



**Kandidatnummer 901**

---

# **En studie av KI-genererte beskrivelser i et fotoarkiv**

**A study of AI-generated descriptions in a photo archive**

**Bacheloroppgave 2024**

**Bachelor i arkivvitenskap**

**OsloMet- Storby-universitetet, Institutt for arkiv-, bibliotek- og informasjonsfag**

# Innhold

1. Innledning.....	4
1.1 Problemstilling .....	4
1.2 Disposisjon .....	5
1.3 Avgrensninger .....	5
1.4 Metode.....	6
2. Litteraturgjennomgang.....	7
2.1 Arkivbeskrivelse .....	7
2.1.1 Metadata .....	8
2.2 Fotoarkiv og forholdet til arkivbeskrivelse .....	9
2.3 Kunstig intelligens .....	11
2.3.1 KI og fotoarkiv .....	13
3. Eksperimentelle studiet.....	15
3.1 TAM-arkiv.....	15
3.2 Parametere for eksperimentet .....	16
3.2.1 Valg av bilder .....	17
3.2.2 Valg av modeller.....	17
3.2.3 Dataanalyse.....	19
3.3 Gjennomføring og resultater .....	19
3.3.1 Gjennomføring objektnivå .....	20
3.3.2 Resultater objektnivå .....	22
3.3.2.1 BLIP.....	23
3.3.2.2 Sammenligning av alle modeller objektnivå .....	26
3.3.3 Gjennomføring serie- og samlingsnivå.....	31
3.3.4 Resultater serie- og samlingsnivå.....	33
4. Drøfting.....	36
4.1 Fordeler med KI-genererte beskrivelser i fotoarkiv .....	36
4.2 Ulemper med KI-genererte beskrivelser i fotoarkiv .....	37
4.3 Søk og gjenfinning.....	38
4.4 Begrensninger til generativ KI .....	39
5. Konklusjon .....	41
Litteraturliste .....	44
Vedlegg.....	47

Python skripter .....	47
Kodeblokk 1 - BLIP .....	47
Kodeblokk 2 - CLIP .....	48
Kodeblokk 3 - NLPconnect .....	49
Tabeller.....	50
Tabell 1.....	50
Tabell 2.....	55
Resultater fra ChatGPT.....	61
Forespørsel 1.....	61
Forespørsel 2.....	62
Forespørsel 3.....	63
Forespørsel 4.....	64

# 1. Innledning

Kunstig intelligens (KI) er langt ifra et nytt konsept, men har fått økt oppmerksomhet og betydning de siste årene. Det teoretiske grunnlaget for KI ble lagt allerede på 1930-tallet, men det er først nå de store gjennombruddene har kommet, noe som ofte omtales som en kommende «AI spring». Populariteten til chatboten ChatGPT, som ble lansert av OpenAI i 2022, trekkes ofte frem som utløsende for allmenhetens økte bevissthet rundt KI. (Stanford HAI, 2023) ChatGPT er et eksempel på fremgangen innen det som kalles generativ KI, teknologier som kan skape nytt innhold basert på eksisterende informasjon.

Digitalisering har skapt store utfordringer for arkivene, som nå står ovenfor et enormt volum av digitalt materiale som ofte er vanskelig å gjenfinne og tilgjengeliggjøre. Mange digitale arkiver blir derfor i praksis låst vekk i det Lise Jaillant kaller «dark archives». Ved å automatisere enkelte arbeidsoppgaver, som for eksempel arkivbeskrivelse, kan KI være nøkkelen til å løse denne problematikken. (Jaillant, 2022, s. 7-8)

Et viktig forskningsområde innen KI er *computer vision*, der hovedoppgaven er for maskiner å forstå verden slik den blir oppfattet visuelt. Det vil si å bygge en modell av verden slik den blir oppfattet visuelt. (Chen, 2018, s. 275) Computer vision kombinert med generativ KI betyr at det nå finnes modeller som kan skape nye bilder basert på tekst, og andre som kan generere tekstbeskrivelser av eksisterende bilder. Derfor passer fotoarkiv seg ypperlig til forskning av automatisk generering av arkivbeskrivelser. Utviklingen innen KI skjer raskt, og samfunnets forventninger til hva denne teknologien kan oppnå virker å øke i takt med fremgangen. Derfor er det viktig med praktiske undersøkelser for å komme frem til realistiske implementeringer i arkivene. Denne oppgaven skal undersøke hvordan dette kan gjøres, spesielt rettet mot digitale fotoarkiv.

## 1.1 Problemstilling

Det er nettopp denne enorme utviklingen som denne oppgaven skal se nærmere på. Dermed er problemstillingen for denne oppgaven: ***har utviklingen av generativ KI kommet til det nivået at arkiver kan bruke det for å automatisk generere beskrivelser i fotoarkiv?***

Målet med denne problemstillingen er å undersøke om KI kan brukes til å automatisk generere beskrivelser i et fotoarkiv, både fra et teoretisk og et praktisk standpunkt. Modeller som kan generere tekstbeskrivelse av bilder finnes, men denne oppgaven skal forsøke å

avklare hvordan disse beskrivelsene passer inn i et arkivfaglig perspektiv. Arkivbeskrivelse skal ikke være beskrivelse for beskrivelsens skyld, men har et klart formål. En av de er å bidra til gjenfinning av arkivmateriale. Et gjenfinningsperspektiv er derfor viktig for denne oppgaven.

## 1.2 Disposisjon

For å kunne svare på problemstillingen består oppgaven av to delmål, den teoretiske og den praktiske. Det første delmålet består av å gjøre rede for hvilken rolle arkivbeskrivelse har, hva KI er og de teoretiske diskusjonene rundt mulige applikasjoner innen arkiv, med et spesielt fokus på arkivbeskrivelse. Det andre delmålet består av å analysere hvilken kvalitet og nøyaktighet KI-genererte bilde beskrivelser har igjennom en praktisk undersøkelse. KI er et relativt nytt forskningsområde innen arkivfaget og utviklingen skjer i høyt tempo. Derfor er det viktig med praktiske undersøkelser. Denne undersøkelsen vil ta for seg fotoarkivet til Tjänstemäns & akademiker (TAM) arkiv. Dette fotoarkivet har jeg fått tilgang til via Lars-Erik Hansen fra TAM arkiv og OsloMet. Det er gjort en undersøkelse i denne fotosamlingen tidligere som en del av en bacheloroppgave levert til OsloMet i 2023. Da ble YOLO (You Only Look Once) algoritmen, en objektidentifiserende algoritme, brukt for å identifisere innhold i bilder. Dette kan da brukes som metadata for å øke arkivbeskrivelsen. (Zeng, 2023) Jeg har fått tilgang til både oppgaven og data fra YOLO opprettet i forbindelse med den. Jeg vil bygge på det som er gjort tidligere, men vil i denne oppgaven gå bort ifra objektidentifiserende algoritmer for å fokusere på KI-modeller som genererer deskriptive beskrivelser av bilder. Videre i oppgaven vil jeg ha med drøftinger fra analysen før oppgaven avsluttes med en konklusjon.

## 1.3 Avgrensninger

Selv om denne oppgaven vil ta for seg en praktisk gjennomføring av KI, er det viktig å presisere at dette ikke er en IT-faglig oppgave. Denne oppgaven retter seg heller mot den fremvoksende retningen *Computational Archival Science (CAS)*. Formålet med dette feltet er å bruke kunnskapen innen både informatikk og arkivvitenskap for å utføre arkivfaglig forskning. Slik kan man få en bedre forståelse av hvordan nye teknologier påvirker opprettelsen, bruken, lagringen og bevaringen av arkivmateriale. (Marciano, 2022) Det er nettopp denne forståelsen som er fokuset i denne oppgaven. Derfor vil ikke oppgaven

inneholde en teknisk gjennomgang av hverken de modellene som blir brukt i undersøkelsen. Kriteriene for valg av modeller er diskutert nærmere i kapittel 3.2.2.

#### 1.4 Metode

Jeg vil bruke en blandet metode for denne oppgaven. For å presentere aktuell teori og forskning i kapittel 2, valgte jeg litteraturgjennomgang som metode. For arkivfaglig forskning er litteraturgjennomgangen god metode for å samle, analysere og sammenligne data på en fleksibel måte. Utvelgelse av kilder til gjennomgangen skjer etter visse kriterier relatert til for eksempel en bestemt tid, teori eller praksis. Kilder velges på bakgrunn av sin autoritet innen det aktuelle forskningsområdet. (Duff & Cumming, 2016 s. 459-460) For denne oppgaven vurderer jeg autorative kilder som fagfellevurderte artikler, standarder, eller kilder som er skrevet av arkivfaglige forskere innenfor det aktuelle temaet.

For kilder som forklarer visse teknologier har jeg brukt lære- og fagbøker rettet mot informatikkstudenter. På grunn av den raske utviklingen har det ikke vært lett å finne kilder som beskriver visse teknologiske aspekter innenfor disse kriteriene. Derfor har jeg hentet flere kilder fra nettet. Her har jeg vurdert at autorative kilder er kilder som er pålitelige og har en viss «name recognition» som for eksempel Stanford og Harvard Business Review. Jeg har også lagt vekt på kilder som spesialiserer seg på informasjon om KI som for eksempel Almodels.fyi.

For å gjennomføre undersøkelsen i kapittel 3, var det mye frem og tilbake om dette egnet seg best som et case-studie eller som en eksperimentell studie. Begge er metoder kan brukes for å svare på spørsmålene «hvorfor» og «hvordan». (Yin, 2018 s. 9) Case-studie er en empirisk tilnærming som brukes for å utforske et bestemt fenomen i sin aktuelle kontekst. Slik skiller denne case-studien seg fra den eksperimentelle metoden, hvor det er ønskelig å skille fenomenet fra sin kontekst. Dette gjøres ved å kontrollere det i et laboratoriemiljø. (Yin, 2018 s. 15) Case studier derimot, skal studere fenomenet i sin naturlige setting. (Choempryong & Wildemuth, 2016, s. 51-52)

Hovedtrekket til eksperimenter er kontroll. Det vil si at alle mulige variasjoner er enten kontrollert, eller at variasjonene inntreffer på en systematisk måte. Når eksperimenter utføres, manipuleres uavhengige variabler for å analysere deres effekt på avhengige variabler. (Wildemuth & Cao, 2016, s. 103-104) Denne undersøkelsen skal bestå av en

sammenligning av flere generative KI-modeller (den uavhengige variabelen) og automatisk generering av beskrivelse. Resultatet er kvaliteten, nøyaktigheten og relevansen på disse beskrivelsene (den avhengige variabelen). Derfor konkluderer jeg at en eksperimentell studie er best egnet metoden for denne oppgaven. Detaljene rundt utformingen av eksperimentet kommer i kapittel 3.2.

## 2. Litteraturgjennomgang

### 2.1 Arkivbeskrivelse

I artikkelen *Origin and Development of the Concept of Archival Description*, skrev Luciana Duranti at arkivbeskrivelse handler om å skrive om arkivmateriale. (Duranti, 1993, s. 47) Denne definisjonen går rett i kjernen av hva denne oppgaven er, men sier lite om hvordan og hvorfor. International Council on Archives har utviklet en standard for arkivbeskrivelse, ISAD(G), og der er begrepet definert som «The creation of an accurate representation of a unit of description and its component parts, if any, by capturing, analyzing, organizing and recording information that serves to identify, manage, locate and explain archival materials and the context and records systems which produced it.» (ICA, 2000, s. 10) Duranti argumenterte i 1993 at arkivbeskrivelse aldri har vært en arkivfunksjon i seg selv, men heller et verktøy for å oppnå andre arkivfunksjoner som bevaring og kommunikasjon av arkivdokumentasjon. (Duranti, 1993, s. 52) Duranti mente også at arkivbeskrivelse på det tidspunktet ikke lenger hadde hensikt å forbedre tilgjengeligheten av arkivmateriale, at dette heller var en oppgave for andre verktøy som registre og indekser. (Duranti, 1993, s. 51) Dette er nok et utdatert syn som ikke reflekterer nåtidens heldigitale hverdag.

Nyere kilder presenterer et annet syn på arkivbeskrivelse. I følge Lois Hamill er dette en av arkivarens kjerneoppgaver. (Hamill, 2017, s. 1) Lina Bountouri hevder til og med at det er den aller viktigste oppgaven, ettersom effektiv gjenfinning og bruk av gjenfinningsverktøy, inkludert de registrene og indeksene som Duranti nevner, er helt avhengig av arkivbeskrivelse. (Bountouri, 2017, s. 1) I følge ISAD(G) standarden er det prosessen med arkivbeskrivelse som gjør det mulig å etablere den intellektuelle kontrollen som er nødvendig for at pålitelige, autentiske, meningsfulle og tilgjengelig dokumentasjon kan føres videre. (ICA, 2000, s. 7)

Arkivbeskrivelse er en oppgave basert på to av de viktigste arkivprinsippene: *respect des fonds*, også kjent som proveniens, og opprinnelig orden. (Pitti & Rush, 2018, s. 1425) På grunn av dette er ordning og beskrivelse to oppgaver nært knyttet hverandre, og kanskje til og med helt avhengig av hverandre. Ordning er prosessen hvor en arkivar organiserer arkivmateriale, først og fremst basert på deres proveniens og deretter deres opprinnelige orden. Ordning av digitalt arkivmateriale bør skje uavhengig av format og bør heller baseres på arkivskaper eller funksjon. Hamill påpeker at i praksis så skjer dette uansett, spesielt med fotografier. Å ordne etter format kan være ønskelig fra et bevaringsståsted, men vil ikke nøyaktig reflektere de intellektuelle sammenhengene og konteksten som arkivmaterialet er en del av. Hamill mener at arkivbeskrivelse er prosessen med å lage arkivkataloger, *finding aids* på engelsk, eller andre verktøy som beskriver den fysiske eller intellektuelle strukturen til et arkiv. Slike verktøy kan da brukes som et hjelpemiddel for å finne frem i materialet. (Hamill, 2017, s. 18)

### 2.1.1 Metadata

Digitalt materiale beskrives i stor grad på samme måte som analogt materiale. Fordi digitalt materiale er avhengig av maskiner og programvare, må det opprettes og registreres tilhørende metadata. Å fange opp slike metadata er en del av arkivbeskrivelsen. (Hamill, 2017, s. 107-108) Metadata er som regel definert som «data om data». Dette er et mye brukt begrep, men kan også være et utydelig begrep som kan forstås på ulike måter. (Gilliland, 2016, s. 5) Metadata kan deles inn i fem kategorier basert på deres funksjoner: administrative-, deskriptive-, bevarings-, tekniske- og bruksmetadata. Opprettelsen og administrasjon av metadata er en relativt kompleks prosess som involverer både manuelle og automatiske trinn. Dette kan utføres innen ulike funksjoner og av ulike individer når som helst i livsforløpet til et informasjonsobjekt. (Gilliland, 2016, s. 17-20)

Det er stor overlapp i bruken av begrepene beskrivelse og metadata. I *Recordkeeping metadata and archival description: a revisit* skriver Jinfang Niu om debatten som har oppstått som følge av dette. Ifølge Niu mener arkivviterne David Bearman og Margaret Hedstrom at i et elektronisk miljø vil all informasjon nødvendig for arkivbeskrivelse bli fanget opp mens materialet fortsatt er i bruk. Alt som er nødvendig er å velge ut de metadataene det er behov for og så er det ikke behov for en etterfølgende beskrivelse. Heather MacNeil derimot argumenter for at metadata ikke kan erstatte arkivbeskrivelser. Metadata er som dagbøker



som registrer daglige hendelser i livsløpet til dokumentasjon innenfor bestemte systemer, mens arkivbeskrivelser er som biografier som oppsummerer livsløpet til dokumenter på tvers av alle systemene i en organisasjon. MacNeil mener at dette er grunnen til at metadata vanligvis er detaljerte, men begrenset i omfang. Arkivbeskrivelse er mindre detaljert, men kan avsløre det større mønsteret i dokumentasjonens liv og sammenhengene mellom dokumenter. (Niu, 2013, s. 208)

I sitt arbeid med samme TAM-fotoarkiv, som er utgangspunkt for denne oppgaven, mente Zeng at «Metadata er mer objektiv og nøytralt i forhold til arkivbeskrivelse.» (Zeng, 2023, s. 8) Bruken av metadata er mer standardisert med spesifikke instruksjoner sammenlignet med arkivbeskrivelse, (Bountouri, 2017, s. 2) og dette kan være begrunnelsen til forskjellen i objektivitet. Videre skriver Zeng at metadata er informasjon om selve objektet som eksisterer innenfor en ordnet struktur. Arkivbeskrivelse gir en helhetlig overblikk over strukturen i en samling ved hjelp av nettopp metadata som systematisk beskriver arkivet. (Zeng, 2023, s. 8) Disse perspektivene om skillet mellom arkivbeskrivelse og metadata er viktig å ha med seg, men ikke nødvendig for å svare på problemstillingen i denne oppgaven. Jeg opplever at fotografier og deres beskrivelser eksisterer litt «utenfor» vanlig arkivmateriale, noe som kommer opp i neste delkapittel. Beskrivelse, slik vi definerer det i vår dagligtale, kan være en del av deskriptive metadata som en skildring av et informasjonsobjekt. Samtidig kan slike metadata være en del av arkivbeskrivelse som definert av arkivfaget.

## 2.2 Fotoarkiv og forholdet til arkivbeskrivelse

Joan Schwartz mener i *Coming to terms with photographs* at hierarkiske beskrivelser av visuelt materiale er ofte dårlig forstått. I et fotoarkiv kan de forskjellige nivåene i et hierarki eksistere mer som intellektuelle konsepter, ofte uten direkte materielle motparter. Et fotoalbum kan både fysisk og intellektuelt tilhøre mer enn en beskrivelse, og flere album kan være en del av en beskrivelse. (Schwartz, 2002, s. 156) Siden TAM er knyttet til fagbevegelsen, kunne et mulig eksempel fra dette arkivet være at fotografier fra ulike fagforeningsmøter er samlet i en mappe, mens fotografier fra en streik kan være delt inn i to separate mapper. Burcak Senturk skriver i *Photographs as archival material* at beskrivelse av fotografier kan skje på tre forskjellige nivåer: samlingsnivå, serienivå eller objektnivå. For å vurdere hvilket nivå som skal beskrives må arkivaren ta hensyn til kvaliteten og kvantiteten av fotografiene, og søkeadferden til brukere. Senturk mener at det å beskrive på samlingsnivå

sikrer bedre fysisk og intellektuell kontroll, men at arkivaren må selv vurdere hvilket nivå det passer best å beskrive fotografier. (Senturk, 2013, s. 625)

I *The Archival Photograph and Its Meaning: Formalisms for Modeling Images* presenterer Allen Benson to fremgangsmåter for beskrivelse av fotoarkiv: beskrivelse på objektnivå eller arkivkatalogen. Beskrivelse på objektnivå blir referert til som «The Bibliographic Record Model» (Benson, 2009, s. 154) på engelsk. Etter denne metoden blir fotografier indeksert etter emne og som regel med annerkjennelse til fotograf, men ofte uten kontekstuell informasjon om deres opprinnelse. Fotoarkiv blir derfor mer som «stock photo» bibliotek hvor bildene blir redusert til sine synlige elementer. Schwartz mener at konsekvensene av dette forverres av digitaliseringsprosjekter med formål å gi tilgang til bilder på internett. (Schwartz, 2002, s. 157)

Theodore Schellenberg mente i sin tid at arkivarer burde beskrive arkivene igjennom arkivkataloger som presenterer serier innenfor større samlinger. Arkivarer burde også unngå å beskrive på objektnivå inntil man har en omfattende oversikt og beskrivelse av arkivbestanden. (Benson, 2009, s. 150) Arkivkataloger er den mest brukte beskrivelsesmetoden, også for å beskrive fotoarkiver, men metoden er ikke uten kritikere. Richard Cox, som sitert av Benson mente at det er usannsynlig at en slik metode «... can represent all the layers, details, nuances, and vagaries that constitute records.» (Benson, 2009, s. 158) Benson mener også at det kan være store utfordringer å bruke arkivkataloger for å finne frem i fotoarkiv, spesielt på nett, og bruker University of Pittsburgh Library System som et eksempel.

De har en samling på rundt 150 000 fotografier og negativer. For å finne frem i samlingen må man enten oppsøke fysiske permer som inneholder utskrifter av arkivbeskrivelsene eller bruke PITTcat, bibliotekets digitale katalogsystem. Denne gir lenker til beskrivelser i EAD-format. (Benson, 2009, s. 159) Encoded Archival Description (EAD) er en internasjonal standard for å strukturere og formatere arkivbeskrivelse i XML. Slik kan arkivmateriale deles og tolkes på en standardisert måte. EAD standardiserer ikke selve innholdet i beskrivelsen, dette brukes ISAD(G) standarden til. (Pitti & Rush, 2018, s. 1423) For å søke i PITTcat-systemet må det brukes boolske operatorer, (Benson, 2009, s. 159) noe en tilfeldig arkivbruker kanskje ikke er like godt kjent med. Benson stiller derfor spørsmål til hva de som bygger systemer basert på arkivkataloger egentlig fokuserer på og siterer Cox igjen: «Is there

anything natural in how archivists construct finding aids, except for the millenia old tradition of making lists that date back nearly to the beginning of writing?» (Benson, 2009, s. 159)

### 2.3 Kunstig intelligens

Kunstig intelligens kan være et utfordrende begrep å definere. John McCarthy en pioner innen KI, erklærte i 1955 at formålet til KI, som sitert av Wolfgang Ertel i *Introduction to Artificial Intelligence*, var å « ... develop machines that behave as though they were intelligent.» (Ertel, 2017, s. 1) En annen definisjon fra Encyclopedia Britannica, igjen sitert av Ertel, lyder som følgende: «AI is the ability of digital computers or computer controlled robots to solve problems that are normally associated with the higher intellectual processing capabilities of humans ... ». (Ertel, 2017, s. 2) Ertel mener at begge disse definisjoner har svakheter. Den første er mangelfull fordi det er lett å tolke atferd fra maskiner som intelligente uten at det trenger nødvendigvis å være det. Ertel peker her til Braitenberg-kjøretøyene som et eksempel, et konsept innenfor robotikk som illustrer hvordan enkle forbindelser mellom sensorer og motorer kan resultere i tilsynelatende kompleks atferd. I følge Ertel kan man nesten definere enhver datamaskin, med sin kapasitet for komplekse beregninger og muligheten til både å memorere og innhente lange tekster, som intelligente under den andre definisjonen. (Ertel, 2017, s. 1-2)

For å virkelig etterligne menneskelig intelligens må maskiner kunne forstå resonnementet som ligger bak atferden vår og de valgene vi tar. Derfor landet Ertel til slutt på Elaine Rich sin definisjon som mest passende: «Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.» (Ertel, 2017, s. 2) Ertel mener at denne definisjonen best oppsummerer hva KI forskere har prøvd å oppnå de siste 50 årene, og den vil også være like aktuell 50 år frem i tid. (Ertel, 2017, s. 2)

Uansett hvilken definisjon som brukes krever forskning på KI kunnskap fra forskjellige fagfelt, som nevrovitenskap og lingvistikk. KI kan dermed ses på som en fellesbetegnelse for forskjellige fremgangsmåter, noe som kan skape forvirring og mangel på forståelse for hva KI faktisk er. Dette er en forvirring som kan ses både i allmenheten og blant forskere. (Jordan, 2019 ) Alle de forskjellige begreper som blir brukt som eksempler på KI kan også bidra til denne forvirringen. På engelsk blir blant annet begreper som *machine learning*, *neural networks*, *large language models*, *natural language processing* brukt som eksempler på KI.

Mennesker er tilpasningsdyktige. Gjennom læringsprosesser er vi i stand til å omstille oss til miljøet vi befinner oss i og endre vår atferd etter behov. Dette er en ferdighet hvor mennesker er helt klart overlegne maskiner, derfor er maskinlæring et sentralt område innen KI forskning. (Ertel, 2017, s. 3) Det er blitt såpass sentralt at Michael I. Jordan i artikkelen *Artificial Intelligence- The Revolution Hasn't Happened Yet*, mener at maskinlæring har bare blitt omdøpt til KI. (Jordan, 2019 ) Maskinlæring går utpå at datamaskiner skal læres opp til å gjenkjenne mønstre ved å anvende datavitenskapelige teorier og algoritmer. Dette er et område som er blitt forsket på over lengre tid og maskinlæring er nå et ganske modent underfelt av KI. (Chen, 2018, s. 272-273)

Modeller som er basert på maskinlæringsalgoritmer kan skille mellom *discriminative* og *generative*. *Discriminative* modeller er presentert med en inngangsvariabel, som for eksempel et røntgenbilde, som skal kobles mot et målrettet utgangsvariabel, som i dette tilfelle kan være en diagnostisk klassifisering. *Generative* modeller har kapabiliteten til å generere helt nye data, som for eksempel tekst, musikk eller bilder. (Babcock & Bali, 2021, s. 3) ChatGPT kan for eksempel produsere lange tekster på et blunk. Det er til og med de som mener at skoleoppgaver generert av ChatGPT kunne fått toppkarakter. Modellen DALL-E kan produsere alle mulige slags bilder på forespørsel. (McKinsey & Company, 2023) Dette er bare noen få eksempler på de generativ KI produktene som er nå tilgjengelig.

Tidligere måtte alle maskinlæringsalgoritmer bygges fra bunnen. Modellene ble bygd basert på data som var tilgjengelig, som ble betegnet som «Ground Truth» av forskere, altså det som kan fastslås med sikkerhet å være sant. (Hodel, 2022, s. 161) Generativ KI-modeller er kjent som *pre-trained models*, eller forhåndstreinte modeller. De er trent på forhånd med et stort datasett og det er det som gjør dem i stand til å utføre de oppgavene de skal gjøre. (AIModels.fyi, 2023) Dette kan by på bias problemer i det materialet som generativ KI produserer. Disse modellene blir trent opp på alle mulige slags kilder skapt av mennesker. Det er en fare for at disse modellene også «arver» fordomsfulle holdninger fra disse, som for eksempel rasisme, sexisme eller homofobi. Safiya Umoja Noble poengterer i sin bok *Algorithms of Opression* at det er lett å anta at slike teknologier er nøytrale. Men det er viktig å forstå at disse er skapt av mennesker og vil da alltid ha menneskelige fordommer. Diskriminering kan være innbakt i selve programvarekoden. (Noble, 2018, s. 1-2) Selv om det kan virke sånn for brukeren, produserer ikke disse modellene materiale fra intet. Grunnlaget

de er trent på er konstruert av programvare som prosesserer enorme samlinger av bilder og tekst. Da reiser det seg et naturlig spørsmål om opphavsrett også gjelder for KI produsert materiale. (Harvard Business Review, 2023) Hvis du spør en chatbot om å generere en tekst, vil mange gjøre det uten å henvise til sine kilder. Selv om den gjør det, er det ikke gitt at forfatterne av materialet har gitt samtykke for at det skal brukes på den måten.

Selv om det har vært en enorm fremgang, er det er usikkert om det er riktig å omtale maskiner som intelligente på dette tidspunktet. Spesielt i sammenligning til menneskers evne til å lære og tilpasse seg. (Hodel, 2022, s. 161) De fremgangene vi har sett til nå er bygget på mønstergjenkjenning, mens maskiner som resonnerer og tenker på høyt nivå er enda ikke utviklet. (Jordan, 2019 ) Det er derimot et stort ønske at maskiner skal være intelligente. Mange ser for seg en utopisk fremtid der maskiner utfører arbeidet, mens vi mennesker tar livet med ro. (Ertel, 2017, s. 12) Det er kanskje disse forventningene som ligger bak KI som et «buzzord». Det er kanskje dette som også har ført til at gamle konsepter som maskinlæring har blitt relansert som KI.

### 2.3.1 KI og fotoarkiv

Når man diskuterer KI er det naturlig å reflektere over intelligensens natur, og knytte KI med evnen til å utføre komplekse og utfordrende oppgaver. Litt paradoksalt er det da at KI sin verdi for arkivene er ikke i å løse vanskelige oppgaver. I stedet ligger verdien i å automatisere oppgaver som, selv om de ikke er kompliserte, er for tidskrevende for mennesker å utføre. (Jaillant, 2022, s. 14) Arkivarene står ovenfor et massivt omfang av digitalt materiale som skal beskrives. Arbeidsmengden kan ramme kvaliteten på beskrivelser og føre til etterslep. For å håndtere dette etterslepet er det mange arkivarer som lener seg på More Product, Less Process (MPLP) metoden. MPLP begrenser beskrivelse til kun det aller nødvendigste. (Waugh, 2019, s. 20) Under slike forhold er det er naturlig å vurdere KI som en løsning.

Hvis arkivarer skal bruke KI for å erstatte tradisjonelle beskrivelses metoder i fotoarkiv, må det vurderes hva som egentlig skal erstattes og hva som skal beskrives. Benson stiller spørsmål til hva som er den viktigste referansen, er det fotografiet selv, teksten som beskriver fotografiet, eller subjektet(ene) representert i fotografiet? Den mest nøyaktig representasjonen av et objekt er selve objektet, og kopier og surrogater vil alltid være uperfekte. Videre mener Benson at arkivarer bør tenke på nytt hvordan de jobber med å representere fotografier og peker til arkivkataloger som bare angir at «mange fotografier» er

en del av en samling. Da er det stor risiko at for at fotografiene ikke blir nøyaktig representert. (Benson, 2009, s. 171) Akkurat hvordan arkivarer kan bruke KI for å oppnå nøyaktige representasjoner er veldig usikkert siden det er gjort lite forskning på dette området.

Det er gjennomført noen prosjekter som bruker KI i fotoarkiv. Arkivverket har nylig lansert en søkeløsning for bilder i Digitalarkivet som kombinerer både treff i tekstlige beskrivelser og KI som indentifiserer innhold i bilder. Søk i digitale fotoarkiv har tidligere vært helt avhengig av tilhørende metadata, som ofte kan være av dårlig kvalitet eller mangelfulle. Med denne kombinasjonsløsningen kan brukere få treff selv på bilder som mangler metadata. (Aksess, 2024) Arkivverket omtaler dette som «starten på det nye Digitalarkivet», og vil etterhvert utvide tjenesten med søk i dokumenter, lyd og video. (Arkivverket, 2024) Arkivverket har her valgt å bygge systemer der det viktigste referansepunktet er subjektet(ene) i bildesamlingen for å gi brukere treff på objektnivå. Løsningen kan eksistere sammen med tradisjonelle metoder for beskrivelse, uten å erstatte det helt. Samtidig er dette kanskje et eksempel på de digitale løsningene som Schwartz har kritisert. Det er veldig positivt at slike løsninger gjør det mulig for brukere å få bedre tilgang til fotoarkiv, men brukerne mister tilgangen til kontekstualiserende informasjon. Geoffrey Yeo på sin side mener at slike tilnærminger med treff på objektnivå ikke bare er i tråd med den nåværende teknologiske utviklingen, men også med moderne arkivbrukere sine ønsker. Fremtidens arkivbrukere ønsker nok ikke arkivene fremstilt som statiske samlinger i faste aggregater. (Yeo, 2015, s. 179)

Forskere ved Cornell University, Stanford University og University of Toronto har samarbeidet med Frick Art Reference Library (FARL) for å utforske hvordan KI kan forbedre tilgjengeligheten og gjenfinning i deres fotoarkiv. FARL har en samling på rundt 1,2 millioner bilder bestående av reproduksjoner av kunstverk. De ønsket å utnytte *deep convolutional neural networks* for automatisk annotering av bilder i samlingen basert på overskriftene som brukes i arkivets klassifiseringssystem. (Han et al., 2022, s. 29) Resultatet av denne forskningen viser at slike nettverk kan tilpasses for å klassifisere bilder i samsvar med spesialiserte hierarkiske klassifikasjonssystemer. Dette viser også fordelene med samarbeid mellom KI-forskere og kulturminneinstitusjoner. (Han et al., 2022, s. 57) FARL har nå potensielt et verktøy som kan håndtere en voksende mengde bilder som er digitalisert, men ennå ikke klassifisert. (Han et al., 2022, s. 29)

Som nevnt i innledningen, er det allerede gjort en undersøkelse i TAM sitt fotoarkiv tidligere som en del av en bacheloroppgave levert i 2023. Formålet med oppgaven var å undersøke automatisk produksjon av metadata ved hjelp av YOLO-algoritmen. Dette er også en *convolutional neural network* som ble først utviklet i 2015-16 og har blitt oppdatert siden. YOLO opererer ved å separere objektene på et bilde i bokser og deretter assosiere de i klasser. Dessverre skaper denne metoden begrensninger fordi den kan bare identifisere en type objekt for hver boks. (Zeng, 2023, s. 16) De objektene som ble identifisert hyppigst var person(er), slips, kopper, vinglass, bord, stoler, bord, bøker, tv, potteplanter, sykler, motorsykler, koffert og vesker. Algoritmen identifiserte en eller flere personer i 85% av bildene. Selv om den ikke ble brukt for ansiktsgjenkjenning i dette tilfelle, er det den oppgaven den er trent opp til og vil derfor være god på å gjenkjenne personer. (Zeng, 2023, s. 21)

Zeng fant både potensiale og svakheter i bruken av slik teknologi for å produsere metadata. Først og fremst er det tidsbesparende, og det er mulig å tillegge metadata til materiale selv uten kontekstuell informasjon. Dette kan også gjøres på hvilket som helst tidspunkt i arkivprosessen. De eksisterende metadataene til TAM-arkivet er menneskeskapte deskriptive beskrivelser. Dette er en tidskrevende prosess, men gir inngående kunnskap om hendelsene og historiene rundt bildene som de automatiske genererte metadataene ikke kan. (Zeng, 2023, s. 24-25) Ved å bruke de resultatene fra tidligere undersøkelser som er gjort i samme fotoarkiv kan man få en idé om utviklingen av KI som har skjedd i det siste. Slike *convolutional neural networks* er god på mønstergjenkjenning som kan brukes for klassifisering og objektidentifisering. Generativ KI kan produsere helt nytt innhold som for eksempel deskriptive beskrivelser av bilder. Jeg har ikke lyktes med å finne forskning som bruker generativ KI i arkiv, antageligvis på grunn av at teknologien er så ny.

### 3. Eksperimentelle studiet

#### 3.1 TAM-arkiv

Tjänstemäns & akademikers (TAM) arkiv er en ideell organisasjon som har som oppgave å bevare og tilgjengeliggjøre sine medlemsorganisasjoners arkivmateriale. Medlemmer av TAM er fagforeninger på nasjonalt nivå innenfor tjenestemannsområdet og akademikerorganisasjoner. (TAM-Arkiv) TAM har et fotoarkiv på 6403 digitale bilder i en

Microsoft Access database kalt *PICTOR* som jeg har fått tilgang til. Denne har vært svært utfordrende å jobbe med, den er veldig treg og henger seg ofte opp. Derfor ble det foretatt en databasedump med programmet *Database Preservation Toolkit*, og metadataene i databasen ble migrert til to CSV filer. Den ene filen inneholder metadataene knyttet til bildene og den andre er en liste av arkivskapere med hver sin tallkode. Metadataene tilhørende bildene består av arkivskaper, bildenavn, negativer (true/false), sted, årstall, beskrivelse og publikasjon. Bildenavnene består av tallkoden knyttet til arkivskaper etterfulgt av bindestrek og fire tall. Et eksempel på dette er bilde 009-0051, der 009 er koden til arkivskaperen, i dette tilfelle Finansförbundet. Siste delen av bildenavnet, 0051, er den unike koden som tilhører dette spesifikke bildet. I undersøkelsen ble bare CSV filen med metadata tilhørende bildene brukt, siden denne også identifiserer arkivskaper for hver bildepost. Bildene fikk jeg tilgang til via en separat filoverføring.

Beskrivelsen av bildene gir utfyllende informasjon der konteksten rundt som regel er godt forklart, og sted og personer er ofte identifisert med navn. Dette gir et godt sammenligningsgrunnlag for undersøkelsen. Bildearkivet består av totalt 28 serier organisert etter arkivskaper. 18 av seriene har under 10 bilder, to av seriene har under 100 bilder, 6 av seriene har hundrevis av bilder og to av seriene har over tusen bilder. Den største serien er fra arkivskaperen Svenska Industrijänstemannaförbundet (SIF) med 2093 bilder. Bildene innenfor en serie kan strekke seg over flere år og ha et variert innhold. Metadataene og beskrivelsene er på objektnivå, det er ingen overordnet beskrivelse av dette arkivet i sin helhet. Hver serie i arkivet fremstår heller som en løs samling av bilder skapt av en arkivskaper over tid.

### 3.2 Parametere for eksperimentet

Siden både den forrige undersøkelsen og de opprinnelige beskrivelsene er på objektnivå vil jeg også først og fremst ha fokus på det. Det forrige studiet brukte en blanding av kvantitativ og kvalitativ metode. Kvantitativ analyse ble brukt for å identifisere mønstre i metadata generert av YOLO-algoritmen, mens kvalitativ analyse ble brukt for å tolke fenomener og temaer som oppstår fra algoritmen og dens resultater. (Zeng, 2023, s. 5-6) Dette er en praktisk tilnærming for å håndtere store datasett på en effektiv måte. For å svare på problemstillingen i denne oppgaven må jeg gå mer i dybden for å si noe om kvaliteten på KI-genererte beskrivelser. En kvalitativ analyse er mest hensiktsmessig i dette tilfellet.



Metodene for en slik analyse kan variere avhengig av metodologi, forskningstradisjon og fagfelt. Det er ikke nødvendigvis en korrekt måte å utføre en kvalitativ analyse på; det viktigste er å finne metoder som lar deg reflektere over dataene. Eksempler på slike metoder inkluderer å sortere og analysere materialet for å identifisere mønstre og temaer, isolere disse for videre datainnsamling, gradvis utarbeide generaliseringer fra dataene og deretter konfrontere disse generaliseringene med etablerte teorier. (Lester et al., 2020, s. 96) En analyse av beskrivelser av over seks tusen bilder ville være for tidskrevende for omfanget av denne oppgaven. Derfor valgte jeg heller å plukke ut et antall bilder som er representative for hele fotoarkivet. Jeg konkluderte med at 40 bilder var et passende antall for analysen. Dette antallet gir et helhetlig blikk på bildene uten at det blir for tidskrevende å analysere.

### 3.2.1 Valg av bilder

For å kunne velge ut bilder som skal representere hele fotoarkivet, skaffet jeg meg først et overblikk av alle bildene. Arkivet består for det meste av fotografier med unntak av noen avisutklipp, pamfletter, tegninger og avbildning av dokumenter. Det er i hovedsak svart-hvitbilder, med et knippe fargebilder. Hovedmotivene i bildene er mennesker. Det er mange portretter, gruppebilder, mennesker i arbeid, bilder fra konferanser, taler, o.l. Det er også noen bilder av landskap, bygninger og båter. Bildene spenner helt fra slutten av 1800-tallet til 1990-tallet. Jeg prøvde å fange tidsspennet i valget av bilder, men ville unngå å bruke bilder der individer som fortsatt lever i dag kunne bli identifisert. Av personvern hensyn kan det derfor være litt skjevhet av eldre bilder. Jeg vurderte også å velge ut bilder basert på arkivskaper, slik at alle arkivseriene er representert. Det er stor forskjell i antall bilder og en del serier har bare 1-3 bilder. Dette kan føre til et ubalansert utvalg som ikke er representativt for hele arkivet. Derfor baserte jeg valget kun på innholdet i bildene. Av de 40 bildene valgte jeg kun ett fargebilde, resten er i svart-hvit. Som i resten av bildearkivet er hovedmotivet i de utvalgte bildene mennesker i ulike sammenhenger. 21 av bildene er bilder av to eller flere mennesker og 7 av bildene er av enkeltindivider. Resten av fotografiene er av objekter og landskap. Jeg valgte også ut noen ikke-fotografier, en tegning, et avisutklipp og et avbildet dokument.

### 3.2.2 Valg av modeller

I valg av modeller var Hugging Face den viktigste ressursen. Dette er en plattform og et felleskap som utviklere kan bruke i oppbyggingen, distribueringen og treningen av

maskinlæringsmodeller. Brukere kan bla igjennom modeller og datasett som andre har lastet opp. Fordi Hugging Face lar utviklere dele og teste arbeidet sitt åpent, blir plattformen ofte sammenlignet med GitHub, at det er «GitHub for maskinlæring». (TechTarget) Det er flere hundre *image-to-text*, eller bilde-til-tekst, modeller tilgjengelig på Hugging Face, derfor måtte jeg først etablere noen kriterier. Som nevnt tidligere, dette er ikke en IT-faglig oppgave, derfor er ikke valg av modeller basert på tekniske krav. For at jeg som arkivvitenskapstudent med noe programmeringskunnskap skal klare å gjennomføre denne undersøkelsen, er det viktig å velge modeller med god dokumentasjon og grundig instruksjoner/koder for implementering. Jeg baserte meg til en viss grad på «name recognition», at det er selskaper knyttet til utviklingen som jeg kjenner igjen. Dette gir modeller en større grad av troverdighet. Det siste kriteriet er at programvare som brukes er gratis.

Etter disse kriteriene var det tre modeller som pekte seg ut: BLIP (Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation), CLIP (Contrastive Language-Image Pre-training) og NLPconnect. Dette er modeller bygd på Computer Vision. BLIP er knyttet til Salesforce, et godt kjent amerikansk IT-selskap og CLIP er utviklet av OpenAI, selskapet bak ChatGPT, for KI-forskning. Det ble også publisert vitenskapelige artikler av utviklingen av begge disse modellene. NLPconnect er utviklet av en privat person og er ikke knyttet til et større selskap. Allikevel er dette en modell med grundig informasjon om både oppbygning og implementering. Valget av disse tre betyr ikke at det er de eneste modeller som passer innenfor disse kriteriene, men det er nødvendig å begrense antall modeller for å kunne gjøre en grundig analyse. Bruken av flere modeller handler ikke om å gjøre en vurdering av hvilken som er best, men for å kunne gjøre generalisering om det nåværende nivået på bilde-til-tekst modeller.

Kildene brukt i litteraturgjennomgangen i kapittel 2 la vekt på beskrivelser i de forskjellige nivåene innenfor et fotoarkiv. Dette bildearkivet består av tre nivåer: objektnivå (individuelle bildene), serienivå (alle bildene til en arkivskaper), samlingsnivå (hele fotoarkivet til TAM-arkiv). Selv om fokuset i oppgaven er beskrivelser på objektnivå, er jeg nysgjerrig på om det er mulig å bruke disse beskrivelsene videre for å lage beskrivelser av de øvrige nivåene i dette arkivet. Dette kan også være en oppgave for generativ KI. Her pekte OpenAI sine GPT modeller ut, spesielt siden webapplikasjon deres ChatGPT er veldig enkel å bruke og er god på tekstgenerering. I begynnelsen av denne undersøkelsen var det bare GPT3.5 som var

tilgjengelig gratis for brukere. Denne har begrensninger på antall tegn som kan brukes i en forespørsel og større tekster må brytes ned i flere forespørsler. 13. mai 2024 ga OpenAI begrenset adgang til den kraftigere GPT4o for gratis brukere. Disse brukere kan nå laste opp filer som en del av forespørselen, som kan analyseres av ChatGPT. (OpenAI, 2024b) Dette betyr at hvis jeg har en fil med alle beskrivelsene generert, kan jeg be ChatGPT om å gi meg en overordnet beskrivelse av serien eller bildearkivet basert på det.

### 3.2.3 Dataanalyse

Analysen på serie- og samlingsnivå vil bare bestå av å trekke generelle konklusjoner ut ifra de forespørslene som jeg gir til ChatGPT. Hovedfokuset vil være på objektnivå. Ved å basere analysen på et utvalg av bilder har jeg muligheten til å gå nøye igjennom hvert enkelt bilde. Jeg vil sammenligne innholdet i hvert bilde med de opprinnelige beskrivelsene, YOLO-metadataene og beskrivelsene fra denne undersøkelsen. Dette gir meg mulighet til å studere utviklingen innen maskinlæring og sammenligne YOLO-algoritmen og bilde-til-tekst modeller. Samtidig kan jeg gjøre en grundig sammenligning mellom de opprinnelige beskrivelsene og KI-genererte beskrivelser. Dette kan forhåpentligvis gi et godt grunnlag for å vurdere nivået til de generativ KI-modellene jeg har valgt. Dette kan gi et generelt bilde av generativ KI sitt potensiale, slik den er i dag, for å generere beskrivelser i et fotoarkiv.

### 3.3 Gjennomføring og resultater

Skriptene brukt for å kjøre modellene er kodeblokker hentet fra Hugging Face. For å implementere modellen CLIP brukte jeg også kode hentet fra Google Colab. Det var nødvendig med noen modifikasjoner, blant annet for å angi riktig mappesti for bildene og for å sørge for at dataene ble eksportert til en CSV fil. Skriptene er i programmeringsspråket Python. Bruken av Python er utbredt på grunn av de forskjellige bibliotekene, også kalt pakker, tilgjengelig med kode utviklet for å løse spesifikke problemer. Det er et stort antall pakker tilgjengelige gratis for ulike formål, inkludert maskinlæring. (Horstman & Necaie, 2016, s. 6)

Alle tre kodeblokker er inkludert som vedlegg til denne oppgaven, sammen med nettadressen de er hentet fra. Planen var opprinnelig å gjennomføre undersøkelsen på min private PC. Dette er en bærbar som kjører operativ systemet Windows. Den kan best beskrives som en datamaskin i mellomsjiktet som er god på å håndtere helt vanlige oppgaver. Med lite RAM og uten grafikkort er den ikke det beste valget for oppgaver som krever

intensiv grafikk eller kraftig prosessering. Etter hvert fikk jeg tilgang til en annen, kraftigere PC. Dette er en stasjonær PC som kjører Linux, med mer RAM og et NVIDIA grafikkort i mellomstjiktet. Med et slik grafikkort er det mulig å kjøre skriptene på GPU istedenfor CPU, noe som vil spare tid.

Det som var mest utfordrende med denne undersøkelsen var å sette opp miljøet riktig for hver modell. Flere modeller kan bruke de samme bibliotekene, men i forskjellige versjoner. For å løse denne problematikken ble verktøyet Poetry brukt. Poetry lar brukere presisere hvilket Python bibliotek, og hvilken versjon, et prosjekt er avhengig av. Verktøyet vil da administrere disse bibliotekene for deg. Ved bruk av Poetry er det viktig at det installeres i et dedikert virtuelt miljø for å isolere det fra resten av systemet. (Poetry)

Alle modellene produserer beskrivelser på engelsk. Jeg har valgt å ikke oversette disse i analysen både for effektivitetens skyld og for et mer objektivt resultat. For å analysere resultatene organiserte jeg de i tabeller. Disse tabellene, Tabell 1 og Tabell 2, er også inkludert som vedlegg til denne oppgaven. Der de opprinnelige beskrivelsene fra PICTOR databasen er med i tabellene, har jeg tatt de direkte fra filen fra databasedumpen. Der enkeltpersoner er identifisert med navn har jeg tatt bort hele navnet og erstattet med initialer. For bedre lesbarhet har jeg også ryddet opp i beskrivelsene der det har vært nødvendig. Dette var i tilfeller der det manglet en komma mellom navn eller manglet mellomrom mellom ord.

For siste steg i undersøkelsen måtte jeg ta i betraktning at ChatGPT har begrensninger i bruken av GPT4o modellen for gratis brukere. Dette gjelder spesielt antall ganger det er mulig å laste opp filer innen en bestemt tidsperiode. Derfor måtte jeg gjøre noen avgrensninger. Dette steget var mest for å se på mulighetene, og det er dermed ikke nødvendig med et stort datagrunnlag for å gjøre sammenligninger på tvers av modeller. Derfor valgte jeg å bruke data fra kun en modell for å få en beskrivelse av hele arkivet og på serienivå valgte jeg kun en serie. Resultatene generert av ChatGPT er også inkludert som vedlegg.

### 3.3.1 Gjennomføring objektnivå

Første steget i undersøkelsen var å generere beskrivelser med de tre modellene jeg hadde valgt på utvalget av 40 bilder. Hugging Face har utviklet et eget Python-bibliotek

*Transformers*, som forenkler nedlastningen av modellene som er tilgjengelig på sin plattform. (TechTarget) Dette betyr at så lenge jeg har lastet ned de pakkene som er nødvendig, inkludert Transformers, trenger jeg bare å spesifisere hvilken modell som skal brukes i koden. Her er et eksempel fra Kodeblokk 1, skriptet brukt for å kjøre modellen BLIP:

```
import os
import csv

from PIL import Image
from transformers import BlipProcessor, BlipForConditionalGeneration

processor = BlipProcessor.from_pretrained("Salesforce/blip-image-captioning-base")
model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-image-captioning-base")
```

#### Utdrag fra Kodeblokk 1

BLIP er den første modellen som ble testet. Koden for implementering var bygd opp slik at man får to beskrivelser: en «conditional» og en «unconditional», altså en betinget og en ubetinget beskrivelse. Et utdrag fra koden:

```
#conditional beskrivelse
text = "an image of"
inputs = processor(raw_image, text, return_tensors="pt")
out = model.generate(**inputs)
beskrivelse1 = processor.decode(out[0], skip_special_tokens=True)
print (img_file + " " + beskrivelse1)

#unconditional beskrivelse
inputs = processor(raw_image, return_tensors="pt")
out = model.generate(**inputs)
beskrivelse2 = processor.decode(out[0], skip_special_tokens=True)
print(img_file + " " + beskrivelse2)
```

#### Utdrag fra Kodeblokk 1

I den første linjen av dette kodeutdraget står det «text = «an image of», det er dette som er betingelsen i den genererte beskrivelsen. I den opprinnelige koden stod det egentlig «a photography of». For å undersøke om en endring av betingelsen ville gi et annet resultat, kjørte jeg tre forskjellige versjoner av dette skriptet på utvalget av 40 bilder. Det eneste jeg endret var betingelsen. Jeg brukte «a photography of», «an image of» og «an archival image of». Det er sikkert flere betingelser jeg kunne ha testet, men det ville vært for tidskrevende å analysere. Resultatene er organisert i Tabell 1, her er et utdrag:

Bildenavn	Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
007-0033	I. G, SIF, S. A., SIF, B. H., SIF och H. F., SIF under PTK-kongressen 2-3 juni 1983	two men sitting in chairs looking at something	a photography of the first session of the united states congress committee, in the united states, july	an image of two men sitting in a lecture hall	an archival image of two men sitting in a lecture hall
009-0051	SBmfs representantskap. Middagen.	a group of people sitting at a table	a photography of a group of men and women seated at a table	an image of a group of people sitting at a table	an archival image of a group of men sitting at a table
009-0070	SBmfs representantskaps middag. I.B. opptråder med ett nummer ur Stockholmarnas bankrevy av banktjänsteman G W. W.	a woman in a white dress and hat playing a guitar	a photography of a woman playing a guitar	an image of a woman playing a guitar	an archival image of a woman in a sailor's uniform playing a guitar

#### Utdrag fra Tabell 1

Som det er mulig å se fra dette utdraget er det forskjeller basert på hvilken betingelse som BLIP blir gitt. For gjennomkjøringen av BLIP ble min egen PC brukt. Det tok omtrent 15 minutter for å generere beskrivelser for utvalget på 40 bilder. For de neste modellene CLIP og NLPconnect brukte jeg den kraftigere stasjonære PC-en. Den brukte bare noen minutter på å generere beskrivelser for dette utvalget. For dette stadiet har jeg da de tre gjennomkjøringene med BLIP. Dette har gitt meg fire sett med beskrivelser fra denne modellen, tre betinget og en ubetinget. Jeg har også beskrivelsene av utvalget generert av CLIP og NLPconnect. Resultatene blir analysert i neste delkapittel.

#### 3.3.2 Resultater objektnivå

På grunn av variasjonen i betinget beskrivelsene generert av BLIP har jeg tatt med en egen analyse av resultatene fra denne modellen. Analysen er basert på resultatene i Tabell 1. Det er på bakgrunn av denne analysen at jeg velger de beskrivelsene som skal videre i sammenligningen av alle modellene. Den analysen er basert på resultatene i Tabell 2.

### 3.3.2.1 BLIP

For å undersøke forskjellene mellom de fire settene med beskrivelser generert av BLIP, organiserte jeg alle beskrivelsene i tre kategorier: ingen eller ubetydelig forskjell, beskjeden forskjell og betydelig forskjell. Her er det forskjellen mellom beskrivelsene, og hvor mye de ligner på hverandre, som jeg er interessert i. Ikke nøyaktigheten i beskrivelsen. 12 bilder falt under kategorien ingen eller ubetydelig forskjell. Her kommer et par eksempler på dette:



**Bilde 055-0021**

Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
Klubbordförande för Sv. Philips. PM för december i SIF-tidningen. D. H.	a man sitting on a bench in front of a television	a photography of a man sitting on a bench in front of a television	an image of a man sitting in front of a television	an archival image of a man sitting on a bench in front of a television

[Utdrag fra Tabell 1](#)

Her er beskrivelsene alle en variasjon av «a man sitting on a bench in front of a television». En beskrivelse manglet ordet «bench», dette er et eksempel på det jeg anser som en ubetydelig forskjell. Bare fordi BLIP beskriver et bilde ganske likt i alle fire tilfeller, betyr ikke at den er korrekt i sin tolkning av bildet. Et eksempel på dette er følgende bildet av en stor gruppe mennesker samlet rundt en maistang eller midsommarstang. I alle beskrivelsene fra BLIP ble stangen identifisert som et tre:



**Bilde 404-0578**

Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
Skollovskoloni vid midsommarstången.	a group of people standing around a tree	a photography of a group of people gathered around a tree	an image of a group of people standing around a tree	an archival image of a group of people standing around a tree

[Utdrag fra tabell 1](#)

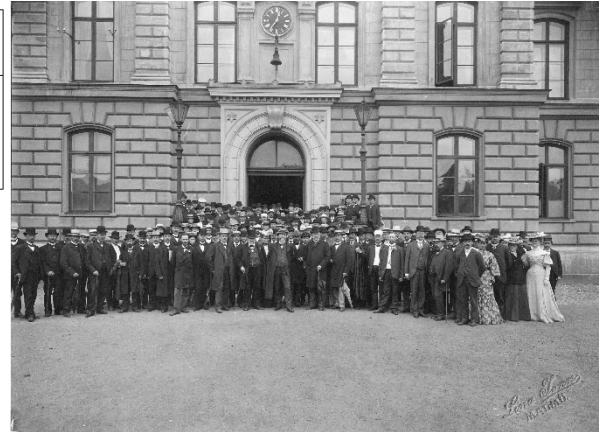
Den største kategorien var beskjeden forskjell, 18 bilder ble plassert i denne kategorien. Typiske eksempler er bilder som involverer objekter eller andre ting som modellen kan ha vanskeligheter å tolke, som for eksempel store grupper av flere kjønn. Beskrivelser som

egentlig handler om det samme, men er formulert ganske annerledes faller også under her.

Et par eksempler fra denne kategorien:

Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
Svenska lärarmötet i Malmö. I mitten längst fram E. H.	a group of people standing in front of a building	a photograph of a group of men and women in front of a building	an image of a group of people standing in front of a building	an archival image of a group of men in front of a building

[Utdrag fra Tabell 1](#)



Bilde 404-0129



Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
SIF:s 40-års jubileum. SIF 100.000 medlemmar. G. T., sekr SIF, Stockholm ställer in ett stämpelur från sekelskiftet.	a woman is putting a clock on the wall	a photograph of a woman working on a clock	an image of a woman holding a clock	an archival image of a woman working on a clock

[Utdrag fra Tabell 1](#)

Bilde 055-0789

10 bilder ble klassifisert i kategorien betydelig forskjell. Et av disse bildene er av et dokument, som BLIP hadde vanskeligheter med å tolke. Med resten av bildene er det snakk om tilfeller der BLIP har funnet på ting. Dette er et kjent problem med generativ KI og blir ofte referert til som «hallusinasjoner». Dette skjer fordi slike modeller tolker mønstre eller objekter som ikke eksisterer, eller mennesker ikke kan oppfatte. Dette fører til unøyaktige resultater, eller resultater som er rent sludder. Dette kan skyldes blant annet overtilpasning, skjevhet i treningsdata eller høy modellkompleksitet. (IBM) Et eksempel på hallusinerer fra undersøkelsen:

Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
SIFs Riksstämman Från vänster: N. K.-E., L. P., K. B., A. H., E.S.	a group of men sitting at a table with flags	a photograph of the first meeting of the first world war in the united states, june 1945	an image of a group of men sitting at a table	an archival image of the first world war conference

[Utdrag fra Tabell 1](#)



Bilde 055-0090



Selv hvis ikke de opprinnelige beskrivelsene hadde vært tilgjengelig, vet jeg med en gang at beskrivelsen «meeting of the first world war in the united states, june 1945» eller «the first world war conference» ikke stemmer. Dette er ikke det eneste tilfellet der BLIP forsøker å sette en tidskontekst på bildene. Noen andre eksempler fra denne kategorien:

Bildenavn	Opprinnelig beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
055-1041	Tidigare byggnad på samma fastighet som det nya SIF-huset fördes upp och invigdes 1954.	a black and white photo of a building	a photography of the building at the corner of the street in the early 1900s	an image of a building in the snow	an archival image of the building at the corner of the street
149-0709	Demonstration för högre löner 22 maj 1969. Sjuksköterskelever demonstrerar för bättre löner.	a group of people holding signs on the street	a photography of a demonstration by the british labour labour movement in london, 1968	an image of a group of people holding signs	an archival image of a demonstration by the british labour labour party in london, england, in the

[Utdrag fra tabell 1](#)

Det er interessant at BLIP forsøker å sette bildene i en tidskontekst, fordi Zeng oppdaget at YOLO-algoritmen identifiserte objekter som svært usannsynlig fantes i den tidsperioden bildet ble tatt. Dette gjaldt spesielt begrepene «bærbar pc» eller «mobil». (Zeng, 2023, s. 22) Det virker som tidskontekst er en særlig svakhet for KI.

Jeg fant en viss grad av hallusinasjoner i alle de tre kategoriene. Det som skiller seg ut i den siste kategorien er at det er ganske grove tilfeller. I de andre kategoriene kunne BLIP for eksempel gjette seg til at en gruppe sykepleiere står foran et usynlig sykehus, eller at noen drikker vin fra et glass selv om det er umulig å tyde hva slags væske er i glasset. I de grove tilfellene prøver BLIP å gjette seg frem til et tid og sted som ikke stemmer. Det virker fra disse eksemplene som betingelsene «a photography of» og «an archival image of» er spesielt utsatt for hallusineringer. Jeg har ikke funnet eksempler av grove hallusineringer i «unconditional» og «an image of» beskrivelsene.

For å kunne sammenligne beskrivelsene fra BLIP med de andre modellene, bestemte jeg meg for å ta med «unconditional» eller ubetinget beskrivelsene. Jeg måtte også jeg velge et sett med betinget beskrivelser å ta med videre. Jeg bestemte meg for å ta med beskrivelsene med «an image of» betingelsen, siden dette arkivet inneholder både fotografier og andre typer bilder. Med tanke på at 30 av de 40 bildene falt i kategoriene ingen forskjell/ubetydelig

forskjell eller beskjeden forskjell, anser jeg det som uproblematisk å fortsette analysen med kun disse to beskrivelsene fra BLIP. Jeg vil uansett ta med meg problematikken med hallusinerer med videre i analysen.

### 3.3.2.2 Sammenligning av alle modeller objektnivå

For å kunne vurdere kvaliteten og nøyaktigheten til beskrivelsene generert av de tre modellene har jeg samlet de generert av YOLO og de opprinnelige beskrivelsene for sammenligning. For å kunne si noe om nivået på beskrivelsene generert av de tre modellene i undersøkelsen er det tre egenskaper jeg vil se etter: stemmer beskrivelsene overens med innholdet i bildet, hvordan er kvaliteten sammenlignet med YOLO beskrivelsene og hvordan er kvaliteten sammenlignet med de opprinnelige beskrivelsene.

Det er 9 tilfeller hvor alle tre modeller har gitt korrekt beskrivelse av innholdet i bildet. Her er det viktig å merke seg at korrekt betyr at beskrivelsen generert ikke inneholder feil, ikke nødvendigvis at modellen har klart å beskrive alle momentene i bildene. Dette gjelder spesielt gruppebilder, hvor modellene klarer å gi generiske beskrivelser som er egentlig helt korrekt, men sier veldig lite om bildet. Et eksempel på dette:



**Bilde 055-0480**

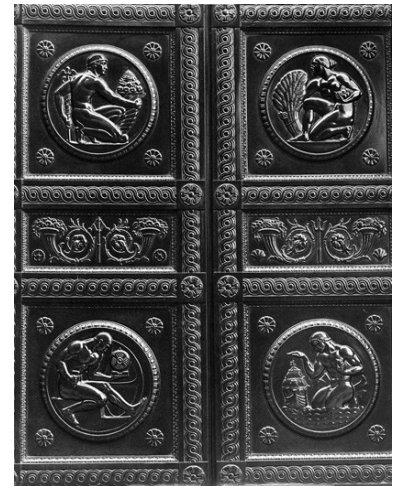
Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Personal på FK och FA 16.	11 persons, 11 ties	a group of people posing for a picture	an image of a group of people posing for a picture	a group of people sitting and standing together .	a large group of people posing for a picture

#### Utdrag fra Tabell 2

Det er bare 3 tilfeller der alle tre modellene har generert beskrivelser som ikke stemmer med innholdet i bildet i det hele tatt. En av de var bilde 404-0578, vist tidligere i delkapittel 3.3.2.1, av en gruppe samlet rundt en midsommerstang. BLIP klarte ikke å beskrive dette bilde, og det lykkes ingen av de andre modellene heller. CLIP beskrev den som «a group of people standing around a cross», mens NLPconnections beskrev den som «a crowd of people standing around a fire hydrant». En av de andre bildene av modellene klarte å beskrive riktig er bilde 009-0478:

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Exteriör Portalen till Sv. Handelsbanken Huvudkontoret Stockholm	2 clocks	a black and white photo of a metal ceiling	an image of a black and white photo of a set of four metal plaques	a black and white photo of a metal panel .	a clock that is on top of a wall

[Utdrag fra Tabell 2](#)



Bilde 009-0478

Jeg synes det var utfordrende å forstå ved første blick hva dette bilde inneholder, så det er ikke så overraskende at ingen av modellene klarte det heller. Ifølge de opprinnelige beskrivelsene er dette et bilde av den ytre porten til den Svenske Handelsbanken sitt hovedkontor i Stockholm. Det siste bilde som ble tolket feil av alle modellene er bilde 404-0364:

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Hyllningsadress till Danmarks lærerforening.	(no detections)	a book with a picture of a woman and a man	an image of a book with a picture of a man in the middle ages	a close up of two pages of an old book	a sign that is on top of a pile of books

[Utdrag fra Tabell 2](#)



Bilde 404-0364

Dette kan enten være et scannet dokument eller en annen type avbildning av et dokument. Ifølge den opprinnelige beskrivelsen er dette en hyllest til Danmarks lærerforening. Begge BLIP beskrivelsene identifiserer personer i dette bildet, og den betingede beskrivelsen fra BLIP har til og med beskrevet det som «an image of a book with a picture of a man in the middle ages». Det er kanskje den gammeldagse skriften som gjør at BLIP identifiserer dette som et bilde fra middelalderen. Både CLIP og NLPconnections beskriver en bok eller bøker i

bildet, noe ikke stemmer helt. Det er ikke umulig at dette er scannet fra en bok, men det er i så fall ren gjetting.

Beskrivelsene på resten av bildene kan klassifiseres som delvis korrekte. I mange tilfeller er det snakk om gruppebilder der modellene sliter med å identifisere riktig antall personer. De eneste tilfeller det blir korrekt er når modellene genererer en variasjon av «a group of people» eller «a bunch of people». Gruppebilder var også en utfordring for YOLO-algoritmen. (Zeng, 2023, s. 22) Ellers er det mange tilfeller her hvor alle bortsett fra en klarer å beskrive bildet korrekt. Her utmerker NLPconnections seg spesielt, jeg fant 15 tilfeller der den ikke klarte å generere en korrekt beskrivelse, mens de andre modellene klarte det. Et par eksempler av dette her:



**Bilde 009-0070**

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
SBmfs representantskaps middag. I.B. opptråder med ett nummer ur Stockholmnas bankrevy av banktjånsteman G W. W.	(no detections)	a woman in a white dress and hat playing a guitar	an image of a woman playing a guitar	an old photo of a woman holding a guitar	a woman holding a tennis racquet on a tennis court

[Utdrag fra Tabell 2](#)

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Diverse bilder från SIF:s konstmuseum.	1 chair	a room with a bunch of old telephones	an image of a room with a bunch of telephones	a group of old phones hanging on the wall .	a painting of a woman in a dress and a man in a suit

[Utdrag fra Tabell 2](#)

**Bilde 055-0544**



Alle modellene har eksempler på feil, fra små feil til grove feil. Det er mange små feil der beskrivelsene fanger essensen av bildet, men modellen gjetter en eller flere detaljer som

enten er feilaktig eller umulig å vite ved å bare se på bildet. Det er for eksempel flere forekomster der en modell prøver å gjette innholdet i et glass eller en flaske. Sammenlignet med beskrivelsene fra YOLO-algoritmen, er disse tre generative KI-modellene en høyere grad av nøyaktighet. Av de 40 bildene var det 7 bilder der YOLO ikke klarte å identifisere noen objekter. Bare 8 av bildene fikk beskrivelser fra YOLO som samsvarte med innholdet i bildene. Dette var som regel portretter eller bilder med en eller to personer. Som tidligere nevnt, YOLO ble først trent på ansiktsgjenkjenning, så den er god på å identifisere personer. Den er derimot ikke noe god på å identifisere korrekt antall personer. Det var bare ett bilde der YOLO klarte dette i et bilde med flere enn tre til stede. I slike tilfeller kan det være bedre med generiske beskrivelser fra generative KI-modeller, som gir beskrivelser som «a group of people» eller «a bunch of people». Selv i de tilfellene der YOLO er korrekt, er jeg usikker på hvor stor nytte slike objektidentifiserende beskrivelser har. Som i disse eksemplene:



Bilde 447-0014

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Konferens ang. utdanningsfrågor. T. H., TCO, handläggare i utbildningsfrågor, ombudsman, Stockholm	1 person, 1 tie, 2 bottles	a man in a suit and tie sitting at a table	an image of a man sitting at a table	an old photo of a man wearing glasses	a man sitting at a table with a glass of wine

[Utdrag fra Tabell 2](#)



Bilde 226-0036

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
G. N., medlem av centralstyrelsen, kassör 1910 -1912, sekreterare 1912-1925, ordförande 1926-1931, redaktör för Lärarinneförbundet från starten 1914-1933	1 person	a woman in a black and white photo	an image of a woman in black and white	an old black - and - white photo of a woman .	a woman in a black and white photo

[Utdrag fra Tabell 2](#)

Det første bildet inneholder en person med slips og to flasker, men dette er en veldig unaturlig måte å beskrive et bilde. Den ubetinget beskrivelsen fra BLIP er et mye bedre



eksempel på en mer naturlig beskrivelse: «a man in a suit and tie sitting at a table». Det andre bildet inneholder en person slik YOLO sier, men dette er også en unaturlig og mangelfull måte å beskrive dette bildet. Selv om det ikke er så mye innhold i dette bildet, gir de tre modellene mer deskriptive beskrivelser som har litt mer substans.

Det er sjeldent er de opprinnelige beskrivelsene skildrer innholdet i bildene. De forklarer heller konteksten rundt eller gir utfyllende informasjon om personen(e) eller stedet avbildet. Ta bilde 226-0036 fra det forrige eksempelet, et portrett som hverken YOLO eller generativ KI modellene kunne si så mye om. I de opprinnelige beskrivelsene får vi vite denne personens navn og hennes bakgrunn i fagforeningen. Når det gjelder bilde 447-0014, også fra det eksempelet, får vi vite både identiteten til personen i bildet og konteksten rundt fra den opprinnelige beskrivelsen. Det er sjeldent de opprinnelige beskrivelsene ikke gir en god forklaring av bildet, men det er noen få tilfeller av dette. Som for eksempel bilde 055-0480, et gruppebilde vist tidligere: «Personal på FK och FA 16.». Eller beskrivelsen til et annet gruppebilde, bilde 055-1970 som også er vanskelig å forstå: «SIF:s kongress. Fr. höger?, SSAB, SIF, kongressombud, FS-ledamot, FA24, ASO(?), TKO(?),».

Ellers virker beskrivelsene fra generativ KI svært generisk i forhold til de opprinnelige. Jeg vil spesielt trekke frem dette eksempelet:



Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Cerebral angiografi av hjärna på Södersjukhuset i Stockholm. Operationa. Överläkaren dr S. L. punkterar. Syster E. P. assisterar, dr R. G. ger narkos och en elev instrumenterar. b. Samma tillfälle som vid fotografi a men endast system och eleven syns på fotografiet	1 person	a black and white photo of a woman in a hospital room	an image of a hospital room with a doctor and patient	a black - and - white photo of two doctors in a room .	a woman working in a hospital room

[Utdrag fra Tabell 2](#)

**Bilde 149-3495**

Ingen av modellene klarte å telle antall personer riktig i dette bildet, men alle klarte å identifisere at det var leger til stede eller at bildet var tatt på et sykehus. Fra den opprinnelige beskrivelsen får vi vite navnene på de som er avbildet, og hva slags rolle de har. Vi får også vite hva slags behandling pasienten får, og hva legen gjør akkurat når bildet blir tatt. Dette er

et godt eksempel på hvor mye informasjon som finnes i selve bildet, og hvor mye som avhenger av konteksten. Et annet eksempel på dette:

Opprinnelige beskrivelse	YOLO	BLIP unconditional	BLIP conditional: «an image of»	CLIP	NLP connections
Behandling av diaréer hos spedbarn. Mat är gott. Bilden är publicerad i samband med en artikel om "Morötter vid spädbarnsdyspepsi".	1 person, 2 bottles	a baby eating a peanut butter butter	an image of a baby eating food	a baby in a high chair eating food .	a baby sitting at a table with a spoon

[Utdrag fra Tabell 2](#)



Bilde 149-1214

De aller fleste KI-genererte beskrivelser er jo korrekte her, selv om peanøttsmøret er ren gjetting fra BLIP. Den opprinnelige beskrivelsen sier ingenting om babyen som spiser, men gir heller informasjon om hvilket sammenheng bildet ble tatt i. Dette er kunnskap man ikke får ved å bare se på innholdet i bildet.

### 3.3.3 Gjennomføring serie- og samlingsnivå

Neste stadiet i undersøkelsen handler om å bruke beskrivelsene av objektene i arkivet for å beskrive de to øvrige nivåene, altså serie- og samlingsnivå. Beskrivelse på disse nivåene skal gi bedre intellektuell kontroll over et arkiv og vil vanligvis inneholde en overordnet fremstilling av arkivstrukturen. Vi vet ingenting om strukturen til TAM sitt fotoarkiv, vi vet bare at seriene er ordnet etter arkivskaper. Det kan godt hende at bildene var organisert i mapper på et tidspunkt, men dette er ikke informasjon som fremkommer av databasen. Jeg ønsker å finne ut om objekt beskrivelsene er tilstrekkelig for å gi detaljert informasjon om en arkivserie eller hele arkivet, ved å bruke ChatGPT til å generere beskrivelser basert på de tidligere genererte beskrivelser.

For å få data til dette stadiet i undersøkelsen, brukte jeg de tre modellene for å generere beskrivelser for alle 6403 bilder i bildearkivet. Min egen PC brukte omtrent 48 timer på å generere beskrivelsene med BLIP, inkludert stopp over natta. For CLIP og NLPconnect brukte jeg den stasjonære PC-en. På denne brukte CLIP rundt 45 minutter på å generere beskrivelser for hele bildearkivet, og NLPconnect brukte rundt 15 minutter. Dette er en enorm forbedring på de 48 timene som min PC brukte. Dataene fra alle disse gjennomføringene ble eksportert

til separate CSV filer, med 6403 beskrivelser i hver fil. Dette er enorme mengder med data og gitt begrensningene for gratisbrukere av GPT4o, modellen som kan analysere vedlagte filer, måtte jeg gjøre avgrensninger. På samlingsnivå bestemte jeg meg for å velge data fra en modell, jeg valgte tilfeldig og landet på CLIP. På serienivå brukte jeg bare data fra en serie, Sveriges Folkskolläraryrkesförbundet. Jeg valgte denne siden dette var en av de mellomstore seriene med 157 bilder. Siden alle bildene starter med en kode tilsvarende arkivserie/arkivskaper kunne jeg lett klippe ut de dataene jeg trengte fra denne serien til en separat fil for å brukes i forespørselene.

For å generere de beskrivelsene jeg trengte brukte jeg webapplikasjonen til ChatGPT. Her kan man laste opp filer sammen med en forespørsel direkte i en chat. Jeg forsøkte å bruke regnearkene, men fikk raskt problemer med dette. Det hjalp heller ikke å konvertere fra CSV format til Excel format. Jeg opplevde ofte at jeg ble stengt ut av chatten før jeg i det hele tatt fikk et resultat. OpenAI har ikke gitt grundig informasjon om hvor mange filopplastninger gratis brukere kan gjøre ved bruk av GPT4o. Dette er avhengig av flere faktorer som for eksempel format og filstørrelse. (OpenAI, 2024a) Jeg bestemte meg derfor å forsøke i et annet format og konverterte filene til Word dokumenter. Jeg begynte på samlingsnivå og jeg fikk lov til å gjøre to forespørsler med filopplastninger før dette ble stengt i en del timer.

Resultatene fra slike chatboter vil alltid være avhengig av forespørselen som gis. Det er mulig å bruke uendelig tid på å spisse formuleringen for å se forskjellen i resultatet. På grunn av begrensningene i antall forsøk, hadde jeg ikke muligheten til å prøve forskjellige formuleringer. Jeg måtte bare skrive en forespørsel og holde meg til det. Det var viktig å holde forespørselene så lik som mulig, men jeg er interessert i å se om resultatene ville endre seg dersom jeg ga tilleggsinformasjon. Som for eksempel at arkivet tilhører TAM eller hvem skaperen av arkivserien er. Derfor skrev jeg to versjoner av samme forespørsel. Hver forespørsel ble foretatt i hver sin chat.

### Serienivå

#### Forespørsel 1:

«Vedlagt er et dokument med beskrivelser av bilder i en arkivserie. Serien er en del av en større samling. Beskrivelsene er kun basert på innholdet i bildet. Bruk disse beskrivelsene for å lage en overordnet beskrivelse av denne arkivserien.»



Forespørsel 2:

«Vedlagt er et dokument med beskrivelser av bilder tilhørende Svensk Sjuksköterskeförning. Dette er en arkivserie som er en del av en større samling. Beskrivelsene er kun basert på innholdet i bildet. Bruk disse beskrivelsene for å lage en overordnet beskrivelse av denne arkivserien.»

#### Samlingsnivå

Forespørsel 3:

«Vedlagt er et dokument med beskrivelser av bilder i et fotoarkiv. Beskrivelsene er kun basert på innholdet i bildet. Bruk disse beskrivelsene for å lage en overordnet beskrivelse av dette arkivet.»

Forespørsel 4:

«Vedlagt er et dokument med beskrivelser av bilder i Tjänstemäns & akademikers fotoarkiv. Beskrivelsene er kun basert på innholdet i bildet. Bruk disse beskrivelsene for å lage en overordnet beskrivelse av dette arkivet.»

Resultatene fra disse forespørlene blir diskutert nærmere i neste delkapittel.

#### 3.3.4 Resultater serie- og samlingsnivå

Alle beskrivelsene generert av ChatGPT følger så og si samme format: en rask innledning, identifisering av temaer og motiver og en kort konklusjon. Dette utfallet var forventet, ettersom forespørlene var svært like. Innledningene var svært generelle og manglet utfyllende informasjon, unntatt i de tilfellene hvor kontekst om arkivet eller arkivskaperen ble gitt:

«Den vedlagte arkivserien består hovedsakelig av fotografier som dokumenterer ulike samlinger og møter i forskjellige sosiale og offentlige settinger. Basert på bildenes beskrivelser, kan vi identifisere flere gjentakende temaer og motiver som gir et innblikk i arkivets overordnede innhold og betydning.»

#### [Utdrag fra Forespørsel 1](#)

«Arkivserien tilhørende Svensk Sjuksköterskeförning, basert på de medfølgende bildebeskrivelsene, gir et omfattende innblikk i foreningens aktiviteter og medlemmer. Bildene spenner over en rekke situasjoner og settinger som reflekterer det daglige livet og spesielle anledninger innenfor organisasjonen.»

#### [Utdrag fra Forespørsel 2](#)

«Fotoarkivet som beskrevet i dokumentet inneholder en omfattende samling av hovedsakelig svart-hvitt fotografier.»

#### [Utdrag fra Forespørsel 3](#)

«Tjänstemäns & akademikers fotoarkiv er et omfattende fototek, som hovedsakelig inneholder svart-hvitt bilder som viser ulike profesjonelle og akademiske aktiviteter. Bildene fanger et bredt spekter av scener, inkludert møter, forelesninger og sosiale sammenkomster, ofte med grupper av menn og kvinner som sitter rundt bord eller står i formelle poseringer.»

#### Utdrag fra Forespørsel 4

Neste del av de genererte beskrivelsene gikk utpå å identifisere temaer og motiver. Selv om alle forespørslene fikk tilgang til de samme dataene, var ikke resultatene alltid like. På serienivå resulterte forespørslene fem ulike klassifikasjoner for temaer og motiver.

Forespørsel 1 genererte: Sosiale samlinger, formelle møter, portretter og poseringer, kulturelle og kunstneriske hendelser, kvinnegrupper. Forespørsel 2 genererte: Gruppebilder og sosiale samlinger, møter og arbeidsmiljø, seremonier og offentlige arrangementer, kulturelle og fritidsaktiviteter, portretter og personlige interaksjoner. Selv om det er forskjeller mellom de to, er det også en del likheter. Begge legger vekt på flere av de samme elementene som sosiale samlinger, møter, portretter og kultur.

Forespørslene på samlingsnivå resulterte også i fem ulike klassifikasjoner for temaer og motiver. Forespørsel 3 genererte: Natur og landskap, by- og bygdeomgivelser, maritime bilder, personer og portretter, historiske øyeblikk og gjenstander. Forespørsel 4 genererte: Profesjonelle møter og konferanser, offentlige taler og presentasjoner, sosiale interaksjoner, arbeidsplasser, historiske og kulturelle øyeblikk.

Det er mye større forskjell mellom de to beskrivelsene på samlingsnivå, enn de på serienivå. Jeg mistenker at størrelsen på filen matet til ChatGPT kan ha en påvirkning. Resultatet til Forespørsel 3 var den minst detaljerte av alle beskrivelsene, og i denne forespørselen fikk ChatGPT ingen kontekst og måtte analysere en fil med over seks tusen beskrivelser. Det er mulig at dette store antallet gjør det utfordrende for ChatGPT å gjøre en grundig analyse, dermed blir resultatet mer generisk. Kanskje dette er grunnen til at beskrivelsene på serienivå var generelt mer detaljert enn de på samlingsnivå. Forespørsel 1 på serienivå genererte til og med en beskrivelse som brukte spesifikke bilder som eksempler:

«1. **Sosiale Samlinger:**»

- Mange bilder viser grupper av mennesker samlet i forskjellige kontekster, som foran bygninger, i kirker, og i rom. Eksempler på dette er beskrivelse 226-0001 (en gruppe mennesker foran en bygning) og 226-0035 (en stor folkemengde i et rom).

2. **Formelle Møter:**»

- En betydelig del av bildene dokumenterer formelle møter og diskusjoner rundt bord. Dette inkluderer grupper av mennesker som sitter rundt bord med dokumenter og papirer, som beskrevet i 226-0010 (to kvinner ved et bord med papirer) og 226-0097 (en mann ved et bord med bøker).»

#### [Utdrag fra Forespørsel 1](#)

Ingen av de andre beskrivelsene gjorde det samme. Konklusjonen i denne beskrivelsen er også den mest detaljerte og informasjonsrike:

«Denne arkivserien gir et rikt innblikk i sosiale og formelle interaksjoner i en bestemt periode. Bildene fremhever viktigheten av både formelle møter og uformelle samlinger, med en spesiell fokus på kvinners deltakelse i samfunnet. Gjennom en blanding av portretter, gruppebilder, og dokumentasjon av kulturelle arrangementer, tilbyr denne serien en omfattende visuell dokumentasjon av menneskelige aktiviteter og samspill.»

#### [Utdrag fra Forespørsel 1](#)

De andre konklusjonene kan ikke annet enn å beskrives som mangelfulle:

«Denne arkivserien er dermed en rik visuell dokumentasjon som kan brukes til å forstå både det formelle og det uformelle livet i Svensk Sjuksköterskeförning, og gir en dyptgående innsikt i deres historie og kultur.»

#### [Utdrag fra Forespørsel 2](#)

«Samlet sett fungerer Tjänstemäns & akademikers fotoarkiv som en rik visuell dokumentasjon av profesjonelt og akademisk liv, og fremhever interaksjonene, arbeidet og begivenhetene som definerer disse samfunnene.»

#### [Utdrag fra Forespørsel 4](#)

Konklusjonen til Forespørsel 3 gjør det ganske åpenbart at GPT4o ikke klarer å analysere filene med større mengder data:

«Dette fotoarkivet gir en rik visuell dokumentasjon av både det naturlige og menneskelige miljøet gjennom tidene, med et sterkt fokus på maritime aktiviteter og hverdagsliv i ulike samfunn.»

#### [Utdrag fra forespørsel 3](#)

Det stemmer at noen av bildene har maritimt innhold, med motiver som båter og fyrstårn. Disse bildene utgjør en svært liten minoritet; jeg har bare funnet et fåtall slike bilder i en serie, og dette er ikke den serien som ble analysert på serienivå. Det store flertallet av bilder i dette fotoarkivet er jo av mennesker så det er veldig rart at beskrivelsen velger å fokusere på det maritime. Det er vanskelig å komme med en utdypende analyse når resultatene er så mangelfulle. Selv om denne delen av undersøkelsen funket best på samlingsnivå, var det likevel varierende resultater på dette nivået. Grunnen til at det funket bedre på samlingsnivå var mest sannsynligvis på grunn av mindre datasett i forespørlene. Forskjellene mellom

beskrivelsene på disse to nivåene viser hvor avhengig resultatet fra ChatGPT er av den informasjonen som den blir matet med. Dette vil være et problem uansett hvor avansert slike chatbotene eventuelt blir.

Denne undersøkelsen viser at dette ikke er en egnet metode for å beskrive serie- og samlingsnivå i et fotoarkiv. Beskrivelser av arkiv på serie- og samlingsnivå skal gi bedre intellektuell kontroll over arkivet, beskrivelsene generert er alt for mangelfulle til å kunne oppnå dette. Brukere kan kanskje få en idé om innholdet i noen bilder, men ikke hva fotoarkivet handler om. Dette kan være på grunn av mangler i modellen, den hadde vansker med grundig analyse av større filer. Det kan også være fordi beskrivelser av objektene ikke er passende for å beskrive helheten i dette arkivet.

## 4. Drøfting

### 4.1 Fordeler med KI-genererte beskrivelser i fotoarkiv

Den største fordelen med å bruke slik teknologi er tidsbesparelse. Selv med bruk av mindre egnet utstyr tok det betydelig kortere tid å generere beskrivelser for flere tusen bilder enn om dette skulle gjøres manuelt. Det er også mulig å gjøre andre oppgaver i den tiden, noe som ikke er mulig ved manuell generering av beskrivelser. En annen fordel er muligheten for å tilegne metadata der det er ingen fra før av, eller ved mangelfulle metadata. Dette virker å være et gjennomgående problem i fotoarkiv, som påpekt i Aksess sin artikkel om Arkivverkets nye KI bildesøk løsning. Flere av kildene i litteraturgjennomgangen i kapittel 2 peker på viktigheten av arkivbeskrivelse for søk og gjenfinning. KI kan derfor tilgjengeliggjøre arkiv der det tidligere hadde vært umulig.

Det kan være avvik i beskrivelser generert av KI, men jeg mener at det er lettere å identifisere disse enn feil gjort av mennesker. KI-genererte beskrivelser kan kun baseres på innholdet i bildene, noe som gjør det lett å kontrollere for feil. Metadataene i TAM sitt fotoarkiv er menneskeskapte, noe som vil gi en høy grad av tillit for mange. De blir kanskje ikke møtt med lik skepsis som KI-genererte beskrivelser. Allikevel kan de inneholde feil, menneskelig hukommelse kan være veldig upålitelig. De opprinnelige beskrivelsene skildret heller konteksten rundt enn innholdet i bildene, noe som er umulig for en uten slik bakgrunnskunnskap å kontrollere. Jeg har for eksempel flere ganger oppdaget uoverensstemmelser mellom metadatapostene for årstall og de beskrivelsene som inneholdt

dato for når bilde ble tatt. I disse tilfellene kan jeg ikke annet enn å gå ut ifra at det som står i beskrivelsene er korrekt.

#### 4.2 Ulemper med KI-genererte beskrivelser i fotoarkiv

Feilene, eller hallusinasjonene, som forekommer i KI-genererte beskrivelser er svært problematiske. Når jeg begynte på denne oppgaven var jeg mest opptatt av søk og gjenfinning, men arkivbeskrivelse har også andre formål. I følge ISAAD(G) standarden skal arkivbeskrivelse også sørge for at pålitelige og autentisk arkivmateriale kan føres videre. (ICA, 2000, s. 7) Hvis feilinformasjon blir lagt til i arkivet, kan det påvirke brukeres tillit til arkivet og det kan bli satt spørsmål til materialets autenticitet. Det blir heller ikke særlig tidsbesparende hvis det blir nødvendig å sjekke hver eneste beskrivelse for feil i etterkant.

Et av problemene med generativ KI er at den vil alltid gi deg et svar. Det høres kanskje ikke ut som en ulempe, men konsekvensen er at den vil bare finne på ting selv om den ikke vet hva den skal identifisere, noe undersøkelsen viser kan føre til grove feil. Hvis YOLO-algoritmen ikke klarte å identifisere objekter ville den returnere «no detection» som svar. I beskrivelsene av bildene 404-0364, hyllesten til Danmarks lærerforening, og 009-0478, porten til Handelsbanken sitt hovedkontor begge vist tidligere i delkapittel 3.3.2.2, var det tydelig at ingen av modellene klarte å tolke bildet korrekt. Det er mer ideelt at modellene hadde generert en type «no detection» svar istedenfor. Arkivarer kan eventuelt fylle på slike beskrivelser manuelt i etterkant. Feilene som generativ KI gjør kan blant annet skyldes på skjevhet i materialet den er trent på. Kunnskapen til KI-modeller, uansett hvor avansert, vil alltid være begrenset av det materialet som utviklerne har tilgang til.

I litteraturgjennomgangen i delkapittel 2.3, fant jeg at kunstig intelligens er et vanskelig begrep å definere. Dette kan være fordi den vil alltid sammenlignes med menneskelig intelligens. Med sin kompleksitet, er også menneskelig intelligens utrolig vanskelig å definere. Den er heller ikke lett målbar, selv om vi ofte forsøker. Den innebærer blant annet kreativitet og fantasi, noe som det er vanskelig å forestille seg at maskiner kan gjenskape. Det er også viktig å ta i betraktning at det er forskjell mellom kunnskap og intelligens, og uansett nivået på en persons intelligens vil de ha tilgang til kunnskap som vil være utenfor rekkevidde for maskiner fordi de er begrenset til skriftlig informasjon. Eksempler på dette kan være en persons minner eller kulturen de er vokst opp i. Som en norsk person med nærhet til svensk

kultur kunne jeg identifisere midsommerstangen i bilde 404-0578, vist tidligere i delkapittel 3.3.2.1, med en gang. Ingen av modellene klarte det samme.

Det kan kanskje argumenteres at de tilfellene med småfeil der modellene gjetter seg frem til detaljer i et bilde, er eksempler på å gjenskape menneskelig intelligens igjennom resonnement. Det er kanskje ikke så usannsynlig at innholdet i noe som ligner på et vinglass er vin. Eller at en gruppe i sykepleieruniform står foran et sykehus, selv om det ikke er synlig på bildet. Samtidig har mennesker evnen til å forstå sine egne begrensninger og vet at dette er ting som ikke kan fastslås med sikkerhet. Mennesker ville sannsynligvis unnlate slike detaljer fra sine beskrivelser.

I et mer brukersentrert perspektiv, kan man stille spørsmålet om hvor farlig alt dette er? Hvis en bruker ønsker et bilde, spiller det noen rolle hvis beskrivelsen inneholder feil så lenge det bilde blir tilgjengelig for brukeren? Spesielt hvis beskrivelsen ikke inneholder grove feil, eller hvis feilen er så innlysende at det er lett for brukeren å ignorere. Da er det fare for å falle i fellen Schwartz advarte mot, og fotoarkivet blir heller et «stock photo» bibliotek. Ved å redusere bilder i et fotoarkiv til kun sitt innhold vil brukere gå glipp av kontekstualiserende informasjon. Jeg mener dette kan minske materialets verdi for visse arkivbrukere som forskere, historikere og andre som søker en dypere innsikt og forståelse. Beskrivelser generert av bilde-til-tekst modeller kan aldri gi mer innsikt enn hva den mener bildet inneholder.

#### 4.3 Søk og gjenfinning

Beskrivelse er en nødvendighet for å kunne gjenfinne arkivmateriale. Selv om mesteparten av arkivbeskrivelser skildrer heller arkivstrukturen, viser Arkivverket sin nye bildesøk løsning på Digitalarkivet at søk på objektnivå er en metode som fungerer for å tilgjengeliggjøre fotoarkiv. Slike løsninger ligner veldig på søkemotorer som mange er godt vant med på forhånd. I en analyse av flere brukerundersøkelser fant Jennifer Schaffner at brukere forventer å finne frem i arkivmateriale på nett ved hjelp av de samme søketeknikkene de bruker i andre sammenheng. De vil finne frem til emner, men bruker som regel nøkkelordsteknikker fremfor strukturert terminologi i sine søk. (Schaffner, 2009, s. 4-5) Schaffner mener det er usikkert om den tradisjonelle arkivkatalogen samsvarer med arkivbrukerens forventninger og forkunnskaper, og bemerker at mange vet ikke engang at de eksisterer. (Schaffner, 2009, s. 11)

Utfordringer med gjenfinning i digitale arkiv skyldes ikke bare brukergrensesnittene som brukes, men også de metadataene som er tilknyttet. Når nettbaserte løsninger tas i bruk mister arkivarene kontrollen over gjenfinning. Metadataene vi oppretter er derfor viktigere enn noen gang. Som Schaffner skriver, «... the metadata *is* the interface». (Schaffner, 2009, s. 4) Beskrivelsene som produseres av generativ KI kan fungere som et godt utgangspunkt for å generere metadata som møter brukeres forventninger til søk i arkivene. Analysen av de 40 utvalgte bildene viste at generativ KI produserte mer detaljerte og mer nøyaktige beskrivelser enn YOLO-algoritmen. De er også beskrevet på en mer naturlig måte enn det objekt identifiserende algoritmer klarer. Jeg mener at informasjon om antall personer, klokker, slips, stoler o.l., som er typisk metadata generert av YOLO, ikke er så nyttig i søk.

Likevel kan det hende at beskrivelsene som ble generert av de tre modellene i denne undersøkelsen ikke er detaljerte nok til å være relevant i søk de heller. Det må også tas i betraktning feilene som disse modellene til tider gjør. Selv i de tilfellene der beskrivelsene ikke er feilaktige er det mange elementer fra bildene som ofte mangler. Som for eksempel bilde 055-0544 vist tidligere i delkapittel 3.3.2.2, der et rom med gamle telefoner er avbildet. Alle modellene, bortsett fra NLPconnect, identifiserte telefonene som hang på veggen, men kunne ikke si noe om resten av innholdet i bildet. De opprinnelige beskrivelsene var for det meste ganske detaljerte, men beskrev ofte konteksten rundt innholdet i bildene.

Beskrivelsen for bilde 055-0544 forklarer at dette er «Diverse bilder från SIF:s konstmuseum.» Det er også viktig å ta i betraktning at den opprinnelige PICTOR databasen har også andre metadataposter som kan hjelpe til i søk dersom beskrivelsen er mangelfull.

#### 4.4 Begrensninger til generativ KI

Litteraturgjennomgangen i delkapittel 2.3 viste hvordan forhåndstrente modeller er spesielt utsatt for bias. Mange stiller også spørsmål til hvordan disse modellene generer materiale ved å låne fra eksisterende materiale. Moderne arkiver er bygd på prinsipper om åpenhet og transparens, dersom bruken av KI kan gi skjevheter i resultater må det være åpenhet om dette. Noen arkivforskere har begynt å bruke begrepet paradata for å beskrive en metode som støtter behovet for ansvarlighet ved bruk av KI i arkivene. Paradata kan brukes for å dokumentere KI-prosesser i både tekniske og organisatoriske kontekster. (Cameron et al., 2023, s. 76) Foreslåtte elementer til paradata inkluderer ytelsesinformasjon, versjonsinformasjon, samt trenings-, test- og valideringsdata. Davet et al. mener komplette

kopier av disse datasettene burde inkluderes i siste kategorien og gjøres være tilgjengelig for enkel undersøkelse og tolkning. Dette kan også inkludere informasjon om hvordan datasettene ble konstruert fra rådata, metodologisk dokumentasjon, utviklerforklaringer på datasettenes omfang, og informasjon om automatiserte systemer dedikert til opprettelse av datasettene. (Davet et al., 2023, s. 285)

Når det gjelder modellene brukt i denne undersøkelsen, har utviklerne av både BLIP og CLIP dokumentert informasjon om datasettene brukt i treningen av sine modeller. Utviklerne av CLIP brukte både tidligere etablerte datasett som MS-COCO og bilder hentet fra internett. Totalt sett ble rundt 400 millioner bilder brukt. (Radford et al., 2021, s. 3) Utviklerne av BLIP brukte også noen av de samme etablerte datasettene som CLIP sammen med andre som Visual Genome og Conceptual Captions for et totalt datasett på 115 millioner bilder. (Li et al., 2022, s. 4) Alle datasettene brukt består av bilder med tilhørende beskrivelse. For begge sett med utviklere er det antall bilder som var hovedkriteriet for valg av datasett, vi får ikke noe særlig mer utfyllende informasjon om det. Jeg tror ikke det er særlig nyttig å gi vanlige arkivbrukere tilgang til hundrevis av millioner av bilder for å gjøre sine egne undersøkelser og tolkninger. Jeg mener heller at det er viktigere å informere brukere om mulige begrensninger ved bruken av slike KI modeller og de potensielle konsekvensene av disse begrensningene, som for eksempel skjevheter i resultatene.

Molly White, forsker, systemutvikler og skaperen av nyhetsbrevet *[citation needed]*, har sammenlignet blesten rundt KI med den rundt blockchain teknologi og kryptovaluta for en del år siden. Den allmenne interessen for de sistnevnte har så å si forsvunnet fullstendig. I artikkelen *AI isn't useless. But is it worth it?* utførte hun sine egne undersøkelser for bruksområdene til generativ KI og fant at områdene der slike modeller fungerer bra er ofte mer trivielle enn det som forventes. Som i generering av korte tekster til for eksempel eposter, i sammendrag av lengere tekster og til hjelp i koding. Dette er langt ifra det utviklerne lover. Da er det naturlig å stille seg spørsmålet om det er verdt det, spesielt når kostnadene assosiert tas i betraktning. Dette inkluderer de enorme økonomiske kostnadene som ligger bak utviklingen av disse modellene, samt miljøkostnadene knyttet til datakraften som trengs. White påpeker at KI innebærer store menneskelige kostnader også. KI er avhengig av «ghost labor», altså skjult arbeidskraft som ser ut til å utføres av datamaskiner, men er gjort av mennesker. Arbeidet består av datakompilering og -merking, ofte uten hensyn til opphavsrett



eller samtykke. Mange av de som gjør slikt arbeid jobber på underbetalte kontrakter uten noe arbeidsvern. White konkluderer at kostnadene ikke er verdt nytten, spesielt når brukerne som får mest nytte ut av generativ KI er kun de som forstår begrensningene. (White, 2024)

Det er ikke lett for enkeltbrukere å forstå disse begrensningene når det er såpass mye oppmerksomhet og skyhøye forventninger. Det hjelper heller ikke at selskaper «jukser» når de presenterer sine produkter. Som for eksempel når Google presenterte Gemini, deres konkurrent til ChatGPT, med en video demonstrasjon. I etterkant har det vist seg at denne demonstrasjonen er manipulert for å gi inntrykket av at den fungerer raskere. Videoen viste også instruksjoner gitt muntlig, når den egentlige demonstrasjonen ble gjort med tekstbaserte forespørsler. (Techradar, 2023) OpenAI presenterte Sora i februar 2024, et KI-produkt som kan generere video fra tekst. Produktet er ikke tilgjengelig for allment bruk enda, men klippene som ble presentert har allikevel fått mye oppmerksomhet. Det var spesielt et klipp som mange var imponert av kalt «Air Head». Dette er en kort video på rundt to minutter som følger Sunny, en mann som har en ballong istedenfor et ansikt. Produksjonsselskapet bak videoen har avslørt i etterkant at det var mer bak videoen enn en tekst instruksjon, de måtte gjøre en god del manuelle redigeringer av videoen som Sora produserte. (The Byte, 2024) Dette opplyste ikke OpenAI om når de presenterte videoen, noe som kan gi inntrykket av at Sora er mye mer avansert enn det den egentlig er. Skal generativ KI være nyttig for arkivene, må vi se forbi interessen og forventningene som skapes av både samfunnet og av utviklere. Og viktigst av alt, vi må forstå begrensningene før det tas i bruk.

## 5. Konklusjon

Formålet med denne oppgaven var å undersøke gjennombruddene som har skjedd innen generativ KI og hvorvidt denne teknologien kan brukes for å automatisk generere beskrivelser i et fotoarkiv. Dermed ble problemstillingen for denne oppgaven: ***har utviklingen av generativ KI kommet til det nivået at arkiver kan bruke det for å automatisk generere beskrivelser i fotoarkiv?***

Fra et teoretisk standpunkt virker generativ KI som et ypperlig verktøy for beskrivelse. Digitaliseringen har ført til enorme volum av digitale arkiv som det er utfordrende å tilgjengeliggjøre. Arkivbeskrivelse i form av arkivkatalogen har tidligere fungert som et gjenfinningsverktøy, men det er usikkert hvor godt de tradisjonelle metodene fungerer i en

digital hverdag. Fotoarkiv er spesielt avhengig av god beskrivelse for å kunne tilgjengeliggjøres. Likevel er dette noe som ofte er mangelfull. Det er absolutt den riktige tiden å ta i bruk nye teknologiske virkemidler for å møte denne problematikken. Vi må synliggjøre «dark archives».

Den praktiske undersøkelsen viser tydelig at generativ KI kan være tidsbesparende i tidkrevende oppgaver som beskrivelse. Generativ KI produserer også mer nøyaktige, detaljerte og naturlige beskrivelser sammenlignet med objektidentifiserende algoritmer som YOLO. Samtidig viser også undersøkelsen at det er store svakheter i bilde-til-tekst modeller sine beskrivelser. Slike modeller gjør ofte feil, såkalte hallusinasjoner. Dette kan være alt fra småfeil, et typisk eksempel fra denne undersøkelsen var å gjette på innholdet i glass og flasker, til grove feil som gjettinger på tid og sted som virker å komme fra intet. For å unngå feilinformasjon i arkivene er det viktig at arkivarer forstår slike begrensninger som kommer med generativ KI før det tas i bruk. Selv når modellene er korrekte, er det usikkert om beskrivelsene er tilstrekkelige detaljerte til å være relevante for søk og gjenfinning.

Arkivbeskrivelse har mange formål og skal dekke ulike behov. Det skal sørge for at pålitelig og autentisk arkivmateriale blir ført videre. Samtidig skal det sørge for fremfinning av arkivmateriale. Noen brukere ønsker dypere forståelse og innsikt av konteksten til materialet. Andre ønsker kun lett tilgjengelighet. Uansett hvor avanserte KI-modeller kan bli i fremtiden er det usikkert på om dette er redskapet for å dekke alle disse behovene, spesielt i fotarkiv. Slike teknologier vil alltid redusere bilder til kun deres innhold, kun det som kan ses. Arkivbrukere går dermed glipp av mye kontekstualiserende informasjon.

Slik kontekstualiserende informasjon ble ofte gitt av arkivkatalogen, som ble brukt for å vise strukturen på arkivet og den intellektuelle sammenhengen mellom arkivmateriale. Undersøkelsen viser også at det å bruke beskrivelser generert av KI på objektnivå, ikke vil gi en god forståelse av verken en arkivserie eller arkivet i sin helhet. Dette kan skyldes modellenes begrensninger eller at beskrivelsene av objektenes innhold ikke er tilstrekkelig for å dekke de øvrige nivåene. Dette viser forskjellen på beskrivelser av det som er *i* arkivet og beskrivelser *av* arkivet.

«Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.» (Ertel, 2017, s. 2) Selv om jeg ikke har testet alle bilde-til-tekst

modeller, mener jeg at det er rimelig å konkludere at for nå er mennesker bedre til å beskrive bilder enn maskiner. Med tanke på den raske utviklingen på dette feltet, er ikke denne konklusjonen ment å være endelig. Algoritmer blir mer effektive og modeller stadig mer komplekse. Forskjellen mellom YOLO-algoritmen og generativ KI viser dette. Bruken av KI i arkivene er derfor et område som må konstant forskes på. I fremtiden kan slike modeller bli bedre på ansiktsgjenkjenning og kan for eksempel gruppere bilder av samme personer sammen. En del av bildene inneholder også tekst, så kanskje disse modellene kan kombineres med tekstgjenkjenning for å gi mer detaljerte beskrivelser. Drøftingskapittelet viste også at det er etiske betraktninger knyttet til bruken av KI. EU har nå introdusert lovgivning for å regulere dette området, basert på en risikovurdering. Generativ KI er foreløpig ikke klassifisert som høy risiko, men må fortsatt overholde krav til åpenhet og EUs opphavsrettslovgivning. (European Parliament, 2023) Det vil uansett være viktig for arkivarer å følge med på utviklingen av denne lovgivningen.

## Litteraturliste

- AIModels.fyi. (2023, 25. mars). *Demystifying AI: What Is a Pre-trained AI Model?* Hentet 19.03.24 fra <https://notes.aimodels.fyi/what-is-a-pretrained-ai-model/>
- Aksess. (2024, 12. jan). *KI + metadata = bedre treff*. Hentet 12.03.24 fra <https://www.aksess-tidsskrift.no/aktuelt/arkivdata/arkivverket/bildesok/178130>
- Arkivverket. (2024, 09. jan ). *Finn bilder enklere*. Hentet 15.03.24 fra [https://www.arkivverket.no/nyheter/finn-bilder-enklere?fbclid=IwAR2HgYmlyi0UYnRn6SOmj1WoP6u7pyaXNB7iCdd2ARP6d\\_g1jY\\_j\\_HYIZnE](https://www.arkivverket.no/nyheter/finn-bilder-enklere?fbclid=IwAR2HgYmlyi0UYnRn6SOmj1WoP6u7pyaXNB7iCdd2ARP6d_g1jY_j_HYIZnE)
- Babcock, J. & Bali, R. (2021). *Generative AI with Python and TensorFlow 2: Create Images, Text, and Music with VAEs, GANs, LSTMs, Transformer Models*. Packt Publishing, Limited.
- Benson, A. C. (2009). The Archival Photograph and Its Meaning: Formalisms for Modeling Images. *Journal of archival organization*, 148-187.
- Bountouri, L. (2017). *Archives in the digital age: standards, policies and tools* (1. utg.). Elsevier.
- Cameron, S., Franks, P. & Hamidzadeh, B. (2023). Positioning Paradata: A Conceptual Frame for AI Processual Documentation in Archives and Recordkeeping Contexts. *Journal on computing and cultural heritage, Volume 16*(4), 1-19 <https://doi.org/https://doi.org/10.1145/3594728>
- Chen, J. (2018). Artificial Intelligence IJ. D. McDonald & M. Levine-Clark (Red.), *Encyclopaedia of Library and Information Sciences* (4. utg., s. 269-278). CRC Press.
- Choempryong, S. & Wildemuth, B. M. (2016). Case Studies IB. M. Wildemuth (Red.), *Applications of Social Research Methods to Questions in Information and Library Science* (2. utg., s. 51-59). Bloomsbury Publishing USA.
- Davet, J., Hamidzadeh, B. & Franks, P. (2023). Archivist in the machine: paradata for AI-based automation in the archives. *Archival Science, Vol.23*(2), 275-295. <https://doi.org/info:doi/>
- Duff, W. & Cumming, K. (2016 ). Respect My Authority IA. Gilliland, S. McKemmish & A. Lau (Red.), *Research in the Archival Multiverse* (s. 456-478). Monash University Publishing.
- Duranti, L. (1993). Origin and Development of the Concept of Archival Description. *Archivaria*, 35(35), 47-54.
- Ertel, W. (2017). *Introduction to Artificial Intelligence* (2. utg.). Springer International Publishing
- European Parliament. (2023, 19.12.23). *EU AI Act: first regulation on artificial intelligence*. Hentet 19.05.24 fra <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- Gilliland, A. (2016). Setting the stage IM. Baca (Red.), *Introduction to Metadata* (3. utg., s. 5-37). <http://www.getty.edu/publications/intrometadata/>
- Hamill, L. (2017). *Archival arrangement and description: analog to digital*. Rowman & Littlefield.
- Han, X. Y., Papyan, V., Prokop, E., Donoho, D. L. & Johnson, C. R. (2022). Artificial Intelligence and Discovering the Digitized Photoarchive IL. Jaillant (Red.), *Archives, Access and Artificial Intelligence : Working with Born-Digital and Digitized Archival Collections* (s. 29-60). Bielefeld University Press. <https://doi.org/10.1515/9783839455845>
- Harvard Business Review. (2023, 07. april ). *Generative AI Has an Intellectual Property Problem*. Hentet 19.03.24 fra <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>
- Hodel, T. (2022). Supervised and Unsupervised: Approaches to Machine Learning for Textual Entities IL. Jaillant (Red.), *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections* (s. 157-177). Bielefeld University Press. <https://doi.org/10.1515/9783839455845>
- Horstman, C. & Necaie, R. (2016). *Python for Everyone* (3. utg.). Wiley.
- IBM. *What are AI hallucinations?* Hentet 20.05.24 fra <https://www.ibm.com/topics/ai-hallucinations>
- ICA. (2000). *ISAD(G): General international standard archival description, second edition* ICA.

- Jaillant, L. (2022). Introduction IL. Jaillant (Red.), *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections* (s. 7-28). Bielefeld University Press. <https://doi.org/10.1515/9783839455845>
- Jordan, M. I. (2019 ). Artificial Intelligence - The Revolution Hasn't Happened Yet *Harvard Data Science Review* <https://doi.org/10.1162/99608f92.f06c6e61>
- Lester, J. N., Cho, Y. & Lochmiller, C. R. (2020). Learning to Do Qualitative Data Analysis: A Starting Point. *Human resource development review*, 19(1), 94-106. <https://doi.org/10.1177/1534484320903890>
- Li, J., Li, D., Xiong, C. & Hoi, S. (2022). BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation. *arXiv.org*. <https://doi.org/10.48550/arxiv.2201.12086>
- Marciano, R. (2022). AFTERWORD: Towards a new Discipline of Computational Archive Science IL. Jaillant (Red.), *Archives, Access and Artificial Intelligence : Working with Born-Digital and Digitized Archival Collections* (s. 207-218). Bielefeld University Press. <https://doi.org/10.1515/9783839455845>
- McKinsey & Company. (2023, 19. jan ). *What is generative AI?* . Hentet 19.03.24 fra <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>
- Niu, J. (2013). Recordkeeping metadata and archival description: a revisit. *Archives and manuscripts*, 41(3), 203-215. <https://doi.org/10.1080/01576895.2013.829751>
- Noble, S. U. (2018). *Algorithms of oppression: how search engines reinforce racism*. New York University Press.
- OpenAI. (2024a). *File uploads FAQ*. Hentet 20.05.24 fra <https://help.openai.com/en/articles/8555545-file-uploads-faq>
- OpenAI. (2024b). *Introducing GPT-4o and more tools to ChatGPT free users*. Hentet 19.05.24 fra <https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/>
- Pitti, D. & Rush, M. (2018). Encoded Archival Description. I J. D. McDonald & M. Levine-Clark (Red.), *Encyclopaedia of Library and Information Sciences* (4. utg., s. 1423-1432). CRC Press.
- Poetry. *Introduction*. Hentet 19.05.24 fra <https://python-poetry.org/docs/>
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G. & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *arXiv.org*. <https://doi.org/10.48550/arxiv.2103.00020>
- Schaffner, J. (2009). *The Metadata is the Interface. Better Description for Better Discovery of Archives and Special Collections, Synthesized from User Studies*. OCLC Research. <https://www.oclc.org/content/dam/research/publications/library/2009/2009-06.pdf>
- Schwartz, J. M. (2002). Coming to Terms with Photographs: Descriptive Standards, Linguistic "Othering," and the Margins of Archivry. *Archivaria*, 54(54), 142-171.
- Senturk, B. (2013). Photographs as Archival Material. *Türk kütüphaneciliği*, 27(4), 619-632.
- Stanford HAI. (2023, 17. mars ). *AI Spring? Four Takeaways from Major Releases in Foundation Models*. Hentet 12.03.24 fra <https://hai.stanford.edu/news/ai-spring-four-takeaways-major-releases-foundation-models>
- TAM-Arkiv. *Om TAM-Arkiv*. Hentet 08.05.24 fra <https://tam-arkiv.se/om-tam-arkiv/>
- Techradar. (2023, 08.12.23). *That mind-blowing Gemini AI demo was staged, Google admits* Hentet 26.05.24 fra <https://www.techradar.com/computing/artificial-intelligence/that-mind-blowing-gemini-ai-demo-was-staged-google-admits>
- TechTarget. *Definition Hugging Face*. Hentet 19.05.24 fra <https://www.techtarget.com/whatis/definition/Hugging-Face>
- The Byte. (2024). *Turns out that extremely impressive Sora demo wasn't exactly made with Sora*. Hentet 26.05.24 fra <https://futurism.com/the-byte/openai-sora-demo>
- Wagh, D. (2019). Description and delivery. I A. D. Purcell (Red.), *The Digital Archives Handbook* (s. 19-44). Rowman & Littlefield.
- White, M. (2024, 17.04.24). *AI isn't useless. But is it worth it?* Hentet 26.05.24 fra <https://www.citationneeded.news/ai-isnt-useless/>

- Wildemuth, B. M. & Cao, L. L. (2016). Experimental Studies IB. M. Wildemuth (Red.), *Applications of Social Research Methods to Questions in Information and Library Science* (2. utg., s. 103-113). Bloomsbury Publishing USA.
- Yeo, G. (2015). Contexts, Original Orders, and Item-Level Orientation: Responding Creatively to Users' Needs and Technological Change. *Journal of archival organization*, 12(3-4), 170-185.  
<https://doi.org/10.1080/15332748.2015.1048626>
- Yin, R. K. (2018 ). *Case Study Research and Applications* (6. utg.). SAGE.
- Zeng, L. (2023). *En utforskning av automatisk generering av metadata for å øke arkivbeskrivelsen i en fotosamling* [Bacheloroppgave, OsloMet, bruk av denne oppgaven godkjent av forfatteren].

## Vedlegg

Python skripter

Kodeblokk 1 - BLIP

```
import os
import csv
from PIL import Image
from transformers import BlipProcessor, BlipForConditionalGeneration

processor = BlipProcessor.from_pretrained("Salesforce/blip-image-captioning-
base")
model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-image-
captioning-base")
mappe_sti = '\\Bilder\\tam\\'
fil_liste = os.listdir(mappe_sti)

fieldnames = ['filnavn', 'beskrivelse1', 'beskrivelse2']
with open("BLIPallebilder.csv", "w", newline='') as csvFile:
    writer = csv.DictWriter(csvFile, fieldnames=fieldnames, quotechar='')
    writer.writeheader()

    for fil in fil_liste:
        print('Skal prosesere ' + mappe_sti + '\\' + fil)
        img_file = '\\Bilder\\tam\\' + fil
        raw_image = Image.open(img_file).convert('RGB')

        #conditional beskrivelse
        text = "an image of"
        inputs = processor(raw_image, text, return_tensors="pt")
        out = model.generate(**inputs)
        beskrivelse1 = processor.decode(out[0], skip_special_tokens=True)
        print (img_file + " " + beskrivelse1)

        #unconditional beskrivelse
        inputs = processor(raw_image, return_tensors="pt")
        out = model.generate(**inputs)
        beskrivelse2 = processor.decode(out[0], skip_special_tokens=True)
        print(img_file + " " + beskrivelse2)

        writer.writerow({'filnavn': img_file, 'beskrivelse1': beskrivelse1,
'beskrivelse2': beskrivelse2})
        print("ok")

print('Finished')
```

Kode hentet fra: <https://huggingface.co/Salesforce/blip-image-captioning-large> med noen modifikasjoner

## Kodeblokk 2 - CLIP

```
import open_clip
import torch
from PIL import Image
import os
import csv

device = "cuda" if torch.cuda.is_available() else "cpu"
model, _, transform = open_clip.create_model_and_transforms(
    model_name="coca_ViT-L-14",
    pretrained="mscoco_finetuned_laion2B-s13B-b90k",
    device=device
)

bilde_path = "./tam"
csv_file = open("ouut.csv", 'w', newline='')
fieldnames = ['filnavn', 'beskrivelse']

with open("clip-output-gpu.csv", "w", newline='') as csvFile:
    writer = csv.DictWriter(csvFile, fieldnames=fieldnames, quotechar='')
    writer.writeheader()
    count = 0
    for item in os.listdir(bilde_path):
        count += 1
        if (count % 100) == 0:
            print('.')
        image = Image.open(bilde_path + "/" + item).convert("RGB")
        im = transform(image).unsqueeze(0)
        if device == "cuda":
            im = im.cuda()

        with torch.no_grad(), torch.cuda.amp.autocast():
            generated = model.generate(im)

        beskrivelse =
open_clip.decode(generated[0]).split("<end_of_text>")[0].replace("<start_of_text>", "")
        # print("Beskrivelse: ", beskrivelse)
        writer.writerow({'filnavn': item, 'beskrivelse': beskrivelse})

print('Finished')
```

Kode hentet fra <https://huggingface.co/openai/clip-vit-large-patch14> og [https://colab.research.google.com/github/mlfoundations/open\\_clip/blob/master/docs/Interacting\\_with\\_open\\_coca.ipynb](https://colab.research.google.com/github/mlfoundations/open_clip/blob/master/docs/Interacting_with_open_coca.ipynb) med noen modifikasjoner



### Kodeblokk 3- NLPconnect

```
import open_clip
import torch
from PIL import Image
import os
import csv

device = "cuda" if torch.cuda.is_available() else "cpu"
model, _, transform = open_clip.create_model_and_transforms(
    model_name="coca_ViT-L-14",
    pretrained="mscoco_finetuned_laion2B-s13B-b90k",
    device=device
)

bilde_path = "./tam"
csv_file = open("ouut.csv", 'w', newline='')

fieldnames = ['filnavn', 'beskrivelse']
with open("clip-output-gpu.csv", "w", newline='') as csvFile:
    writer = csv.DictWriter(csvFile, fieldnames=fieldnames, quotechar='')
    writer.writeheader()
    count = 0
    for item in os.listdir(bilde_path):
        count += 1
        if (count % 100) == 0:
            print('.')
        image = Image.open(bilde_path + "/" + item).convert("RGB")
        im = transform(image).unsqueeze(0)
        if device == "cuda":
            im = im.cuda()

        with torch.no_grad(), torch.cuda.amp.autocast():
            generated = model.generate(im)

        beskrivelse =
open_clip.decode(generated[0]).split("<end_of_text>")[0].replace("<start_of_text>", "")
        # print("Beskrivelse: ", beskrivelse)
        writer.writerow({'filnavn': item, 'beskrivelse': beskrivelse})

print('Finished')
```

Kode hentet fra <https://huggingface.co/nlpconnect/vit-gpt2-image-captioning> med noen modifikasjoner

## Tabeller

Tabell 1

Bildenavn	Opprinnelige beskrivelse	BLIP unconditional	BLIP conditional: «a photography of»	BLIP conditional: «an image of»	BLIP conditional: «an archival image of»
007-0033	I. G, SIF, S. A., SIF, B. H., SIF och H. F., SIF under PTK-kongressen 2-3 juni 1983	two men sitting in chairs looking at something	a photography of the first session of the united states congress committee, in the united states, july	an image of two men sitting in a lecture hall	an archival image of two men sitting in a lecture hall
009-0051	SBmfs representantskap. Middagen.	a group of people sitting at a table	a photography of a group of men and women seated at a table	an image of a group of people sitting at a table	an archival image of a group of men sitting at a table
009-0070	SBmfs representantskaps middag. I.B. opptræder med ett nummer ur Stockholmarnas bankrevy av banktjænsteman G W. W.	a woman in a white dress and hat playing a guitar	a photography of a woman playing a guitar	an image of a woman playing a guitar	an archival image of a woman in a sailor's uniform playing a guitar
009-0201	FS- internat 4-6 oktober 1973 Val till arbeidsutskott-B.H.	a man standing in front of a blackboard	a photography of a man writing on a blackboard	an image of a man writing on a blackboard	an archival image of a man writing on a blackboard
009-0343	Svenska bankmannaföreningens centralstyrelse och revisorer. Följande personer är med S. C. (v.o.), N. H., V. von Z., G. H., G. H., M. B., E. H., A. B., H. H., C. W., G. L., J. D., H. R., K. A., N. L., A. L., J. K., T. O., R. S., B. W., S. C..	a book with many different pictures of men	a photography of a group of men in formal attire	an image of a group of men in formal attire	an archival image of the first president of the united states, george h whitte, and
009-0478	Exteriör Portalen till Sv. Handelsbanken Huvudkontoret Stockholm	a black and white photo of a metal ceiling	a photography of a set of four black cast iron ceiling tiles	an image of a black and white photo of a set of four metal plaques	an archival image of the cast iron ceiling panels
009-0542	Kongress Deltagarbild	a classroom full of students	a photography of a classroom at the university of chicago	an image of a classroom with students	an archival image of a classroom at the university of chicago
055-0021	Klubbordförande för Sv. Philips. PM för december i SIF-tidningen. D. H.	a man sitting on a bench in front of a television	a photography of a man sitting on a bench in front of a television	an image of a man sitting in front of a television	an archival image of a man sitting on a bench in front of a television

055-0090	SIFs Riksstämman Från vänster:N. K.-E., L. P., K. B., A. H., E.S.	a group of men sitting at a table with flags	a photograph of the first meeting of the first world war in the united states, june 1945	an image of a group of men sitting at a table	an archival image of the first world war conference
055-0104	Riksstämman. Karrikatyren.	three sketches of people sitting at a table	a photograph of a sketch of a man and a woman	an image of three sketches of people	an archival image of a sketch of a man and a woman
055-0180	Ingenjöskonferens i Vaxholm. Diverse personer.	a newspaper article with a photo of a man in a suit	a photograph of a meeting of the german parliament	an image of a newspaper with a picture of a man in a suit	an archival image of a meeting of the german parliament
055-0480	Personal på FK och FA 16.	a group of people posing for a picture	a photograph of a group of people posing for a picture	an image of a group of people posing for a picture	an archival image of a group of people
055-0544	Diverse bilder från SIF:s konstmuseum.	a room with a bunch of old telephones	a photograph of a telephone booth with a telephone and a telephone	an image of a room with a bunch of telephones	an archival image of a telephone booth
055-0639	Från vänster: G. T., förh.chef SIF, Stockholm och P.-E. B., 1:e omb.man SIF, Stockholm	a man standing next to a statue of a man	a photograph of a statue of a man and a woman	an image of a statue of a man and woman	an archival image of a man and a woman standing next to a monument
055-0789	SIF:s 40-års jubileum. SIF 100.000 medlemmar. G. T., sekr SIF, Stockholm ställer in ett stämpelur från sekelskiftet.	a woman is putting a clock on the wall	a photograph of a woman working on a clock	an image of a woman holding a clock	an archival image of a woman working on a clock
055-1041	Tidigare byggnad på samma fastighet som det nya SIF-huset fördes upp och invigdes 1954.	a black and white photo of a building	a photograph of the building at the corner of the street in the early 1900s	an image of a building in the snow	an archival image of the building at the corner of the street
055-1970	SIF:s kongress. Fr. höger?, SSAB, SIF, kongressombud, FS-ledamot, FA24, ASO(?), TKO(?),	a group of people sitting in a room	a photograph of a crowd of people in a lecture hall	an image of a crowd of people in a lecture hall	an archival image of a crowd of people in a lecture hall
128-0056	Båthamnen i Vinga	a black and white photo of a boat in the water	a photograph of a boat in the water	an image of a boat in the water	an archival image of a boat in the water
128-0096	Sammansatt bild av lotsstation i Malmö. Bilden föreställer lots- och fyrpersonalen på stationen mellan år 1850-1926 samt stationens lotsbåtar. Små bilder överst från vänster: J.M D., J. S., S.N D., H.P J.,	a black and white photo of a group of men	a photograph of the first world war naval fleet	an image of a group of people in a circle	an archival image of the first world war naval ship, the uss nipi, which was the

	P.L.P., O.P.A., A.P.B., L.O.K., J.F.K., C.G.E.. Små bilder i imtten: M. M., J.R.L.D., W.F.N.. Små bilder nederst från vänster: M. S., S.P.B., J.A.L., F.J.O.P., J.H.B. (LF:s ombudsman&redaktör 1920-1939 resp 1916-1939), A.O.H. (LF:s ordf 1914-1915), P.A.P.T. Samuelsson Stora bilder överst från vänster: Lots- och fyrfartyg nr 27 1893-1921, Lotsfartyget Odin 1876-1893, Lotskontor 1840-1874, Lotskontor uppfört 1874, Lotsbåt 1850-1870, Lotsbåt 1870-1909				
149-0173	SSFs fullmäktige 6-8 juni 1952. Glada systrar.	nurses on the stairs of the hospital	a photograph of nurses on the stairs of the hospital	an image of nurses on the stairs	an archival image of nurses on the stairs of the hospital
149-0176	SSFs fullmäktige 10-12 juni 1956. A. H., centralstyrelsen, hälsar gästerna välkomna.	a woman drinking wine while another watches	a photograph of a woman drinking wine	an image of a woman drinking wine	an archival image of a woman drinking wine
149-0458	Kurs i arbetsledning 5-9 februari 1962. Kurs för avdelningssköterskor. D. A. L., kursens huvudlärare.	a man in a white shirt and tie sitting at a desk	a photograph of a man in a lab	an image of a man sitting at a desk	an archival image of a man in a lab
149-0659	SSFs 25 000 medlem april 1958. Eleven vid Sophiahemmet sjuksköterskeskola M. G.	a woman holding a baby in her arms	a photograph of a nurse holding a baby	an image of a woman holding a baby	an archival image of a nurse holding a baby
149-0709	Demonstration för högre löner 22 maj 1969. Sjuksköterskelever demonstrerar för bättre löner.	a group of people holding signs on the street	a photograph of a demonstration by the british labour movement in london, 1968	an image of a group of people holding signs	an archival image of a demonstration by the british labour party in london, england, in the
149-1214	Behandling av diarréer hos spädbarn. Mat är gott. Bilden är publicerad i samband med en artikel om "Morötter vid spädbarnsdyspepsi".	a baby eating a peanut butter	a photograph of a baby eating a peanut butter	an image of a baby eating food	an archival image of a baby eating a peanut butter
149-1579	Sjuksköterskebrocher. Uniform.	a bunch of different badges on a white background	a photograph of a bunch of badges	an image of a bunch of buttons	an archival image of a bunch of badges
149-3495	Cerebral angiografi av hjärna på Södersjukhuset i Stockholm. Operationa.	a black and white photo of a	a photograph of a patient in a hospital room	an image of a hospital room	an archival image of a

	Överläkaren dr S. L. punkterar. Syster E. P. assisterar, dr R. G. ger narkos och en elev instrumenterar. b. Samma tillfälle som vid fotografi a men endast system och eleven syns på fotografiet	woman in a hospital room		with a doctor and patient	woman in a hospital room
226-0036	G. N., medlem av centralstyrelsen, kassör 1910 -1912, sekreterare 1912-1925, ordförande 1926-1931, redaktör för Lärarinneförbundet från starten 1914-1933	a woman in a black and white photo	a photograph of a woman	an image of a woman in black and white	an archival image of a woman in a black and white photo
404-0033	18 juni SF:s 50 års jubileum. Banketten skall börja fr.v. Förbundsordf. D. R. statsrådet I. P., statsrådet U. L.	a man and woman sitting in front of a large black and white painting	a photograph of the president and his wife, with the president of the united states, and the	an image of a man and woman sitting in front of a large black and white painting	an archival image of the president and his wife, with the president
404-0072	S. N., Skolkökslärarinna levde mellan 1840 - 1923, insåg vikten av undervisning i hushållskunskap.	a black and white photo of a man in a suit	a photograph of person	an image of a man in a black and white photo	an archival image of person, who was the first person to be elected in the united states
404-0129	Svenska lärarmötet i Malmö. I mitten längst fram E. H.	a group of people standing in front of a building	a photograph of a group of men and women in front of a building	an image of a group of people standing in front of a building	an archival image of a group of men in front of a building
404-0141	Sammanträdesrummet, detalj med originalteckningar.	a table with a lamp and pictures on the wall	a photograph of a table with a bunch of pictures on it	an image of a table with pictures on it	an archival image of a table with a candle and a picture of a man
404-0151	SAF:s 100 års jubileum, 29 oktober Stiftelsen SAF:s ordf. G. L.-U. i talarstolen.	a woman standing at a podium with a microphone	a photograph of a woman speaking at a podium	an image of a woman speaking at a podium	an archival image of a woman speaking at a podium
404-0243	Jordfästning	a black and white photo of a church	a photograph of the interior of the st mary's church, st mary's church,	an image of a church with people sitting in pew pews	an archival image of the interior of the church
404-0248	Middag i samband med avtäckandet av Jöns Franzens porträtt. Fr.v. Okänd, O. F., okänd, H. E., H. L., överlärare K., H. N., J	a group of people sitting at a table	a photograph of a group of people sitting at a table	an image of a group of people sitting at a table	an archival image of a group of people sitting at a table

	P J., Okänd, R. W., J. W., okänd, F. H., J P. L., G. N., O. H., R. E., S. W., okänd, E. K.. Fr.h. A. J., S. B., E., A. O., H. B., okänd, S. N., J. F., E. S..				
404-0364	Hyllningsadress till Danmarks lärarförening.	a book with a picture of a woman and a man	a photograph of a book with a picture of a woman in a garden	an image of a book with a picture of a man in the middle ages	an archival image of a book with a picture of a man in a hat
404-0365	Föreningen Nordens Kurs, kursdeltagare framför Sigtunastiftelsens byggnad.	a group of people are posing for a photo	a photograph of a group of people in front of a house	an image of a group of people posing for a picture	an archival image of a group of people in front of a house
404-0578	Skollovskoloni vid midsommarstången.	a group of people standing around a tree	a photograph of a group of people gathered around a tree	an image of a group of people standing around a tree	an archival image of a group of people standing around a tree
404-0591	Möte mellan representanter för nordiska folkskollärarsammanslutnin gar ordf. i Danmarks Laererforening N. N. fram för porträtter av E. H.	a man sitting in front of a painting	a photograph of a man in a suit sitting in front of a portrait of a man	an image of a man sitting in front of a painting	an archival image of a man in a suit sitting in front of a portrait of a man
447-0014	Konferens ang. utbildningsfrågor. T. H., TCO, handläggare i utbildningsfrågor, ombudsman, Stockholm	a man in a suit and tie sitting at a table	a photograph of a man in a suit and tie sitting at a table with a bottle of beer	an image of a man sitting at a table	an archival image of a man in a suit and tie sitting at a table with a bottle of

Tabell 2

Bildenavn	Opprinnelig beskrivelse	YOLO	BLIP unconditional	BLIP conditional «an image of»	CLIP	NLP connections
007-0033	I. G, SIF, S. A., SIF, B. H., SIF och H. F., SIF under PTK-kongressen 2-3 juni 1983	5 persons	two men sitting in chairs looking at something	an image of two men sitting in a lecture hall	a black and white photo of a group of men looking at something .	a large group of people sitting on a bench
009-0051	SBmfs representantskap. Middagen.	3 persons	a group of people sitting at a table	an image of a group of people sitting at a table	a black and white photo of a group of people sitting at a table .	a large group of people sitting at a table
009-0070	SBmfs representantskaps middag. I.B. opptræder med ett nummer ur Stockholmarnas bankrevy av banktjænsteman G W. W.	(no detections)	a woman in a white dress and hat playing a guitar	an image of a woman playing a guitar	an old photo of a woman holding a guitar	a woman holding a tennis racquet on a tennis court
009-0201	FS- internat 4-6 oktober 1973 Val till arbetsutskott- B.H.	2 persons	a man standing in front of a blackboard	an image of a man writing on a blackboard	a chalkboard that is written in white chalk .	a man standing in front of a wall with a sign
009-0343	Svenska bankmannaföreningens centralstyrelse och revisorer. Följande personer är med S. C. (v.o.), N. H., V. von Z., G. H., G. H., M. B., E. H., A. B., H. H., C. W., G. L., J. D., H. R., K. A., N. L., A. L., J. K., T. O., R. S., B. W., S. C..	3 persons, 2 clocks	a book with many different pictures of men	an image of a group of men in formal attire	an old photo of a bunch of men in suits	a collage of photographs of a woman in a suit
009-0478	Exteriör Portalen till Sv. Handelsbanken Huvudkontoret Stockholm	2 clocks	a black and white photo of a metal ceiling	an image of a black and white photo of a set of four metal plaques	a black and white photo of a metal panel .	a clock that is on top of a wall
009-0542	Kongress Deltagarbild	1 person	a classroom full of students	an image of a classroom with students	a large group of people sitting at a table .	a large group of people sitting at a table
055-0021	Klubbordförande för Sv. Philips. PM för december i SIF-tidningen. D. H.	1 person, 1 tie	a man sitting on a bench in front of a television	an image of a man sitting in front of a television	a black and white photo of a man sitting in	a man sitting in front of a tv

					front of a television .	
055-0090	SIFs Riksstämman Från vänster:N. K.-E., L. P., K. B., A. H., E.S.	5 persons, 6 ties, 2 bottles, 1 cup	a group of men sitting at a table with flags	an image of a group of men sitting at a table	a black and white photo of six men sitting at a table .	a large group of people sitting at a podium
055-0104	Riksstämman. Karrikatyren.	2 persons, 1 snowboard	three sketches of people sitting at a table	an image of three sketches of people	a drawing of a man in glasses and a suit .	a painting of a man sitting on top of a table
055-0180	Ingenjörskonferens i Vaxholm. Diverse personer.	6 persons	a newspaper article with a photo of a man in a suit	an image of a newspaper with a picture of a man in a suit	a newspaper article with a picture of a man sitting at a desk .	a series of photographs of people sitting in a row
055-0480	Personal på FK och FA 16.	11 persons, 11 ties	a group of people posing for a picture	an image of a group of people posing for a picture	a group of people sitting and standing together .	a large group of people posing for a picture
055-0544	Diverse bilder från SIF:s konstmuseum.	1 chair	a room with a bunch of old telephones	an image of a room with a bunch of telephones	a group of old phones hanging on the wall .	a painting of a woman in a dress and a man in a suit
055-0639	Från vänster: G. T., förh.chef SIF, Stockholm och P.-E. B., 1:e omb.man SIF, Stockholm	2 persons	a man standing next to a statue of a man	an image of a statue of a man and woman	two men standing next to a bust of a man .	a statue of a man and a woman in front of a tree
055-0789	SIF:s 40-års jubileum. SIF 100.000 medlemmar. G. T., sekr SIF, Stockholm ställer in ett stämpelur från sekelskiftet.	1 person, 1 clock	a woman is putting a clock on the wall	an image of a woman holding a clock	a woman standing in front of a clock .	a man standing next to a machine with a clock on it
055-1041	Tidigare byggnad på samma fastighet som det nya SIF-huset fördes upp och invigdes 1954.	5 cars, 1 bench	a black and white photo of a building	an image of a building in the snow	an old photo of an old building in a city .	a building with a clock on the front of it
055-1970	SIF:s kongress. Fr. höger?, SSAB, SIF, kongressombud, FS-ledamot, FA24, ASO(?), TKO(?),	14 persons, 4 ties, 9 chairs, 1 book	a group of people sitting in a room	an image of a crowd of people in a lecture hall	a group of people sitting in a room .	a large group of people sitting in a room



128-0056	Båthamnen i Vinga	(no detections)	a black and white photo of a boat in the water	an image of a boat in the water	an old black and white photo of a sail boat .	a large boat is docked in a body of water
128-0096	Sammansatt bild av lotsstation i Malmö. Bilden föreställer lots- och fyrapersonalen på stationen mellan år 1850-1926 samt stationens lotsbåtar. Små bilder överst från vänster: J.M D., J. S., S.N D., H.P J., P.L P., O.P A., A.P B., L.O K., J.F K., C.G E.. Små bilder i imtten: M. M., J.R.L.D., W.F N.. Små bilder nederst från vänster: M. S., S.P B., J.A L., F.J.O P., J.H B. (LF:s ombudsman&redaktör 1920-1939 resp 1916-1939), A.O H. (LF:s ordf 1914-1915), P.A.P T. Samuelsson Stora bilder överst från vänster: Lots- och fyrfartyg nr 27 1893-1921, Lotsfartyget Odin 1876-1893, Lotskontor 1840-1874, Lotskontor uppfört 1874, Lotsbåt 1850-1870, Lotsbåt 1870-1909	4 clocks	a black and white photo of a group of men	an image of a group of people in a circle	a black - and - white photo of a group of people .	a collage of photos of people in various positions
149-0173	SSFs fullmäktige 6-8 juni 1952. Glada systrar.	1 fire hydrant	nurses on the stairs of the hospital	an image of nurses on the stairs	a black and white photo of a group of nurses .	a number of women posing for a picture
149-0176	SSFs fullmäktige 10-12 juni 1956. A. H., centralstyrelsen, hälsar gästerna välkomna.	1 person	a woman drinking wine while another watches	an image of a woman drinking wine	an old photo of a woman holding a glass of wine	a woman standing next to a man on a microphone
149-0458	Kurs i arbetsledning 5-9 februari 1962. Kurs för avdelningssköterskor. D. A. L., kursens huvudlärare.	1 person, 1 tie	a man in a white shirt and tie sitting at a desk	an image of a man sitting at a desk	a black and white photo of a man at a table .	a man in a suit and tie in a kitchen
149-0659	SSFs 25 000 medlem april 1958. Eleven vid Sophiahemmet sjuksköterskeskola M. G.	1 person	a woman holding a baby in her arms	an image of a woman holding a baby	a woman holding a baby in her arms .	a woman holding a baby in her arms

149-0709	Demonstration för högre löner 22 maj 1969. Sjuksköterskelever demonstrerar för bättre löner.	8 persons	a group of people holding signs on the street	an image of a group of people holding signs	a black and white photo of a group of women holding signs .	a crowd of people standing in front of a sign
149-1214	Behandling av diarréer hos spädbarn. Mat är gott. Bilden är publicerad i samband med en artikel om "Morötter vid spädbarnsdyspepsi".	1 person, 2 bottles	a baby eating a peanut butter	an image of a baby eating food	a baby in a high chair eating food .	a baby sitting at a table with a spoon
149-1579	Sjuksköterskebrocher. Uniform.	1 clock	a bunch of different badges on a white background	an image of a bunch of buttons	a bunch of buttons that are on a table .	a collage of photos of various types of electronic devices
149-3495	Cerebral angiografi av hjärna på Södersjukhuset i Stockholm. Operationa. Överläkaren dr S. L. punkterar. Syster E. P. assisterar, dr R. G. ger narkos och en elev instrumenterar. b. Samma tillfälle som vid fotografi a men endast system och eleven syns på fotografiet	1 person	a black and white photo of a woman in a hospital room	an image of a hospital room with a doctor and patient	a black - and - white photo of two doctors in a room .	a woman working in a hospital room
226-0036	G. N., medlem av centralstyrelsen, kassör 1910 -1912, sekreterare 1912-1925, ordförande 1926-1931, redaktör för Lärarinneförbundet från starten 1914-1933	1 person	a woman in a black and white photo	an image of a woman in black and white	an old black - and - white photo of a woman .	a woman in a black and white photo
404-0033	18 juni SF:s 50 års jubileum. Banketten skall börja fr.v. Förbundsordf. D. R. statsrådet I. P., statsrådet U. L.	2 persons	a man and woman sitting in front of a large black and white painting	an image of a man and woman sitting in front of a large black and white painting	three people standing in front of a painting .	a man and woman standing next to each other
404-0072	S. N., Skolkökslärarinna levde mellan 1840 - 1923, insåg vikten av undervisning i hushållskunskap.	(no detections)	a black and white photo of a man in a suit	an image of a man in a black and white photo	an old black and white photo of a woman .	an old photograph of a man in a suit
404-0129	Svenska lärarmötet i Malmö. I mitten längst fram E. H.	3 persons	a group of people standing in front of a building	an image of a group of people standing in front of a building	a group of people standing in front of a building .	a large group of people standing in front of a building

404-0141	Sammanträdesrummet, detalj med originalteckningar.	1 cup	a table with a lamp and pictures on the wall	an image of a table with pictures on it	a black and white photo of a table with a vase and candles .	a painting on a wall with a vase on top of it
404-0151	SAF:s 100 års jubileum, 29 oktober Stiftelsen SAF:s ordf. G. L.-U. i talarstolen.	1 person	a woman standing at a podium with a microphone	an image of a woman speaking at a podium	an older woman is speaking at a podium .	a woman in a suit talking on a microphone
404-0243	Jordfästning	(no detections)	a black and white photo of a church	an image of a church with people sitting in pew pews	a black - and - white photo of a group of people sitting in a church .	a large room filled with people sitting in chairs
404-0248	Middag i samband med avtäckandet av Jöns Franzens porträtt. Fr.v. Okänd, O. F., okänd, H. E., H. L., överlärare K., H. N., J P J., Okänd, R. W., J. W., okänd, F. H., J P. L., G. N., O. H., R. E., S. W., okänd, E. K.. Fr.h. A. J., S. B., E., A. O., H. B., okänd, S. N., J. F., E. S..	1 person	a group of people sitting at a table	an image of a group of people sitting at a table	a group of people sitting at a long dinner table .	a large group of people sitting at a dinner table
404-0364	Hyllningsadress till Danmarks lärarförening.	(no detections)	a book with a picture of a woman and a man	an image of a book with a picture of a man in the middle ages	a close up of two pages of an old book	a sign that is on top of a pile of books
404-0365	Föreningen Nordens Kurs, kursdeltagare framför Sigtunastiftelsens byggnad.	(no detections)	a group of people are posing for a photo	an image of a group of people posing for a picture	a group of people in front of a large building .	a large group of people sitting in front of a train
404-0578	Skollovskoloni vid midsommarstången.	(no detections)	a group of people standing around a tree	an image of a group of people standing around a tree	a group of people standing around a cross .	a crowd of people standing around a fire hydrant
404-0591	Möte mellan representanter för nordiska folkskollärarsammanslutningar ordf. i Danmarks Laererforening N. N. fram för porträtter av E. H.	1 person, 1 chair	a man sitting in front of a painting	an image of a man sitting in front of a painting	an older man standing in front of a painting .	a man in a suit and tie standing in front of a painting

447-0014	Konferens ang. utbildningsfrågor. T. H., TCO, handläggare i utbildningsfrågor, ombudsman, Stockholm	1 person, 1 tie, 2 bottles	a man in a suit and tie sitting at a table	an image of a man sitting at a table	an old photo of a man wearing glasses	a man sitting at a table with a glass of wine
----------	---	----------------------------	--	--------------------------------------	---------------------------------------	---

## Resultater fra ChatGPT

### Forespørsel 1

Den vedlagte arkivserien består hovedsakelig av fotografier som dokumenterer ulike samlinger og møter i forskjellige sosiale og offentlige settinger. Basert på bildenes beskrivelser, kan vi identifisere flere gjentakende temaer og motiver som gir et innblikk i arkivets overordnede innhold og betydning.

### ### Overordnet beskrivelse av arkivserien

#### \*\*Temaer og Motiver:\*\*

##### 1. \*\*Sosiale Samlinger:\*\*

- Mange bilder viser grupper av mennesker samlet i forskjellige kontekster, som foran bygninger, i kirker, og i rom. Eksempler på dette er beskrivelse 226-0001 (en gruppe mennesker foran en bygning) og 226-0035 (en stor folkemengde i et rom).

##### 2. \*\*Formelle Møter:\*\*

- En betydelig del av bildene dokumenterer formelle møter og diskusjoner rundt bord. Dette inkluderer grupper av mennesker som sitter rundt bord med dokumenter og papirer, som beskrevet i 226-0010 (to kvinner ved et bord med papirer) og 226-0097 (en mann ved et bord med bøker).

##### 3. \*\*Portretter og Poseringer:\*\*

- Flere bilder er portretter eller poserte fotografier av enkeltpersoner eller grupper. For eksempel, beskrivelse 226-0022 (et svart-hvitt foto av en mann i dress) og 226-0039 (en kvinne som gir et kunstverk til en annen kvinne).

##### 4. \*\*Kulturelle og Kunstneriske Hendelser:\*\*

- Noen bilder viser kulturelle aktiviteter som kunstutstillinger eller teaterforestillinger. Dette inkluderer beskrivelse 226-0039 (en kvinne som gir et kunstverk til en annen kvinne) og 226-0169 (en folkemengde som ser på en film).

##### 5. \*\*Kvinnegrupper:\*\*

- En markant del av bildene fokuserer på grupper av kvinner i ulike aktiviteter, fra møter til sosiale sammenkomster. Eksempler inkluderer beskrivelse 226-0002 (en gruppe kvinner i en kirke) og 226-0033 (kvinner som sitter på gresset).

#### \*\*Aktiviteter og Miljøer:\*\*

- **\*\*Innendørs Samlinger:\*\*** Bildene viser ofte folk som sitter eller står i rom, ofte rundt bord eller i stoler, som i beskrivelse 226-0020 (folk rundt et bord) og 226-0066 (folk ved bord i et rom).

- **\*\*Uteområder:\*\*** Noen bilder er tatt utendørs, som folk som sitter på gresset eller foran bygninger, beskrevet i 226-0033 (kvinner på gresset) og 226-0181 (folk på trappene til en bygning).

## **\*\*Mangfold i Demografi:\*\***

- Arkivserien dokumenterer et bredt spekter av mennesker, inkludert menn, kvinner og barn, og gir et variert bilde av samfunnet som er representert. For eksempel, beskrivelse 226-0200 (en gutt med blondt hår) og 226-0087 (folk som poserer foran en bygning).

### **### Konklusjon**

Denne arkivserien gir et rikt innblikk i sosiale og formelle interaksjoner i en bestemt periode. Bildene fremhever viktigheten av både formelle møter og uformelle samlinger, med en spesiell fokus på kvinners deltakelse i samfunnet. Gjennom en blanding av portretter, gruppebilder, og dokumentasjon av kulturelle arrangementer, tilbyr denne serien en omfattende visuell dokumentasjon av menneskelige aktiviteter og samspill.

### Forespørsel 2

Arkivserien tilhørende Svensk Sjuksköterskeförning, basert på de medfølgende bildebeskrivelsene, gir et omfattende innblikk i foreningens aktiviteter og medlemmer. Bildene spenner over en rekke situasjoner og settinger som reflekterer det daglige livet og spesielle anledninger innenfor organisasjonen. Her er en overordnet beskrivelse av denne arkivserien:

### **### Overordnet Beskrivelse av Arkivserien**

#### **\*\*Temaer og Motiver:\*\***

##### **1. \*\*Gruppebilder og Sosiale Samlinger:\*\***

- En betydelig andel av bildene viser grupper av mennesker samlet i forskjellige settinger, som foran bygninger, i kirker, rundt bord, og i ulike rom. Disse gruppene består ofte av kvinner, men også av menn og blandede grupper. Mange av disse samlingene ser ut til å være formelle møter, sosiale sammenkomster, eller spesielle arrangementer.

##### **2. \*\*Møter og Arbeidsmiljø:\*\***

- Flere bilder viser mennesker som sitter rundt bord med papirer, bøker, og blomster, noe som antyder at dette er arbeidsmøter, konferanser eller seminarer. Det er også bilder av folk i klasserom, som peker mot utdanningsaktiviteter eller opplæring.

##### **3. \*\*Seremonier og Offentlige Arrangementer:\*\***

- Bildene inkluderer scener med folk på en scene, ved talerstoler, og ved andre formelle seremonier. Dette indikerer at foreningen også deltar i eller arrangerer offentlige arrangementer og seremonier.

##### **4. \*\*Kulturelle og Fritidsaktiviteter:\*\***

- Det er bilder av folk som sitter på gresset, på benker, eller i teateroppsett, som kan indikere fritidsaktiviteter, sosiale arrangementer eller kulturelle hendelser.

## 5. **Portretter og Personlige Interaksjoner:**

- Det finnes også en del portrettbilder, samt bilder av personer som interagerer med hverandre i mer personlige settinger, som kan være representative for vennskap, samarbeid eller individuelle prestasjoner innen foreningen.

### **Bred Representasjon:**

- Bildene viser en bred representasjon av foreningens medlemmer, hovedsakelig kvinner, som var typisk for sykepleieryrket i historisk kontekst. Det er også tilstedeværelse av menn i noen av bildene, som kan indikere en blanding av kjønnsroller innenfor visse aspekter av foreningens aktiviteter.

### **Historisk og Kulturell Verdi:**

- Samlingen gir verdifull innsikt i arbeidsmiljøet, sosiale strukturer, og kulturelle praksiser innenfor Svensk Sjuksköterskeförening over tid. Det reflekterer utviklingen i sykepleieryrket og kan fungere som et historisk dokument over foreningens rolle i samfunnet.

Denne arkivserien er dermed en rik visuell dokumentasjon som kan brukes til å forstå både det formelle og det uformelle livet i Svensk Sjuksköterskeförening, og gir en dyptgående innsikt i deres historie og kultur.

## Forespørsel 3

Fotoarkivet som beskrevet i dokumentet inneholder en omfattende samling av hovedsakelig svart-hvitt fotografier. Disse bildene kan deles inn i flere kategorier basert på deres innhold:

### 1. **Natur og landskap:**

- Mange bilder viser landskap som vann, sjø, innsjøer, strender og fjell. Det er mange bilder av fyrårn på forskjellige steder som steinete øyer og ved sjøen.

- Det finnes også bilder av trær, fjell, skogsområder og felt med vindmøller, som fanger naturens skjønnhet og variasjon.

### 2. **By- og bygdeomgivelser:**

- Arkivet inneholder bilder av bygninger som hus, kirker, klokketårn, og havner. Mange av disse bildene er tatt fra luften og gir et godt inntrykk av byens struktur og arkitektur.

- Det er også bilder av gater og mennesker som går, samt bygg under konstruksjon.

### 3. **Maritime bilder:**

- En stor del av bildene i arkivet er maritime, med fokus på båter, skip, havner, fyrårn og folk som jobber på eller ved sjøen. Dette inkluderer gamle seilbåter, fiskebåter og moderne skip.

### 4. **Personer og portretter:**

- Det finnes en mengde portretter og bilder av mennesker i ulike settinger. Dette inkluderer alt fra formelle portretter til uformelle bilder av mennesker som smiler, arbeider, eller er i grupper.

- Mange bilder viser mennesker i yrkesrelaterte situasjoner, som leger, sykepleiere, arbeidere på verft, og personer i laboratorier og kontorer.

#### 5. **\*\*Historiske øyeblikk og gjenstander\*\***:

- Arkivet inneholder også bilder av historiske øyeblikk, inkludert gamle bilder av mennesker i uniform, gamle kjøretøy, og bygninger med historisk betydning. Det er også bilder av gjenstander som bøker, dokumenter, og verktøy som er viktige for å forstå fortiden.

Dette fotoarkivet gir en rik visuell dokumentasjon av både det naturlige og menneskelige miljøet gjennom tidene, med et sterkt fokus på maritime aktiviteter og hverdagsliv i ulike samfunn.

#### Forespørsel 4

Tjänstemäns & akademikers fotoarkiv er et omfattende fototek, som hovedsakelig inneholder svart-hvitt bilder som viser ulike profesjonelle og akademiske aktiviteter. Bildene fanger et bredt spekter av scener, inkludert møter, forelesninger og sosiale sammenkomster, ofte med grupper av menn og kvinner som sitter rundt bord eller står i formelle poseringer.

Samlingen fremhever flere tilbakevendende temaer:

1. **\*\*Profesjonelle møter og konferanser\*\***: Mange bilder viser grupper av menn og kvinner som er engasjert i diskusjoner, ofte sittende rundt bord eller stående i formelle omgivelser. Disse bildene antyder en fokus på å dokumentere profesjonelle og akademiske samlinger, og fremhever de samarbeidende og kommunikative aspektene ved disse miljøene.
2. **\*\*Offentlige taler og presentasjoner\*\***: Flere bilder viser individer som holder taler eller presentasjoner, noe som indikerer en vektlegging av lederskap og offentlig kommunikasjon innenfor profesjonelle og akademiske kontekster.
3. **\*\*Sosiale interaksjoner\*\***: Arkivet inneholder mange bilder av mennesker som interagerer sosialt, enten i formelle settinger som konferanser og møter, eller i mer uformelle omgivelser. Denne aspekten understreker viktigheten av nettverksbygging og relasjonsbygging i profesjonelle og akademiske kretser.
4. **\*\*Arbeidsplasser\*\***: Det er også bilder av individer ved sine skrivebord, arbeider på datamaskiner eller engasjert i andre arbeidsrelaterte aktiviteter. Disse bildene gir et innblikk i hverdagslivet til profesjonelle og akademikere, og fremhever mangfoldet av oppgaver og miljøer de navigerer i.
5. **\*\*Historiske og kulturelle øyeblikk\*\***: Noen bilder fanger viktige historiske eller kulturelle øyeblikk, som marsjer, prisutdelinger og bemerkelsesverdige begivenheter, noe som



reflekterer den bredere samfunnsmessige innflytelsen og involveringen til profesjonelle og akademikere.

Samlet sett fungerer Tjänstemäns & akademikers fotoarkiv som en rik visuell dokumentasjon av profesjonelt og akademisk liv, og fremhever interaksjonene, arbeidet og begivenhetene som definerer disse samfunnene .