

**Maskinell Transkribering av Snapchat Skjermopptak**  
**Automatic Transcription of Snapchat Screen Recordings**  
**Visjonsdokument**

**Versjon 4.0**

*Dokumentet er basert på Visjonsdokument utarbeidet ved NTNU. Revisjon og tilpasninger til bruk ved IDER, DATA-INF utført av Carsten Gunnar Helgesen, Svein-Ivar Lillehaug og Per Christian Engdal. Dokumentet finnes også i engelsk utgave.*

## REVISJONSHISTORIE

Dato	Versjon	Beskrivelse	Forfatter
25/jan/22	1.0	Produktets funksjonelle egenskaper	Vilde, Thea, Anne
25/jan/22	1.1	Ikke-funksjonelle egenskaper og andre krav	Vilde, Thea, Anne
27/jan/22	1.2	Sammendrag problem og produkt	Thea
28/jan/22	1.3	Innledning og produktoversikt	Anne
15/feb/22	1.4	Revisjon etter tilbakemelding	Vilde, Thea, Anne
16/feb/22	2.0	Korrigerer etter revisjon	Anne
17/feb/22	2.1	Videre korrigerer etter revisjon	Anne
22/feb/22	2.2	Omskriving og endelig korrigerer	Anne
10/apr/22	3.0	Små endringer til rapport utkast 1	Vilde, Thea, Anne
20/mai/22	4.0	Endringer før siste innlevering	Anne

## INNHALDSFORTEGNELSE

<b>1</b>	<b>INNLEDNING</b> .....	<b>1</b>
<b>2</b>	<b>SAMMENDRAG PROBLEM OG PRODUKT</b> .....	<b>2</b>
2.1	<i>Problemsammendrag</i> .....	2
2.2	<i>Produktsammendrag</i> .....	2
<b>3</b>	<b>BESKRIVELSE AV INTERESSENER OG BRUKERE</b> .....	<b>3</b>
3.1	<i>Oppsummering interessenter</i> .....	3
3.2	<i>Oppsummering brukere</i> .....	3
3.3	<i>Brukermiljøet</i> .....	3
3.4	<i>Sammendrag av brukernes behov</i> .....	4
3.5	<i>Alternativer til vårt produkt</i> .....	4
<b>4</b>	<b>PRODUKTOVERSIKT</b> .....	<b>6</b>
4.1	<i>Produktets rolle i brukermiljøet</i> .....	6
4.2	<i>Forutsetninger og avhengigheter</i> .....	6
<b>5</b>	<b>PRODUKTETS FUNKSJONELLE EGENSKAPER</b> .....	<b>7</b>
<b>6</b>	<b>IKKE-FUNKSJONELLE EGENSKAPER OG ANDRE KRAV</b> .....	<b>9</b>
<b>7</b>	<b>REFERANSER</b> .....	<b>10</b>

# 1 INNLEDNING

Visjonsdokumentet gir en oversikt over problemet som skal løses i bacheloroppgaven. Dette prosjektet går ut på å utvikle et program som transkriberer tekst fra et skjermopptak av en chat-samtale i Snapchat. Den transkriberte teksten skal lagres og være søkbar slik at det er mulig å finne igjen skjermbilder fra videoen med innholdet det søkes etter. Prosjektet er i samarbeid med politiet som for øyeblikket utfører disse oppgavene manuelt.

## **2 SAMMENDRAG PROBLEM OG PRODUKT**

### **2.1 Problemsammendrag**

Manuel transkribering av videoer manuelt krever ressurser og tid, ved å effektivisere kan disse ressursene brukes på andre arbeidsoppgaver. På markedet i dag finnes det ikke produkter som utfører transkribering fra video til tekst med formålet og forutsetningene til prosjektet. Det største problemet med å utføre transkriberingen manuelt er tidsforbruket som kan videre føre til utvidet ventetid på oppgaver som avhenger av transkribert tekst.

Ved å bruke et produkt som automatiserer transkriberingsprosessen vil oppgavene gjennomføres mer effektivt og på en enklere måte for oppdragsgiver.

### **2.2 Produktsammendrag**

Interessenter av produktet er organisasjoner og personer som bruker programvaren og den transkriberte teksten. Programvaren vil brukes av etterforskere i politiet og potensielt i samarbeid med dataetterforskere. Resultatet kan brukes som vedlegg i rapporter som fremlegges for retten i straffesaker. Brukere av det transkriberte resultatet inkluderer etterforskere, dommere, aktorat og forsvarere.

Etterforskere bruker det transkriberte resultatet for å ta beslutninger angående videre steg i etterforskningen, mens dommere, aktorat og forsvarere bruker det som et hjelpemiddel for å avgjøre skyld eller uskyld.

Produktet vil i hovedsak hjelpe politiet med hensyn på effektivitet. Ekstra funksjonalitet i produktet kan implementeres for å forbedre forståelse av teksten. Eksempler på slik ekstra funksjonalitet er oversetting fra andre språk til norsk slik at etterforskere kan bedre forstå innholdet.

Produktet transkriberer tekst fra skjermopptak av chat-samtaler som gjør det mulig å anvende produktet på videoer fra andre chat-applikasjoner enn Snapchat. Chat-samtaler er private, og ved kommersielt bruk av produktet kan det oppstå situasjoner hvor bruk av produktet bryter med personvern regler.

### 3 BESKRIVELSE AV INTERESSENER OG BRUKERE

#### 3.1 Oppsummering interessenter

Navn	Utdypende beskrivelse	Rolle under utviklingen
Etterforskere	Etterforskere i politiet vil være brukere av programvaren for å effektivisere arbeidsoppgaver. Etterforskere vil også kunne bruke resultatet fra programvaren for videre arbeid relatert til saken hvor videoen er hentet fra.	Politiet er oppdragsgiver og vil under utviklingen av systemet komme med innvendinger og krav til systemløsningen
Dommer Aktorat Forsvarere	Interessentene her vil bruke den transkriberte teksten hvis den blir fremlagt i retten.	Disse interessentene har ikke en direkte rolle under utviklingen, men fører til krav på formatet til den transkriberte teksten. Den transkriberte teksten må være oversiktlig og lesbar for flere personer.

#### 3.2 Oppsummering brukere

Navn	Utdypende beskrivelse	Rolle under utviklingen	Representert av
(Digitale) Etterforskere	Etterforskere i politiet vil være de eneste som er direkte brukere av produktet.	Politiet er oppdragsgiver og vil under utviklingen av systemet komme med innvendinger og krav til systemløsningen.	Etterforskere i politiet

#### 3.3 Brukermiljøet

Det er ønskelig at programvaren skal være enkel å bruke slik at det ikke er nødvendig for brukerne å ha teknisk bakgrunn. For at det skal være enklere å utføre endringer i programvaren er det ideelt å unngå unødvendige avhengigheter i systemet. Dette krever at systemet må struktureres og dokumenteres på en oversiktlig måte slik at det er mulig for andre utviklere å tilpasse programvaren etter egne behov.

### 3.4 Sammendrag av brukernes behov

Behov	Prioritet	Påvirker	Dagens løsning	Foreslått løsning
Tolkning av skjermopptak for å produsere en tekstuell representasjon	1	Politiet	Manuell transkribering av etterforskere	Automatisk transkribering utført av programvare
Formatering av transkribert tekst	4	Politiet Rettsaker	Manuell nedskrivning av etterforskere	Automatisk formatering slik at resultater kan eksporteres på en god måte
Transkribering av videoer med ulik oppløsning, FPS eller hastighet på skrolling	3	Politiet	Manuell nedskrivning av etterforskere	Dele skjermopptak opp i skjermbilder på en måte som håndterer disse variasjonene
Mulighet for å søke etter ord i en video og finne skjermbilder hvor ordet er med	2	Politiet	Manuelle søk av etterforskere	Lagre transkribert tekst på en måte slik at det er mulig å hente ut hvilke skjermbilder som har ulik tekst.
Transkribering uten å miste emoji'er som brukes i chat	5	Politiet	Manuell nedskrivning av etterforskere	Bruk av programvare som klarer å håndtere emoji'er

### 3.5 Alternativer til vårt produkt

For å oppnå transkribering av tekst fra video kommer programvaren til å dele skjermopptak opp i skjermbilder og transkribere bilde for bilde. For å transkribere tekst fra bilder brukes populært Optical Character Recognition (OCR).

Det er ikke produkter på markedet som oppnår nøyaktig det som skal oppnås i prosjektet. For å analysere konkurrenter og alternativer til prosjektet deles arbeidet inn i to kategorier; produkter som deler filmer inn i færre bilder og produkter som bruker OCR for å transkribere tekst fra bilder.

Det eksisterer mange verktøy som kan dele videoer opp i mindre videoer eller ulike skjermbilder. Videoredigeringsprogramvare kommer vanligvis inkludert i operativsystemer og kan brukes til dette formålet. Det er da mulig å dele videoer opp i bilder og fjerne skjermbilder som det ikke er ønskelig å bruke.

OCR er en populær teknologi som brukes av flere verktøy. OmniPage Ultimate (WonderShare PDF Element, 2022) er et produkt som har mange funksjonaliteter, inkludert OCR. OmniPage leser av teksten på PDF-baserte dokumenter som inneholder tekst, og konverterer den avleste teksten til ulike formater.

En annen programvare som transkriberer tekst fra videoer er InqScribe (InqScribe, 2022). Dette er en digital media transkriberingsprogramvare som transkriberer tekst fra video ved å gi en tekstuell representasjon av teksten slik den representeres i videoen. Denne programvaren sitt bruk er i hovedsak relatert til undertekster.

Alternativene er gode programvarer som dekker deler av behovet til oppdragsgiver. Bruk av disse vil kreve bruk av flere programvarer for å oppnå de totale kravene. Produktet som utvikles i prosjektet vil kombinere de to konseptene slik at brukere kun trenger å forholde seg til en programvare.



## **4 PRODUKTOVERSIKT**

### **4.1 Produktets rolle i brukermiljøet**

Produktet vil brukes av etterforskere på dedikerte maskiner i saker hvor skjermopptak av chat-samtaler brukes som bevismateriale. Hos politiet brukes i hovedsak Microsoft Windows som operativsystem på dedikerte maskiner. Samtidig må produktet utvikles på en måte slik at det er uavhengig av operativsystem grunnet ulike operativsystem blant utviklerne.

For at etterforskere med ulik teknisk bakgrunn skal kunne bruke programvaren vil den utvikles slik at den er intuitiv å bruke og å forstå. Dette krever et godt brukergrensesnitt som tydelig viser funksjonalitet.

Programvaren vil være avhengig av ulike pakker og det vil da opprettes en installasjonsfil slik at installasjon av disse utføres automatisk.

### **4.2 Forutsetninger og avhengigheter**

Det finnes flere forutsetninger for å utvikle et godt produkt. Språket på chat-samtalene og alfabetet i dette språket er viktig for å kunne få korrekt transkribering. Det vil tas utgangspunkt i norsk hvor man må ta hensyn til de norske bokstavene æ, ø og å, samtidig som programvaren vil fungere på andre språk som bruker det latinske alfabetet.

Grunnet bruk av skjermopptak må det tas hensyn til kvalitet, bilder per sekund og hastighet på skrolling i videoene. Skrolling utføres manuelt og kan variere mellom videoer og innad i videoen. Det er viktig at programvaren fungerer til tross for disse variasjonene.

Chat-applikasjoner kan variere i design ut fra merke på telefon eller maskin, eller versjon av applikasjonen. Programvaren må ikke være avhengig av design for å kunne transkribere.

## 5 PRODUKTETS FUNKSJONELLE EGENSKAPER

Funksjonalitet beskrevet i listen under er noen ganger store oppgaver som igjen krever mer detaljert funksjonalitet fra programvaren.

1) *Automatisk installasjon og innlasting av pakker*

Krever en kjørbær installasjonsfil og at pakker importeres i kildekoden der de er nødvendig

2) *Importering av video som skal transkriberes*

Mulighet for brukeren å velge video som skal lastes opp

3) *Importering av mappe med bilder som skal transkriberes*

Mulighet for brukeren å velge at det er en mappe som skal laste opp, og velge hvilken mappe

4) *Oppdeling av video til bilder*

Oppdelingen må skje på en måte slik at innhold i videoen ikke går tapt og med begrenset overlapp i innhold

5) *Transkribering av tekst på bilder*

Teksten som sendes mellom personer i samtaler må kunne transkriberes med god nøyaktighet

6) *Transkribering av dato og tidsstempler*

Tidsstempler i chat er ofte i annet format enn resten av teksten, programvaren må kunne lese disse av selv om de gjerne med vanskeligere å lese

7) *Lagring av endelig tekst*

Den transkriberte teksten må kunne lagres slik at brukere har tilgang på teksten fra videoen.

8) *GUI for å laste opp fil/mappe*

Må ha et brukergrensesnitt hvor det er mulig for bruker å velge filer/mapper med annerledes plassering enn programvaren

9) *Automatisert kjøring av programvaren*

Etter at filer/mapper er valgt skal programvaren gjennomføre bilderedusering og transkribering automatisk

10) *Lagring av skjermbilder som representerer video*

Programvaren vil redusere antall bilder som representerer video til færrest mulig uten å miste innhold. Disse bildene må også lagres.

11) *Filtrering av bilder i mappe*

Hvis det skal transkriberes for bilder i mappe må disse gå gjennom filtrering for å redusere antallet bilder som transkriberes

12) *Filtrering av bilder fra video*

Programvaren må kunne filtrere bilder hentet ut fra video for å finne innhold uten mye overlapp eller manglende informasjon

13) *GUI for innskriving av søkeord*

Brukeren må kunne gi et søkeord som skal lokaliseres i bildene

14) *Søkefunksjonalitet i programvaren*

Programvaren må automatisk kunne søke gjennom den transkriberte teksten for å finne bilder hvor søkeordet er med

15) *Eksportering av søkeresultat*

Bilder som inneholder søkeordet må kunne eksporteres som PDF

16) *Mulighet for å velge områder av bilder som skal transkriberes*

I skjermopptak er det ofte unødvendig informasjon som ikke skal transkriberes. Bruker kan da velge området hvor teksten er slik at bildene kan beskjæres og transkriberes raskere.

## **6 IKKE-FUNKSJONELLE EGENSKAPER OG ANDRE KRAV**

Listen under beskriver ikke-funksjonelle egenskaper og andre krav satt til produktet.

- Enkelt å bruke uavhengig av bruker sin tekniske bakgrunn
- Effektivt å bruke i forhold til slik oppgavene løses nå
- Tilpasningsdyktig for variasjoner i chat-medium design
- Mulighet for bruk uavhengig av operativsystem
- Leselige resultater
- For å bruke programvaren er det et krav på at python3 må være installert på maskinen

## **7 REFERANSER**

InqScribe. (2022). *InqScribe: Simple Software for Transcription and Subtitling*. Hentet fra InqScribe: <https://www.inqscribe.com/>

WonderShare PDF Element. (2022). *Edit and Extract text from images with OCR*. Hentet fra Wondershare PDF Element: <https://pdf.wondershare.net/ocr-page.html>