



# Helsearkivregisteret åpner nye dører for forskningen

---

## FRA FAGMILJØENE

### ELIN BJØRNSTAD-TUVENG

Elin Bjørnstad-Tuveng er sykepleier med mastergrad i folkehelsevitenskap, seniorrådgiver i Helsearkivregisteret og ph.d.-student, der første del av prosjektet fokuserer på kunstig intelligens og identifikasjon av obduksjonsrapporter i Helsearkivregisteret.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

### KARI NYTRØEN

karnyt@arkivverket.no

Kari Nytrøen er fagdirektør i Helsearkivregisteret og har en ph.d. i preventiv kardiologi og hjerterehabilitering. Hun har profesorkompetanse med erfaring fra forskning, veiledning, undervisning og prosjektledelse.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

---

Helsearkivregisteret inneholder snart én million komplette pasientjournaler, og det vokser fort. Registeret tar nå i bruk kunstig intelligens for å gjøre det lettere å generere data til forskningsbruk.

Norsk helsearkiv forvalter Helsearkivregisteret som åpnet i 2019 og er et sentralt helseregister bestående av avdøde pasienters journaler (1,2). Alle virksomheter i spesialisthelsetjenesten plikter å avlevere sine journaler dit. Elektroniske journaler mottas digitalt, mens eldre papirjournaler blir maskinlest og digitalisert. Komplette journaler fra avdøde personer utgjør store mengder data som kan tilgjengeliggjøres. Til forskjell fra vanlige registerdata kan de beskrive hele pasientforløp.

## Innsikt i ustrukturerte data

Helsearkivregisteret har strukturerte metadata, dvs. lett søkbare data som overordnet beskriver innholdet i hver pasientjournal, f.eks. personnummer, dødsdato og ICD-koder. I tillegg inneholder registeret store mengder ustrukturerte data i form av fritekst i pasientjournaler. Slike data kan være tidkrevende å finne frem i. Hvis en forsker ønsker spesifikke data som ikke inngår i metadataene, kreves det vanligvis manuell leting.

Helsearkivregisteret har derfor inngått et innovasjonspartnerskap med Anzyz Technologies og videreutvikler et forskningsstøtteverktøy basert på en språkalgoritme (Natural Language Processing) som åpner nye muligheter for å søke i ustrukturerte data. Den kunstige intelligensen vil gi enklere og mer presis fremfinning av både kvantitativt og kvalitativt innhold som forskere søker etter. For eksempel kan man finne frem til hvilke journaler som inneholder klokketester, som en forsker nå søker etter som grunnlag for en prediksjonsmodell for demens.

Våren 2022 ble forskningsstøtteverktøyet testet og evaluert for sin evne til å identifisere obduksjonsrapporter. 86 % av treffene var faktiske obduksjonsrapporter, og de falskt positive besto i all hovedsak av standarddokumenter som det bør være mulig å luke vekk. Det var ingen falskt negative obduksjonsrapporter. Verktøyet lyktes i å identifisere like mye som ved manuell leting, bare mye raskere (3). Forskningsstøtteverktøyet videreutvikles og testes i prosjekt høsten 2022, og rundt årsskiftet vil det være klart til bruk.

Kunstig intelligens kan også bedre personvernet. Slike verktøy reduserer behovet for manuell gjennomgang av journaler og dermed unødvendig eksponering av helsedata. Det kan også spare ressurser ved utlevering og mottak av data. Forskere kan søke om data fra Helsearkivregisteret på [helsedata.no](https://helsedata.no). Utlevering av personidentifiserbare helseopplysninger krever at nødvendige godkjenninger er på plass. Enkelt dokumenter med spesifikt innhold kan hentes ut og enkelt anonymiseres for utlevering til forskning, i motsetning til anonymisering av hele pasientjournaler, som i praksis er umulig. Kunstig intelligens kan også gjøre det mulig å generere retrospektive kontrollgrupper til planlagte forskningsprosjekter.

Helsearkivregisteret har store mengder data som inneholder over 100 år med erfaring og kunnskap. Ikke bare inneholder registeret andre typer data enn det som tradisjonelt rapporteres til helseregistre, men også all den finmaskede informasjonen som finnes i pasientjournalene. Kunstig intelligens øker tilgjengeligheten til det rike datamaterialet slik at det kan brukes på nye måter.

---

## REFERENCES

1. Arkivverket. Norsk helsearkiv. <https://www.arkivverket.no/om-oss/norsk-helsearkiv/> Lest 9.9.2022.
2. Helsedata. Helsearkivregisteret. <https://helsedata.no/no/forvaltere/arkivverket/helsearkivregisteret/> Lest 9.9.2022.
3. Bjørnstad-Tuveng E, Helstad G, Abouzeid A et al. Identifying autopsy-reports in unstructured, digitized patient records: A pilot-study from the Norwegian Health Archives Registry. hAIst Conference – Health-related Artificial Intelligence in Stavanger. <https://ebooks.uis.no/index.php/USPS/catalog/book/195/> Lest 22.9.2022.

---

Publisert: 7. november 2022. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.22.0570

Mottatt 5.9.2022, første revisjon innsendt 9.9.2022, godkjent 22.9.2022.

© Tidsskrift for Den norske legeforening 2022. Lastet ned fra [tidsskriftet.no](https://tidsskriftet.no) 2. desember 2022.