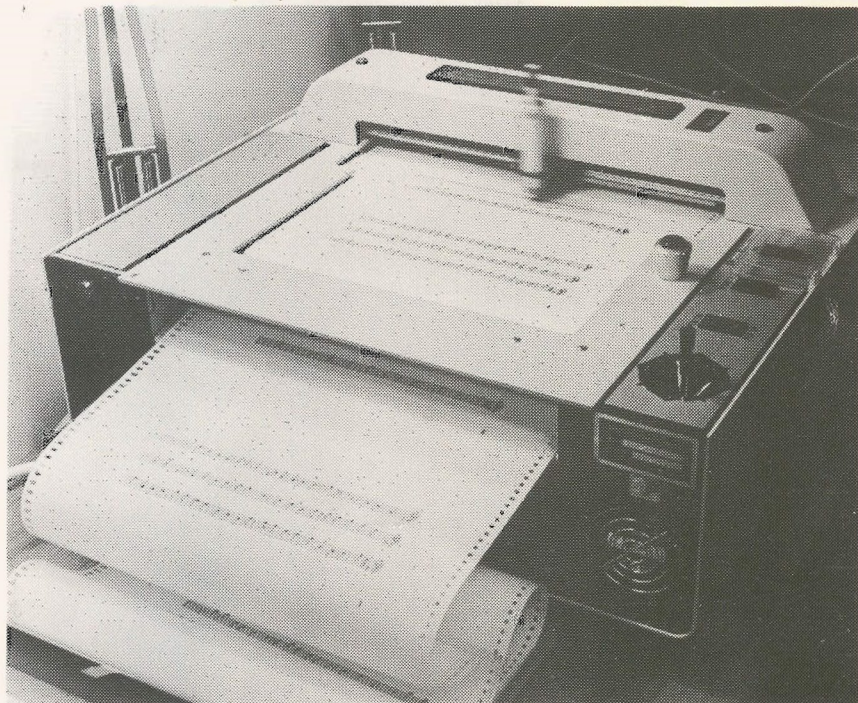


# humanistiske data 1-1981



**NAVF**

ARTIKLER  
RAPPORTER  
MELDINGER  
SUMMARY

NAVF's EDB-senter  
for humanistisk  
forskning

The Norwegian  
Computing Centre for  
the Humanities.



# Innhold

## Artikler

EDB som hjelpemiddel ved oversettelse . <i>Gulbrand Alhaug</i> .....	4
Musikus — en datamaskinbasert musikkanalyse. <i>Tor Sverre Lande</i> .....	11
EDB-prosjekt ved Statsarkivet i Bergen. <i>Egil Øvrebø</i> .....	14
Current work on the LOB Corpus. <i>Stig Johansson</i> .....	19
Forskningsarbeid og den teknologiske utvikling.	
Utdrag fra rapporten Datateknologi og humanistisk forskning .....	22
Mikroprosessoren, arbeidet og mennesket. <i>Anfinn Stigen</i> .....	32

## Rapporter

EDB-kurs for tilsette i arkivverket .....	43
First International Workshop on Natural Communication with Computers, Warszawa 9. — 12. sept. 1980 .....	44
COLING 80. Tokyo, 30. sept. — 4. okt. 1980 .....	47
Statistikk og EDB i arkeologisk forskning. ....	52
ALLC, København 11. — 12. des. 1980. ....	54

## Meldinger

NAVF og Datatilsynet .....	57
Internasjonale konferanser .....	59
Automatisk syntaktisk analyse .....	60
Datateknologi og humanistisk forskning. ....	62
EDB-prosjekter i de humanistiske fag .....	62
Senterets rapportserie. ....	63
Konsulenttjeneste ved Universitetet i Tromsø .....	65
Prosjektsamarbeid som betalte oppdrag .....	66
Hovedpunkter i Senterets arbeid i 1981 .....	66
Forskerstipend .....	69
Programmeringskurs 24. aug. — 4. sept. 1981 .....	71
Summary .....	72
Medarbeidere .....	79

# EDB som hjelpemiddel ved oversettelse

*Gulbrand Alhaug*

Alt fra de første datamaskinene ble produsert, har en tumlet med planer om å bruke EDB ved oversettelse fra et språk til et annet. I USA ble det f.eks. satset store beløp på forskjellige oversettelsesprosjekter. Særlig var en interessert i automatisk oversettelse av russisk faglitteratur til engelsk. Kvaliteten på oversettelsene kunne ikke betegnes som tilfredstillende, men for fagfolk var det ofte viktigere å få oversettelsen raskt enn å vente i lengre tid på en mest mulig korrekt oversettelse. Sammen med den ufullkomne oversettelsen fulgte det gjerne en «bruksanvisning» som gjorde greie for forskjellige typer språkfeil, f.eks. ordstillsingsfeil. Det vesentlige for fagfolk var altså at teksten i det store og hele var forståelig - ikke at språket var 100% korrekt.

I Norge har det i liten grad vært arbeidet med automatisk oversettelse, men det kan nevnes at *Mike Gillow* alt i 60-åra arbeidet med oversettelse fra tysk til engelsk. En rapport om dette arbeidet er utgitt av EDB-avdelingen, Universitetet i Bergen (1965).

På det internasjonale markedet finnes det nå firmaer som på rent forretningsmessig basis tar på seg oversettelsesoppdrag. Dette gjelder f.eks. det amerikanske firmaet Weidner (Utah) som bl.a. kan tilby maskinell oversettelse fra engelsk til spansk, fransk og arabisk.

## EDB-støttet oversettelse

De vanligste betegnelse på oversettelse gjort med datamaskin har vært «maskinell oversettelse» eller «automatisk oversettelse». Som følge av den datatekniske utviklinga de siste åra har en imidlertid begynt å bruke betegnelsen «EDB-støttet oversettelse». I motsetning til tidligere kan nemlig mennesket gripe mer aktivt inn under oversettelsen, f.eks. fra terminal. I betegnelsen «EDB-støttet oversettelse» kommer altså samarbeidet menneske maskin klarere fram.

Et eksempel på dette samarbeidet er behandlingen av flertydige ord (homografer). Hvis vi f.eks. hadde et dataprogram som oversatte fra norsk til engelsk, ville et ord som *fyr* by på problemer (1. på havet, 2. på peisen, 3. på gata). Med en homograf-rutine innebygd i programmet, som bl.a. gjelder *fyr*, vil programmet stoppe opp ved dette ordet og be den som overvåker oversettelsen fra terminal, å velge det riktige oversettelsesalternativet på grunnlag av konteksten.

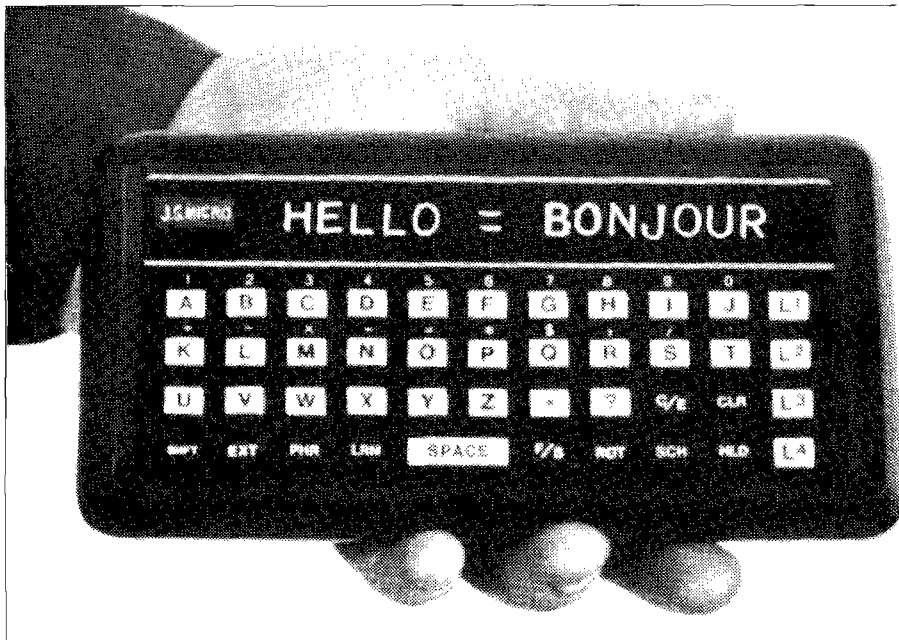
Flertydighet (homografi) har vist seg å være et av de største problemene ved

EDB-oversettelse. Men i den grad jeg har syslet med slik oversettelse - fra bokmål til nynorsk - har flertydighet vært et mindre vesentlig problem. Her er det nemlig ofte slik at et flertydig ord i den ene målforma har de samme betydningsalternativene i den andre målforma, f.eks. *fyr*. Dette kan også gjelde ord som har forskjellig form i de to målformene, f.eks. *bare* (1.adverb, 2.flertall av adj. *bar*) og *berre* (1. adverb, 2. flertall av adj. *berr*). I så nærstående «språk» som bokmål og nynorsk vil altså homografiproblemet være mindre enn i språk som ligger lenger fra hverandre.

Det som særlig kan gjøre EDB-støttet oversettelse aktuelt i Norge, er at lærebøkene i skolen skal foreligge på både bokmål og nynorsk. Som oftest har originalversjonen vært på bokmål, og det har dermed vært nødvendig å oversette til nynorsk. I 1980 bevilget Stortinget 18 mill.kroner til nynorske parallellutgaver.

En viktig *teknisk* forutsetning for at EDB-oversettelse er mer aktuell i dag enn tidligere, er at de fleste bøker som trykkes i dag under sjølve trykkeprosessen foreligger i et maskinlesbart format. En slipper overføringsutgifter, f.eks. til punching eller optisk lesing.

EDB-støttet oversettelse vil ventelig bli vesentlig rimeligere enn tradisjonell oversettelse. Dessuten vil parallellutgavene foreligge raskere enn før.



Som prøvemateriale har jeg brukt utdrag fra læreboka «Aktiv biologi, for grunnskolen - ungdomstrinnet» av *Haakon Olsen* o.fl. (Oslo 1975). Både bokmålversjonen og nynorskversjonen av denne læreboka er maskinlesbare, noe som har vært nyttig ved forskjellige sammenligninger av de to versjonene f.eks. av ordfrekvenser.

Grovt sett kan en skille mellom disse fasene ved EDB-støttet oversettelse fra bokmål til nynorsk:

### **A.Oversettelse på ordplanet.**

1.Oversettelse av frekvente ord, f.eks.*ikke* til *ikkje*

2.oversettelse av endingsapparatet

a)Bøyningsendinger f.eks.-*ende* til -*ande* (*nærgående* til *nærgåande*)

b)Suffikser f.eks. *lig* til *leg* (*vanlig* til *vanleg*)

c)Kombinasjonen av suffiks og bøyningsending f.eks. -*erne* til -*arane* (*brukerne* til *brukarane*)

3.Oversettelse av frekvente prefikser/røtter i ord *sammen* til *saman* (*sammenheng* til *samanheng*)

*navn* til *namn* (*mannsnavn* til *mannsnamn*)

### **B. Omforming på det syntaktiske plan.**

Eks. vårt land → landet vårt

### **C.Kontroll og forbedring av oversettelsen.**

Det meste av den videre beskrivelsen gjelder fase A. Fase B er mer på forsøksstadiet - særlig fordi det her trengs mer kompliserte dataprogrammer enn i fase A. I fase C er det mennesket som skal kontrollere og forbedre det oversettelsesresultatet som er oppnådd etter de to foregående fasene.

Ved såkalt setningsbasert oversettelse ville det ikke være nødvendig å skille mellom fase A og B (og heller ikke mellom underpunktene 1, 2 og 3 i fase A). En setningsbasert oversettelse vil si at dataprogrammet utfører oversettelsesinstruksjonene i en setning før det går over til neste.

Ved såkalt ordbasert oversettelse, som det er lagt opp til her, vil derimot dataprogrammet oversette alle forekomster av et ord i teksten i samme operasjon, f.eks. *ikke* til *ikkje*. Deretter går programmet over til neste ord i registreret

over ord som skal oversettes, f.eks. *bare* til *berre*. Det er først og fremst data-økonomiske grunner til at en ordbasert oversettelse nok er å foretrekke framfor en setningsbasert. Dette vil gå klarere fram av den følgende beskrivelsen.

### Ordbasert oversettelse

Før den ordbaserte oversettelsen begynner, blir hele teksten organisert som en indeks (alfabetisk ordliste med indeksnumre som viser til linjenummer og posisjon i linja for hvert enkelt ord). Hvis f.eks. *ikke* er brukt 200 ganger i en tekst,rekker det altså med å oversette *ikke* til *ikkje* bare 'en gang i indeksen. I fortløpende tekst derimot måtte en ha oversatt *ikke* 200 ganger. Etter at alle oversettelsene i indeksen er avsluttet, blir så teksten reorganisert på grunnlag av indeksnumrene. Men denne gangen kommer altså *ikkje*, *frå* osv. i de posisjonene *ikke*, *fra* osv. sto før.

Det er særlig med tanke på oversettelse av endringsapparatet at en slik organisering av teksten har dataøkonomiske fordeler.Mye av forskjellen mellom bokmål og nynorsk ligger nemlig i endingsapparatet, og ved baklengssortering av indeksen vil en få oversatt i store jafser ord med samme ending, f. eks.:

lig → leg (vanlig, kvinnelig, farlig osv.)

ende → ande (nærgående,tiltalende,påfallende osv.)

erne → arane (brukerne,borgerne,avvikerne osv.)

En stor del av de mest vanlige ord. er identiske i bokmål og nynorsk, f.eks. *er*, *kan*, *som*, *vi*, *på*, *at*. Disse orda skal jo ikke oversettes, og dataprogrammet gir beskjed om at slike ord skal stå uforandret. Når en tar for seg de 260 mest frekvente orda i begge lærebøkene, viser det seg at omtrent 150 av disse er identiske i begge målformer. I løpende tekst utgjør disse 150 orda så mye som 46% i de to lærebøkene.

Også av de orda som skal oversettes, utgjør noen få høgfrekvente ord en vesentlig del av løpende tekst. Således utgjør ca. 90 slike ord omtrent 13% av løpende tekst i de to lærebøkene (ikke-ikkje, fra-frå, være-vere, bare-berre osv.)

### Norsk ordregistrant

Sjøl om en kan oppnå en forholdsvis bra oversettelseseffekt ved bare å bruke frekvente ord (fra - frå osv.), kan en få en mye større effekt ved å koble til den maskinlesbare ordboka Norsk ordregistrant. Denne ordboka er bygd opp ved Prosjekt for datamaskinell språkbehandling, Universitetet i Bergen, og inneholder nå omtrent 50 000 oppslagsord på begge målformer. I denne ordboka er korresponderende oppslagsord i de to målformene knyttet sammen,

f.eks. *heder-heider*, *sang-song*. Med denne ordboka knyttet til oversettelsessystemet vil f.eks. *heder* i en tekst bli oversatt til *heider*.

På grunnlag av et kodesystem kan Norsk ordregistrant lett utvides til å omfatte bøyningsformer av oppslagsorda, f.eks. *båter-båtar*, *blader-blad*. Dermed vil f.eks. *båter* bli oversatt til *båtar*, *blader* til *blad*. Men denne utvidelsen trenger en bare gjøre for bokmålsendinger som ikke gir et entydig oversettelsesalternativ, f.eks. -er. Endinger som gir en entydig oversettelse, kan heller oversettes ved generelle regler, f.eks. -ende til -ande (jfr. pkt. 2a i fase-oversikten). I et fullt utbygd oversettelsessystem vil det i tillegg til Norsk ordregistrant være aktuelt å benytte seg av ordregistrene i Norsk termbank. Denne ordbanken for fagspråk er nå under oppbygging ved Universitetet i Bergen (Prosjekt for datamaskinell språkbehandling). I tillegg til bokmål og nynorsk vil denne ordbanken bl.a. inneholde fagord på svensk, dansk og engelsk. Ved oversettelse av f.eks. ei biologilærebok kan en da koble til biologiregisteret for bokmål og nynorsk og få oversatt *sivspurv* til *sivsporv*, *starrmyr* til *storrmyr* osv.

Sjøl om en bruker både Norsk ordregistrant og Norsk termbank, vil det i en tekst finnes ord som ikke står i disse ordregistrene. Dette gjelder særlig sammensetninger, f.eks. *rekordtall* og *tall-kolonne*. Slike sammensetninger kan en få oversatt ved å koble til et maskinlesbart morfemregister bygd opp av de mest vanlige morfemene (orddelene) i bokmål og nynorsk, f.eks. *tall-tal*, *navn-namn*. Dermed vil sammensetningene ovafor bli oversatt til *rekordtal* og *talkolonne*. I enkelte tilfelle må en riktignok stille opp unntaksregler slik at en unngår feiloversettelser som *metallbit* til *metalbit*. Sekvensen *tall* skal altså stå uforandret når den kommer etter sekvensen *me*.

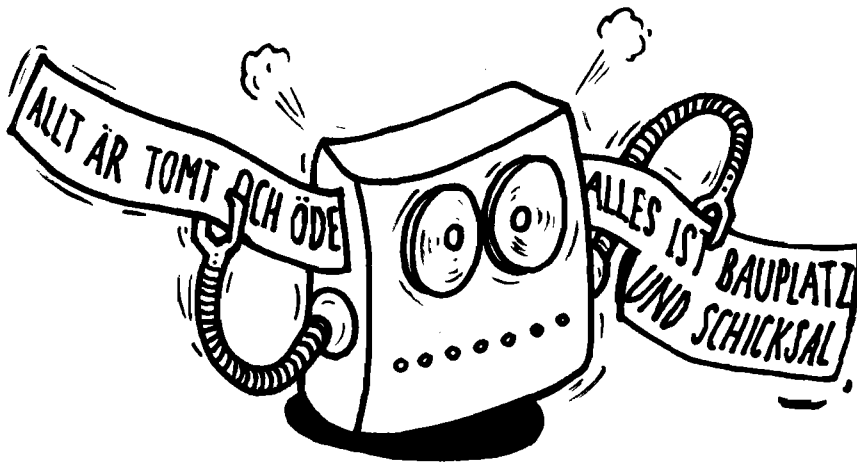
### Syntaktiske forandringer

Ved ord-oversettelsene kan en i det store og hele benytte seg av standardprogrammer. Men om en vil gjøre syntaktiske forandringer (fase B) må en normalt regne med å skrive spesialprogrammer. Et eksempel på dette kan være systematisk flytting av foranstilt eiendomspronomen:

*vårt land — landet vårt*

Eiendomspronomenet flyttes altså bak substantivet, og substantivet får bestemt form. I det oversettelsessystemet som det amerikanske firmaet Weidner tilbyr, er det gitt eksempler på enkelte syntaktiske omforminger, men det er opplagt begrensninger for hvor langt en kan komme med datamaskinell oversettelse på dette punktet. Som eksempel på dette kan vi nevne genitivkonstruksjoner i bokmål. Disse kan i noen tilfelle gjengis med *til* i nynorsk (Per's sykkel — sykkelen til Per). Men i andre tilfelle er det andre preposisjoner eller helt andre uttrykksmåter som skal brukes. Det finnes altså ingen automatisk måte å gjengi genitivkonstruksjoner på. I det hele tatt er det viktig at mennesket kommer inn i sluttfasen av oversettelsen (fase C). Den mer





«Alt kan ikke overlates til maskinene.»

---

trivielle delen av oversettelsen er da unnagjort av dataprogrammet (*ikke til ikkje, vanlig til vanleg*), og det gjenstår å få en fullgod nynorsk oversettelse.

De som tidligere har arbeidd med oversettelse fra bokmål til nynorsk, bør fortsette med det. Men istedenfor å begynne oversettelsesarbeidet helt fra grunnen av kan de nå få tilsendt datautskrifter der den mest trivielle delen av oversettelsen alt er gjort, f.eks. oversettelse av formverket. Det som gjenstår, er den mest utfordrende delen av oversettelsen - å få en fullgod nynorsk stiltone i læreboka.

Oversetteren kan i datautskriftene få bokmålsteksten plassert på linja over den EDB-oversatte teksten. Dermed kan han/hun lett vurdere den foreløpige oversettelsen i forhold til bokmålsteksten. Oversetteren fører inn sine rettinger i datautskriftene som så sendes tilbake til forlaget/trykkeriet. Her blir rettingene innført ved hjelp av EDB-baserte redigeringsystemer.

Med stadig synkende priser på EDB-utstyr bør oversetterne på lengre sikt kunne sitte hjemme ved en terminal og føre inn rettingene. Når dette er gjort, er læreboka stort sett klar for trykking. Det som nå normalt vil gjenstå, er å få inn spesiell grafisk informasjon på de rette stedene i læreboka, f.eks. kursiv, halvfete typer, varierende spaltebredde osv. Men det er nå utviklet spesielle systemer som overfører den grafiske informasjonen direkte fra

originalversjonen til oversettelsesversjonen. Dette gjelder f.eks. systemet Compuset som er utviklet i USA.

## Begrensninger

En kan naturligvis reise en rekke innvendinger mot EDB-støttet oversettelse av lærebøker. Mange vil rent følelsesmessig reagere mot at datamaskiner skal tukle med språket, men en må da huske på at dataprogrammet bare gjør grovarbeidet, dvs. den mer trivielle delen av oversettelsen. Det er oversetteren som med sin stilsans skal forbedre den EDB-baserte oversettelsen slik at språket kan betegnes som god nynorsk. Noen tester som jeg har gjort på det lærebokmaterialet jeg har arbeidd med, viser at litt over 90% av ordformene i løpende tekst er riktige etter den EDB-baserte oversettelsen (i denne testen var Norsk ordregistrant ikke tilkoblet). I tillegg til å rette gjenstående feil i ordformene må altså oversetteren sørge for å få en god nynorsk syntaks. Hvor mye arbeid oversetteren må legge ned i dette, vil i høy grad avhenge av syntaksen i bokmålsversjonen. Hvis denne følger et folkelig mønster, blir det lite å gjøre. Derimot kan det bli nokså arbeidskrevende å få god nynorsk ut av en kronglete og danskpreget syntaks i bokmålsversjonen.

Enkelte forlagsfolk har gitt uttrykk for at lærebokforfattere burde få opplæring i å skrive bedre. Hvis språket i bokmålsversjonen ikke er godt nok, vil dette naturlig nok også kunne nedfelle seg i oversettelsen.

Bruk av EDB kan gi uheldige utslag når det gjelder de valgfrie hovedformene, f.eks. *gjev* - *gir*, *vart* - *blei*, *samstundes* - *samtidig*. Hvis f.eks. dataprogrammet inneholder oversettelsesalternativene *gjev*, *vart* og *samstundes*, vil altså alle lærebøkene få disse orda. EDB innebærer således en fare for uniformering av språket i lærebøkene. Dette har også språkpolitiske implikasjoner i den forstand at en i dataprogrammet bevisst kan legge opp til en tilnærmingsnynorsk (*gir*, *blei* osv.) eller en mer tradisjonell nynorsk (*gjev*, *vart* osv.). En måte å unngå slik standardisering på kan være å la den enkelte oversetter sjøl velge ut de valgfrie alternativene han/hun finner mest naturlige. Oversetteren kan få lagt inn sitt personlige sett av valgfrie hovedformer i dataprogrammet hver gang han/hun skal oversette ei lærebok.

Ellers vil noen kanskje hevde at EDB innebærer en fare for tilnærmingsnynorsk i den forstand at det er enklere å bruke ei form som er identisk med bokmål (*gir*), enn ei avvikende form (*gjev*). Fra et datamaskinelt synspunkt koster det imidlertid ikke mer å forandre *gir* til *gjev* enn å la *gir* stå uforandret. En mer detaljert beskrivelse av EDB-støttet oversettelse finnes i rapporten EDB-støttet oversettelse av lærebøker fra bokmål til nynorsk (utgitt i serien Norske språkdata).

# Musikus - en datamaskinbasert musikkanalyse

*Tor Sverre Lande*

Kan man i musikkanalyse, som er en disiplin med vesentlige estetiske elementer, bruke datamaskin? MUSIKUS - et samarbeidsprosjekt ved Universitetet i Oslo, støttet av NAVF - har vist at spørsmålet kan besvares bekreftende.

Formålet med prosjektet er å utvikle metoder og programmer for analyse av musikk med EDB. Som en del av prosjektet er det utviklet programmer for tegning av noter ved hjelp av en datamaskin, og det er bygget et spesielt orgel, hvor en mikroprosessor kan registrere det som spilles på orgelet.

Forfatteren har lagt vekt på å presentere de mest forståelige og lettest tilgjengelige deler av prosjektet, på bekostning av det vitenskapelig mest interessante.

Datamaskinbasert musikkanalyse forutsetter at

*musikken er lest inn og lagret i datamaskinen*

*og at*

*analysemetoden er utviklet og programmert.*

Med disse forutsetninger kan vi ved hjelp av datamaskinen utføre forskjellige analyser. For at disse data skal bli meningsfylte, må de tolkes av musikkforskeren. Dette er kanskje den mest interessante og vanskeligste delen av datamaskinbasert musikkanalyse: Hvordan formulere en metode som finner og måler musikalsk sett interessante egenskaper? Musikkanalyse er en disiplin med vesentlige estetiske elementer. Det er viktig å bevare disse også i en datamaskinbasert analyse. Nettopp derfor har vi lagt stor vekt på å unngå fremmedgjøring, både når det gjelder formulering av metode og presentasjon av resultater.

Til formulering av analysemetoden har vi utviklet et begrepsapparat som er rimelig enkelt å bruke for musikkforskere. Disse begrepene brukes også i analyseprogrammene. Programmering av analysemetoden skjer ved et samarbeid mellom musikkforskeren og en programmerer. Videre har vi lagt vekt på å bruke vanlige musikologiske termer i presentasjonen av resultatene. Denne fremgangsmåten har vært brukt til programmering av flere analyser: Fra harmonisk analyse av tidlig barokkmusikk og koraler, via melodisk analyse av norske hardingfeleslåtter og barnesanger til samklangsanalyse av Fartein Valens musikk. Ved Institutt for musikkvitenskap er det til nå levert to hovedoppgaver som begge bruker datamaskinell analyse av musikk.

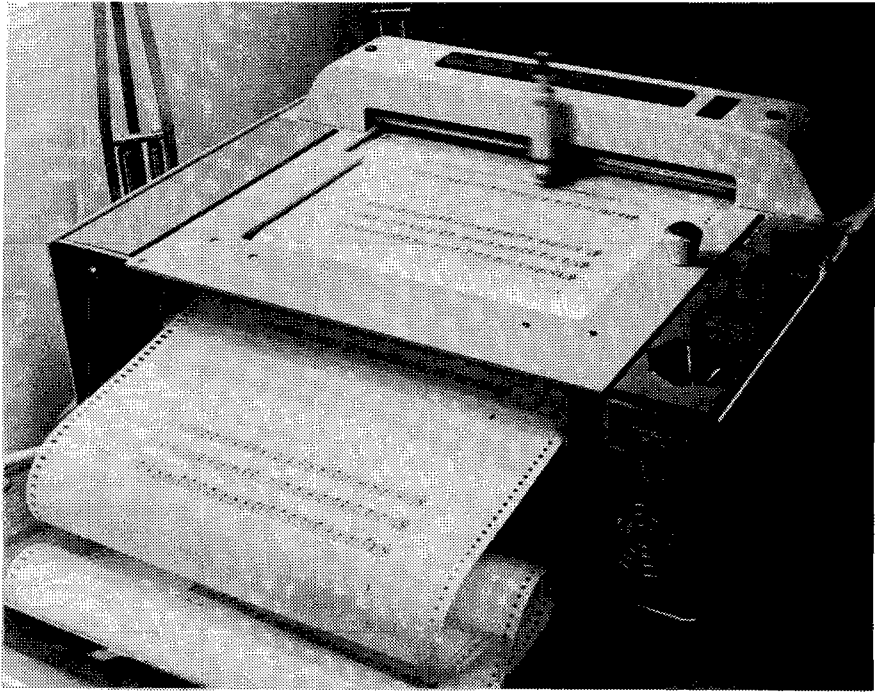
En av forutsetningene for at EDB-basert musikkanalyse skal lønne seg, er at musikken som er lagret i datamaskinen inneholder all informasjon i notebildet, slik at de samme data kan brukes til forskjellige typer analyse. Den tradisjonelle måten å tilføre musikk til datamaskinen, er å omskrive notene til en rekke bokstaver og andre tegn, kalt kode. Denne koden bør ta vare på all informasjon i notebildet. Koden som brukes i MUSIKUS-prosjektet heter MUSIKODE.

### **Orgel og mikroprosessor**

Innkoding er en forholdsvis langsom og kostbar prosess og legger en vesentlig begrensning på størrelsen av analysematerialet. Vi har derfor prøvd å redusere denne kostnaden ved hjelp av et spesialbygget orgel koplet til en mikroprosessor. Ideen er at musikken som spilles på orgelet, registreres av datamaskinen. Orgelet er bygget som en hovedoppgave i fysikk av Harald Støp *Bowitz*. Programmer for inn- og avspilling av musikk på orgelet lages av Jens Harald *Kvam*, som en del av hans hovedoppgave i Informatikk. I skrivende stund er programmene for avlytting av innkodet musikk ferdige, mens innspillingsprogrammene er under utvikling.

Innspilling fra orgelet er det vanskeligste. I de fleste analysesammenhenger ønskes informasjon om stemmeinndelingen i musikken. Foreløpig forsøker vi å spille inn en stemme av gangen. Datamaskinen spiller så av den/de stemmene som er innspilt, samtidig som neste stemme spilles inn osv. Den største vanskeligheten er å bestemme varigheten av tonene. For å få inn variasjoner i rytmen, har vi tatt i bruk en «trampe-takten» pedal. Vi har tro på at innspillingen fra orgelet vil redusere innkodningskostnaden betydelig.

Musikk noteres vanligvis som noter. Det er derfor naturlig at innkodet og innspilt musikk kan presenteres som vanlige notetegn. For å gjøre dette, bruker vi en tegnemaskin (plotter) og en grafisk skjerm. Geir M. *Ekre* har som en del av sin hovedoppgave i informatikk, utviklet programmer for mikromaskin som tegner noter på plotter og skjerm. Det er lagt vekt på at kvaliteten skal være høy. Et produkt av MUSIKUS-prosjektet er altså et ganske avansert notetegnesystem. For å prøve om kvaliteten av notene tegnet med datamaskin er god nok, tegnet vi notene til en barnesangbok utgitt på J.W.Cappelens forlag: «Barnas egen sangbok», redigert av Jon Roar *Bjørkvold*. Konklusjonen er at kvaliteten er meget god. Det er derfor rimelig å anta at et slikt utstyr kan redusere tegnekostnaden til noter betydelig. Det skal presiseres at programmene for notetegning foreløpig er utviklet for en-stemt musikk, og et betydelig programmeringsarbeid gjenstår. Vi har derfor satt igang en ny hovedoppgave for utvidelse og videreføring av tegneprogrammene. Det skal videre nevnes at vi i disse dager prøver å få igang en norsk industri for trykking av noter basert på resultater fra MUSIKUS-prosjektet.



*«Musikken som spilles inn på et spesialbygd orgel, registreres i en mikroprosessor. Derfra kan den hentes ut igjen i form av noter, enten på papir via en tegnemaskin (plotter) som vist på bildet, eller via skjerm.»*

---

Om andre oppgaver som er aktuelle, kan nevnes et program for overføring av noter til blindeskrift (Braille). Dette ville bety en betydelig lettelse for blinde musikkstudenter og musikere, da manuell overføring er et kostbart og tidkrevende arbeid.

Alt dette bygger på notert musikk. Vårt mål er å kunne analysere egenskaper ved klingende musikk (Fourier-analyse). Arbeidet med slike analyser, særlig i forbindelse med folkemusikk, er startet - men er foreløpig helt i begynnerfasen. En hovedoppgave i informatikk er innlevert på området.

Det er å håpe at prosjektet vil gi resultater som kan komme norsk musikkliv til gode i fremtiden.

# EDB-prosjekt i Statsarkivet i Bergen

Egil Øvrebo

Siden januar 1974 har Statsarkivet i Bergen samarbeidd med Historisk institutt ved Universitetet i Bergen om EDB-registrering av nokre kjelder i Statsarkivet. Vi har fått dei første resultatata og vil gjerne dele røynslene så langt med andre interesserte. Vår EDB-kompetanse er beskjeden nok: 3 av arkivarane har følgd elementærkurs på NAVF's EDB-senter i Bergen, den eine har vore medlem av eit EDB-utval for Arkivverket. Elles er røynslene av praktisk og prosaisk art. Eg trur dei kan vere nyttige når vi no ser ein gryande debatt om bruken av EDB i Arkivverket.

Det har vore arbeidd med følgjande prosjekt:

Emigrantprotokollane for Bergen 1874-1924 . . . . .	99.900 personar
Folketeljinga for Bergen 1875 . . . . .	34.400 personar
Skifteregistret for Bergen 1821-1899 . . . . .	11.400 personar
Skifteregistret for Sunnfjord 1727-1845 . . . . .	10.800 personar
Skifteregistret for Nordhordland 1805-1862 . . . . .	7.900 personar
Kyrkjebøkene for Bergen 1816-, nyleg påbegynt.	

*Emigrantprotokollane 1874-1924* i Bergen politimeisterarkiv var utan register. Ettersom vestlendingane ikkje emigrerte i alfabetisk rekkjefølgje, måtte vi *bla* oss fram til personane, anten det galt arvesaker, etterlysingssaker eller dei mange spørsmål frå etterkommarar i Amerika. Dessutan samla det seg sterke forskningsinteresser om materialet, og kvar registrerte for seg. Protokollane hadde store slitasjeskader.

Prosjektopplegget vart forma i samarbeid med professor Knut Mykland og amanuensis Jan Oldervoll. Mykland var også aktiv i finansieringsprosessen, medan Oldervoll heile tida har stått for det EDB-faglege. Protokollane, som hadde frå 8 til 15 rubrikkar, skulle totalregistrerast, det vart rekna 1 1/2 årsverk til skrivning, 1/2 årsverk til korrektur, det ga eit lønsbudsjett på kr. 70.000 som vart skaffa utanfrå (Norsk Kulturråd, Nordmanns-Forbundet og Landslaget for bygde- og byhistorie). Driftsutgiftene på kr. 30.000 kom på Statsarkivet sitt budsjett (I tillegg til dette har vi seinare disponert mindre summar til optisk lesing og framstilling av mikrokort.)

Skrivearbeidet kom i gang i januar 1975, og kalkylen heldt. Derimot vart korrekturarbeidet langt meir krevjande enn venta, ein viktig grunn var dei mange rubrikkane. Statsarkivet påtok seg førstekorrekturen, arbeidet vart fordelt på 6 tomannslag og tok oss nær 2.000 timar over 3 år, 1976-1978. Opprettinga vart utført av instituttet og er av Oldervoll rekna til rundt 1.000 timar, medan han set sitt eige arbeid med prosjektet til rundt 500 timar.

Vi har no fått ut følgjande, alt på mikrokort:

Fullstendig utskrift, 1-2 kort pr. årgang, i alt 58 kort. Alfabetisk førenamnsregister på 10 kort. Alfabetisk etternamnsregister på 10 kort. Materialet ordna etter fødestad og opphaldsstad: 1-2 kort for kvart prestegjeld i Hordaland og Sogn og Fjordane, vidare kort for kvart fylke eller landsdel, og for utlandet, i alt 83 kort. Desse inneheld fullstendig utskrift for vedkommande distrikt, kronologisk ordna, og alfabetiske etternamns- og førenamnsregister.

Endeleg har vi magnetbandversjonen som skal gi papirutskrifter til historielag, bygdeboknemnder og andre interesserte, og som skal vere tilgjengeleg for vidare forskning i materialet.

Statsarkivet har hatt korta i bruk siste halvåret og kan no sperre originalprotokollane. Indeks og koordinatsystem på korta gjer bruken lett, det heile fungerer, og vi sparar tid dagleg. Enno er der feil i materialet: ein prøve peikar mot ein svært låg feilprosent og viser dessutan at 1/3 av dei skriv seg frå originalprotokollane. Prøven galdt rett nok dei første rubrikkane (registerkvaliteten) der vi la mest arbeid under korrekturen. Elles meiner vi å ha fått den lærdommen at prosjektet var for stort, - det gjekk for lang tid før vi såg resultat av investeringane.

Mikrokort vil no bli sende til arkiva og sentralbiblioteka i landet, papirutskrifter til historielag og bygdeboknemnder. Det står att å fastleggje ein ekspedisjonsrutine og å gjere tilbodet kjent, sidan skal alle interesserte kunne dra nytte av arbeidet.

*Folketeljinga for Bergen 1875* (34.400 personar, ordna etter rode- og husnummer, 14 rubrikkar): Fordi byen var stor, og fordi vi manglar adressekalendarar, har desse eldre teljingane nærmast vore utilgjengelege. Då Statsarkivet påtok seg førstekorrekturen på emigrantprotokollane, nytta vi derfor den ledige kapasiteten på prosjektet til å skrive denne teljinga. Resten av arbeidet fell på Historisk institutt, og korrekturen har nådd 20.ode (av 24).

*Skifteregistret for Bergen 1821-1899* (11.400 personar, etternamn, førenamn, dato, rode- og husnummer) er eit originalregister ført ved skifteretten og delvis alfabetisert. Høg bruksfrekvens og slitasjeskader gjorde det ønskeleg å

fornye registret. Det vart sett prøve på registerkvaliteten, eitt blad som mangla vart rekonstruert etter registreringsprotokollen, sidan vart heile registret skrive i Statsarkivet med OCR-typar etter mønster oppsett av Oldervoll. Manuskriptet på 547 sider vart lese optisk, og vi fekk eit korrekturksemlar, dessutan ei alfabetisert utgåve som vart merkt «ukorrigert» og straks teken i bruk på lesesalen. Førearbeid og korrektur tok 83 timar, opprettinga på terminal nærmar seg no avslutting, og vi får ut alfabetiske førenamns- og etternamnsregister på mikrokort eller papir etter ønskje.

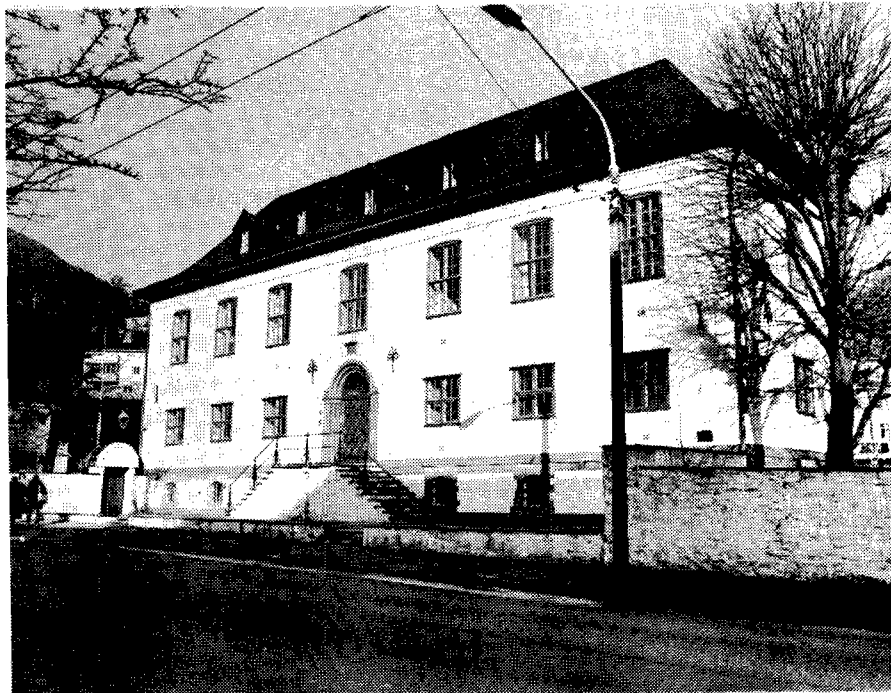
*Skifteregister for Sunnfjord 1727-1845* (10.800 personar, namn, skipreide, protokoll, folio): Vi hadde frå før eit handskrive kronologisk register som dekte kring halvparten av materialet. Til resten hadde vi originale register av ulike slag i protokollane. Alt dette heldt på å bli bladd i filler. Registreringsopplegget vart det same her som for Bergens-registret, men førearbeid og korrektur tok meir tid: på kopiar av registra sette vi først til skipreidenamnet der det mangla, og vi las korrekturen mot skifteprotokollane, ikkje mot det gamle registret som viste seg å vere mangelfullt. Korrekturen er ferdig, og etter opprettinga tek vi ut alfabetiske førenamns og etternamnsregister, dessutan vil bygdeboknemnder og andre interesserte kunne få ut registra distriktsvis. Det må nemnast at vi her hadde eit mindre tilskot frå Landslaget for bygde- og byhistorie til skrivearbeidet.

*Skifteregister for Nordhordland 1805-1862* (7.900 personar): Også her måtte eit eldre handskrive register fornyast på grunn av slitasje. Opplegget var det same som ovanfor, og det nye registret er under korrektur.

*Kyrkjebøker for Bergen 1816-*. Utgangspunktet er det same her som for folketeljningane: i så stor by som Bergen tek det urimeleg lang tid å skulle bla seg fram til opplysningane. I kyrkjebøkene vert det likevel gjort, og vi får slitasjeskadene. - Frå 1816 har bøkene nokolunde faste rubrikkar, og målet vert å registrere døypte, vigde og døde fram til 1912 då Bergen fekk sitt folkeregister. Historisk institutt har nyleg komme i gang med registrering av døypte fram til 1850 i den første av 4 kyrkjelokner. Statsarkivet bør ta del så snart ressursane tillet det, i mellomtida leverer vi dei kopiane som trengst. Kopiane går seinare attende til Statsarkivet og inn i rekkja av brukskopiar på lesesalen.

Statsarkivet og Historisk institutt har hatt ulike mål med desse registreringsarbeida. For instituttet gjeld det å opne eit stadig større materiale for differensiert forskning, til hovudfagsoppgåver, til samkøyring med 1801-teljninga. Sett derfrå vil skifteregistra eller den topografiske ordninga av emigrantane vere mindre interessante. - Statsarkivet skal sikre utsette originalkjelder, spreie EDB-versjonar og dermed minske pågangen hos oss. Samtidig skal vi skape profesjonell arbeidsreidskap, raskare og sikrere rutiner som set oss i stand til å møte publikumspågangen. - Langt på veg fell interessene saman, vansken ligg meir i at arbeidet begge stader må utførast i konkurranse med andre viktige gjeremål, slik at registreringa tek for lang tid. Her har vi fått den lærdommen,





*Statsarkivet i Bergen.*

---

at så lenge ressursane er knapt tilmålte, er vi i Statsarkivet heilt avhengige av samarbeid med ein EDB-ekspertise utanfor huset.

NAVF's utval for EDB og primærkjelder har i innstillinga frå 1977 merkt seg at mange arkivinstusjonar ikkje har god nok oversikt over kjeldene, og at vi manglar kapasitet til å tilretteleggje tenlege register og katalogar. Utvalet går sterkt inn for at NAVF yter hjelp her, og som eit første resultat av dette får vi no, ved førstearkivar Nysæter, utarbeidd eit oversyn over kva som finst av kataloger og register i offentlege norske arkivinstusjonar. Kartleggjinga er viktig, men løyser ikkje problem for arkiva. Derfor er det gledeleg at utvalet går vidare og rår til at NAVF - i prosjektsamanhang - yter hjelp til ordning, katalogisering og registrering i arkiva. Med eller utan slik hjelp vert dette ei av dei viktigaste oppgåvene for Statsarkivet i Bergen dei kommande åra, - som det har vore det dei siste åra.

Framleis må bruksfrekvens og brevjournal vere retningsgivande ved prioriteringa, derfor vert opprusting av åstadsakregistret ei første - rett nok mindre forskarvenleg - oppgåve, sidan har vi fritt val mellom stiftamtmannsarkiv, futearkiv, bispearkiv, prestearkiv, o.s.v. Her er det ikkje nok å publisere eksisterande katalogar og avleveringslister, vi vert tvinga til å begynne i rette enden og først skape fagleg forsvarlege tilstandar på hyllene og i katalogane. Vi skal ikkje snakke om perfeksjonisme. Men vi skal skilje mellom stiftsskrivar og stiftamtstueskrivar, mellom verjemålssaker og verjerådssaker, mellom stolteste i Korskirken og jordfeste i Kinn. Eg meiner å ha den røynsla at både forskarar og advokatar helst vil ha svar på det dei spør om, - dei har beintfram krav på det, og så får vi innrette prioritering og arbeid i arkiva deretter.

I dette opprustningsarbeidet vil det vere rimeleg å følgje katalogmønster som tek sikte på EDB-presentasjon, slik Nysæter nemner i eit innlegg i Arkivposten (oktober 1980). Om det i mellomtida vert høve til fleire EDB-prosjekt av same slag som ovanfor, det vert eit spørsmål om pengar og arbeidslyst.

# Current work on the LOB Corpus

Stig Johansson

The Lancaster-Oslo/Bergen (LOB) Corpus of British English was completed in 1978 as the result of cooperation between the University of Lancaster, the University of Oslo, and the Norwegian Computing Centre for the Humanities (NCCCH). The Corpus, which is distributed at cost from the NCCCH, is now being used by researchers in many countries. A second stage of Lancaster Oslo-Bergen cooperation has now started. Aided by grants from the Social Science Research Council in England and the Norwegian Research Council for Science and the Humanities, work is in progress on the grammatical analysis of the LOB Corpus. This will lead in two years time, we hope, to a grammatically annotated version of the Corpus, which can be used to study grammatical constructions in contemporary British English. A similar project has been completed for the American counterpart to the LOB Corpus, the Brown Corpus. The two projects will make possible a detailed comparison of a wide range of aspects of British and American English as well as provide valuable material for more general linguistic studies.

The LOB Corpus project is now gaining attention outside narrowly academic circles, presumably because of the great interest among the general public in the relationship between British and American English. In connection with the award of the generous grant from the Social Science Research Council, Geoffrey *Leech* (University of Lancaster) was interviewed by the press. Reports have appeared in a large number of newspapers (125 newspapers in the U.S. alone, we are told), such as *The Observer*, *The New York Times*, *The Straits Times* (Malaysia), etc.

As the LOB Corpus has been quite widely distributed, it is difficult to survey the research which is in progress based on the material. At the University of Oslo we are using the material for 'hovedfag' theses as well as for term papers and other more limited studies. A thesis (by Mette Cathrine *Jahr Sørheim*) on the genitive in British and American English was completed in 1980. Another thesis, which is now being completed, is Inger *Krogvig's* study of *shall*, *will*, *should*, and *would*. Other topics being studied at the moment are the use of the perfect and the subjunctive. Detailed word frequency studies based on the LOB Corpus, including a comparison with its American counterpart, have been completed and will be published shortly. A list of some current and forthcoming publications is given below. For further information, see *ICAME NEWS*.

## BIBLIOGRAPHY:

Coates, Jennifer and Geoffrey N. Leech. 1980. 'The Meanings of the Modals in Modern British and American English', *York Papers in Linguistics* 8, 23-34.

# ICAME NEWS

Newsletter of the International Computer  
Archive of Modern English (ICAME)

Published by: The Norwegian Computing Centre for the Humanities, Bergen  
The Norwegian Research Council for Science and the Humanities



Machine-readable  
texts in  
English language  
research

**No. 4**  
Sept. 1980

*«Bladet ICAME NEWS gir opplysninger om engelske tekstsamlinger klargjort for datamaskinell analyse. Tekstene kan bestilles hos NAVFs EDB-senter for humanistisk forskning. Interesserte kan abonnere på bladet ICAME NEWS gratis. Stig Johansson er redaktør for bladet.»*

Hofland, Knut and Stig Johansson. 1979. Microfiche concordance of the Lancaster-Oslo/Bergen Corpus. Bergen: Norwegian Computing Centre for the Humanities.

*ICAME NEWS*. Newsletter of the International Computer Archive of Modern English. Bergen: Norwegian Computing Centre for the Humanities.

Johansson, Stig. 1980. 'Corpus-Based Studies of British and American English'. In S. Jacobson, ed., *Papers from the Scandinavian Symposium on Syntactic Variation, Stockholm, May 18-19, 1979*. Stockholm Studies in English, 52. Stockholm: Almqvist & Wiksell. 85-100.

Johansson, Stig. 1980. 'Word Frequencies in British and American English: Some Preliminary Observations'. In J. Allwood and M. Ljung, eds., *ALVAR. A Linguistically Varied Assortment of Readings. Studies Presented to Alvar Ellegård on the Occasion of His 60th Birthday*. Stockholm Papers in English Language and Literature, 1. Department of English, University of Stockholm. 56-74.

Johansson, Stig. 1980. 'The LOB Corpus of British English Texts: Presentation and Comments', *ALLC Journal* 1, 25-36.

Johansson, Stig, Leech, Geoffrey N. and Helen Goodluck. 1978. *Manual of Information to Accompany the Lancaster-Oslo/Bergen Corpus of British English, for Use with Digital Computers*. Department of English, University of Oslo.

Johansson, Stig and Knut Hofland. (forthcoming) *Word Frequencies in British and American English*. Bergen: Norwegian Computing Centre for the Humanities.

Krogvig, Inger. The Use of *Shall, Will, Should, and Would* in British and American English. 'Hovedfag' thesis, Department of English, University of Oslo. To be completed in February 1981.

Leech, Geoffrey N. and Jennifer Coates. 1980. 'Semantic Indeterminacy and the Modals'. In S. Greenbaum, G. Leech, and J. Svartvik, eds., *Studies in English Linguistics: For Randolph Quirk*. London: Longman. 79-90. :

Sørheim, Mette-Cathrine Jahr. 1980. The s-Genitive in Present-Day English. 'Hovedfag' thesis, Department of English, University of Oslo.

Sørheim, Mette-Cathrine Jahr. 1981. 'The Genitive in a Functional Sentence Perspective'. In S. Johansson and B. Tysdahl, eds., *Papers from the First Nordic Conference for English Studies, Oslo, September 17-19, 1980*. Department of English, University of Oslo. 405-423.

# Forskningsarbeid og den teknologiske utvikling

*Utdrag av rapporten Datateknologi og humanistisk forskning.*

Datateknologi har allerede i betydelig grad endret vilkårene for å drive forskningsarbeid. Det er god grunn til å tro at datateknologien i større og større utstrekning vil bli integrert i forskningsarbeidet i tiden som kommer. Nye anvendelser av datateknologi vil oppstå, og utstyrstilbudet kommer til å øke. Det vil både foreligge mer generell og brukervennlig programutrustning og et større arsenal av avansert applikasjonsrettet programvare.

Datateknikk vil kunne få en gjennomgripende virkning for forskningsarbeidet, og dette vil gjøre den forskjellig fra det tekniske utstyr som tidligere har vært tatt i bruk for forskningsformål. Frem til i dag har det i økende omfang vært utviklet spesialutstyr til bruk i enkelte fag for å muliggjøre spesielle typer forskningsoppgaver, eksempelvis innen fysikk, kjemi, medisin o.fl. og også i begrenset grad innenfor de humanistiske fag, som i arkeologi og fonetikk. Datateknologi vil derimot bli et så generelt og dyptgripende redskap for forskning *i sin allminnelighet* at det vil skape et tidsskille også innen forskningen.

Enkelte forskere har trukket sammenligninger mellom datateknologi og det tekniske gjennombrudd som kom med elektrisiteten. Sammenligningen har åpenbare svakheter, men gir likevel et uttrykk for hvilke universelle virkninger den nye teknologi ventes å få.

På den annen side vil bruk av datateknologi kunne medføre at det oppstår et teknologisk gap i forskningsverdenen mellom dem som gjennom aktivt og eksplorerende arbeid utnytter datateknologi og det forskningspersonalet som ikke eller kun i begrenset grad gjør bruk av datateknikk.

Gruppedannelsen er kanskje ikke i dag blitt særlig sterk og entydig, men man kan oppleve at det er et skille mellom «EDB-forskere» og en større «konservativ» gruppe som mer eller mindre klart avviser det nye verktøy og blir ved sin lest. Til dels opplever gruppene også kommunikasjonsproblemer, idet den nye gruppen utvikler et metodespråk som ikke er umiddelbart tilgjengelig for de uinnvidde. Dette er en lite heldig utvikling for en teknikk som skulle fungere som et redskap for og integreres i den eksisterende forskning og åpne for nye veier og nye løsninger på gamle problemer.

Polariseringen viser seg også ved en viss tendens til at nye prosjekter defineres med utgangspunkt i de anvendelsesmuligheter som datateknikken byr på og ikke først og fremst som resultat av en kritisk vurdering av det foreliggende materialet. Teknikken kan bli bestemmende, ikke bare for den metode som anvendes ved problemløsningen, men også for fastsettelse av selve problemstillingen. Dermed fjerner «EDB-forskeren» seg ytterligere fra sin mer tradi-

sjonelle kollega, og polariseringen og mangelen på integrering trer tydeligere frem.

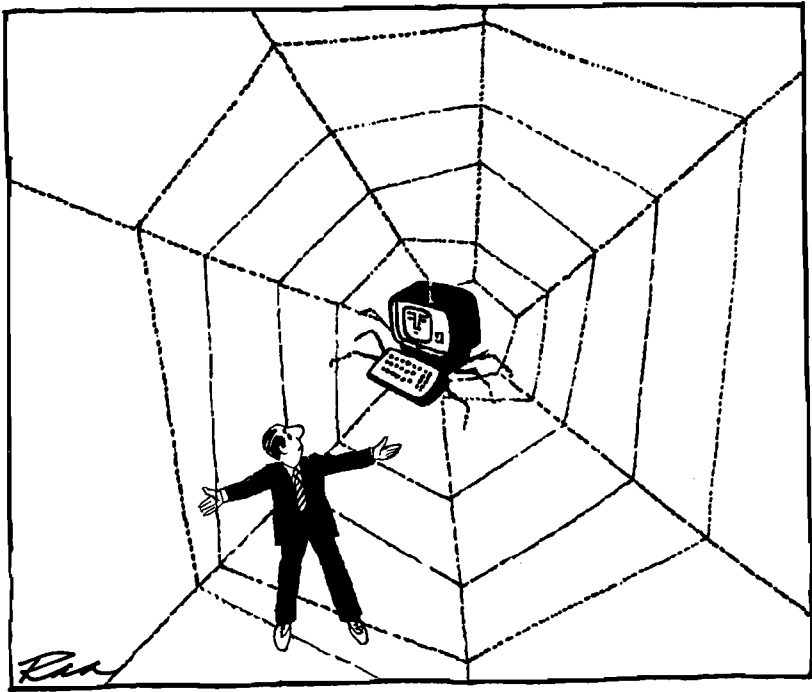
Dette leder naturlig over til spørsmålet om prioritering av forskning og forskningsprosjekter. Det er vel neppe fremmed for noen at den elektroniske databehandlings økonomiske og teknologiske dimensjoner vil kunne skaffe oss problemer også på dette felt. Hva skjer når store investeringer er gjort i avansert utstyr, og nye stillinger er fylt med nye mennesker med ny dyktighet i nye teknikker? Vil da nye forskningsprosjekter bli utvalgt og foretrukket, ikke på grunn av sin betydning for fagområdet som sådant, men på grunn av at de er egnet for databehandling? Eller for å trekke det ned på det individuelle plan: Når forskeren har fullført sitt dataprojekt, finner han da frem til neste prosjekt ut fra sitt kritiske kjennskap til fagområdet, eller ut fra sin nyvunne kunnskap om datateknikker? Kort sagt: vil EDB overta styringen i prioriteringsprosessen?

En positiv virkning av den datateknologiske utvikling vil være at den nåværende situasjon med knapphet på datakraft i mange forskningsmiljøer og konsentrasjon av EDB-ekspertisen ved de store forskningssentra vil bli avløst av en situasjon hvor datakraft vil bli lett tilgjengelig selv på mindre forskningssteder og hvor kunnskap om dette hjelpemiddelet vil bli langt mer utbredt enn tidligere.

Dette vil skape en større likhet i forskningsvilkårene f.eks. mellom forskere innen- og utenfor universitetsmiljøene. Forskere i de mindre forskningsmiljøer, som f.eks. ved de regionale høyskoler, vil få bedre vilkår til å løse forskningsoppgaver som innbefatter bruk av datateknikk. En langt større andel forskere enn i dag vil i fremtiden få anledning til å utnytte datakommunikasjon og ved det få mulighet til å komme i kontakt med forskningsdata i andre miljøer innen- og utenlands.

Det blir ofte hevdet at det nå ikke lenger spiller noen rolle hvor de store dataanlegg er plassert. Terminaler, telefonlinjer og/eller moderne datanett gjør det mulig å anvende en datamaskin hvor den enn måtte befinne seg. På denne måte blir datatjenestene desentralisert, hevdes det, og brukerne får det samme tilbud om maskinen er nærmeste nabo, eller om avstanden er 100 mil eller mer.

Dette er en sannhet med meget betydelige modifikasjoner. Den kan være holdbar for institusjoner som bare ønsker å nyttiggjøre seg bestemte tilbud som finnes på et anlegg, og som ikke har interesse av å påvirke utviklingen av tjenestetilbudet og ikke har behov for spesialtjenester. Men ønsker man å utarbeide egne tilbud på maskinen eller delta i utviklingen av de tjenester som tilbys, er det av avgjørende betydning å ha gode muligheter for nær kontakt med det systempersonalet som er knyttet til dataanlegget. Erfaringen tilsier at slik kontakt blir klart mest tilfredstillende når gruppene befinner seg i nærheten av



hverandre, og det blir da også færre praktiske problemer knyttet til bruk av dataanlegget.

Personell som kan tilrettelegge tjenester på en datamaskin blir en stadig viktigere ressurs etter hvert som anvendelsesområdet utvides og utstyrsprisene synker. For forskningsmiljøene, som alltid vil ha behov for spesieltjenester, er denne type personellressurser en forutsetning for anvendelse av EDB. Slikt personell må finnes i det enkelte forskningsmiljø, og de må ha lett tilgang til det aktuelle dataanlegg og til nøkkelpersonellet rundt dataanlegget. I dag finnes det EDB-tjenester for de humanistiske fag ved Universitetene i Oslo, Bergen og Trondheim og en tilsvarende tjeneste i Tromsø er under etablering. Disse EDB-organer har kun en EDB-konsulent hver. Dette er allerede i dag altfor lite for å møte de behov for tjenesteytelser som de humanistiske fagmiljøer har.

Interessante perspektiv for forskningsarbeidet gir også fremveksten av det elektroniske kontor, som gir anledning til en nærmest direkte kontakt (og faglig interaksjon) mellom forskere via datatekniske hjelpemiddel. Det vises her til den kommunikasjon og meldingstjeneste som EDB-folk i dag driver via terminalutstyr. Slike tjenester vil i fremtiden bli utviklet i retning av en superteleks, etterhvert også kombinert med billedoverføring. Det vises her dessuten til den rivende utvikling som pågår innen faksimileteknikken. Forskere har



påpekt at databehandling brukt i forskningsarbeid også kan øke samvirket mellom forskere i den forstand at det nå i større utstrekning enn før blir anledning til å dokumentere forskningsresultater og gi anledning til reell etterprøving av de analyseresultater som fremlegges. Mens det tidligere i praksis sjelden lot seg gjennomføre å etterprøve forskningsresultater gjennom å skaffe seg kopi av eller nyregistrere og behandle på nytt et stort forskningsmateriale, vil det nå bli langt bedre muligheter til å dokumentere et grunnlagsmateriale og stille det til disposisjon for andre forskere. Slik vil også datasituasjonen kunne endre seg dramatisk for samarbeidsprosjekter der en tar sikte på en systematisk, akkumulerende forskningsinnsats i grupper som arbeider i ulike miljøer. Et eksempel fra våre felt kan her være det samarbeid som er utviklet mellom bl.a. norske, engelske og amerikanske fagmiljøer i forbindelse med språklig tilrettelegging og analyse av store tekstmengder.

## Databanker

Gjennom utvikling av ny informasjonsteknologi basert på datateknikk vil også omløpstiden for forskningsresultater gå ned. I mange fag er det i dag en stor hemsko for fagutviklingen at det går for lang tid før ny viten i forskning kommer frem til kolleger eller får praktisk utnyttelse i anvendt vitenskap eller i utviklingsarbeid innen næringsliv og samfunnsliv. Her vil de kommende IoD-tjenester basert på datakommunikasjon kunne få stor effekt (jmf. utviklingsprogrammene for Scannet og Euronet) og den store aktivitet som i dag pågår (også kommersielt) for å skape internasjonalt rettede faktabaser for vitenskap og teknikk og databanker med vitenskapelig grunnlagsmateriale.

Til slike databanker vil det vanligvis være knyttet et informasjonssystem som gjør det mulig å trekke ut alle former for opplysninger som finnes i banken. Dette gjør det ofte mye lettere for brukerne å finne den informasjon de er ute etter enn om den samme informasjonsmengde var lagret i et tradisjonelt arkiv, i bokform e.l. Adgang til databanken kan man få via en terminal (f.eks. data-skjerm), og ved hjelp av nasjonale og internasjonale datanett kan man koble seg til fjerntliggende databanker. Allerede i dag utnytter forskere, særlig innenfor mat.nat. fag, medisin og samfunnsfag slike databanker i betydelig utstrekning.

Selv om slike databanker foreløpig har et forholdsvis beskjedent omfang innenfor de humanistiske fag, kan dette bli viktige datakilder allerede i forholdsvis nær fremtid. Pr. i dag er ca. 15 humanistiske databaser operative over terminal innenfor systemene til Lockheed og SDC i USA. Institute of Scientific Information, Philadelphia, melder at deres Arts and Humanities database vil bli åpnet for interaktiv søking i 1981. Slike utsikter gir interessante -men etter manges mening også skremmende - perspektiver.

Fra et norsk humanistisk forskningssynspunkt er det særlig to forhold ved emnet som bør tillegges vekt. Det ene er at vi sikrer oss tilstrekkelig kompetanse til å dra nytte av den utvikling som finner sted. Oppbygging av egne databan-

ker gir store muligheter, men forutsetter kunnskap om denne spesialiserte form for informasjonsformidling. Utnyttelse av andre databanker kan også gi store gevinster og krever normalt mindre kompetanse. Men skal man dra full nytte av data, vil det ofte være aktuelt å bearbeide dem videre på egenhånd.

Det andre forholdet er de muligheter slike databanker og deres lett tilgjengelige informasjonssystemer vil ha for å påvirke hva slags data forskerne tar i bruk og hva slags informasjon de benytter seg av. Det er rimelig å tro at lett tilgjengelige databanker vil kunne påvirke forskernes valg av forskningstema og -opplegg. Det er derfor viktig at spørsmålet om oppbygging av databanker vurderes også i forskningspolitisk sammenheng. En norsk politikk på dette området bør ses i internasjonal sammenheng. Det er f.eks. neppe ønskelig at utenlandske databanker (særlig amerikanske) skal bli en hovedkilde til norske forskeres vitenskapelige informasjon. Selv om databasene for bibliografisk informasjon og prosjektinformasjon kan virke imponerende ved sin størrelse og effektive oppdatering, må en ikke glemme at de også i høg grad er selektive i sin datafangst.

Direkte virkning for forskningsarbeidet vil også de nye former for resultatpresentasjon få. Som tidligere omtalt vil publiseringsfeltet undergå store forandringer i den kommende tid. Det er trolig at en stadig mindre andel (relativt sett) av vitenskapelige avhandlinger og andre publikasjoner vil bli utgitt og spredd i vanlig forstand. I stedet for utgivelse av bøker vil det bli etablert dokumentasjonssentra som tilrettelegger forskningsdokumentasjon og den vitenskapelige litteratur i datamaskinell form (evt. kombinert med mikroformer). Det vil bli utviklet EDB-tjenester for ulike sektorer av forskningsfeltet, og brukerne vil få mulighet til å motta aktuell litteratur via flere medier. Det kan nevnes at bl.a. Riksbibliotek tjenesten og Universitetsforlaget for tiden vurderer ulike opplegg for et service-organ for vitenskapelig, upublisert informasjon.

Fra forskerhold har det vært pekt på at den nye publiseringspraksis nok kan by på atskillige problemer for forskerne. Vi vil få to grupper: de som er så betydelige at de kan vise til at deres avhandlinger kommer ut i bokform, og de som bare kan henvise til at deres avhandlinger er tilgjengelige i maskinlesbar form eller på mikrofiche. Muligens vil dette kunne føre til at forskningen beveger seg i retning av det populærvitenskapelige, men en slik utvikling kan igjen bli en trussel for grunnforskningen.

### **Virkninger for humanistisk forskning.**

Ideelt sett bør det være slik at de datateknologiske hjelpemidler er tilpasset bruken i humanistisk forskning og er forenlig med målsettinger, tradisjoner og metoder innenfor denne forskningsgren. Dette er et viktig prinsipp å holde frem for å unngå at de humanistiske fag påvirkes - kanskje umerkelig - av datateknologi, på en måte som få ønsker. Teknologi skal tilpasses de faglige formål - ikke omvendt.

På dette punkt er det grunn til å være mer optimistisk enn tidligere. Mulighetene fremover til å tilpasse EDB-teknikk til de enkelte faglige formål vil bli større og ikke mindre - jfr. det som tidligere er sagt om de sentrale utviklingstendenser når det gjelder utstyr og programvare.

Det bør fremholdes at de humanistiske fag også selv vil ha et ansvar for hvilken art og kvalitet av datateknikk som vil være tilgjengelig. Det vil være opp til fagmiljøene selv og de nasjonale fagorganer gjennom engasjement, kunnskap og kreativt utviklingsarbeid å tilrettelegge bruk av denne teknikk slik at den fremmer den faglige utvikling. I så måte kan det vise seg å være mer gunnlig å ha en offensiv holdning også overfor leverandører av datateknologi for å få dem til å interessere seg for humanistisk databehandling enn å forsøke å skjerme seg fra den innflytelse som nye databehandlingsmetoder kan representere.

Oversikten foran har vist at ikke alle fag ligger like godt til rette for bruk av EDB. Her skal vi likevel være forsiktige å spå om fremtiden. Utviklingen går generelt raskt, og vi må regne med teknologiske sprang eller gjennombrudd som åpner for helt nye anvendelser av teknikken.

På den annen side vil det være slik at det vil foreligge en rekke rammevilkår som vil begrense innslag av nye hjelpemiddel i humanistisk forskning. De viktigste vil gjelde de generelle økonomiske betingelser for humanistisk forskning i årene som kommer. Ifølge de siste offisielle utredninger om veksten i norsk økonomi i 80-årene, vil det sannsynligvis bare kunne bli meget beskjeden økonomisk vekst i samfunnet som helhet. En slik lav vekstrate vil bety at den økonomiske ekspansjon innen humanistisk forskning, når det gjelder stillinger og driftsmidler, også vil bli meget begrenset. Dette vil i sin tur bety at det neppe er duket for en omfattende, teknologisk basert innovasjon i 80-årene dersom dette betyr store ekstraavgifter for institusjonene.

Ikke minst i de første faser når ny teknologi tas i bruk, vil det bli nødvendig med parallellføring av nye og tradisjonelle opplegg, f.eks. på arkivsektoren. Dette vil kunne gi ekstrakostnader inntil de nye opplegg er rutinisert og får full effekt.

### **Fritt valg?**

Det er likevel også grunn til å påpeke at situasjonen fremover kan bli slik at de humanistiske fag vil ha liten anledning til selv å velge fritt om nye data- og analysemetoder skal tas i bruk. Vi ser nå at det vil bli en langt mer åpen debatt enn tidligere om bruken av samfunnets økonomiske midler. Denne debatt vil også innbefatte forskningsområdet, og det vil mer enn før bli tale om å svare for forskningens mål og virkemidler. For humanistisk forskning kan dette bety at det vil bli lagt et ytre press på forskningsinstitusjonene for å få dem til å ta i bruk ny teknologi dersom det kan påvises at nye hjelpemidler setter forskningen i stand til å bruke de økonomiske midler på en mer fornuftig måte.

Den situasjonen kan da oppstå at det blir krevd at de humanistiske fag innfører nye metoder dersom de ønsker å opprettholde nivået av de økonomiske bevilgninger. Dersom forskningsmiljøene finner at en slik situasjon *kan* oppstå i fremtiden, bør det føre til en fornyet debatt om fag- og forskningspolitiske mål og midler.

Det er vel lite realistisk å tro at humanistisk forskning vil forsøke å skjerme seg fra den generelle tekniske utvikling i samfunnet ved å henvise til sin egenart.

Humanistisk forskning har heller ikke forsøkt å unngå bruk av telefon, skrivemaskin, regnemaskin eller spesialisert teknikk i forskningsarbeidet. Sett fra forskningssynspunkt er datamaskinen kun et nytt ledd i en lang, ubrutt teknologisk utvikling hvor f.eks. oppfinnelsen av trykkeriteknikk representerer en annen viktig milepæl. På den annen side representerer datateknologien farer, og dette kommer vi tilbake til nedenfor.

Forskning som baseres på EDB-bruk vil ofte kreve andre organisatoriske løsninger enn de vanlige innenfor humanistisk forskning og medføre andre typer kostnader. Større EDB-prosjekter vil gjerne foregå som gruppeprosjekter hvor det bl.a. kan inngå personale med EDB-kunnskap. Selv i mindre prosjekter vil det være behov for regelmessig konsulentassistanse.

Det er sjelden hensiktsmessig at forskeren selv setter seg tilstrekkelig inn i databehandling til å anvende datamaskinen helt selvstendig. Han trenger assistanse fra en EDB-kyndig person, ofte på et nivå hvor man i fellesskap kan utforme metoder og deler av forskningsopplegget. Man har altså en prosjektsituasjon hvor to eller flere personer samarbeider på et vitenskapelig nivå og utfyller hverandre med ulike spesialiteter.

Et av problemene her er at det nesten ikke finnes stillinger for denne type EDB-kyndig personell, jfr. den tidligere omtale av EDB-tjenestene for de humanistiske fag ved universitetene. Forskningsmiljøene har tradisjonelt bare rekruttert personer som kan forske på egenhånd innen vedkommende fag. De har foreløpig i liten utstrekning erkjent noe behov for å rekruttere kompetanse innenfor databehandling. Dette medfører naturligvis at det også er få som satser på å kvalifisere seg innen dette området. Dette har også sammenheng med at de tradisjonelle kriterier for meriterende arbeid ikke passer særlig godt for å vurdere tilrettelegging av EDB-metoder for humanistisk forskningsarbeid. Utvikling av EDB-metoder og forskningsopplegg i tilknytning til andres prosjekter kan resultere i prosjektrapporter og pen omtale i forord eller fotnoter. Innsatsen blir neppe vurdert på linje med forskernes artikler og avhandlinger, noe som representerer et problem for rekrutteringen. Dette er et problem som for øvrig ser ut til å gjelde prosjektforskning mer generelt.

Ut fra den kortfattede gjennomgåelsen som tidligere er gitt av de humanistiske fags forskningstilfang og arbeidsmåter, skulle det fremgå at en god del av de humanistiske fag er basert på bruk av talt og skrevet språk, ulike typer historisk primærkildemateriale og fremstilling av omfattende systematiserte materialsamlinger. Utviklingen av mer effektive redskaper for behandling av denne type informasjon gir håp om at det i fremtiden kan bli frigjort mer tid til analytisk behandling av forskningsmaterialet enn nå. Til i dag har det undertiden vært slik i mange forskningsmiljøer at en uforholdsmessig stor del av tiden har gått med til å bygge oversikt over et stort og komplisert datagrunnlag.

Heller ikke i fremtiden vil behov for vitenskapelig innsats ved oppbyggingen av kildene bortfalle, men mer effektive informasjonsbehandlingssystemer vil kunne korte avstanden fra materialtilrettelegging til forskningsfasen.

Som tidligere nevnt vil ikke minst automatisk kvantitativ databehandling kunne gi oversikt over og innsyn i et materiale på en langt mer rasjonell måte enn tidligere.

For de humanistiske fag vil også moderne datateknologi kunne gi opphav til nye faglige aktiviteter ved at bruken av de nye redskaper bringer humanistene



*Datamaskiner og edb vil snart være et almenfag både i ungdomsskolen og videregående skole. Bildet er fra Christianslund videregående skole i Fredrikstad der over halvparten av de ca. 700 elevene har edb som fag. (Norsk Datatidende, 2-81.)*

i kontakt med andre fagområder. Undertiden vil det utkrystalliseres nye fagdisipliner, slik som i det tidligere omtalte feltet kognitiv vitenskap. Språkforskerne som arbeider med datamaskiner, samarbeider her med psykologer, filosofer og EDB-forskere bl.a. i utforskningen av menneskets evne til lagring, kombinasjon og utnyttelse av kunnskap. Et av målene er å simulere språkforståelse i en datamaskin.

I det hele vil det være store utfordringer til forskere innen språkvitenskap til å gå med i utviklingen av nye datamaskinelle hjelpemiddel. Med sin innsikt i språkets systematiske og funksjonelle sider vil de kunne gi et viktig bidrag til utvikling av mer brukervennlige systemer for kommunikasjon i naturlig språk mellom menneske og maskin og gjennom utvikling av f.eks. informasjonssøkesystemer som bevarer de karakteristiske trekk som spørsmål og svar personer imellom inneholder. Utvikling av mer menneskevennlige systemer på disse områder vil kunne medvirke til å redusere det teknologiske gap som mange frykter kan oppstå mellom dem som behersker og dem som ikke kan bruke den nye teknikk.

### **Farer for den humanistiske forskning**

Mange humanistiske forskere er opptatt av de farer som ny datateknologi kan representere for de humanistiske fag. Nedenfor vil vi ta opp en del av disse. Det hevdes f.eks. at bruk av EDB kan føre til at de humanistiske fagmiljøer som har gjenstander eller tekster o.l. som sitt forskningsgrunnlag, vil komme til å fjerne seg fra dem ved bruk av de nye metoder. Frykten har som utgangspunkt at en ved bruk av EDB-metoder får langt større muligheter enn tidligere til å lagre og behandle data om f.eks. gjenstander, mens gjenstandene selv og de egenskaper ved dem som vanskelig er formaliserbare, kommer i bakgrunnen. Dette er en frykt som må tas alvorlig og som griper noe vesentlig ved den humanistiske forsknings egenart. Det ville være til uboteleg skade om bruk av datamaskin førte til at kun de deler av den humanistiske forskning som kan gjøres operasjonelle innenfor et hypotetisk-deduktivt analysebegrep, fikk tilfredsstillende levevilkår i årene som kommer.

Humanistisk forskning tar tradisjonelt i bruk en serie arbeidsmetoder. Noen av de viktigste innebærer bruk av intuisjon, medinnlevelse og tolkningsvirksomhet. Ved å redusere det humanistiske metodearsenal til et naturalistisk vitenskapbegrep kutter en samtidig over livsnerven i humanistisk forskningsarbeid. I den grad datamaskinelle metoder fører til en slik forarming av den humanistiske forskning, må det ropes et varsko!

For å unngå slike virkninger nytter det likevel ikke å forsøke å kapsle seg inne av frykt for datateknologi. Muligheten til å mestre de utfordringer som datateknologi stiller, skjer best gjennom en bevisst satsing på viten om de nye hjelpemidler og deres potensiale og ved å utvikle et høyt refleksjonsnivå omkring deres bruk i humanistisk forskning.

Av flere grunner vil den humanistiske forsker rett og slett være nødt til å interessere seg for EDB selv om han/hun personlig stiller seg kritisk til det nye og gjerne vil skjerme seg mot det. En av grunnene er at selv den forskning som springer ut av lokale forhold, vil komme i stadig sterkere kontakt med forskningsmiljøer i utlandet. Det er grunn til å tro at en rekke ledende forskningsmiljøer i utlandet vil utnytte muligheten til å bruke EDB-teknikk minst like mye som vi, også fordi arbeidsvilkårene for humanistiske fag i mange land er vanskeligere enn hos oss slik at de humanistiske forskere kan føle seg nødt til også gjennom bruk av ny teknikk å hevde seg i det totale forskningsmiljø.

Skal våre forskere være i stand til å følge med i forskningen ute, vil de i stigende grad bli nødt til å interessere seg også for de hjelpemidler og metoder som forskningen i utlandet gjør nytte av. Allerede i dag kan det i en rekke fagdisipliner være vanskelig å følge med for norske forskere på grunn av den mer utstrakte bruk av datamaskinelle og kvantitative metoder i sentrale forskningsmiljøer ute.

Slike problemer vil også oftere enn før kunne oppstå i forskerens hverdag. Som tidligere nevnt, er det som regel den unge generasjon forskere og personer i rekrutteringssituasjon som griper de utfordringer moderne databehandling gir. Men ved dette får også den etablerte del av forskersamfunnet nye utfordringer. Det er som oftest forskere i toppstillinger som skal veilede forskningsrekrutter eller evaluere forskningsarbeid ved eksamener og stillingsbesetninger. Uten innsikt i databehandling vil de i mange forskningsmiljøer ha vanskelig for å skjønne sine funksjoner på disse felt på en tilfredsstillende måte i 80-årene.

# Mikroprosessoren, arbeidet og mennesket

*Anfinn Stigen*

Hovedhensikten med dette foredraget\* kunne gjerne vært uttrykt i spørsmåls form: «Lever vi for å arbeide, eller arbeider vi for å leve?» Noen av oss, de privilegerte, vil kanskje svare at arbeidet er vårt liv. Men de aller aller fleste vil nok svare, iallfall etter å ha tenkt seg om, at selvfølgelig arbeider vi for å leve. Ikke desto mindre lever mange *som om* de levde for å arbeide.

Bakgrunnen for at det i dag kan være nyttig å stille dette spørsmålet, gi et svar og tenke gjennom hvorfor man vil svare som man gjør, er visse forhold som har å gjøre med det man gjerne kaller informasjonsteknologien, hvor bl.a. mikroprosessoren er et ledd. Innføringen av de automatiske styringsmekanismer som den nye teknologien gjør mulig, vil i visse virksomheter i ganske stor grad kunne erstatte menneskelig arbeidskraft. Dette gjelder bl.a. for store deler av industrien, kontorarbeid, administrasjon, tjenesteytende næringer, men naturligvis i mye mindre grad for helsestellet og andre omsorgsfunksjoner, der personen er like viktig som selve det arbeid som utføres. Det er to hovedgrunner til at man må regne med at den nye teknologi vil bli tatt i bruk i økende omfang utover i 80-åra. For det første er den nye teknologi pålitelig i drift; den reagerer bokstavelig talt med lynets hastighet, og den behøver ingen pauser som mennesket; den kan fungere under ekstremt ugunstige forhold, under meget høye temperaturer, i støv og skitt, dag og natt; den kan programmeres slik at den øyeblikkelig oppdager feil i arbeidsprosessen, og den kan også rette på disse feil, om den er programmert for det. For det andre er prisutviklingen på den nye teknologi nå slik at det for mange arbeidsoperasjoners vedkommende vil være direkte tapsbringende ikke å investere i det nye. Den menneskelige arbeidskraft prises ut av arbeidsmarkedet. I NOU 1980: 33 («Perspektiver på den teknologiske og økonomiske utviklings betydning for sysselsetting og arbeidsmiljø», Kommunal- og arbeidsdepartementet, avg. i august i år. Heretter ref. til som NOU.) gis et illustrerende eksempel på prisutviklingen (s.14). I 1975 tilsvarte prisen på en liten datamaskin 1000 ganger en gjennomsnittlig månedslønn i industrien. I dag koster en tilsvarende datamaskin *mindre* enn en månedslønn!

De fleste er enige om at dette vil føre til reduksjon i antall arbeidsplasser i de tradisjonelle næringer. Arbeidsløshet vil bli resultatet dersom det ikke skaffes nye arbeidsplasser på andre områder av samfunnslivet. Spørsmålet er om dette er mulig. I alle tilfeller må man regne med stor «strukturarbeidsløshet», dvs. problemer ved at de arbeidsledige (f.eks. fra industrien) ikke uten videre

---

*\*(Artikkelen bygger på et foredrag ved Chr. Michelsens Institutt Humanistiske Seminar på Nesbyen, 9. oktober 1980)*



kan fylle nyopprettede arbeidsplasser (f.eks. i helsestellet). Når det gjelder hvor stor arbeidsløshet man må regne med, varierer anslagene sterkt, ikke bare fordi forholdene er ulike fra land til land, men også fordi man vurderer de økonomiske utsikter for 80-årene sterkt forskjellig.

I en bok med den talende tittel «The Collapse of Work» (1979) vurderer Clive Jenkins og Barrie Sherman utsiktene for Storbritannias vedkommende og kommer med dystre spådommer. Hvis britene holder fram som de stemner og avviser den nye teknologi, sier de, så vil man i løpet av 15-20 år i Storbritannia få en arbeidsløshet på 5,5 millioner, dvs. at hver fjerde arbeidssøkende ikke vil finne arbeid. Grunnen er at en rekke konkurranseutsatte næringer vil være utkonkurrert av land som har tatt teknologien i bruk. Om britene derimot mottar utfordringen og anvender det nye, vil man i Storbritannia likevel få en arbeidsløshet på omkring 5 millioner. «Samme hvilken vei vi velger», konkluderer forfatterne, «så vil arbeidsmarkedet bryte sammen»(s.123).

NOU tegner et atskillig mer håpefullt bilde, iallfall for Norges vedkommende. Den peker på at svært mye vil avhenge av hvilken økonomisk politikk man velger. Det innrømmes at de aller fleste industriland nok vil oppleve betydelig arbeidsløshet i 80-åra. «Utsikten for den internasjonale økonomien i 80-årene synes dyster», og man må regne med «en omverden preget av svak økonomisk vekst og arbeidsløshet». Norge har imidlertid spesielle muligheter til å opprettholde relativ høy sysselsetting «selv i omgivelser preget av arbeidsløshet». (S.10 og 61-62.)

De spesielle omstendigheter som stiller Norge gunstigere enn de fleste andre land, er de store oljeinntektene som midt i 80-årene kan ventes å bringe overskudd i statsbudsjettet. Et alternativ som NOU peker på, er å bruke oljepengene til en utbygging av den offentlige sektor, bl.a. helsevesenet, slik at nye arbeidsplasser skapes til erstatning for dem som går tapt på grunn av innføring av ny teknologi. Dette alternativet, som NOU har gitt overskriften «solidaritet for kvalitet», er uttrykk for en fordelingsstrategi, bl.a. av arbeidsplasser, og er motivert av «omtanke for kollektivets interesse». (s.81, 91.)

### **Hvordan ble arbeid for alle det overordnede politiske mål?**

Dette overordnede politiske mål, som synes å være apriorisk godtatt ikke bare av NOU, men av samtlige politiske partier og politikere i vårt land, er «arbeid for alle» eller «full sysselsetting». Om politikeren blir presset, vil han si at naturligvis er ikke full sysselsetting noe mål i seg selv, men at det synes å være

et så uomgjengelig nødvendig middel til oppnåelse av de primære menneskelige verdier og goder, dvs. til det som gjør et liv rikt og godt, at det i praksis blir dette, full sysselsetting, man må kjempe for å sikre.

Og politikeren har rett. Slik samfunnet har vært og framleis er, er full sysselsetting ikke bare ønskelig, men også nødvendig. Men vi er nå på grunn av den nye teknologi på vei inn i en ny situasjon hvor svaret ikke er så opplagt. Hvis innføringen av den nye teknologi kan sikre opprettholdelsen av vår levestandard, eller endog øke den, samtidig med at behovet for menneskelig arbeidskraft går ned, hvorfor skulle vi da finne det påkrevd å opprettholde full sysselsetting? Er det nødvendig for mennesket å arbeide - for å kunne leve et meningsfylt liv? Kan bare et liv i arbeid være rikt og meningsfylt?

La oss stoppe opp et øyeblikk og betrakte vår situasjon. Her har menneskene i årtusener betraktet arbeidet som en forbannelse. Vi synes nå å stå på terskelen til realiseringen av den eldgamle drøm om gullalderen: muligheten til å kunne leve godt uten å måtte arbeide. Og så gripes vi av fortvilelse! Det verste er at vår fortvilelse synes berettiget. Men det er noe paradoksalt i hele situasjonen.

For arbeidet i seg selv er neppe en kilde til særlig glede for de fleste. For noen få er arbeidet en glede, men ikke for det store flertall. Og ingen skal innbille meg at mange foretrekker å stå opp i grålysningen en kald norsk vinterdag framfor å bli liggende. Ennå færre gleder seg til reisen til og fra arbeidsstedet. Og selve arbeidet er for de fleste monotont, kjedelig, sløvende og krafttappende. I en undersøkelse nylig svarte 85% at arbeidet var belastende og lite tilfredstillende.

Konklusjonen må bli at de fleste mener arbeidet *ikke* er et gode, snarere et nødvendig onde. De innfødte på Haiti, er det blitt meg fortalt, sier at hvis arbeidet hadde vært er gode, så ville de rike allerede for lenge siden ha sikret dette gode for seg selv! Det ligger mye visdom i dette. De rike og de mektige, de arbeider ikke. De lar i stedet andre arbeide for seg.

Arbeidsetikken - som går ut på: «Du skal arbeide, og det er *godt* for deg!» - er nettop utviklet for å få folk til å arbeide for seg. Den ble oppfunnet i en tid da det ble særlig lønnsomt å ha folk i sin tjeneste og ble lansert av dem som ville høste fruktene av andres arbeid. Den ble gjentatt fra prekestolen og fremstilt som del av Guds plan med menneskene. I folks bevissthet ble arbeidsomhet forbundet med alle gode karakteregenskaper: ærlighet, beskjedenhet, sindighet, pålitelighet. Det ikke å arbeide ble forbundet med uærlighet, upålitelighet, drikkfeldighet - med det å vite å klare seg på «annen måte», underforstått: på uærlig vis.

På 1700 og 1800 tallet ble arbeidet etter hvert så ufyselig, så tappende, så farlig at det gikk på helsa og på alle virkelige menneskelige goder løs. Arbeidsetikken måtte sprites opp for overhodet å hindre folk i å forlate arbeidet eller rett og slett i å gjøre opprør. En rekke slagord ble lansert, såsom «Arbeidet

adler mannen», «Arbeit macht das Leben süß» og «Lediggang er roten til alt ondt» - det siste særlig effektivt fordi det har en bibelsk klang, selv om det ikke forekommer i Bibelen. Marx syntes synd på arbeiderne, men hyllet arbeidet og hevdet at arbeidet er kilden til all verdi, dvs. at alle goder kommer fra arbeidet!

Det var på denne tid at arbeidet begynte å erstatte religionen som opium for folket. På liknende måte som religionen for mange tidligere hadde vært en tilflukt til trøst for livet i verdens jammerdal, ble nå selve arbeidet noe man kunne flykte til for å komme unna de egentlige livsproblemer, og arbeidsnarkomani ble et honnørord, slik gudfryktighet tidligere hadde vært en hedersbetegnelse. De aller fleste blant de eldre generasjoner bærer enda i dag arbeidsetikken i sitt hjerte.

I betraktning av arbeidsetikkens sterke grep om oss, er det bemerkelsesverdig at de fleste likevel, når de skal være ærlige, vil si at *arbeidet i seg selv* snarere er et onde enn et gode, men altså et nødvendig onde. Men likevel -når et menneske står i fare for å miste sin *arbeidsplass*, er fortvilelsen ekte nok.

### **Hva er det som gjør en fast arbeidsplass til et gode?**

Vi må derfor skjelle mellom arbeidet i seg selv og det å ha en arbeidsplass, en jobb. Det å ha en trygg arbeidsplass er utvilsomt et gode. Men hva er det som gjør det til et gode? Hva er det som er så viktig i dette? Hvilke primære menneskelige verdier er man sikret i og med at man har en trygg arbeidsplass?

Med utgangspunkt i dette spørsmål kan vi se at det hovedspørsmål vi har reist i det foregående - «kan vi oppnå et rikt og godt liv og en meningsfylt tilværelse uten å arbeide i tradisjonell forstand» kan besvares i tre etapper: 1) Hvilke menneskelige goder følger med det å ha en trygg arbeidsplass? 2) Kan disse verdier sikres på annen måte enn gjennom å arbeide, dvs. uten å ha en fast arbeidsplass i tradisjonell forstand? 3) Hvis dette er mulig, er det da grunn til fortvilelse? For i så fall er ikke sysselsetting i tradisjonell forstand et nødvendig middel til oppnåelse av disse verdier.

Hva er så de verdier, de goder, som sikres gjennom en trygg arbeidsplass? For det første gir den oss inntekter i form av lønning, som i vårt samfunn er et praktisk talt nødvendig middel til sikring av *livsnødvendigheter*, såsom mat, klær, hus osv. For det andre er denne lønn en fortjent belønning for vår arbeidsinnsats, og det at vi fyller en arbeidsplass gir oss en følelse av at vi har en nyttig funksjon i samfunnet. Disse forhold gir oss, på bakgrunn av vår sterke innforlivede arbeidsetikk, et grunnlag både for *andres aktelse* og egen *selvaktelse* og selvtillit. For det tredje gir bevisstheten om å ha en fast arbeidsplass oss *trygghet*, ikke bare i øyeblikket, men også for fremtiden, og bidrar på denne måten til å fjerne bekymringer og usikkerhet. For det fjerde vil de fleste arbeidsplasser bringe en i *nærhet til andre* og ikke sjelden skape grobunn for *vennskap*.

Disse fire momenter vil sannsynligvis de fleste være enige om å knytte til en arbeidsplass, og jeg går ut fra at de fleste også vil regne dem som utvilsomme og fundamentale goder. I tillegg vil nok mange legge til slike momenter som at arbeidet bringer en viss regelmessighet, en viss rytme, inn i deres liv, som altså oppleves som noe positivt. Andre vil ta med slike ting som at arbeidet og livet på arbeidsplassen bringer forandring og omskifte inn i en tilværelse som ellers ville være ensformig, det å komme bort fra det daglige mas med familie, barn osv. oppleves utvilsomt også av mange som positivt.

### **Hva er de primære menneskelige goder?**

Hva det er som gjør disse momenter til verdifulle og viktige for mennesker, er et spørsmål som det er viktig å tenke over, men som vi ikke her kan gå nærmere inn på. Men man bør iallfall være oppmerksom på at «verdi» ofte brukes i uegentlig betydning, f.eks. om penger eller verdisaker, som derfor heller burde kalles verdimidler eller verdimålere, fordi de egentlig er verdifulle bare i den grad de kan brukes som midler til å skaffe egentlig menneskelige verdier. En egentlig verdi er et gode for mennesket, dvs. noe som direkte bidrar til at et menneske opplever sitt liv som bedre og rikere. En verdi er noe jeg strever etter å oppnå hvis jeg ikke har det, og som jeg strever etter å beholde om jeg alt har det.

Filosofer, psykologer og sosiologer har prøvd å sette opp lister over primære menneskelige verdier (goder). Det synes å være forholdsvis bred enighet om at en slik liste iallfall bør inneholde følgende momenter, forsøksvis prioritert :

Livsnødvendigheter

Trygghet

Nærhet (fortrolighet) til andre

Kjærlighet

Vennskap

Andres aktelse

Selvrealisering

Følelsen av å skape noe verdifullt

En slik liste er beheftet med mange uklårheter, især på grunn av vagheten i betegnelsene. Også når det gjelder rekkefølgen av momentene - hva er viktigst å sikre seg - kan man ikke vente enighet. Den som har noen å være glad i, men ikke har realisert seg selv og f.eks. ikke har funnet seg en yrkesmessig tilfredstillende plass i livet, vil kanskje mene at selvrealisering er viktigere enn kjærlighet og derfor sette selvrealisering høyere enn kjærlighet. Likevel synes det å



være ganske stor enighet om at et menneskeliv oppleves som rikere, bedre og mer meningsfylt for så vidt og i desto større grad som disse momenter er til stede i livet.

### **Kan de primære goder sikres uten å arbeide?**

Hvis vi nå sammenlikner denne liste med den liste av verdier en god arbeidsplass gir oss, og som så slik ut:

Livsnødvendigheter

Trygghet

Nærhet (fortrolighet) til andre

Vennskap

Andres aktelse

så ser vi noen interessante ulikheter. (Selvaktelse er her utelatt fra begge lister, ikke fordi jeg ikke regner det som et primært gode kanskje viktigst å oppnå, men vanskeligst å beholde - men fordi jeg betrakter det som et resultat av oppnåelsen av de andre godene.) Kjærlighet er et primært menneskelig behov, og å få dette behovet tilfredsstilt er et primært menneskelig gode, men få får det tilfredsstilt på arbeidsplassen. To andre momenter savnes også i arbeidsplassverdilisten, nemlig selvrealisering og følelsen av å skape noe verdifullt, to momenter som henger nøye sammen.

Dette siste forhold er meget viktig. Skapertrangen, behovet for å gjøre eller lage noe som man finner verdi i, synes å ligge dypt i hver av oss. Denne gir seg de forskjelligste uttrykk: fra det å dyrke jorda og bygge seg en hytte, over til de forskjelligste former for kunstnerisk virksomhet, til det å utvikle sin egen personlighet, utvikle sine psykiske og fysiske anlegg og evner, med eller uten tanke på å fylle en bestemt funksjon i samfunnet. Det er dette siste som vanligvis forbindes med «selvrealisering».

Men foregår ikke denne selvrealisering også i arbeidslivet, vil man spørre. Det er dessverre slik at oppdelingen av arbeidsprosessen har gjort det vanskelig å oppleve arbeidet som en form for selvrealisering og vanskelig å oppleve skaperglede i arbeidet, forhåpentligvis med unntak av noen få primærnæringer. Skapergleden er knyttet til bevisstheten om å gjøre eller lage noe som er av verdi, noe som er godt, noe som kan ha betydning for noen og som beriker tilværelsen.

Om ens arbeid faktisk gjør livet bedre for noen, er et spørsmål som en vanlig arbeider sjelden stiller, og vanskelig kan stille. En arbeider har ikke råd til å spørre om det han eller hun lager, er av verdi for menneskene, om det som produseres gir menneskene et rikere liv. Selve arten av arbeidsprosessen har



gjort spørsmålet latterlig - ja, umulig. De fleste bryr seg ikke en døyt om hva de produserer: hus eller pyntelister på biler, vaskemaskiner eller spilleautomater, brød eller tobakksartikler, legemidler eller bakterier til bakteriologisk krigføring.

### **Er verdiinnsats viktigere for mennesket enn arbeid?**

Her er jeg ved mitt viktigste poeng. Det perspektiv vi nå har anlagt, gjør det naturlig å erstatte det etter hvert verdinøytrale begrep «arbeid» med det positive begrep «skaping (og bevaring) av verdier» og å hevde at et kvalitativt bedre samfunn er avhengig av *verdiskapende (-bevarende) innsats*, ikke av det arbeid som utføres. Politisk må vi ta sikte på et samfunn som legger forholdene til rette for verdiskapende innsats («verdiinnsats») av enhver art, organisert eller ikke. Og det er slik innsats, hva enten den faller inn under det tradisjonelle arbeidsbegrep eller ikke, som må oppmuntres, honoreres, verdsettes og belønnes. Denne idè danner grunnlag både for en vurdering av lønnsarbeid i tradisjonell forstand og for å søke å finne fram til et *system for å verdsette og belønne enhver innsats etter dens verdi* for menneske og samfunn.

Det tradisjonelle arbeidsbegrep er samtidig for snevert og for vidt. Følgen av at det er for snevert, er at mye av det som *burde* gjøres fordi det bidrar til et rikere liv, ikke blir gjort ettersom det ikke kommer inn under vanlig lønnsarbeid og derfor ikke belønnes eller oppmuntres på annen måte. Følgen av at det er for vidt, er at mye gjøres som *ikke* er verdifullt, dvs. ikke bidrar til et rikere liv, men altså gjøres likevel fordi det belønnes.

La meg konkretisere påstanden om at det tradisjonelle arbeidsbegrep er for snevert og derfor ikke omfatter så mye som det burde omfatte. Når jeg leker med mine egne barn i hjemmet, regnes det ikke som arbeid, men hvis jeg som barnehagelærer leker med barn, selv om det var mine egne, ville det være arbeid. Hvis jeg pleier en syk i familien, er det ikke arbeid, men ganske annerledes hvis det gjøres av en pleier på et sykehus. Hvis jeg gir gode råd til en venn i en vanskelig livssituasjon, er det ikke arbeid, men saken stiller seg ganske annerledes om jeg er psykisk utdannet rådgiver i funksjon. Om jeg på personlig basis hjelper en elev med leksene eller en student med studiene, eller overhodet er med på å utvikle en persons psykiske eller fysiske evner, er det ikke arbeid. Annerledes hvis jeg gjør det samme som lærer eller idrettsinstruktør. Hvis jeg bygger mitt eget hus, blir det ikke verdsatt av samfunnet, noe ganske annet om jeg overlater det til en byggmester. Tendensen er at verdiinnsats ikke blir verdsatt av samfunnet før den er profesjonalisert. Og når innsatsen først er profesjonalisert, blir den verdsatt hva enten det skapes virkelige verdier eller ikke. I stedet burde all innsats som skaper menneskelige og samfunnsmessige verdier, honoreres eller verdsettes på en eller annen måte. All virksomhet som tar sikte på å gjøre en ting mer verdifull gjennom



bearbeidelse, danning, dyrking, må verdsettes. Slik verdiskapende eller verdibevarende innsats er «kulturinnsats» i vid forstand. En person som utvikler sine egne evner gjennom studier, er kulturarbeider like fullt som en bonde som kultiverer jorda.

Naturligvis kan det være vanskelig i de enkelte tilfeller å avgjøre om en innsats er verdiskapende eller ikke, hva enten innsatsen skjer i eller utenfor det organiserte arbeidsliv. Er produksjon av krigsmateriell menneskelig og samfunnsmessig verdifullt? Ja, hvis et sterkt forsvar er nødvendig for å bevare vår livsform og våre fundamentale menneskerettigheter. Er det å utvikle sin sangstemme eller lære å spille et instrument verdifullt? Ja, hvis det beriker tilværelsen. Viktigere enn å kunne gi klare svar er det å venne seg til å stille slike spørsmål.

### **Et system for å verdsette en innsats etter dens verdi for menneske og samfunn?**

Tilbake til vår situasjon. For første gang i menneskehetens historie er muligheten kanskje til stede til å skape et samfunn hvor en god del mennesker ikke er opptatt med lønnsarbeid i vanlig forstand, men i stedet kan være opptatt med uorganisert verdiinnsats, dvs. med å bevare verdier og å skape verdier ved personlig innsats, med «kulturvirksomhet» i vid forstand. Behovet for slik innsats er naturligvis umettelig.

Det står igjen å finne fram til et praktisk gjennomførbart system for å belønne innsatsen etter den verdi som ligger i den. Det er vanskelig å tenke seg andre praktiske belønningssystemer enn det økonomiske, dvs. pengesystemet. Folk må sikres en form for lønn for sin verdiinnsats. Politikere har vært inne på tanken om en folkegrunnlønn. Kunne en slik lønn, sikret alle over 16 år og svarende til minstepensjonen, danne grunnlaget? Og kunne man tenke seg denne grunnlønn forhøyet etter et poengsystem for å honorere de verdiinnsatser man kunne dokumentere å ha gjort? Innviklet? Ikke mer innviklet enn det skattesystemet vi har nå. Et nytt ligningsvesen som på grunnlag av de enkeltes verdiinnsatsselvangivelser utlignet belønninger i stedet for skatter? Dette er naturligvis bare lite gjennomtenkte løse ideer.

Har vi midlene til å gjennomføre et slikt system? Neppe i dag, men om oljeinntektene blir av en slik størrelse som det er forventet, ville vi iallfall kunne begynne overgangen til et slikt samfunn. Det dreier seg ikke om en overgang i løpet av et par år, naturligvis, men om en gradvis reform over tiår. Og for enkelthets skyld har jeg her snakket som om en del av befolkningen var opptatt med organisert arbeid, en annen del opptatt med uorganisert verdiinnsats. Naturligvis kan en og samme person dele seg på de to slags virksomheter, slik i dag mange kunstnere og mange studenter har et arbeid ved siden av.

## Hva kan vi slå fast allerede nå?

Rent bortsett fra hvordan disse praktiske problemer skulle kunne løses, er det en rekke viktige konklusjoner vi kan trekke allerede nå. For det første: Menneskene kan leve et rikt og verdifullt liv uten å ha et arbeid i tradisjonell forstand. Dette er det viktig å ha klart for seg når vi nå etter alt å dømme står overfor det faktum at mange ikke vil få arbeid, eller at de må innstille seg på et liv med lite arbeid og mye fritid. For det andre: Det avgjørende er at fritiden ikke blir ørkesløshet..For hvis det er riktig at skapertrangen, eller behovet for å gjøre noe verdifullt, er nødvendig for ens egen selvaktelse, for ikke å snakke om andres aktelse, så kan ikke mennesket oppleve sin tilværelse som meningsfylt uten å være *aktiv*. Det viktige blir å hindre at fritiden blir passivitet.

For det tredje: Det er ikke likegyldig hva aktiviteten består i. Fritiden må fylles med verdiskapende aktivitet, virksomhet som tilfører menneskene og samfunnet goder. Holdningen til fritiden må forandres. Folk er ikke opplært til og ikke vant til å *bruke* fritiden, men *forbruke* den. Her bærer skolen et stort ansvar, men naturligvis foreldrene et større. Til tross for mange fine ord om at skolen skal utvikle elevens personlighet, er pensum innrettet til å føre eleven fram til å gjøre en innsats i arbeidslivet, slik at man kan utføre det arbeid man blir satt til, helst uten å spørre etter om det man produserer og skaper, er godt eller dårlig for menneske og samfunn. Skolesystemet har vært innstilt på å sikre at man lykkes i ens karriere, ikke i ens liv. At vi er aktive, at vi gjør og lager noe i vårt arbeid og i vår fritid, det er viktig. Men dette er ikke det viktigste. Det viktigste er *hva* vi gjør og lager i vårt arbeid, i vår fritid, i vårt liv - om det vi skaper, er godt eller ondt for mennesket.

# RAPPORTER

## **EDB-kurs for tilsette i arkivverket**

*Egil Øvrebø*

Ved NAVF's EDB-senter for humanistisk forskning i Bergen vart det i tida 25. august-5. september 1980 halde eit EDB-kurs for tilsette i Arkivverket. Kurset var arrangert ved samarbeid mellom EDB-senteret og Riksarkivaren, med i alt 12 deltakarar frå Riksarkivet og dei seks statsarkiva i landet.

På to felt møter Arkivverket EDB-utfordringa: Vi skal ta mot eldre EDB-materiale frå statsforvaltninga til liks med anna arkivmateriale. Og vi treng å gjere oss nytte av EDB i vårt eige arbeid med arkivsakene. Det første punktet vert tema for eit seinare kurs, - no galdt det bruken av EDB internt, i registrerings- og framfinningsarbeidet. Kurset ga elementær innføring i EDB-teknikk, orientering om prosjektarbeid og om tilgjengeleg programutrusting. Det siste var lagt opp omkring eit aktuelt prøvemateriale (folketeljinga for Tromsø 1885), og med øving på terminal om ettermiddagane. Deltakarane møtte ein venleg og effektiv lei- og lærarstab.

Kurset var så avgjort nyttig. Arkivverket opplever ein sterkt aukande pågang frå alle brukargrupper, - det er kraftutbygging her, nye lærestolar der, det er nærsamfunn og det er «roots» utan ende. Med knappe ressursar skal vi prøve å nå ein katalog- og registerstandard som kan gjere det mogeleg å møte denne pågangen. Vi har ikkje hatt råd til å eksperimentere særleg mykje med EDB i arkiva, vi har heller ikkje råd til å la vere. Slik kurset var lagt opp ga det oss røymsler og rikeleg av impulsar, positive og negative (garbage in - garbage out). Kurset ga oss også kontakt med eit fagmiljø som vi nok skal komme til å plage etter kvart som nyvunnen kunnskap skal prøvast i praksis, på eldre folketeljingar, panteregister, kyrkjebøker, katalogar eller kva det meir måtte verte.

# First International Workshop on Natural Communication with Computers, Warszawa, 9-12 september 1980.

*Knut Hofland*

Konferansen ble arrangert av en internasjonal komite med professor *Leonard Bolc* fra Institutt for informatikk ved Warszawa universitet som formann. Professor Bolc er redaktør av en ny bokserie «Natural Communication with Computers», utgitt på Hanser Verlag, hvor det hittil har kommet ut 4 bind. Flere av prosjektene som ble presentert på konferansen er nærmere beskrevet i denne serien. Konferansen samlet 70 deltagere fra 14 land og av disse kom ca. 1/3 fra Vest-Europa og Amerika. På grunn av den spente politiske situasjonen var deltagerantallet en del lavere enn det som var påmeldt. Konferansen ble holdt i det store kultur- og vitenskapspalasset.

Formålet med konferansen var å bidra til spredning av informasjon om det stadig økende antall prosjekter innen fagfeltet. Ved åpningen av konferansen ble det delt ut skriftlig omtale av ca. 40 prosjekter og 2/3 av disse ble presentert i foredrag. Prosjektene var delt inn i 3 grupper:

- 1) Kommunikasjon ved naturlig tekst
- 2) Kommunikasjon ved tale
- 3) Digitale billedbehandlingsystem

Nesten 2/3 av foredragene var fra gruppe 1 og bare noen få fra gruppe 3. Foredragene fra gruppe 2 og 3 var gjennomgående rettet mot en teknisk beskrivelse av maskinutstyret, mens foredragene fra gruppe 1 beskrev metodene og oppbyggingen av systemene. Det kunne virke som om emnegruppe 3 lå noe til siden for de to andre områdene. Men i framtiden vil utvilsomt behandling av tekst, tale og bilder/grafiske former bli integrert i så forskjellig teknologi som for eksempel kontorarbeidsplassmaskiner og roboter.

I foredragene kom det fram en liten motsetning mellom de som hadde en spesiell lingvistisk modell som utgangspunkt for systemene og de som hadde et praktisk utgangspunkt. For de sistnevnte gjelder det å lage et system som fungerer bra, men der de metoder som blir brukt, ikke kan sies å representere en enhetlig lingvistisk modell.

Mange av systemene som ble presentert hadde som utgangspunkt å gjøre det enklere for brukere å få ut opplysninger fra en eller annen database. Disse systemene skilte seg lite fra hverandre. Et gjennombrudd på denne fronten vil



*Kultur- og vitenskaps-  
palasset i Warszawa.*

trolig skje når maskinlevarandørene videreutvikler slike systemer og markedsfører disse. Noen av systemene som ble presentert var utviklet eller ble støttet av leverandører. Dette gjaldt CONDOR (Siemens), POL (Hewlett Packard) og USL (IBM).

#### **Momenter fra en del av foredragene.**

*N. Banerjee m. fl. (Siemens): CONDOR - A Natural Language Oriented Database - Information System*

Dette systemet utvikles av Siemens med støtte av den tyske stat og arbeidet har pågått siden 1973. Systemet skal kunne brukes av en vid gruppe mennesker fra EDB-spesialisten til brukere uten noe bakgrunn i databehandling. Systemet er foreløpig på eksperimentstadiet. Det inneholder prosedyrer både for å behandle naturlig språk og strukturert informasjon. Databasen kan også inneholde ustrukturert tekstinformasjon. Systemet finner stammer til ord i teksten og bruker statistiske metoder til å plukke ut ord som skal brukes som indekser. Ved spørring kan strukturerte opplysninger kombineres med informasjon i fri tekst. Systemet inneholder også prosedyrer for redigering, sortering og utskrift.

*D. Kaminska-Kepa m. fl. (Warszawa): The Natural Language Question-Answering System DIALOG.*

DIALOG er et eksperimentssystem for spørsmål-svar, dokumentgjenfinning og «story understanding». Systemet består av komponenter for behandling av naturlig språk, deduksjon og databasehåndtering. Den deduktive komponent består av to deler, en basert på predikatlogikk og en som bruker ramme(frame)lignende kunnskapsrepresentasjon. Systemet gjør bruk av en ATN grammatikk og er skrevet i LISP. Et ekstra problem ved analyse av polsk språk er den frie ordstillingen. Systemet prøves ut i to konkrete prosjekter. Det ene bruker en database over alle lover og forordninger innen miljøvern som er vedtatt de siste 50 år. Det andre prosjektet er innen medisinske diagnoser og behandlingsopplegg. Systemet ble demonstrert på terminal under et besøk på et av universitetes dataanlegg.

*H.-J. Schneider m. fl. (Vest-Berlin): What has a SIREN got to do with a BEAST.*

Foredraget ga en oversikt over et prosjekt der målet er å foreta en automatisk konstruksjon av semantiske nettverk fra en naturlig tekst, BACON (Berlin Automatic Construction Of semantic Network). Systemet består av 3 hoveddeler. Den første, BEAST (Berlin Automatic Semantic oriented Translation), oversetter naturlig språk til et semantisk representasjonsspråk basert på operator calculus etter Montague. Denne delen gjør bruk av en dialog for å løse tvetydigheter. Den andre delen, SIREN (SemantIc REpresentation and Net evaluator), konstruerer et semantisk nettverk og tolker spørsmål med hensyn til nettverket. Systemet kan inneholde fakta og indisier. Disse siste kan ved en slutningsprosess gi nye fakta. En siste del av systemet genererer svar i et tilnærmet naturlig språk. Systemet er programmert i SIMULA og det er planer om å bruke dette i et informasjonssystem innen medisin.

*J. Mariani m. fl. (LIMSI, Orsay, Frankrike): Man-Machine speech communication upon different tasks.*

Foredragsholderen startet med å beskrive et forsøk ved John Hopkins universitetet i USA. Det ble her testet forskjellige kommunikasjonsformer mellom to personer der den ene visste løsningen på et problem og der den andre skulle løse problemet. Det viste seg at ved å bruke tale ble tiden for å løse problemet redusert til det halve. Resten av foredraget ga en oversikt over de prosjekter som ble drevet innen talesyntese og -analyse. Det ble gjort utstrakt bruk av mikroprosessorer. Et system for gjenkjenning av isolerte ord kunne i sann tid skille 100 ord. Dette systemet ble blant annet brukt i roboter, vokal FORTRAN programmering og pilotkommandoer i en cockpit. Innen kontinuerlig talegjenkjenning opererte en i 5 ganger sann tid og ca. 50% av fonemene ble korrekt gjenkjent. Dersom det ble gitt flere alternativer til hvert fonem, var det rette fonem innen de 4 beste i 80% av tilfellene. Dette gjalt for

dediserte brukere. Et eksperiment for å segmentere en (perfekt) fonemstreng til grafemer i sann tid basert på en ordliste på 170 000 ordformer ga 5% feil eller tvetydigheter.

*B. H. & F. B. Thompson (Caltech, USA): Introducing POL: A Problem Oriented Language System.*

Foredragsholderne har tidligere vært med på utviklingen av REL systemet der en ved naturlig språk har tilgang til en database og der brukeren kan utvide og modifisere data og definisjoner. REL systemet er skrevet i assembler for IBM. På bakgrunn av de erfaringer som ble gjort med dette systemet er en nå i ferd med å utvikle et nytt system, POL, med utvidete muligheter. POL systemet skrives i Pascal for en minimaskin fra Hewlett Packard med fast disk og primærlager på 400 Ktegn. Utstyret koster i USA ca. 50.000. Alle hjelperutiner skrives separat slik at systemet i høy grad er flyttbart til andre maskiner. Til hjelp i oppbyggingen av analysesystemet for naturlig språk er det blitt gjort forsøk med å studere dialogen som blir brukt ved problemløsning i forskjellige former for kommunikasjon. De 3 formene som ble studert var direkte kommunikasjon ansikt til ansikt, terminal til terminal og menneske til maskin. Ved den siste formen ble REL systemet brukt. Over 80 personer deltok og dialogene utgjorde ca. 80 000 ord og varte i 50 timer. Ved disse forsøkene fikk en fram karakteristiske trekk ved dialogen og dette ble brukt i analysesystemet. Systemet bruker Martin Kays metode for parsing. POL- systemet inneholder et fullstendig relasjonsdatabasesystem og det er mulig å ha flere nivåer av en database. En database kan da bygges på en annen og forandringer i den ene blir gjort gjeldende i den andre men ikke omvendt. Systemet ventes å være ferdigutviklet i løpet av 2 år.

Den lokale arrangementskomite gjorde en utmerket jobb og det ble invitert til en ny workshop i 1983.

---

## COLING 80

**The 8th International Conference on Computational Linguistics, Sept 30 - Oct 4 Nippon Center Hall, Tokyo.**

*Sigbjørn Århus*

COLING 80 er den 8. internasjonale konferanse i datamaskinell lingvistikk som har vært holdt. Konferanseserien startet i 1965, og den forrige ble holdt i Bergen i 1978. Før 1965 var en sesjon på IFIP(International Federation for Information Processing) det eneste internasjonale samlingssted for drøfting av emner innenfor datamaskinell lingvistikk.

Japan er i lengre tid blitt anmodet om å arrangere en konferanse i COLING-serien, men på grunn av den store geografiske avstanden fra Europa og Amerika har det tidligere ikke latt seg gjøre.

I 1980 var imidlertid Japan og Australia valgt som arrangørsted for IFIP, og siden deltakerne på COLING også var potensielle deltakere på IFIP, ble årets COLING lagt til Tokyo umiddelbart før denne.

COLING 80 ble arrangert i Nippon Toshi Senter, som ligger sentralt i Tokyo. Konferansen hadde samlet ca. 230 deltakere fra en lang rekke land, men med en klar overvekt av japanere. En av dagene var avsatt til en felles sightseeing hvor det bl.a. ble gitt en omvisning på bilfabrikken Nissan.

Konferansedagene var delt inn i to økter, og tre parallelle sesjoner, men p.g.a. ulik lengde på foredragene, var det ikke mulig å skifte fra en sesjon til en annen. Titlene på de tre hovedsesjonene var Linguistics, Logic and Informatics.

Innenfor lingvistikk ble det gitt ca. 30 foredrag fordelt på syntaks, semantikk og poetikk, mens det i informatikkseksjonen ble gitt en rekke presentasjoner som tok opp problemer med inndata, utdata, datamaskinell oversettelse, talegjenkjenning, informasjonssystemer, programmeringsspråk o.l.

I det følgende vil jeg referere fra noen av presentasjonene, også med det formål å vise bredden i de emner som ble tatt opp.

### **Dataregistrering av japansk tekst.**

Japan har et stort og komplisert tegnsystem som omfatter ca. 150 fonetiske symboler, kalt kana. De foreligger i to varianter, hirakana og katakana. I katakana brukes mest navn og ord som har vestlig opprinnelse, hirakana for alt annet som skrives fonetisk. Det brukes også store mengder ideografiske tegn, dvs. tegn som uttrykker et helt begrep. Ca. 50 000 ideografiske tegn- kalt kanji- eksisterer, men rundt 1000 kanji er tilstrekkelig for bruk i dagliglivet.

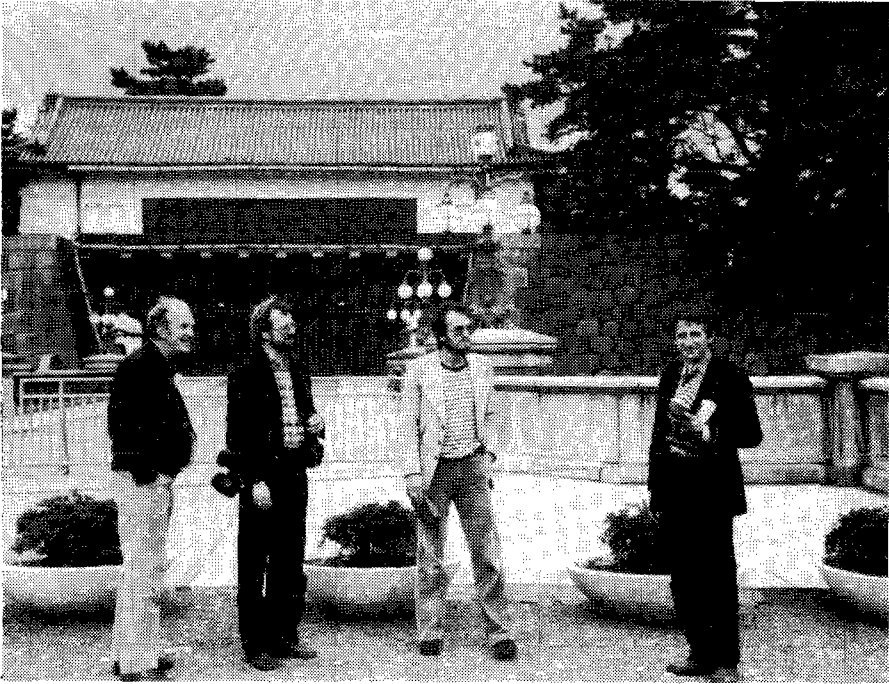
Det store antall tegn skaper store problemer for japansk databehandling, og et vesentlig problem er knyttet til registrering av data. På utskriftsiden kommer man langt med matriseskriverne og skjermterminaler.

Det ble gitt flere interessante foredrag hvor man tok opp problemer knyttet til registrering av japansk tekst.

### **Analyse av det engelske standardtastaturet på skrivemaskiner.**

*Yuzuru Hiraga* betraktet i sitt foredrag maskin-/terminalskrivning som en komplisert prosess, som forutsetter såvel mental aktivitet som fysisk bevegelse av fingrene.





*Fire norske deltakere: fra v. Kolbjørn Heggstad (UiB), Geir Berge (UiB), Sigbjørn Århus (Senteret), Per Bjørn Pedersen (Rogaland DH).*

---

I foredraget ble det presentert resultater fra en grundig analyse av plasseringen av tegnene på et vanlig engelsk skrivemaskintastatur, kalt QWERTY-tastaturet. Dette oppsettet av tastaturet brukes også på skrivemaskiner i Norge.

Grunnen til interessen for en optimalisering av dette tastaturet er utviklingen av et japansk tastatur ved dataregistrering. Det ble antatt at tiden for å skrive to tegn på maskinen var beskrevet ved funksjonen

$$t = F(h_1, h_2, r_1, r_2, f_1, f_2) + E(\text{frek}) + e$$

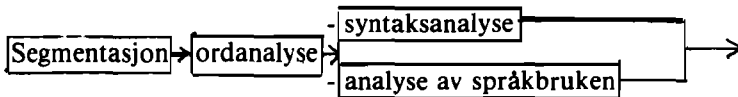
hvor h og f står for henholdsvis hånd og finger mens r indikerer plasseringen av tegnene på tastaturet. E er en korreksjon for hvor ofte de to tegnene opptrer samtidig og e står for faktorer som ikke lar seg beskrive ved hjelp av leddene foran. Ved hjelp av bl.a. multivariat lineær regresjon, ble koeffisientene i funksjonen bestemt. De kunne vise til analyseresultater som

stemte med resultater fra tilsvarende forskning. Det var dessuten foretatt en analyse av feil som oppstår under skriving, og som henger sammen med den fysiske bevegelsen av fingrene. Eks. på slike feil er ulike skrivemåter av det engelske ordet «little», f.eks. litle, liitle, littl, litl. Analysen fortalte at over 60% av feilene skyldes at man under innskriving hopper over karakterer. Konklusjonen var følgelig at tegnene på QWERTY-tastaturet på langt nær har en optimal plassering.

### Lingvistisk basert feilretting av japansk tekst.

I foredraget beskrev *Tsutomu Kawada* et nytt lingvistisk feilkorrigeringsystem. Utgangspunktet for prosjektet var at innlesing av data ved hjelp av optisk leser normalt forutsetter assistanse fra en operatør som retter opp symboler maskinen ikke forstår. Systemet som her ble beskrevet, foretar en lingvistisk analyse av data som kommer ut av den optiske leseren.

Selve feilkorrigeringsystemet kan skisseres som følger



I tilfelle man finner en tvetydig bokstav, vil ordanalyseprogrammet søke i en ordbok for å finne bokstaver som gir et grammatisk og semantisk akseptabelt resultat. Dersom ordanalyseprogrammet gir et utilfredsstillende resultat, blir programmet for syntaksanalyse aktivisert. Dette programmet gjør bruk av informasjonen i konteksten.

Dette feilkorrigeringsystemet klarer å rette over 50% av feilene og tvilstilfeller («rejects») som oppstår ved bruk av utstyr for optisk lesing av data.

### Kommunikasjon menneske - datamaskin i naturlig språk.

Det ble gitt flere foredrag hvor dette emnet ble knyttet til bruk av informasjonssøkesystemer. I foredraget «Embedded sublanguages and natural language processing», la *Richard Kittredge* vekt på at mange systemer for automatisk behandling av tekster i naturlig språk er konstruert for bare å akseptere en begrenset del av språket. Han mente at det ville være mulig å beskrive en relativt presis grammatikk (subgrammatikk) for en slik undermengde av språket, selv om angrepsstrategien i dag er heller uklar. Man har f.eks. ingen stringent maskin-baserte prosedyrer for å bestemme om vilkårlige tekster hører til samme undermengde. Tendensen har vært at man antar at tekster som blir produsert i analoge kommunikasjonssituasjoner er fra samme undermengde.

Om angrepsmåten for forskningen her synes uklar, var formålet klart nok: å skape avanserte systemer for automatisert og «intelligent» teksttolkning.

## **Maskin-genererte thesauri i ulike språk.**

Forfatteren *F.J. Devadason*, Documentation Research and Training Centre, India, gikk detaljert gjennom et system for automatisk generering av en thesaurus. Devadason hevdet at dersom man hadde konstruert en thesaurus på ett språk, kunne den transformeres til et annet språk ved hjelp av en tabell som inneholder ekvivalente termer på det nye språket. Han kom bl.a. inn på begrepet «information retrieval thesaurus» som han gav følgende definisjon: «a controlled dynamic vocabulary of semantically related terms offering comprehensive coverage of a domain of knowledge».

Bruken er spesielt relatert til automatisk lagring og fremhenting av dokumenter. I en søkeprosess bør et spørsmål ved hjelp av en thesaurus kunne utvides, slik at man får en økning i treffsikkerheten («recall»).

Foredragsholderen kom inn på de tradisjonelle måtene for automatisk å lage en thesaurus ved hjelp av statistiske opplysninger over nøkkelordene, beregning av såkalt Tanimoto koeffisienter, konstruksjon av likhetsmatriser og cluster analyse.

Det ble opplyst at til tross for bruk av disse cluster-metoder, hadde man ikke fått noen nevneverdig forbedring av søkesystemets recall. Devadason konkluderte derfor med at en «information retrieval thesaurus» er noe mer enn bare en liste av grupperte nøkkelord.

### **Oppsummering:**

De japanske deltakerne sto for rundt halvparten av de presentasjonene som ble gitt. Den språklige og pedagogiske fremføring var dessverre ikke alltid like bra, men det faglige innholdet i arbeidene var høyt. Japan satser mye på feltet EDB-basert språkanalyse, og dette var en enestående anledning til å få et innblikk i de mange interessante prosjektene som pågår i dag. Den tekniske utstillingen gav også et fascinerende overblikk over den datateknologiske utvikling anno 1980. Den viste med all tydelighet at Japan er i ferd med å komme i fremste rekke som produsent også av data-teknologisk utstyr.

## Statistikk og EDB i arkeologisk forskning

*Jostein H. Hauge*

NAVFs EDB-senter for humanistisk forskning arrangerte i tiden 24. - 26. november et seminar hvor bruken av EDB-baserte statistiske metoder i arkeologiske forskningsprosjekter ble drøftet.

Til seminaret, som ble holdt på Bryggens Museum, Bergen, var det invitert 12 arkeologer fra Oslo, Trondheim og Bergen hvorav de aller fleste for tiden anvender statistiske metoder i sitt forskningsarbeid. I tillegg var det invitert 5 fagstatistikere og EDB-konsulenter som samarbeider med arkeologiske forskere. Seminaret var det første nasjonale tiltaket om dette temaet i vårt land.

Mens databehandling i den første delen av 70-årene hovedsakelig ble brukt til registrering av store gjenstandsmengder, har det i de siste årene vært en klart økende interesse for å ta i bruk EDB og statistiske metoder i andre typer forskningsoppgaver, f.eks. ved analyse av gjenstandsgrupper eller i boplassundersøkelser. Arkeologene i vårt land har imidlertid ikke i dag de ønskelige faglige forutsetninger for å ta i bruk de nye hjelpemidlene, som i mange andre land allerede har åpnet helt nye perspektiver for arkeologisk forskningsarbeid. For eksempel gis det for tiden ikke undervisning i EDB eller statistikk i det arkeologiske fagstudiet ved våre universiteter, utover en viss forsøksvirksomhet ved Universitetet i Oslo.

Seminaret var det første av flere tiltak som NAVFs EDB-senter for humanistisk forskning vil sette i verk for å heve den statistiske kompetanse i de arkeologiske forskningsmiljøene. Ett av disse er at det er opprettet et forskerstipend for en arkeolog som vil spesialisere seg i EDB og statistikk i sitt fag. Ved utløpet av søknadsfristen den 15. februar, hadde det meldt seg flere arkeologer som søkere.

På seminaret ble ulike statistiske metoder presentert av stipendiat *Leiv Breivik* og førstelektor *Erik Bølvigen* fra Universitetet i Tromsø, universitetslektor *Tom Backer Johnsen*, Universitetet i Bergen, førstesekretær *Ole Lauvskar* og førstekonsulent *Sigbjørn Århus*, NAVFs EDB-senter for humanistisk forskning, Bergen.

I alt ble det gitt 9 presentasjoner av pågående arkeologiske prosjekter de tre dagene seminaret varte. Universitetslektor *Reidar Bertelsen*, Universitetet i Tromsø drøftet grunnlaget for bruk av lausfunn i bosetningshistoriske undersøkelser, og forskningsstipendiat *Ericka Helskog*, fra samme universitet, redegjorde for faktoranalyse av et materiale fra steinalderen. Fordeling av gjenstandstyper ved gårdshaugutgravninger var temaet til konservator *Inger Marie Holm-Olsen*, også fra Tromsø. De to foredragsholderne fra Historisk museum, Universitetet i Bergen, førstekonservator *Bente Magnus* og forsk-



En del av deltakerne på arkeologi-seminaret

---

ningsstipendiat *Bergljot Solberg* tok for seg henholdsvis seriasjon av spennformete leirkar og statistisk analyse av et spydmateriale fra yngre jernalder. Utgravningsleder *Erik Schia*, Riksantikvaren, orienterte om Gamlebyprosjektet i Oslo. Vit. ass. *Petter B. Molaug*, som er knyttet til samme prosjekt, drøftet mengdeberegning av et keramikkmateriale. Dosent *Stig Welinder*, Oldsaksamlingen, Universitetet i Oslo, presenterte strukturen i et jernaldergravfelt og orienterte om en programpakke for arkeologi som han har satt sammen. Programpakken er bygd på den generelle statistikkpakken SPSS.

I tilknytning til de enkelte presentasjonene ble det ført en fruktbar diskusjon mellom arkeologer, statistikere og EDB-spesialister om verdien og anvendelsen av ulike statistiske metoder i de enkelte forskningsprosjektene. Seminaret ga en nyttig oversikt over situasjonen i dag og behovene fremover innenfor statistisk orientert arkeologisk forskning hos oss.

Ved slutten av seminaret ble det orientert om det nordiske forskerkurs som NAVFs EDB-senter og en arbeidsgruppe sammensatt av arkeologer og statistikere i Tromsø vil arrangere i juni i år. For arrangøren representerte seminaret en verdifull kontakt med en viktig brukergruppe og ga ideer til planlegging av nye tiltak innenfor feltet statistikk og arkeologi. En konferanse-rapport er under utgivelse.

# **ALLC (Association for Literary and Linguistic Computing) International Meeting & Annual General Meeting**

**København, 11-12 desember 1980.**

*Knut Hofland*

Møtet ble arrangert av en komite med professor Karl Hyldgaard-Jensen som formann. I tillegg til årsmøtesaker var det foredrag innen to emner, maskinoversetting og datamaskinell leksikografi.

*Sergei Perschke* fra EF-kommisjonen orienterte om Kommisjonens bruk av SYSTRAN oversettingsystem og utviklingen av det nye systemet EUROTRA. Han startet med å spørre hvorfor Kommisjonen har engasjert seg i SYSTRAN, et system som ikke er utfordrende hverken fra den datamaskinelle eller den lingvistiske siden. Da Kommisjonen i 1975 laget en oversikt over de eksisterende systemer for maskinoversetting, ble SYSTRAN vurdert som det beste. Et viktig argument var at systemet var i praktisk bruk. Kommisjonen overtok systemet og har videreutviklet det i løpet av 3 år. Systemet kan i dag brukes for 3 språkpar: engelsk-fransk, fransk-engelsk og engelsk-italiensk. SYSTRAN er tatt i bruk innen to områder:

## *1) EURONET.*

EURONET er et datanett mellom ca. 40 maskiner hvor det fins tilgjengelig ca. 100 databaser. De fleste av disse er på engelsk. Bruken av disse databasene vil bli større dersom informasjonen i databasen kan oversettes til brukerens språk. Et problem er at SYSTRAN er et satsvis system mens søking i databasene er interaktiv. Det mangler videre programvare mellom SYSTRAN, søkesystemene og nettet. I løpet av et år vil dette bli utprøvet.

## *2) Intern oversetting i EF.*

EF har verdens største produksjon av oversetting med 1/2 million sider i året. En tredjedel av Kommisjonens budsjett går til oversetting. Ved bruk av SYSTRAN kan en oversetter doble sin produksjon fra 8 til 16 sider pr dag. Det er generelt vanskelig å få tak i kvalifiserte oversettere.

Kommisjonen har satt i gang arbeid med å utvikle et nytt oversettingsystem EUROTRA. Uten at Kommisjonen hadde startet med SYSTRAN og fått konkrete resultater her, ville ikke motivasjonen vært stor nok til å sette i gang EUROTRA. Budsjettet for prosjektet er 15-20 millioner dollar i løpet av en 5 års periode. Medarbeidere kommer fra 9 universiteter i EF. Uansett om dette prosjektet gir et resultat som svarer til forventningene, så har det gitt en kraftig stimulans til utviklingen av fagfeltet ved de universiteter som deltar.

*Gerald C. Keil* fra Manchester universitet holdt et foredrag om programvare for maskinoversetting. Han skilte mellom systemer skrevet i *lavnivåspråk*, som f. eks. SYSTRAN der separasjon av program og underliggende modell er uklar, og systemer skrevet i et *høynivåspråk*, som f.eks. det fremtidige EUROTRA og systemet BEDE ved Manchester Universitet. Han kom også inn på forskjellen mellom et statisk og dynamisk lingvistisk system.

*Bente Maegaard* og *Hanne Ruus* fra København Universitet kom inn på grensesnittstrukturen i EUROTRA. Oversettingen skjer i 3 trinn, analyse og generering som er en-språklig og overføringen som er to-språklig. Grensesnittet mellom analyse og overføring og mellom overføring og generering er et tre der nodene har merkelapper i 4 nivåer: morfo-syntaktisk, syntaktisk funksjon, logisk semantisk relasjon og semantisk relasjonnivå.

*Margaret King* fra Universitetet i Geneve pekte på 3 kjente problemer som viser at det trengs semantikk i forbindelse med oppbygging av EUROTRA, nemlig preposisjonsledd, ordklasse homografi og pronomenreferanser. Det er ikke nok med semantiske markører, det trengs også strukturert informasjon om relasjoner. For visse typer pronomenreferanser er det behov for mer informasjon (world knowledge) enn det er mulig å ta inn i et slikt system.

*Christian Gram* fra Danmarks tekniske høyskole gav en orientering om ADA, et nytt programmeringsspråk som er under utvikling med sterk støtte fra det amerikanske forsvarsdepartement. Språket tilhører Algol-Pascal familien og har også trekk fra SIMULA. Det er lagt vekk på å gjøre språket egnet for strukturert og modulær programmering og å kunne oppnå en bedre standardisering og overførbarhet av programmer.

*W. Martin* holdt et foredrag om hvilke leksikalske opplysninger som kan trekkes ut av en tekst ved hjelp av ulike programmer. Hjelpemidlene som kan brukes er de tradisjonelle lister som indeks, KWIC-konkordans og frekvensliste med eller uten lemmatisering og baklengsordliste. Videre kan en bruke en ordbok over pseudo-synonymer samt en liste over signifikante ordsammenstillinger.

*Maria Bonner* fra Universitetet i Saabruken presenterte et system for å trekke ut ord fra en lemmatisert ordliste over gammelislandske tekster. Forskjellige spsialsorterte lister kunne tas ut etter behov.

*Winfried Lenders* fra Universitetet i Bonn orienterte om et prosjekt for å skaffe oversikt over datamaskinlesbare ordbøker for tysk som fins i Vest-Tyskland. Prosjektet har analysert 12 forskjellige ordbøker. Disse er svært forskjellig i oppbygging, omfang og med hensyn til struktur, klassifikasjon og underliggende lingvistisk teori. Målet til prosjektet er å finne ut om det er mulig å lage en integrert ordbok som kan brukes i forskjellige systemer for språklig databehandling. Et eksempel er EUROTRA.

*Arne Zettersten* gav en oversikt over et prosjekt for å lage en ordbok over engelsk uttale fra år 1500 til 1800. Dette er et samarbeidsprosjekt mellom 7 land. Publiserte ordbøker som omhandler tidsperioden er utgangspunktet for prosjektet. Materialet overføres til et fast skjema. Noe av materialet er gjort maskinlesbart ved Stanford Universitetet. Det er planer om å ta i bruk EDB til fotosetting og produksjon av indekser ol.

På årsmøtet i ALLC ble det gitt rapporter fra formann, sekretær, kasserer og redaktører i ALLC Bulletin og Journal. Rapporter fra formennene i de enkelte spesialistgrupper var oppslått utenfor auditoriet. En ønsket flere annonsører til ALLC-publikasjonene for å styrke disse. Videre ble det vedtatt å holde neste årsmøte i Oxford i desember 1981. For året etter flyttes årsmøtet til Symposiet i Pisa i juni 1982 slik at årsmøtet holdes annenhvert år sammen med ALLC Symposiet. Et nytt årsmøte og internasjonalt møte holdes i San Fransisco i juni 1983.



# MELDINGER

## NAVF og Datatilsynet

NAVF har i den senere tid vært opptatt av å finne fram til ordninger som kan gi forskerne den best mulige adgang til data innenfor de regler som Lov om personregistre m.m av 9. juni 1978 og forskriftene til loven setter.

Det har særlig vært aktuelt for NAVF å vurdere

*NAVFs rolle som kontaktledd mellom forskningsinstitusjoner/eventuelt forskere og Datatilsynet når det gjelder konsesjonssøknader, og*

*etablering av ordninger for oppbevaring av data for eventuell gjenbruk.*

Rådet har som ledd i sitt arbeid utredet de vilkår som forskningssektoren må oppfylle ifølge loven, og hvordan de retningslinjer for konsesjonsbehandling som skal etableres, vil virke inn på Rådets egne rutiner, f.eks. ved behandling av søknader om forskningsstøtte.

Dessuten er det blitt utarbeidet konkrete forslag til kontaktformer med Datatilsynet og forskerne i personvernsaker.

NAVFs styre har som en prøveordning vedtatt å opprette et saksbehandlende organ innenfor NAVF, som i særlig grad skal vurdere de søknader til NAVF som medfører bruk av persondata og hvor konsesjonsbehandling er nødvendig.

Organet skal også tjene som et datafaglig kontaktorgan overfor Datatilsynet, og skal være rådgiver for de enkelte fagråd i NAVF og for forskerne i spørsmål i tilknytning til konsesjonsbehandling.

De fagetsiske spørsmål som bruk av persondata medfører, forutsettes å bli behandlet som en integrert del av saksforberedelsen ved den ordinære søknadsbehandling i NAVFs fagråd.

Organet for behandling av de datafaglige spørsmål er lagt til *Norsk samfunnsvitenskapelig datatjeneste* i Bergen, hvor det for tiden er tilsatt en konsulent og en kontormedarbeider.

Det datafaglige organet vil være behjelpelig med saksforberedelse og samordning av enkeltsaker innenfor arbeidsfeltet. Blant de første oppgaver som må tas opp er dessuten å avklare med Datatilsynet muligheten av og vilkårene for å oppnå generelle konsesjoner for forskningsinstitusjoner.

## **Lov om personregistre m.m. Orientering til forskningsmiljøene**

Det minnes om at personregistre som er opprettet for forskningsformål før 1. januar 1980 må søke Datatilsynet om konsesjon innen 31. mars 1981, dersom registrene fortsatt skal brukes eller oppbevares etter denne dato.

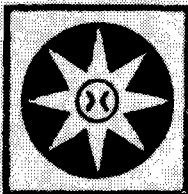
Dette gjelder alle slike personregistre som:

1. gjør bruk av EDB, eller
2. inneholder opplysninger om
  - rase eller politisk eller religiøs oppfatning
  - at en person har vært mistenkt, tiltatt eller dømt i straffesak
  - helseforhold eller misbruk av rusmidler
  - seksuelle forhold, eller
  - opplysninger om familieforhold som ikke gjelder sektskap eller familiestatus, formuesordningen mellom ektefeller og forsørgelsesbyrde.

Søknader skal sendes på eget skjema til

**Datatilsynet**  
**Postboks 7643, Skillebekk**  
**Oslo 2.**

Nærmere opplysninger fås ved henvendelse til Datatilsynet, tlf. (02) 44 70 22, eller til Norsk Samfunnsvitenskapelig Datatjeneste, tlf. (05) 21 00 40.



**NAVF**

**NORGES ALMENVITENSKAPELIGE  
FORSKNINGSRÅD**

I samarbeid med andre organ innenfor NAVF, vil det datafaglige sekretariatet også arbeide for tilfredstillende ordninger når det gjelder lagring og gjenbruk av persondata for forskningsformål. Målsettingen fra NAVFs side er å oppnå avtaler som gir mulighet for lagring av prosjektdata i den opprinnelige form og å finne fram til retningslinjer og praktiske ordninger for gjenbruken. Ved vurdering av ny bruk av data og utlevering fra de institusjoner som har fått konsesjon til å lagre persondata fra avsluttede forskningsprosjekter, bør det gjennom en ny konsesjonsbehandling vurderes hva som bør slettes av personidentifiserende opplysninger.

Det er av stor verdi for forskersamfunnet, både praktisk og økonomisk, at ferdigbehandlet materiale blir betryggende lagret med tanke på eventuell senere bruk. Slik NAVF ser det, er det naturlig at Norsk samfunnsvitenskapelig datatjeneste får ansvar for sikring og arkivering av data fra prosjekter under Rådet for forskning for samfunnsplanlegging og Rådet for samfunnsvitenskapelig forskning. NAVFs EDB-senter for humanistisk forskning får tilsvarende oppgaver knyttet til prosjekter som faller inn under fagområdet til Rådet for humanistisk forskning. NAVFs EDB-tjeneste for medisinsk forskning har etter dette mønsteret ansvar for data knyttet til prosjekter under Rådet for medisinsk forskning.

Vi vil til slutt minne om at de personregistre som er opprettet for forskningsformål før 1. januar 1980 må meldes til Datatilsynet innen 31. mars 1981. Dette gjelder de registre som fortsatt brukes eller oppbevares. Det er utarbeidet et eget søknadsskjema med veiledning som skal brukes ved søknad om konsesjon.

Se for øvrig egen melding fra NAVF.

---

## **Internasjonale konferanser**

### **University of Michigan**

I mai 1981 blir det ved *University of Michigan* arrangert 3 større EDB-konferanser med tilknytning til de humanistiske fag.

*The Fifth International Conference on Computers and the Humanities* , 17. - 20. mai.

*The Third International Conference on Data Bases in the Humanities and Social Sciences* , 20. - 23. mai.

*The Conference on Systems and Their Users* , 20. - 22. mai.

## **Universite de Liege.**

Laboratoire d'Analyse Statistique des Langues Anciennes feirer sitt 20 års jubileum med en internasjonal konferanse kalt *Automatic Data Processing in the Humanities*, Konferansen holdes i tiden 18. - 21. november. Følgende tema oppgis:

Archives

History-Demography

Philology and Linguistics: ancient and modern languages

Automatic and automatized translation

Philosophy

Economics

Social Sciences

Law Data processing and retrieval

## **Centre National de Reserche Scientifique, Paris**

Det andre internasjonale symposium om bruk av datamaskin i musikk-forskning blir holdt i Paris i begynnelsen av juli 81.

Senteret kan gi interesserte nærmere opplysninger om konferansene ovenfor.

## **European Science Foundation.**

Standing Committee for the Humanities har invitert sine medlemsorganisasjoner til å sende to representanter hver til «*Workshop on the possibilities and limits of the computer in producing and publishing dictionaries* ». Konferansen holdes i dagene 20. - 22. mai i Pisa. NAVF er norsk medlem i ESF.

---

## **Senterrapport nr. 17**

**Automatisk syntaktisk analyse.**

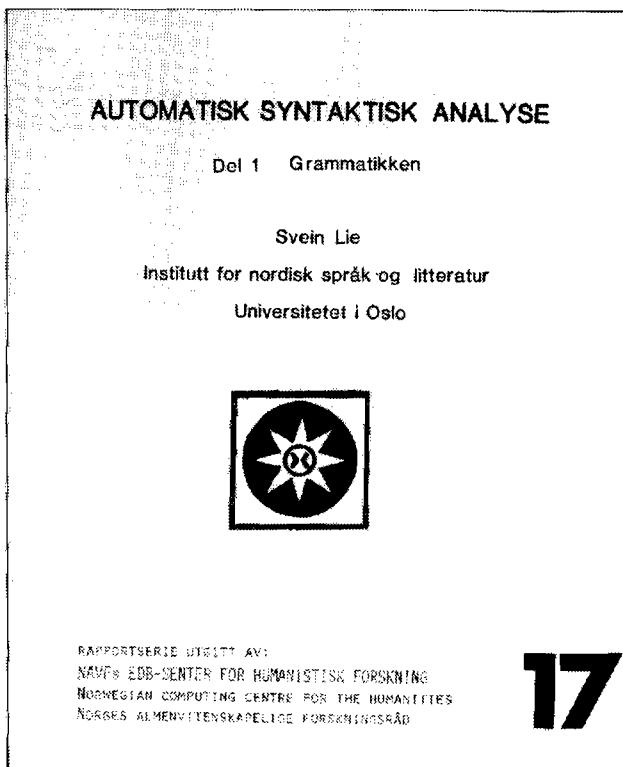
*Del 1: Grammatikken.*

*Svein Lie, Universitetet i Oslo*

Rapporten er basert på et samarbeidsprosjekt mellom *Knut Hofland*, NAVFs EDB-senter for humanistisk forskning, Bergen og *Svein Lie*, Universitetet i Oslo. Knut Hofland har hovedansvaret for programmeringen og Svein Lie for grammatikken. Prosjektet er finansiert av Norges almenvitenskapelige forskningsråd.

Det system som beskrives for syntaktisk analyse av norske setninger ved hjelp av datamaskin, bygger på en modell av prof. Martin Kay. Modellen ble presentert på forskerkurs i København i 1974 og Pisa i 1974 og 1977. Systemet består av tre deler: 1. Ordliste. 2. Grammatikk. 3. Program.

Fremstillingen gir en oversikt over hvordan de enkelte delene i systemet er bygd opp. Det har vist seg at de praktiske løsningene en har funnet på de problemer som oppstår i datamaskinell syntaktisk analyse, har teoretiske implikasjoner for teorier innen allmenn lingvistikk og psykolingvistikk. Dette blir også diskutert. I denne rapporten blir det lagt mest vekt på grammatikken. Programmet og programmeringen vil bli behandlet av Knut Hofland i del 2. Den utgis som egen rapport senere.



## **Senterrapport nr. 18**

### **Datateknologi og humanistisk forskning.**

#### *Bidrag til en NAVF utredning.*

Det er nylig utgitt en rapport om Datateknologi og humanistisk forskning i Senterets rapportserie. Rapporten er et bidrag til NAVFs utredning om datateknologiens konsekvenser for forskning og samfunnsliv. Den er i hovedsak forfattet av Senterets leder, *Jostein H. Hauge* (se Humanistiske data 1-80)

I rapporten blir det bl.a. gitt en beskrivelse av de sider ved den generelle datateknologiske utvikling som har særlig interesse for de humanistiske fag, og anvendelsene av EDB i de ulike fag blir gjennomgått.

I egne kapitler blir de samfunnsmessige konsekvenser av moderne datateknologi tatt opp, i første rekke med tanke på å klarlegge de sosiale og kulturelle virkninger som bør studeres av forskere i de humanistiske fag.

Avslutningsvis blir det gitt en presentasjon av ulike tiltak som Rådet for humanistisk forskning kan sette i verk for å møte de utfordringer som den datateknologiske utvikling vil gi i årene som kommer.

Vi håper at enkelte av våre lesere vil finne anledning til å kommentere eller utdype temaer som tas opp i rapporten, og vi tar gjerne imot bidrag for publisering. Det vises for øvrig til den artikkelen i dette nummer som er hentet fra rapporten.

---

## **Senterrapport nr. 20**

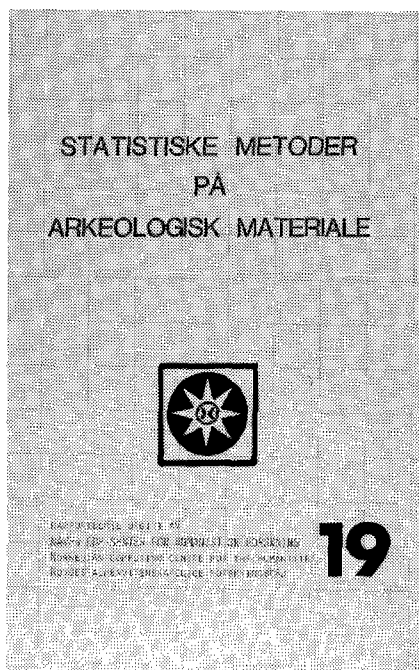
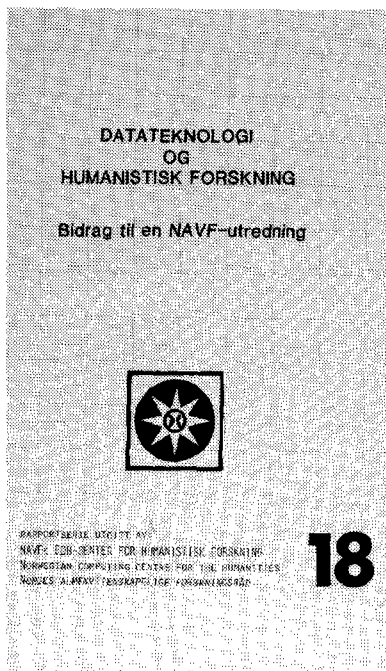
### **EDB-prosjekter i de humanistiske fag.**

Som det ble redegjort for i Humanistiske data 1-80, har Senteret kartlagt EDB-bruken i de humanistiske fag gjennom en spørreskjemaundersøkelse ved årsskiftet 1979 - 80.

Gjennom direkte henvendelse til prosjektlederne høsten 80 har vi sikret oss deres samtykke til å offentliggjøre resultatene i en egen prosjektrapport og samtidig fått opplysningene om prosjektene ajourført.

Rapporten om de ca. 80 prosjektene er for tiden under trykking og vil utkomme som nummer 20 i Senterets rapportserie. Rapporten selges til kostpris. De som har levert bidrag, vil få rapporten gratis tilsendt.

Samtidig vil vi benytte anledningen til å nevne at Rådet for humanistisk forskning i samarbeid med NAVFs Utredningsinstitutt og Senteret vil utgi en prosjektkatalog i løpet av våren som dekker all humanistisk forskning i Norge i dag. Prosjektkatalogen er en ajourføring av *Humanistisk forskning 1977*. Den vil på flere punkter kunne utdype det bilde av EDB-virksomheten i de humanistiske fag som Senterets prosjektoversikt gir.



## Senterets rapportserie

RAPPORT nr. 1. *EDB i gjenstandsfagene*. Rapport fra en konferanse i Bergen, 18. og 19. april 1978. September 1978. Pris kr. 15,

RAPPORT nr. 2. *Et norsk datamaskinelt tekstkorpus*. Rapport fra en konferanse i Bergen, 19. og 20. oktober 1978. Febr. 1979. Pris kr. 20,

RAPPORT nr. 3. *Rapport fra den nasjonale konferanse om EDB i språk og litteraturforskning*, 4. og 5. januar 1979. Mars 1979. Pris kr. 25,

RAPPORT nr. 4. *Oppbygging av EDB-katalog for folkemusea i Hordaland og Kulturgeografisk registrering på Vestlandet*. April 1978. 2. opplag oktober 1979. ISBN 82-7283-000-0. Pris kr. 15,-.

RAPPORT nr. 5. *Rapport fra NKKM's EDB-komite*. August 1979. ISBN 82-7283-001-9. Pris kr. 15,-.

RAPPORT nr. 6. *Prøveprosjekt med EDB ved Norsk Folkemuseum*. Oktober 1979. ISBN 82-7283-002-7. Pris kr. 15,-.

RAPPORT nr. 7. *Ivar Fønnes: Norsk landbruksordbok. Prosjektrapport om databehandling og tilrettelegging for trykking*. September 1979. ISBN 82-7283-008-6. Pris kr. 25,-.

RAPPORT nr. 8. *SEFRAK. Rapport frå prøveprosjekt for databehandling av kulturminneregisteret*. Oktober 1979. ISBN 82-7283-003-5. Pris kr. 20,-.

RAPPORT nr. 9. *Jostein H. Hauge og Sigbjørn Århus: Dataregistrering i humanistiske fag med vekt på optisk lesing*. August 1978. 2. opplag oktober 1979. ISBN 82-7283-004-3. Pris kr. 15,-.

RAPPORT nr. 10. *Roald Skarsten: Innføring i SPSS for humanister*. November 1977. 2. opplag november 1979. ISBN 82-7283-005-1. Pris kr. 15,-.

RAPPORT nr. 11. *Jostein H. Hauge og Knut Hofland: Rapport fra 4 konferanser i USA sommeren 1979*. The 17th Annual Meeting of Computational Linguistics. La Jolla Conference on Cognitive Science. The fourth International Conference on Computers in the Humanities. Data Bases in the Humanities and Social Sciences. November 1979. ISBN 82-7283-007-8.

RAPPORT nr. 12. *EDB og manuskriptregistraturer*. Oktober 1977. 2. opplag november 1979. ISBN 82-7283-009-4. Pris kr. 20,-.

Rapport nr. 13. *Datatjenester for og datasamarbeid mellom kunst- og kulturhistoriske museer*. Februar 1980. ISBN 82-7283-010-8. Pris kr. 30,-.

RAPPORT nr. 14. *NOVA\*STATUS, systemdokumentasjon. Brukerveiledning*. 2. opplag februar 1980. ISBN 82-7283-011-6. Pris kr. 15,-.

RAPPORT nr. 15. *Ivar Fønnes: Tekstsøking på tegnnivå*. Januar 1980. ISBN 82-7283-012-4. Pris kr. 15,-.

RAPPORT nr. 16. *Årsmelding 1979, NAVFs EDB-senter for humanistisk forskning*.

RAPPORT nr. 17. *Svein Lie: Automatisk syntaktisk analyse. Del 1. Grammatikken*. Desember 1980. ISBN 82-7283-014-0. Pris kr. 30,-.



RAPPORT nr. 18. *Datateknologi og humanistisk forskning. Bidrag til en NAVF-utredning.* Desember 1980. ISBN 82-7283-015-9. Pris kr. 30,-.

RAPPORT nr. 19 *Statistikk og EDB i arkeologisk forskning* og RAPPORT nr. 20 *EDB-prosjekter i de humanistiske fag* er under utarbeidelse.

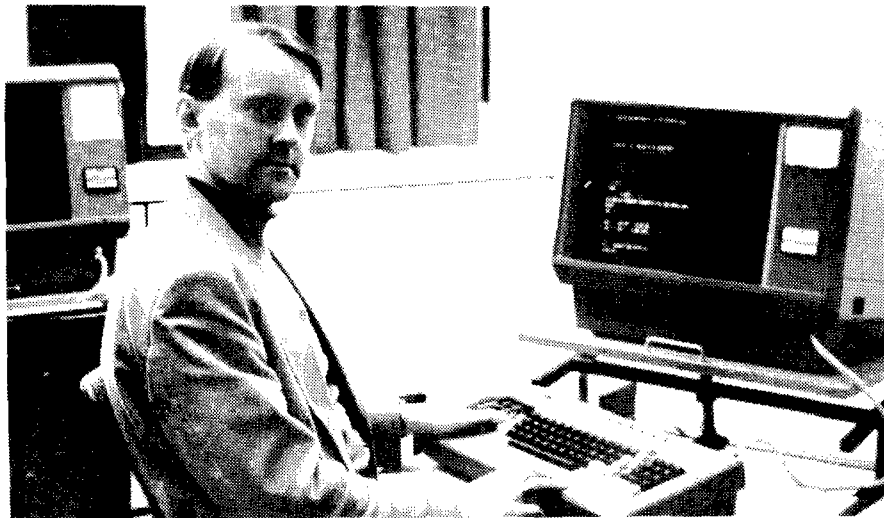
---

## **Konsulenttjeneste for de humanistiske fag ved Universitetet i Tromsø.**

Rådet for humanistisk forskning har fra 1.1.81 opprettet et 3-årig EDB-stipend ved Universitetet i Tromsø. Stipendiaten skal i tillegg til egen forskningsvirksomhet også gi konsulenttjenester til humanistiske forskere ved Universitetet. Stipendiatorordningen er administrativt knyttet til NAVFs EDB-senter for humanistisk forskning i Bergen, og stipendiaten skal også fungere som et bindeledd mellom Senteret og Universitetet i Tromsø.

Cand.philol. *Holger Hagan* ble fra årsskiftet tildelt stipendet. Hagan har engelsk språkvitenskap som forskningsområde, og han vil i det første halvåret særlig sette seg inn i bruken av EDB-anlegget ved Universitetet i Tromsø og den programutrustning som finnes for databehandling i de humanistiske fag.

Hagan har kontor i Brenn-bygget 1. etasje, rom 135 (tlf. 86560, linje 223) og i Breivika, EDB-senteret, underetasjen rom A042 (tlf. 81688, linje 247). Hagan vil gjerne ha kontakt med aktive og potensielle EDB-brukere i de humanistiske fag i Tromsø.



Holer Hagan, EDB-stipendiat ved Universitetet i Tromsø.

## **Prosjektsamarbeid som betalte oppdrag.**

Som en prøveordning har Rådet for humanistisk forskning fra 1.1.81 opprettet en konsulentstilling i Senteret som skal finansieres ved de betalte oppdrag som Senteret påtar seg. Denne nyordningen er kommet i stand fordi Senteret i de siste årene har registrert en stadig økende interesse i humanistiske fag- og forskningsmiljøer for å få utført EDB-virksomhet som betalte oppdrag. Ikke minst har det vist seg aktuelt å gi slik støtte til museer i flere landsdeler.

Senteret vil i løpet av 1. halvår sende ut en separat redegjørelse til de fagmiljøer som ventes å ha spesiell interesse av ordningen. I orienteringen vil det bli redegjort for retningslinjene for oppdragsarbeid og betalingsvilkårene. De som er interessert i å få nærmere opplysninger om denne samarbeidsordningen, kan også ta direkte kontakt med Senteret.

---

## **Hovedpunkter i Senterets arbeid i 1981.**

### *1. Nyorientering.*

Senterets arbeid i år vil på flere måter avvike fra virksomheten i de foregående år. Mens den faste staben tidligere hadde et utstrakt prosjektsamarbeid med ulike fag- og forskningsmiljøer og med enkeltforskere, vil arbeidet i år være mer konsentrert om metode og programutvikling og opplæringstiltak, herunder forskerutdanning.

Vektleggingen på disse felter betyr på den annen side ikke at prosjektarbeidet har stoppet opp. Det vil alltid, også for Senterets egen virksomhet, være viktig at det knyttes prosjektarbeid til Senteret, men det nye er at en hel del av prosjektsamarbeidet heretter vil foregå som betalte oppdrag. Det er ansatt en egen EDB-konsulent til å arbeide med slike oppgaver. Derved frigjøres de tre øvrige EDB-konsulentene til mer konsentrert innsats på andre områder.

### *2. Metode- og programutvikling.*

Av innsatsområder kan nevnes utvikling av en programpakke for arkeologi med vekt på statistiske analyseprogrammer og utarbeidelse av generelle programmer for katalog- og arkivdata.

I forlengelse av det pågående arbeid med EDB-systemer for prosjektinformasjon, vil det bli utarbeidet forslag til permanente EDB-løsninger for denne informasjonstype. Systemet vil særlig rette seg mot NAVFs behov og mot en koordinering med søknadsbehandlingen i Rådet.

Senteret anskaffer våren 81 en mikrodatamaskin. Utstyret vil bli tatt i bruk på en rekke felt, som også forutsetter praktisk prøvevirksomhet og utviklingsarbeid. Stikkordsmessig kan vi nevne:

Bruk i forbindelse med kontor- og informasjonsvirksomhet. Registrering og behandling av vitenskapelige data, herunder arkivdata.

Undervisningsopplegg i forbindelse med egne opplæringstiltak.

Utpøving av norsk og utenlandsk programvare av interesse for de humanistiske fag.

Av andre områder kan nevnes grafisk databehandling. Dette er et lite benyttet anvendelsesområde innenfor de humanistiske fag. Teknikken åpner imidlertid for helt nye måter å visualisere forskningsdata på.

Som betalt utviklingsarbeid vil Senteret fortsette samarbeidet med bl.a. R-direktoratet om utvikling av informasjonssøkesystemet SIFT (Søking I Fri Tekst). En første versjon ventes ferdig i løpet av året.

### *3. Humanistisk prosjektarbeid.*

Det største enkeltprosjekt i år vil være Ibsen-prosjektet, som avsluttes på driftssiden 1. juli, med 1. oktober som slutt punkt for dokumentasjon og rapportering.

I prosjektet, som ledes av prof. *Harald Noreng* og EDB-konsulent *Knut Hofland*, vil det bli utarbeidet konkordanser til alle Ibsens skuespill og dikt med bl.a. grammatisk ordkarakteristikk og individuell kontekstavgrensning. Konkordansene vil bli utgitt på mikrokort og magnetbånd, og det er planlagt et oversiktsbind i bokform. Det er aktuelt å utnytte grunnlagsmaterialet i en rekke prosjektsammenhenger senere.

Senteret vil i år fortsette samarbeidet med University of Lancaster og Britisk institutt, Universitetet i Oslo om en videregående grammatisk behandling av The Lancaster-Oslo/Bergen Corpus. Senterets virksomhet vil i år særlig bli knyttet til metoder for automatisk grammatisk bestemmelse av ordmaterialet.

Flere arkeologiske samarbeidsprosjekter (bl.a. EDB-opplegget for Gamleby-prosjektet) vil bli avsluttet i løpet av første halvår.

### *4. Opplæringsvirksomhet.*

Ulike typer opplæring og forskerutdanning vil stå sentralt i arbeidet i år.

#### a. Hospitantprogram

Senteret gjennomfører hospitantopplegg for forskere og for driftsmedarbeidere i humanistiske EDB-prosjekter. Et hospitantprogram rettet mot museums- og arkivsektoren blir arrangert i april.

#### b. Korttidsstipend i EDB

Det er aktuelt å gjennomføre 4-5 stipendprogram til høsten, enten som et videreutdanningstiltak for humanistiske forskere som er fortrolige med EDB, eller som et program rettet mot en bestemt fag- eller institusjonssektor.

#### c. Programmeringskurs.

I løpet av vinteren vil det bli sendt ut innbydelse til et toukers intensivt programmeringskurs ved Senteret. Kurset tenkes holdt i månedsskiftet august/-september.

#### d. Samarbeid med forskningsstipendiater.

I 1981 vil Senteret ha et utstrakt samarbeid med forskningsstipendiater, ikke minst de stipendiater som inngår i Senterets eget forskningsprogram (språklig databehandling og arkeologi). Også andre stipendiater vil ha permanent eller midlertidig arbeidsplass i Senteret (nomenklaturstudier, teatervitenskap).

#### 5. Kurs og konferansevirksomhet.

Følgende tiltak er planlagt til nå :

##### a. EDB-faglig seminar for humaniora.

Seminaret arrangeres i samarbeid med EDB-tjenestene ved universitetene i tiden 18. - 20. mars. Det er beregnet på EDB-konsulenter i de humanistiske fag og fagpersonale med EDB som arbeidsfelt og omfatter ca. 20 inviterte deltakere. Programmeringsspråk og bruk av mikrodatamaskiner er de viktigste temaer.

##### b. EDB-seminar for filosofi.

Seminaret, som vil ta opp bruk av EDB ved studium av filosofiske forfatterskap, tenkes arrangert i september.

##### c. EDB-seminar for teatervitenskap

Senteret planlegger et seminar høsten 81 for fagmedarbeidere i teatervitenskap og profesjonelle teatermedarbeidere om bruk av databehandling ved studier i norsk teater i samtid og fortid.

d. Anglo-Scandinavian Seminar on «The Use of Computer Corpora in Research and Teaching».

Seminarer arrangeres i samarbeid med British Council for ca. 15 aktive forskere fra Norge, Sverige, Danmark og England.

e. Nasjonal EDB-konferanse.

Senteret planlegger en større EDB-konferanse høsten 81 og overveier å rette arrangementet denne gang mot fagmiljøer utenfor universitetene.

f. Nordisk forskerkurs i multivariate metoder i arkeologi.

Kurset arrangeres i Tromsø 17. - 26. juni i samarbeid med en gruppe arkeologer og statistikere ved Universitetet i Tromsø. Målet med kurset er bl.a. å gi praktisk øving i statistisk analyse på arkeologisk materiale.

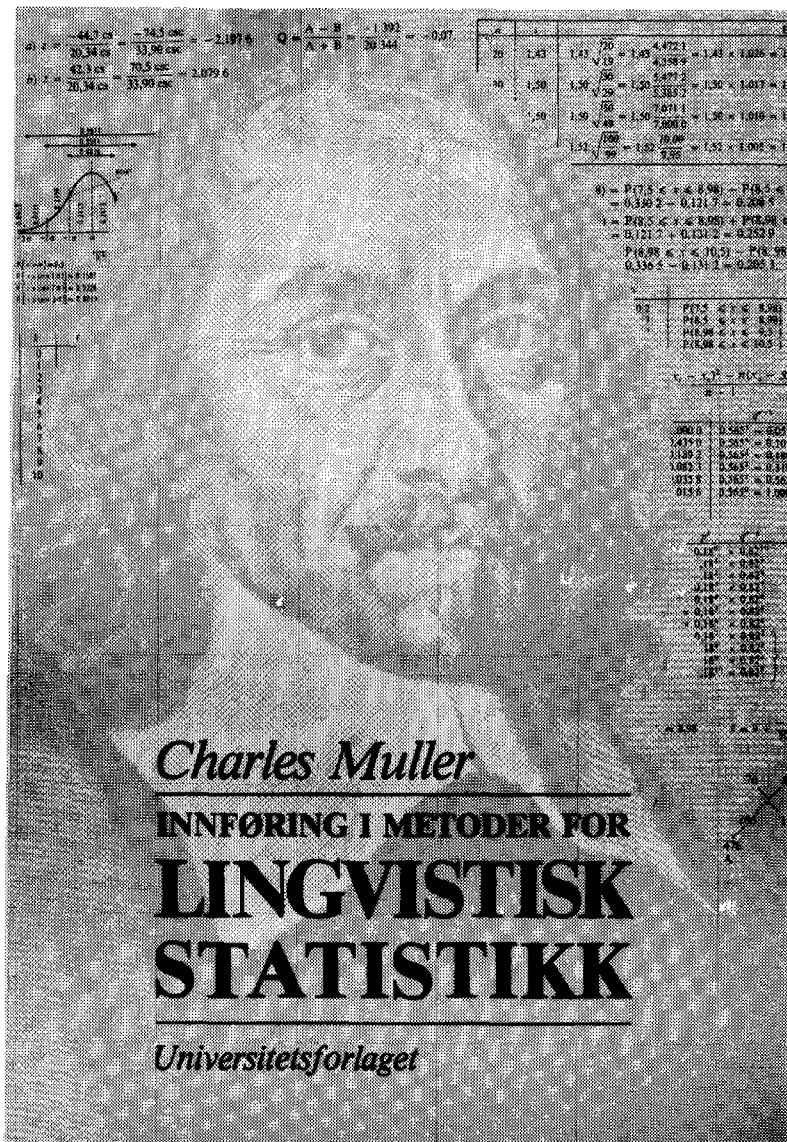
---

## **Forskerstipend knyttet til Senteret.**

Fra 1.8.81 er det opprettet to forskerstipend som faglig og administrativt er knyttet til Senteret. Etter tilråding fra Senteret har Rådet for humanistisk forskning vedtatt å knytte stipendene til språklig databehandling og EDB-metoder i arkeologi. Stipendene har vært utlyst med søknadsfrist 15.2. Det kom inn en rekke søknader som nå er til behandling i NAVF.

Målet med stipendene er å gi kvalifiserte fagfolk mulighet til å fordype seg på utvalgte, sentrale felt innenfor humanistisk databehandling. Stipendene vil bli knyttet til Senteret gjennom fast arbeidsplass der eller ved en særlig tilrettelagt samarbeidsordning dersom stipendiaten også må arbeide utenfor Bergen.

Konsulentene ved Senteret vil samarbeide med stipendiatene og gi dem EDB-faglig støtte ved utvikling av metoder og programutrustning.



Denne boka er nettopp utkommet på Universitetsforlaget. Den er tilrettelagt ved Senteret i samarbeid med EDB-tjenestene ved Universitetet i Oslo, Bergen og Trondheim. Boka er oversatt av cand. philol Kari Fønnes.

# Programmeringskurs for humanister

## 24. august — 4. september 1981

NAVF's EDB-senter for humanistisk forskning vil i tiden 24. august — 4. september 1981 arrangere et intensivkurs i programmering i Bergen (10 deltagere).

Programmeringsspråket blir PASCAL.

Kursleder og hovedlærer blir førstekonsulent Knut Hofland fra Senteret.

Kurset retter seg i første rekke mot forskere og fagmedarbeidere i humanistiske fag. Viderekomne studenter kan også søke.

Det vil bli forelesninger og utstrakt øvingsarbeid med eksempler fra humanistiske problemstillinger.

Maskinen som vil bli brukt, er UNIVAC 1100/82 ved Universitetet i Bergen. Det vil også bli mulighet til å kjøre på Senterets mikrodatamaskin (ALTOS) og på universitetsanleggene i Oslo, Trondheim og Tromsø via UNINETT.

Det er en fordel om deltakerne har elementær kjennskap til databehandling, styrespråk og teksteditor.

NAVF's EDB-senter kan være hjelpelig med å ordne innkvartering, kr. 2.250,— for hele kursperioden (full pensjon). I tillegg kommer seminaravgift på kr. 250,—. De som ønsker å ordne innkvarteringen privat, må gjøre oppmerksom på det i søknaden.

Senteret kan i spesielle tilfelle dekke *oppholdsutgiftene* for deltakere som ikke på annen måte kan skaffe dekning for kursutgiftene.

*Søknadsfrist 1. juni.*

*Søknad sendes NAVFs EDB-senter for humanistisk forskning, Harald Hårfagresgt. 31, Boks 5014 Bergen-U.*

*Tlf. (05) 21 00 40.*

# SUMMARY

## **EDB som hjelpemiddel ved oversettelse (Computer-aided translation)**

Gulbrand Alhaug, lecturer at the Institute of Language and Literature, University of Tromsø, describes in his article a pilot project within the field of computer-aided translation.

In Norway there are two official language variants («bokmål» and «nynorsk») that differ mostly in respect of morphology. When producing Norwegian school books publishers have a legal responsibility of providing schools with texts in both language variants simultaneously. This means that all books have to be translated into one of the variants, until now by ordinary translation. The author has, however, explored ways of automatizing some aspects of this expensive and time-consuming process by using computers. In the first place a method of dictionary look-up of all the words in running texts is applied. Secondly a series of syntagmatic and morphological rules is used to rearrange the word order and substitute phrases special for one language variant.

In his approach to these problems, the author has adopted a purely pragmatic point of view aiming only at automatizing those stages of the translation process that lend themselves easily to computer processing. Preliminary test results show that approx. 90% of the word forms of running texts can be successfully translated from *bokmål* to *nynorsk* by fairly simple computational techniques.

The author claims that in the future the use of computer-based dictionaries, terminology data-bases and methods of computational linguistics will supplement the more conventional tools in the daily work of a majority of professional translators.

## **Musikus - en datamaskinbasert analyse**

### **(Musikus - a computer based system for the analysis of music)**

Tor Sverre Lande, lecturer at the Institute of Informatics, University of Oslo, explains in his article the main features of the project Musikus. The project, started some years ago, demonstrates clearly the positive effects of coope-



ration between humanistic disciplines and the Natural Sciences. The aims of the project have been to develop basic methods and programs for music analysis, including equipment for direct registration of performed music via an organ, and also the automatic production of output in the form of notes. At present work is concentrated on the automatic computer analysis of acoustically transmitted music with a view to the specific research tasks presented by Norwegian folk music played on a fiddle.

### **EDB-prosjekter i Statsarkivet i Bergen**

#### **(Computer Projects of the State Archive of Bergen)**

In his presentation, Egil Øvrebø, Archive Director, lists a series of computer projects carried out during the last few years by the Archive in cooperation with the Institute of History, University of Bergen. The source materials of the projects comprise emigrant registers from the port authorities of Bergen from 1874 till 1924 (covering some 100,000 persons), censuses of the town of Bergen (approximately 35,000 citizens), clerical records of Bergen (registration newly started) and a variety of official records of bequeathed estate from various parts of Western Norway (approximately 60,000 persons). The results of the data processing in the initial stages include a series of sorted lists and indices also linking materials of different origins.

These tools have considerably facilitated basic functions of the archive staff. For one thing the microcard products have thrown open vast and hitherto rather intractable territory for scholarly research and public use. They have also facilitated communication with other archives and other interested parties, for example Local History Societies.

Internally the projects have given better tools for handling inquiries from the public and have also enabled the staff to spare the delicate original source materials which are often in a deplorable physical state. Spurred by the obvious benefits of the computer projects, the Archive wants now to intensify its efforts in the preparing of other materials for the new media. However, lack of financial and computing resources at present seriously limit these efforts.

#### **Current Work on the LOB Corpus**

In his report (in English) Dr. Johansson, Department of English, University of Oslo, sketches how work on the Lancaster/Oslo-Bergen Corpus will proceed in the years to come. Based on the corpus completed in Norway, a project of grammatical analysis is taking shape as a cooperative project financed by the Social Sciences Research Council in England and the Norwegian Research Council for Science and the Humanities. Evidence is

given of the wide interest by researchers and the public for the LOB Corpus project. The report is supplemented by a bibliography.

## **Forskningsarbeid og den tekniske utvikling**

### **(The impact of modern data technology on research activities)**

This article is extracted from a report given by the director of the Centre, Jostein H. Hauge, to the Norwegian Research Council for Science and the Humanities. The aim of the report is to review the present use of computers in the humanities, especially in Norway, and to evaluate possible future uses of computer technology. Some social and cultural consequences of modern computer-based technologies are also discussed, with a special emphasis on aspects that call for the attention of researchers in the humanities.

In the extract presented in the newsletter the author contends that computational technologies will have overall effects on the research community. Many of the effects will be positive, but the computerization of research will also bring in its wake the risk of a technological gap. Moreover, attention is drawn to the possible side-effects of the use of computers, for example the concentration of research projects in the humanities around easily accessible computer-based research material.

The impacts of distributed data bases is also touched upon and the possibilities for researchers to actively explore and re-assess the data underlying publicized research results.

The bulk of the article is concentrated on discussing whether the research communities in the humanities should try and stem the computer wave approaching the humanities, or rather try and ride on it, making the most of the research-expanding powers it offers.

The author claims that since computers ostensibly have come to stay, we must prepare ourselves for the use of the new tools. This does not mean, however, that we should change the research goals of the humanities in a direction more amenable to the methods underlying the new techniques.

The methods of research in the humanities comprise a rich variety of analyses and points of attack. Among these are the use of intuition, interpretation and even empathy. If the use of computers means that we could harm the nerve of life of research in the humanities then we have to be on our guard, but there is no way to encapsulate the humanities in the technological world of tomorrow just out of fear of such a possibility.

## **Mikroprosessoren, arbeidet og mennesket**

### **(Microprocessors, the human being and work)**

In his article, Professor Arnfinn Stigen, Institute of Philosophy, University of Oslo, re-examines a series of beliefs in our culture, notably the notions of work, leisure and happiness. An avalanche of modern books and expert reports try to paint the picture of tomorrow. Most of the pictures of the future describe how the conventional work markets will collapse, creating both new patterns of the producer-consumer relationship and new types of jobs. Also the computerization of society will allegedly boost the rates of unemployment in traditional fields of work.

Professor Stigen presents the following questions: How did it come about that work became a positively valued notion? Is it necessary to work if we can lead a rich and valuable life without working in the normal sense of the word? To Professor Stigen it is obvious that in a majority of cases it is not work in itself that is highly valued, but rather the accessories or values attached to one's work. Among these values he mentions

security,

trust,

friendship,

esteem,

possibility of self-realization and

feeling of creativeness.

If we could create a system of society where these values could be obtained in ways other than through traditional work activities, we would be even better off than today. To create such a society we have to re-define the notion of work and extend it to all kinds of activities of value to the individual and society. This would mean that even non-professional cultural and social activities would be given the status of work and would be paid for according to the values they represent. According to Professor Stigen, to practise a system where such a notion of work prevailed would not perhaps be more difficult than to keep things going as they are today. However, to bring about a new state of affairs we have to change radically our attitudes to a series of cherished notions in our society.

## Reports

Reports are given of a series of international conferences attended by staff of the Centre and of seminars and training programmes conducted by the Centre itself.

Knut Hofland reviews *The First International Workshop on Natural Communication with Computers*, Warsaw, September 9 - 12, and *The Annual Meeting of the Association for Literary and Linguistic Computing*, Copenhagen, December 11 - 12.

Sigbjørn Århus evaluates *The Eighth International Conference of Computational Linguistics*, Tokyo, September 30 - October 4.

Moreover, a report is given of the Centre's training programme Autumn 1980. This time priority of selection was given to officers of the state archives. The two-week intensive course concentrated on the principles of computer-based archival data bases and demonstration through a census test material of the handling of such source material by computational techniques.

On November 23 - 24 the Centre arranged a *Symposium on Statistical Methods in Archaeology* where a group of archaeologists engaged in statistically-oriented fields of study met with a group of statisticians and computer specialists. This mingling of different forms of competence turned out to be most fruitful and also gave a useful «state of the art» overview to be used in the further planning of a Nordic course on the same topic scheduled for June 17 - 28 in Tromsø.

## News

A series of news items of interest for computer-oriented arts scholars is presented. Some of the topics mentioned:

The Centre has set up a computer consultant service for the humanities at the University of Tromsø. This means that all four Norwegian universities have now obtained their own computer services for the humanities.

From this year onwards the Centre has initiated a three-year research programme to promote the development of methods in specific areas of computing in the humanities. Two research scholarships are being offered, within the fields of computational linguistics and quantitative archaeology respectively. The researchers will be affiliated to the Centre, where they will receive extensive computational and methodological back-up services.

The Centre also presents a summary of its planned activities in 1981. As a result of a deliberate re-profiling of the Centre, more emphasis will be given to

program development, instructional services and training programmes for researchers. Cooperation within specific projects will be conducted mostly on the basis of paid commissions.

A series of national seminars and conferences will be held, among them an Anglo-Scandinavian symposium on the use of computer corpora in teaching and research.

For purposes of text handling, teaching and research the Centre will instal a microcomputer during 1981 to supplement the range of its computer facilities.

Special attention is given to the law which came into force on January 1, 1980 which regulates the use of person-related data bases. The Norwegian Research Council for Science and the Humanities has made it an important point to impart the views of the research community to *Datatilsynet*, a newly established state organization for inspection and consultation on these matters.

Research in many disciplines relies heavily on data about individuals, either from external sources or from sources of information collected as part of a project. It is of the utmost importance that the needs of individual researchers and research bodies are sufficiently allowed for and also that data bases giving information about individuals can be retained by secure measures to allow the possible later use of the data files. All computer data bases which give information that may identify individuals have to obtain a formal concession by the State and the same applies for possible second uses of such source materials. To facilitate the handling of these matters by researchers and to create a channel for effective communication with the authorities, the Norwegian Research Council has set up a specialized secretariat for the purposes of guidance and supervision.

The newsletter also provides information on new issues in the report series of the Centre, including an account of automatic syntactic analysis, and a report on present and future uses of data technology in the humanities. Additionally the report refers to the societal consequences of modern technologies of particular interest to the humanities as fields for investigation. Attention is drawn to a forthcoming catalogue of computer projects in the humanities.

All the above-mentioned reports appear in Norwegian.



---

Medarbeidere i dette nr.:

Gulbrand Alhaug, amanuensis, Universitetet i Tromsø.

Tor Sverre Lande, amanuensis, Institutt for informatikk, Universitetet i Oslo.

Egil Øvrebø, statsarkivar, Statsarkivet i Bergen.

Stig Johansson, dosent, Britisk institutt, Universitetet i Oslo.

Jostein H. Hauge, adm. leder, NAVF's EDB-senter for humanistisk forskning.

Anfinn Stigen, professor i filosofi, Universitetet i Oslo.

*Humanistiske Data* blir utgitt av NAVF's EDB-senter for humanistisk forskning. Redaktør: Jostein H. Hauge. De som ønsker å få bladet tilsendt, kan bestille det ved henvendelse til Senteret adresse: Harald Hårfagesgt. 31, Boks 53, 5014 Bergen — Universitetet. Innlegg kan sendes samme adresse. Bladet utkom med 1 nummer i 1980. Redaksjonen for dette nummer avsluttet 25. april 1981.

*Humanistisk Data* is published by the Norwegian Computing Centre for the Humanities. The Editor is: Jostein H. Hauge, Director of the Centre. Issues are free. Contributions are welcome.