formindskning af kompleksitet øger usikkerheden, hvortil det i forsøget på at imødekomme usikkerheden, øger kompleksiteten. Analysens modeller anvendes til at nedbryde LTU's kompleksitet, hvor modellerne belyser forskellige aspekter. Modellerne er dermed en abstrakt repræsentation af LTU. Modellerne supplerer hinanden således at analysen afdækker de usikkerhed som forekommer i forlængelse af modellernes forskellige niveauer af abstraktion.

Analysens modeller har til formål at skabe transparens af LTU, hvor modellerne formidles således, at aspekter af LTU forstås af den almene borger. Modellering supplerer dermed transparens.

Det kan ud fra interviews med almene borgere udledes, at modellerne i hovedtræk er forståelige. Forståeligheden er betinget af en begrebsafklaring og beskrivelse af modellerne. Modeller der har til formål at skabe transparens, kan ikke formidles foruden beskrivelser, og transparens kan ligeledes ikke kan skabes udelukkende med beskrivelser. Dermed skal både modeller og beskrivelser inddrages for at opnå størst mulig forståelse og transparens.

# Litteraturliste

Aarhus Universitet. (n.d.). *Uafhængige og afhængige variabler*. Metodeguiden; Aarhus

Universitet. Hentet 12. Maj 2022, fra

<https://metodeguiden.au.dk/uafhaengige-og-afhaengige-variabler/>

Beskæftigelsesministeriet. (n.d.). *Beskæftigelsesindsats*. Bm.Dk. Hentet 11. Maj 2022, fra

<https://bm.dk/arbejdsomraader/beskaeftigelsesindsats/>

Borgmesterforvaltningen. (2021). *Brug af kunstig intelligens til at målrette*

*beskæftigelsesindsatsen. Orientering om Signaturprojekt*. Odense.Dk. Hentet 30. Maj 2022, fra

<https://www.odense.dk/politik/dagsordner-og-referater/social-og-arbejdsmarkedsudvalget?agendauid=91d90e70-f277-4598-b427-eb7b2af5a430&presentationuid=37b768fb-58a6-4516-9ee6-616177e10774>

Christensen, B. (2009). *forvaltningsret (Forvaltningsretligt skøn)*. Den Store Danske. Hentet

27. Maj 2022, fra

[https://denstoredanske.lex.dk/forvaltningsret\\_\(Forvaltningsretligt\\_skøn\)](https://denstoredanske.lex.dk/forvaltningsret_(Forvaltningsretligt_skøn))

Christiansen, H. (2021). *Vigtige begreber om maskinlæring*. Roskilde Universitet.

Christiansen, P. E., Kjær, H., & Trojel, T. (2001). *Organisation* (2nd ed., s. 89, 99-102, 239-241). Trojka a/s.

Datatilsynet. (2018). *Vejledning om de registreredes rettigheder*. I *Datatilsynet* (s. 47–50).

Hentet 10. Maj 2022, fra

<https://www.datatilsynet.dk/media/6893/registreredes-rettigheder.pdf>

Den Danske Ordbog. (n.d.). *Transparens — den danske ordbog*. Ordnet.Com. Hentet 26. Maj

2022, fra <https://ordnet.dk/ddo/ordbog?query=Transparens>

Den Europæiske Unions Tidende. (2016). *Europa-Parlamentets og Rådets Forordning (EU)*

2016/679. Hentet 4. Maj fra

<https://eur-lex.europa.eu/legal-content/DA/TXT/PDF/?uri=CELEX:32016R0679&from=DA>

Diakopoulos, N. (2020). *Transparency*. In *The Oxford Handbook of Ethics of AI* (s. 197–205). Oxford Handbooks.

Djøf & Tech DK Kommissionen. (2022). *Analyse: Digital forvaltning*. Djøfs TechDK

Kommission. Hentet 31. Maj 2022,

[https://www.djoef.dk/-/media/documents/djoef/t/techdk/techdk\\_analyse\\_digital\\_forvaltning\\_a4\\_0322\\_web.ashx](https://www.djoef.dk/-/media/documents/djoef/t/techdk/techdk_analyse_digital_forvaltning_a4_0322_web.ashx)

Educative. (2021). *Data Science in 5 Minutes: What is One Hot Encoding?* Educative:

Interactive Courses for Software Developers; Educative. Hentet 20. Maj 2022, fra

<https://www.educative.io/blog/one-hot-encoding>

European Commission. (2018). *Guidelines on Transparency under Regulation 2016/679*. In

Europa (s. 4–26). European Commission. Hentet 13. Maj 2022, fra

<https://ec.europa.eu/newsroom/article29/items/622227>

Flügge, A. A., Møller, N. H., Hildebrandt, T. T., & Olsen, H. P. (2022). *Er du grøn?*

*Algoritmer til beslutningsstøtte i det offentlige*. Københavns Universitet. Hentet 10.

Maj 2022, fra

[https://static1.squarespace.com/static/5e3ad7fa73600c394b539f6b/t/6222127f1df65457595e2136/1646400131404/Flügge+et+al+\(2022\)+-+Er+du+Grøn.+Algoritmer+til+beslutningsstøtte+i+det+offentlige.pdf](https://static1.squarespace.com/static/5e3ad7fa73600c394b539f6b/t/6222127f1df65457595e2136/1646400131404/Flügge+et+al+(2022)+-+Er+du+Grøn.+Algoritmer+til+beslutningsstøtte+i+det+offentlige.pdf)

Folketingets Ombudsmand. (n.d.-a). *Afgørelsesbegrebet*. Ombudsmanden. Hentet 27 Maj

2022, fra

<https://www.ombudsmanden.dk/myndighedsguiden/generel-forvaltningsret/afgoerelse-sbegrebet/>

Folketingets Ombudsmand. (n.d.-b). *Begrundelse*. Ombudsmanden. Hentet 11. Maj 2022, fra

<https://www.ombudsmanden.dk/myndighedsguiden/generel-forvaltningsret/begrundelse/>

Folketingets Ombudsmand. (n.d.-c). *Officialprincippet*. Ombudsmanden. Hentet 11. Maj 2022, fra

<https://www.ombudsmanden.dk/myndighedsguiden/generel-forvaltningsret/officialprincippet/>

Frost, S. M. (2022). *Her er de jobs, der er truet af kunstig intelligens – og dem, der aldrig bliver det*. TV 2 DANMARK. Hentet 25. Maj 2022, fra

<https://nyheder.tv2.dk/business/2022-05-05-her-er-de-jobs-der-er-truet-af-kunstig-intelligens-og-dem-der-aldrig-bliver-det>

Ginman, R. F. (2021). *Voldsom vækst i brugen af kunstig intelligens i den offentlige sektor: Vil fordobles over tre år*. Computerworld. Hentet 30. Maj 2022, fra

<https://www.computerworld.dk/art/258586/voldsom-vaekst-i-brugen-af-kunstig-intelligens-i-den-offentlige-sektor-vil-fordobles-over-tre-aar>

Javatpoint. (n.d.). *Regression Analysis in Machine learning - Javatpoint*. Javatpoint.Com. Hentet 12. Maj 2022, fra

<https://www.javatpoint.com/regression-analysis-in-machine-learning>

Justitsministeriet. (n.d.). *Offentlighed i forvaltningen*. Borger.Dk. Hentet 11. Maj 2022, fra

<https://www.borger.dk/samfund-og-rettigheder/Offentlighed-i-forvaltningen>

Kommunernes Landsforening. (2022). *Gentænk velfærden – Kommunernes digitaliseringspolitik* (1st ed., Vol. 1, p. 1-1). Kommuneforlaget A/S.

<https://www.kl.dk/media/50069/gentaenk-velfaerden-kommunernes-digitaliseringspolitik-a5.pdf>

Larman, C. (2005). *Applying UML and patterns: An introduction to object-oriented analysis and design and iterative development* (s. 11, 61-100, 131-171, 221-270). Pearson.

- Mathiassen, L., & Stage, J. (1992). *The principle of limited reduction in software design*. Information Technology & People, 6(2/3). (s. 171–185).  
<https://doi.org/10.1108/eum0000000003550>
- Merriam-Webster. (n.d.). *Definition of black box*. Merriam-Webster Dictionary. Hentet 7. Maj 2022, fra <https://www.merriam-webster.com/dictionary/black%20box>
- Mintzberg, H. (1980). *Structure in 5's: A Synthesis of the Research on Organization Design*. Management Science, 26(3). <https://doi.org/10.1287/mnsc.26.3.322>.
- Monk, A., & Howard, S. (1998). Methods & tools. *Interactions*, 5(2). (s. 21–30).  
<https://doi.org/10.1145/274430.274434>
- Saeed, M. (2021). *A gentle introduction to sigmoid function*. Machine Learning Mastery. Hentet 11. Maj 2022, fra  
<https://machinelearningmastery.com/a-gentle-introduction-to-sigmoid-function/>
- Schultz. (n.d.-a). *Fagsystem til jobcentret*. Schultz.Dk. Hentet 11. Maj 2022, Fra  
<https://schultz.dk/loesninger/schultz-fasit/>
- Schultz. (n.d.-b). *Intelligent ledelsessystem til jobcentret*. Schultz.Dk. Hentet 30. Maj 2022, fra <https://schultz.dk/loesninger/kunstig-intelligens/>
- Seidl, M., Scholz, M., Huemer, C., & Kappel, G. (2015). *UML @ classroom: An introduction to object-oriented modeling* (s. 15-19, 23-39, 49-60, 107-116). Springer.
- Simonsen, J. (1994). *How organizations can be understood in terms of decision processes*. Spring 1994. <https://jespersimonsen.dk/Downloads/Simon-introduction.pdf>
- Styrelsen for Arbejdsmarked og Rekruttering. (n.d.). *Ansvar for beskæftigelsesindsatsen*. STAR. Hentet 11. Maj 2022, fra  
<https://star.dk/om-styrelsen/ansvar-for-beskaeftigelsesindsatsen/>
- Torgersen, E., & Kvittingen, I. (2019). *Hvad er P-værdi, og hvad betyder statistisk signifikans?* Videnskab.Dk. Hentet 15. Maj 2022, fra

<https://videnskab.dk/naturvidenskab/hvad-er-p-vaerdi-og-hvad-betyder-statistisk-signifikans>

Vasudev. (2017). *What is One Hot Encoding? Why and When Do You Have to Use it?*

HackerNoon. Hentet 20. Maj 2022, fra

<https://hackernoon.com/what-is-one-hot-encoding-why-and-when-do-you-have-to-use-it-e3c6186d008f>

Wilson, C. (2014). *Interview techniques for UX practitioners: A user-centered design method.*

Elsevier Inc.

Wood, T. (n.d.). *Sigmoid function.* DeepAI. Hentet 14. Maj 2022, fra

<https://deepai.org/machine-learning-glossary-and-terms/sigmoid-function>