

Fortid & notid for framtid

Arkiv og kulturformidling

Sluttrapport «FORSETE» - Forprosjekt for utvikling
av modell for gjenkjenning av moderne handskrift
på norsk.

IKA Møre og Romsdal IKS

1.	Innleiing.....	3
2.	Gjennomføring	3
3.	Mål	4
4.	Resultat og gevinstar	4
5.	Formidling	7
6.	Læringsrapport.....	8
7.	Økonomi.....	9



1. INNLEIING

Nasjonalarkivets referanse: 2024/13209

Startdato: 03.01.2025

Sluttdato: 02.03.2026

Arkivsektoren har behov for å handsame og å forstå store mengder data som inneheld handskrifa på norsk. Det er i dag fleire pågåande prosjekt som utviklar maskinassistert handskriftgjenkjenning, desse fokusera på trening av modellar for å forstå «eldre» handskrift på norsk. Dette prosjektet hadde som føremål å trene ein modell på «moderne» handskrift for effektiv gjenkjenning og verifisering av innhald i tekst og layout-analyse.

Prosjekt tittel var valt med omhug; «FORSETE» - rettferdsguden i den norrøne mytologien, han som fører forsetet blant dommarane. Prosjektet bring med seg ikkje mindre enn Arkivlovas formålsparagraf: "Føremålet med denne lova er å tryggja arkiv som har monaleg kulturelt eller forskingsmessig verdi eller som inneheld rettsleg eller viktig forvaltningsmessig dokumentasjon, slik at desse kan verta tekne vare på og gjorde tilgjengelege for ettertida."

Prosjektgruppa har bestått av frå Møre og Romsdal politidistrikt, Maria Fausa, leiar dokumentasjonsforvaltning og arkiv ved Møre og Romsdal politidistrikt, med bistand frå rådgjevar Vibecke Beate Valkvæ, rådgjevar Linn Vatlé Wiik, rådgjevar Julie Wren og politifagleg etterforskingseiar Ole Einar Lundanes Alstadsæter.

Frå NTNU i Gjøvik, Fakultet for informasjonsteknologi og elektroteknikk, Institutt for datateknologi og informatikk, Erlend Rømo, Martin Morisbak og Arthur Borger Thorkildsen, alle bachelor studentar i ingeniørfag data med rettleiing frå professor Ivar Farup.

Frå Interkommunalt arkiv for Møre og Romsdal, førstekonsulent Josef Dan Karlsson og prosjektleiing frå rådgjevar Ottar A.B. Anderson ved Fylkessenter for digitalisering av kulturarven (SEDAK).

2. GJENNOMFØRING

I samarbeid med Møre og Romsdal politidistrikt og Fakultet for informasjonsteknologi og elektroteknikk, Institutt for datateknologi og informatikk ved NTNU i Gjøvik, gjennomførte Interkommunalt arkiv for Møre og Romsdal prosjektet innanfor framlagte plan og budsjett i søknad, med justeringar innrapportert ved statusrapportering. Prosjektet fekk tildeling innanfor programområde 1. – Ein innovasjonsarena for arkiv, innsatsområde Arkivspira. Der søkarar kan teste ut, utforske, pilotere eller utvikle ein idé, arbeidsmåte, metodikk eller løysing.

Eit eige dedikert treningssett av data med «moderne» handskrift i «Espås-saka» frå Møre og Romsdal politidistrikt var utgangspunktet for utviklinga av systemet. Espås-saka er ei av Norges mest omtalte og omfattande uopplåste drapssaker. Trude Espås var ei kvinne på 20 år som vart valdteken og drepen i Geiranger i Møre og Romsdal 8. august 1996. Saka har fleire gangar blitt gjenoppteken, utan å bringe nokre konkrete resultat. Saka er under operativ etterforskning og den har ingen foreldingsfrist. I august i år blir det 30 år sidan hendinga.

Nettopp prosjektets unike tilgang til datasettet gav den nødvendige ekstra merksemda til prosjektet, både frå arkivsektoren, academia ved IT-studentane, fagmiljøet i Politiet og sivil- og rettssamfunnet generelt.

Møre og Romsdal politidistrikts arbeid har i stor grad bestått i å beskrive behov og ikkje-behov inn i prosjektgruppa, ofte sett frå en etterforskers perspektiv. Arbeidet har omfatta å gi innspel til kva for funksjonalitet politiet har behov for, og kva for løysningar eller funksjoner som ikkje er nødvendige. Det har vært et mål å holde verktøyet på et minimumsnivå når det gjelder funksjonalitet, slik at det dekker politiets behov utan å innehalde meir er enn det som er nødvendig.

I tillegg har tilsette ved Dokumentasjonsforvaltning brukt betydeleg tid på å omsetje handskrivne tips til maskinskrift. Disse omsettinga vart nytta som grunnlagsmateriale i arbeidet med å trene og vidareutvikle verktøyet (AI-en) til å kjenne igjen og tolke handskrift.

3. MÅL

Prosjektets mål er var å ta tak i ei utfordringa som ennå ikkje var adressert, trening av ein modell for «moderne» handskrift på norsk. Prosjektet vart med dette det fyrste som adresserte denne utfordringa. Sidan det finnes svært store mengder arkiv som inneheld «moderne» handskrift frå periodane 1940 til 1990, både innanfor statelege- fylkeskommunale- og kommunale tenester, så fekk ein her ein vinn-vinn situasjon i samarbeidet med Møre og Romsdal politidistrikt. Arbeidsmetodikk som i stor grad var nytta ved Møre og Romsdal politidistrikt på midten av 90-talet, i den konkrete saka, var å dokumentere tips med handskrift. Det er eit utval av desse digitiserte/mediekonverterte tipsa i saka som er datasettet som vart nytta i prosjektet.

Utan ei effektiv og sikker måte å søke digitalt som ivareteke prinsippa proveniens, autentisitet og rettstryggleik, så får datasetta/arkiva avgrensa nytte. Ei anna utfordringa prosjektet også indirekte belys, er tanken om digitisering/mediekonvertering for kassasjon. Her vil ein heller ikkje oppnå noko gevinst realisering om dei same prinsippa ikkje blir 100% innfridd. Om kvaliteten på mediekonverteringa ikkje er verifisert gjennom bruk av internasjonale standardar som ISO 19264, så vil ein ikkje kunne være 100% sikker på om informasjonen i det fysiske arkivet har blitt digitalt transformert. Ved kassasjon av det fysiske arkivet, så er det som kjent ingen veg tilbake for eit nytt forsøk.

4. RESULTAT OG GEVINSTAR

Ynskja gevinst for prosjektet var det nødvendige «proof of concept» arbeidet som var gjennomført i bacheloroppgåva for å bygge eit kunnskapsgrunnlag for ei potensiell vidareføring i eit hovudprosjekt.

Samarbeidet på tvers av samfunnsinstitusjonar, der ein fylkeskommunal og ein stateleg institusjon samarbeider samstundes som ein bring inn akademia i FoU-arbeid, er også ein viktig gevinst vi vil trekke fram. Det er av prosjektgruppa si oppfatning at slike samarbeid er nødvendige for å løyse samfunnsoppdraget på ein samfunnsøkonomisk og berekraftig måte, innanfor rettsprinsippa. FN-sambandet sine berekraftsmål, nummer 17, Samarbeid for å nå måla er gjennomgåande i prosjektet. Resultata vil potensielt på sikt vil kunne bli eit operativt verktøy for vårt arbeid for sivil- og rettssamfunnet både innanfor arkivsektoren og Politiet.



FN-sambandet, Samarbeid for å nå måla



Bacheloroppgave i ingeniørfag – Data

ved Norges teknisk-naturvitenskapelige universitet (NTNU), Fakultet for informasjonsteknologi og elektroteknikk, Institutt for datateknologi og informatikk

Sammendrag fra Bacheloroppgåva:

«Prosjektet tar for seg et presserende behov innen norsk politiarbeid og arkivforvaltning: rask og personvernvennlig transkribering av moderne håndskrevne dokumenter. Med utgangspunkt i 2 500 digitiserte tips-notater fra den uløste «Espås-saken» fra 1996, satte teamet seg fore å utvikle et ende-til-ende-system som kan kjøre fullstendig innenfor et lukket politinettverk, uten å sende data til eksterne skytjenester. Løsningen som ble utviklet kombinerer en modulær pipeline for håndskriftgjenkjenning (HTR), kalt Forsete-ATR, med en webapplikasjon, Forsete-WEB, som lar etterforskere laste opp PDF- eller bildefiler, gjennomgå transkripsjoner side om side, og eksportere korrigert tekst.

Forsete-ATR styres av HTRflow, som gjør det mulig å definere steg-for-steg-pipelines i YAML. Standard-pipelinen segmenterer først dokumentene i tekstlinjer ved hjelp av en YOLOv9-modell som er finjustert av det svenske Riksarkivet. Deretter gjenkjennes disse linjene med TrOCR-norhand-v3, en transformer-modell tilpasset norsk håndskrift. Til slutt sorteres og eksporteres linjene som ren tekst eller strukturert JSON. Både back end (Go, PostgreSQL) og front end (TypeScript, Tailwind) er utviklet som en tredelt arkitektur. Systemet er pakket med Docker og kan eventuelt provisjoneres med Terraform, slik at hele løsningen enkelt kan distribueres lokalt.

Oppgaven konkluderer med at en lokalt driftet, åpen kildekode-basert ATR-pipeline er både teknisk gjennomførbar og organisatorisk akseptabel i sensitive etterforskninger. Fremtidig arbeid bør fokusere på mer avanserte korrekturverktøy i brukergrensesnittet, automatisk layout-gjenkjenning for komplekse skjemaer, og videre finjustering av modeller etter hvert som bedre annotert treningsdata blir tilgjengelig».

Kjelde:

Bacheloroppgåve 2025 av Erlend Rømo, Martin Morisbak og Arthur Borger Thorkildsen veileder Ivar Farup
«FORSETE – A Locally Hosted Automatic Text Recognition Platfor for the Espås Case»
 Norges teknisk-naturvitenskapelige universitet (NTNU), Fakultet for informasjonsteknologi og elektroteknikk, Institutt for datateknologi og informatikk

Open offentlig tilgjengelig bacheloroppgave:

<https://nva.sikt.no/registration/0198eaf3c763-ed8a7d1-8644-461f-b190-f8d37c0180bf>

Open offentlig tilgjengelig dokumentasjon for Forsete-HTR på GitHub:

<https://github.com/erlendromo/forsete-atr/>

Open offentlig tilgjengelig dokumentasjon for Forsete-WEB på GitHub:

<https://github.com/erlendromo/forsete-web/>



Forside av Bacheloroppgåva FORSETE: A Locally Hosted Automatic Text Recognition Platform for the Espås Case

5. FORMIDLING

Prosjektet var presentert ved Arkivarforeningen vårseminar 2 april 2025 i Oslo av Maria Fausa frå Møre og Romsdal politidistrikt, Erlend Rømo, student ved NTNU i Gjøvik og Ottar A.B. Anderson ved IKA Møre og Romsdal

Lenke til program:

<https://www.arkivarforeningen.no/arkivarforeningens-varseminar-2025/>



Frå venstre Maria Fausa frå Møre og Romsdal politidistrikt, Ottar A.B. Anderson frå IKA Møre og Romsdal og Erlend Rømo, student ved NTNU i Gjøvik presenterer ved Arkivarforeningen vårseminar
Foto: Vibeke Lunheim

Presentasjon ved Arkivarforeningen lokalgruppe Sunnmøre ved Ottar A.B. Anderson 19. Juni 2025 i Ålesund.

TV2 publiserte ein artikkel om prosjektet 2. juli 2025 med tittel «Nytt håp i 29 år gammel drapsgåte» av Arne Rovick.

Lenke til artikkel: <https://www.tv2.no/nyheter/innenriks/nytt-hap-i-29-ar-gammel-drapsgate/17878610/>

Tidsskriftet AKSESS publiserte ein utfyllande artikkel om prosjektet 6. Juli 2025 med tittel «Fra drapsgåte til arkivløsning» av Jeanette Høie
Lenke til artikkel: <https://www.aksess-tidsskrift.no/aktuelt/forsete/htr/handskriftsgjenkjenning/250322>

Presentasjon ved Privatarkivnettverket for Møre og Romsdal ved Ottar A.B. Anderson 29. Oktober ved Nynorsk kultursentrum i Ørsta.

Digital presentasjon ved Privatarkivnettverk Akershus koordinert av Museumstenesten ved Museene i Akershus ved Ottar A.B. Anderson 25. November 2025.

6. LÆRINGSRAPPORT

Manglande tilgang på datakraft for implementering innanfor arkivsektoren

Arkivsektoren har pr i dag ikkje økonomiske verkemidlar for å kunne gjennomføre dei nødvendige innkjøpa av hardware og software for implementering av teknologien. Ein er heilt avhengig av at bevilgande myndigheit gjer grep for at dette kan skje, utan tilgang på grunnleggande verktøy så vil ikkje sektoren ha noko som helst gevestrealisering av dei nye verktøya innanfor maskinassistert handskriftgjenkjenning. Tilknytning til UH-sektoren og andre forskingsmiljø er ei moglegheit for trening av modellar, men ein må også ha grunnleggande verktøy internt i sektoren for å operativt handtere resultat og utføre test trening og bygge kompetanse. Prosjektet har vist at det er eit stort potensiale i å fortsette trening på «moderne» norsk handskrift basert på språkmodellane Nasjonalbiblioteket har gjort tilgjengeleg, desse er som kjent basert på «eldre» handskrift. At det Svenske Nasjonalarkivet opent tilgjengeleg publiserte sin HTR Flow modell danna heile grunnlaget for prosjektet, her ville ein med fordel tenkt at Nasjonalarkivet, gjerne i samarbeid med utviklarane ved Nasjonalbiblioteket kunne dele meir av sine resultat og verktøy med resten av den norske arkivsektoren.

Læringsutbytte for Møre og Romsdal politidistrikt

Leder i Dokumentasjonsforvaltning og arkiv ved Møre og Romsdal politidistrikt har hatt hyppig dialog med Politidirektoratet, særorganene og øvrige politidistrikter gjennom prosjektperioden. Formålet har vært å informere om prosjektet og evt. finne mulige samarbeidspartnere samt skape bevissthet om mulighetene verktøyet kan gi for politi Norge. Videre har det vært et mål å avklare den mest hensiktsmessige veien videre for en eventuell nasjonal innføring.

Tilbakemeldingen har utelukkende vert positive. Politidirektoratet, flere politidistrikter, Nasjonale KI miljøer samt enheten for kalde saker ved Kripas, har gitt uttrykk for at dette er et verktøy det er et klart behov for, og vil kunne gi betydelig nytte i politiets arbeid.

Politidirektoratet har gitt uttrykk for et ønske om at verktøyet videreutvikles og innføres nasjonalt i politi-Norge. Den videre utviklingen og eventuelle innføringen av verktøyet i Politi Norge vil måtte ses i sammenheng med Politidirektoratets øvrige nasjonale utviklingsarbeid og strategiske satsinger. Hvordan arbeidet følges opp videre, vil derfor avhenge av de vurderinger og beslutninger som gjøres i de aktuelle prosessene.

Tilbakemeldingene fra Politidirektoratet har så langt vært svært positive. Det er derfor grunn til å tro at verktøyet vil bli gjenstand for videre satsing og nasjonal innføring.

Prosjektet har gitt verdifull erfaring med tverrfaglig samarbeid. Samarbeidet med Interkommunalt arkiv i Møre og Romsdal har vært en viktig forutsetning for prosjektets gjennomføring, og politiet har vært avhengig av deres kompetanse og erfaring fra tidligere samarbeid med NTNU.

Prosjektet har også gitt innsikt i hvordan tre ulike fagmiljøer kan bidra til å løse konkrete utfordringer. Samarbeidet med Interkommunalt arkiv for Møre og Romsdal og NTNU har vist hvilke muligheter som ligger i å kombinere erfaring og etablert fagkompetanse med ny kunnskap og teknologisk kompetanse.

7. ØKONOMI

IKA Møre og Romsdal interne prosjektnummer i rekneskap: 1075

Dokumentasjon IKA Møre og Romsdal - Sak.:arkiv: 24/834

Budsjett

	Inntekter	Arbeid	Utgifter
Søknad Nasjonalarkivet	250000,-	Lønn prosjektleiar/team	160400,-
		Løn Møre og Romsdal politidistrikt	20050,-
		Reise opphald, møter og formidling	69550
Totale inntekter	250000,-	Totale utgifter	250000,-

Rekneskap

	Inntekter	Arbeid	Utgifter
Tilskot frå Nasjonaarkivet	250000	Lønnsutgifter 2025/2026	207613,-
		Reiser NTNU	19525,-
		Løn Møre og Romsdal politidistrikt	20050,-
		Database – nødvendig sikker sone	2812,-
Totale inntekter	250000,-	Totale utgifter	250000,-

Viser til godkjent omdisponering av søknadsmidler frå reiseaktivitet (69550,-) til bruk på timer pga. at forventa budsjettert reiseaktivitet ikkje var nødvendig. Stadfesta av Nasjonalarkivet 01.07.2025 i sak 2024/13209.

Ottar A.B. Anderson
Prosjektleiar

Ålesund 01.06.2026