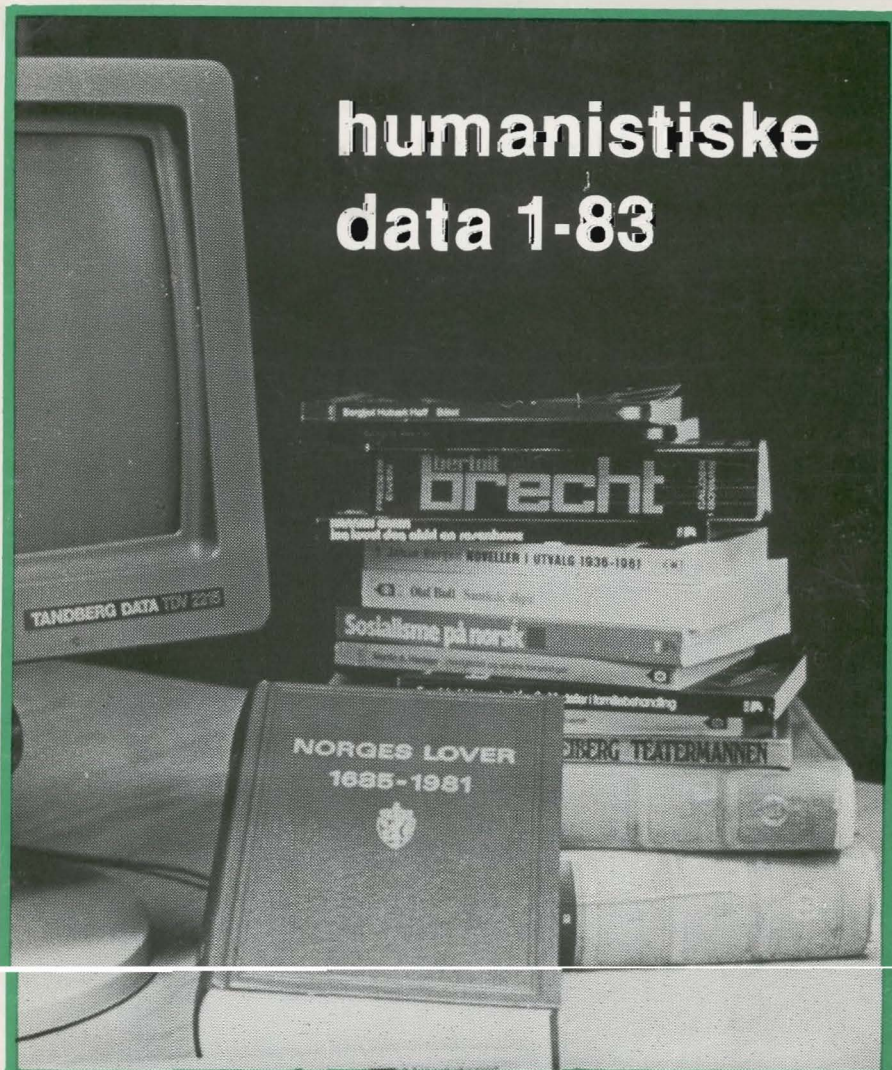


humanistiske data 1-83



NAVF

ARTIKLER
RAPPORTER
MELDINGER
SUMMARY

NAVF's EDB-senter
for humanistisk
forskning

The Norwegian
Computing Centre for
the Humanities

SENTERETS RAPPORTSERIE

- RAPPORT nr. 1. *EDB i gjenstandsfagene*. Rapport fra en konferanse i Bergen, 18. og 19. april 1978. September 1978. 2. opptrykk november 1981. ISBN-82-7283-022-1 Pris kr. 40.
- RAPPORT nr. 2. *Et norsk datamaskinelt tekstkorpus*. Rapport fra en konferanse i Bergen, 19. og 20. oktober 1978. Februar 1979. 2. opptrykk 1981. ISBN 82-7283-016-7 Pris kr. 20.
- RAPPORT nr. 3. *Rapport fra den nasjonale konferanse om EDB i språk og litteraturforskning*, 4. og 5. januar 1979. Mars 1979. 2. opptrykk november 1981. ISBN 82-7283-024-8 Pris kr. 50.
- RAPPORT nr. 4. *Oppbygging av EDB-katalog for folkemusea i Hordaland og kulturgeografisk registrering på Vestlandet*. April 1978. 3. opptrykk november 1981. ISBN 82-7283-000-0 Pris kr. 30.
- 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. 30.
- RAPPORT nr. 9. *Jostein H. Hauge og Sigbjørn Århus: Dataregistrering i humanistiske fag med vekt på optisk lesing*. August 1978. 3. opptrykk januar 1981. ISBN 82-7283-004-3 Utsolgt.
- RAPPORT nr. 10. *Roald Skarsten: Innføring i SPSS for humanister*. November 1977. 3. opptrykk november 1981. ISBN 82-7283-005-1 Pris kr. 30.
- 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 Science. November 1979. ISBN 82-7283-007- 8 Utsolgt.
- RAPPORT nr. 12. *EDB og manuskriptregistraturer*. Oktober 1977. 2. opptrykk november 1979. ISBN 82-7283-009-4 Pris kr. 20.
- RAPPORT nr. 13. *Datatjenester for og datasamarbeid mellom kunst og kulturhistoriske museer*. Februar 1980. 2. opptrykk november 1981. ISBN 82-7283-010-8 Pris kr. 50.

Forts. 3. omslagsside.

humanistiske data 1-83

NAVFs EDB-senter for
humanistisk forskning

The Norwegian Computing
Centre for the Humanities

NAVF NORGES
ALMENVITENSKAPELIGE
FORSKNINGSRÅD

NAVFs EDB-senter for humanistisk forskning ble opprettet av Norges almenvitenskapelig forskningsråd i 1972. Senteret har som oppgave å arbeide på nasjonal basis for utbredelse av edb i forskningsarbeidet i de humanistiske fagene. Det er opprettet en samarbeidsavtale med Universitetet i Bergen som bl.a. gir Senteret adgang til edb-tjenester ved Universitetet.

Av sentrale oppgaver kan nevnes utvikling av programutrustning for humanistiske forskningsoppgaver, konsulenthjelp og informasjonstjenester.

Senteret utgir tidsskriftet *Humanistiske Data* (3 nr. pr. år) og en rapportserie (27 er utkommet pr. 1.3.83).

Senteret er sekretariat for *International Computer Archive of Modern English (ICAME)*, og utgir bladet *ICAME NEWS*.

Senteret driver egne opplæringsprogram for vitenskapelig personale og medarbeidere i den kontor-tekniske gruppen innenfor de humanistiske fag. Det blir også holdt forskjellige kurs og seminar om edb og humanistisk forskning. Tidspunkt og emner blir kunngjort i *Humanistiske Data* og på institusjonene.

Interesserte kan kostnadsfritt bestille årsmelding og *Humanistiske Data* (kr. 40,— for institusjoner).

Humanistiske Data blir utgitt av NAVFs EDB-senter for humanistisk forskning. Redaksjonsgruppe: Jostein H. Hauge (ansv.), Eva Møller, Kristin Natvig, Svein Skotheim.

Senterets adresse: Harald Hårfagesgt. 31, Boks 5014 Bergen-Universitetet.
Tlf. (05) 32 00 40, linje 2956.

Artikler, rapporter, meldinger mottas. Redaksjonen avsluttet 28. februar.

Humanistiske Data is a published by The Norwegian Computing Centre for the Humanities.

Editorial group: Jostein H. Hauge, Eva Møller, Kristin Natvig, Svein Skotheim. The journal can be ordered from the address mentioned above. Contributions are welcome.

Medarbeidere i dette nummer:

Jon Bing, amanuensis, Universitetet i Oslo.

Knut Bjander, høgskolebibliotekar, Rogaland distriktshøgskole.

Anne Golden, hjelpelærer, Universitetet i Oslo.

Håvard Hjulstad, konsulent, Universitetet i Oslo.

Anne Hvenekilde, amanuensis, Universitetet i Oslo.

Egil Kallerud, førstekonsulent, Rådet for humanistisk forskning, NAVF.

Eirik Lien, konsulent, Universitetet i Trondheim.

Trond Løken, konservator, Arkeologisk museum i Stavanger.

Markku Nurminen, professor, Universitetet i Bergen.

Hege Brit Randsborg, vit.ass., Universitetet i Oslo.

Kjell Bjørn Rønning, stud.philol., Universitetet i Bergen.

Roald Skarsten, konsulent, Universitetet i Bergen.

Åse Støgård, førstebibliotekar, Universitetet i Bergen.

Magne Velure, konservator, Hardanger Folkemuseum.

Fra Welinder, dosent, Universitetet i Oslo.

Fra Senteret: *Memund Dalveit*, *Jostein H. Hauge*, *Knut Hofland*, *Rune Johansen*, *Eva Møller*, *Kristin Natvig*, *Øystein Reigem*, *Svein Skotheim*, *Per Vestbøstad*.

Fotosats i kommunikasjon med Univac 1100/82

Sats: Universitetet i Bergen/NAVFs EDB-senter for humanistisk forskning.

Trykk: Bergmanns Boktrykkeri A/S

Forsidebilde: Forfatterrettigheter i en EDB-tid. Se artikkel s. 23. (Foto: Rune Johansen.)

Innhold

Artikler:

Om forholdet mellom EDB og humaniora.

- Markku Nurminen og Kjell Bjørn Rønning. s. 4
Opphavsrett og bruk av tekster i maskinlesbar form. Jon Bing s. 13
EDB og kvinner — en statusrapport. Eva Møller s. 20
LÆREBOKSPRÅK analysert ved hjelp av EDB.
Anne Hvenekilde og Anne Golden s. 27
EDB i fornminneforvaltning ved et arkeologisk museum.
Trond Løken s. 32
EDB og ordbøker. Håvard Hjulstad s. 42
Norsk-vietnamesisk ordbok. Eva Møller s. 47
Automatisk analyse av norsk. Intervju med Per-Kristian
Halvorsen. Jostein H. Hauge s. 51
Tekstbehandling i humanistisk forskning. Svein Skotheim s. 57

Rapporter:

Den andre nasjonale namnegranskarkonferansen,

- Universitetet i Oslo s. 62
Humaniora-uken ved Universitetet i Oslo 1983. s. 63
Seminar og automatisk analyse av naturlig språk s. 69
BIBNETT/BIBNETT 2 s. 71
Dokumentasjonssenter for kulturpolitikk og kulturforskning s. 73
NTAI-prosjektet s. 75
Programpakketet STAR s. 77
Avisinform-prosjektet s. 79
EXTEMPLO — et nordisk elektronisk tidsskrift s. 81
Nytt fra RHF/NAVF s. 83
Fra bokhylla s. 86
Meldinger s. 91
Summary s. 98

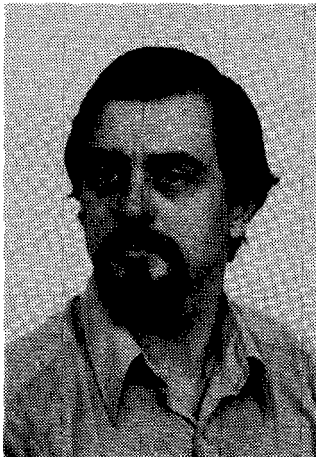
Om forholdet mellom EDB og humaniora

Markku Nurminen og Kjell Bjørn Rønning

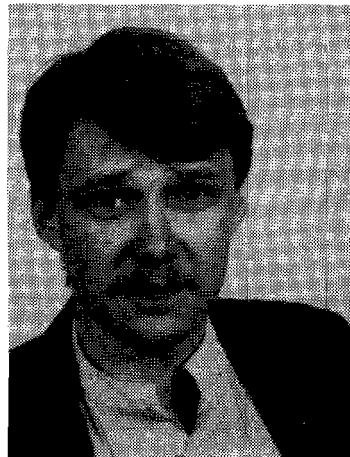
I denne artikkelen har vi valgt å føre sammen to fagområder med helt ulik bakgrunn: edb og humaniora. Bakgrunnen er at edb-teknologien i stadig sterkere grad er blitt et verktøy i den humanistiske forskningsprosessen. Siden vi er tilknyttet Institutt for informasjonsvitenskap, Universitetet i Bergen, faller det naturlig å ta opp denne utviklingen til diskusjon. Instituttet har en eksplisitt målsetning om å utforske mulige relasjoner mellom edb og humaniora. Artikkelen er ment som et diskusjonsinnlegg hvor spørsmålene og problemstillingene er viktigere enn svarene. Vi vil presentere to ulike tilnærminger, som uttrykk for to ulike måter som edb og humaniora kan knyttes sammen på.

De humanistiske fag (humaniora) har en lang tradisjon. Denne tradisjonen representeres ved fag som filosofi, logikk, lingvistikk, historie, osv. Felles for denne tradisjonen er at mennesket og menneskets handlinger, kultur, redskaper, osv. utgjør det sentrale forskningsobjekt. Ofte eksisterer det innen denne tradisjonen ingen målsetning om umiddelbar anvendelse av den resulterende kunnskapen, som middel til en målrasjonell prosess. Tvert imot betraktes humanioras produkt, kunnskap om mennesket og dets handlinger, som verdifull for menneskelig (selv-)forståelse uavhengig av yrker og anvendelsesområder; de humanistiske fag (humaniora) kan identifiseres som del av en *dannelses-tradisjon*. (Per Jacobsen: *Humaniora och teknik - kan de mötas?*, UHÅ-rapport: 1982-6)

Sammenlignet med humaniora er de edb-orienterte fag (informasjonsvitenskap, informatikk, datalogi, osv.) de rene ynglinger. Disse vitenskapene er knyttet til et teknisk utstyr, datamaskinen, og utnyttelsen av denne (som arbeidsredskap). Siden datamaskinen er skapt av mennesker og utnyttes av mennesker i deres arbeidssituasjon, kan den betraktes som en kulturgjenstand. De edb-orienterte fag forholder seg imidlertid



Markku Nurminen.



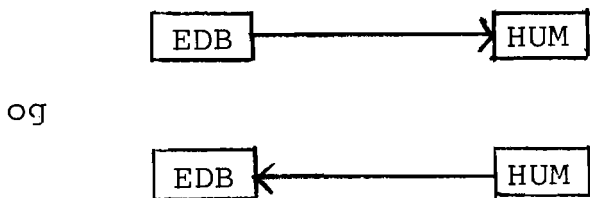
Kjell Bjørn Rønning.

helt annerledes til «sin» kulturgjenstand, enn det som er tilfelle for de humanistiske fag. For de edb-orienterte fag representerer datamaskinen et teknisk potensiale som kan anvendes og utnyttes som verktøy for løsning av konkrete problemer (i en arbeidssituasjon). Den resulterende kunnskap er verdifull kun i den utstrekning den gir opphav til teknikker som er anvendbare. *David Silverman* har gitt en karakteristikk av organisasjonsvitenskapens utgangspunkt og tidlige fase, som vi mener også er dekkende som en beskrivelse av de edb-orienterte fag (i dag):

«Organisation analysis in the first few decades of this century began from the questions asked by those in positions of authority, to whom it sought to offer effective and preferably cheap «solutions» to troubling problems.» (David Silverman: *The Theory of Organisations*, s. 2, Heineman - London 1970)

Mens de humanistiske fag på sett og vis er opptatt av de «evige» spørsmål rundt mennesket og kultur, uten intensjoner om å frembringe teknikker som kan funksjonere som middel i en målrasjonell anvendelsesprosess, er det selve det tekniske anvendelsespotensialet i den resulterende kunnskap som utgjør legitimasjonen for de edb-orienterte fag; problemstillingene er tidsavhengige og knyttet til en teknologisk utvikling som løper parallelt (og ofte foran) utviklingene innenfor fagene. Teknologien er hovedleverandør av premisser og problemstillinger, og fagenes suksess måles i evnen til å omsette denne teknologien til funksjonelle arbeidsredskaper.

Det virker derfor som om disse to fagområdene, edb-orienterte fag og humaniora, har helt forskjellige utgangspunkt, forskjellig tilnærming til problemene, forskjellig tidshorisont og dermed ingen felles plattform. Nettopp derfor kan det kanskje være nyttig å prøve å føre disse to områdene sammen, både for å klargjøre forskjellene bedre, og for å undersøke om det finnes kombinasjonsmuligheter som kan være kimen til tverrvitenskapelig problemhåndtering. For å strukturere fremstillingen velger vi å innføre en distinksjon mellom to ulike måter å kombinere de edb-orienterte fag og humaniora:



Pilene skal preliminært tolkes slik at vi velger å bruke verktøy, metoder og perspektiv fra ett element (fagområde) innenfor det andre. Vi tror at begge retningene i en rendyrket form har en meningsfull og interessant tolkning. Vi vil kalle dem for henholdsvis en «anvendelsesorientert» tilnærming og en «reflekterende» tilnærming. Disse betegnelsene er valgt fordi vi mener de gir en antydning om karakteren av tilnærming til forskningsobjektet innenfor henholdsvis de edb-orienterte fag og humaniora.

Anvendelsesorientert tilnærming

I denne retningen gjør vi bruk av edb-teknologi og kunnskap fra de edb-orienterte fag i den humanistiske forskningsprosessen. Problemstillingen er opprinnelig humanistisk, men vi antar at datamaskinen har visse egenskaper som gjør den til et effektivt forskningsverktøy også for de humanistiske fag. Disse egenskapene er først og fremst datamaskinens:

- mulighet for rasjonell lagring av store datamengder
- mulighet for rask fremhenting, behandling og
- analyse av store datamengder.

Disse egenskapene kan man utnytte med stort utbytte i løsningen av vitenskapelige oppgaver hvor kvantitative aspekter (i form av tid og størrelse) ved undersøkelsesmaterialet utgjør kritiske momenter i gjennomføringen av et forskningsprosjekt. Ved analyse/syntese av musikk eller tale er tidsintervallene ofte så små at disse oppgavene er praktisk talt uløselige uten automatisk analyse- og styreredskaper. Empiriske eller systematiske delmaterialer i f.eks. historiske, lingvistiske eller arkeologiske undersøkelser kan være så store, at manuell kontroll ikke er tilstrekkelig. Innenfor disse områdene har edb-teknologien representert en utvidelse av forskningsmulighetene. I tillegg har denne anvendelsen ledet til en «produktutvikling» i form av utvikling av systemer for syntetisering og reproduksjon av tale, automatisk syntaksanalyse (anvendt i tekstbehandlingssystemer).

Vi kan godt betrakte slike anvendelser av edb-teknologi som utvikling og drift av informasjonssystemer og anvende en analogi med informasjonssystemer i bedrifter og organisasjoner. Denne betraktningssmåte skaper en utvidelse av grensesnittet mellom humaniora og de edb-orienterte fagene. Det åpner for en utnyttelse av et bredere repertoar av styringsteknikker hentet fra edb-fagene; i tillegg til intern-teknologiske styringsteknikker (f.eks. programmering) vil de edb-orienterte fag også kunne bidra med teknikker for modellering av problemsituasjonen, teknikker for styring av innføring av edb-baserte informasjonssystemer, evalueringsteknikker, osv. Dette gir en bredere definisjon av den «anvendelsesorienterte» tilnærmingen enn det som intuitivt synes mulig, men det skjerper også problemaspektet ved en slik sammenføring.

Det er langt fra gitt at denne kombinasjonen av edb og humaniora er uproblematisk. Det er særlig to forhold som kan være vanskelig å styre for den humanvitenskapelige forskeren, og som dermed fortjener en viss oppmerksomhet:

1. i hvilken utstrekning kan edb-teknologien og anvendelsen av edb-faglig kunnskap og teknikker virke styrende på humanioras innhold og kunnskapsstatus
2. hvordan skal den humanistiske forskeren bevare styringen over forskningsprosessen (muligheten for rasjonell rekonstruksjon og kritisk evaluering)

Ad 1. Edb og humanioras innhold og kunnskapsstatus.

Et utgangspunkt for denne problemstillingen er en antagelse om at metoder og teknikker ikke er «utvendige» i forhold til en disiplin. Metoder virker strukturerende på hvordan problemstillinger formuleres og på hvilke problemstillinger som behandles. Metoden er med andre

ord et konstituerende element for disiplinen. Metoden (eller en antisipert oppfatning av metodens anvendelsespotensiale) er med på å forme forskerens «blikk» for tingene. Metoden utgjør et bestemt «grep» på virkeligheten. Problemet kan derfor formuleres slik: i hvilken utstrekning vil edb-teknologien og et utvidet styringsrepertoar fra edb-fagene kunne være med på å styre forskerens perspektiv og virkelighetsoppfatning, og dermed påvirke forskningens innhold? Vil bruk av datamaskinen og edb-faglige styringsteknikker bety en (utilsiktet) overføring av metoder og perspektiv fra tekniske og matematiske disipliner og dermed påtvinge de humanistiske fag en mer teknisk (målrasjonell) tilnærming til forskningsobjektet? Hva med humaniora som (refleksiv) prosess: vil den forandre karakter slik *Sigmund Koch* beskriver:

«A-meaningful thought of inquiry regards knowledge as the result of «processing» rather than discovery; it presumes that knowledge is an almost automatic result of a gimmickry, as an assembly line, a «methodology»; it assumes that inquiry in action is so rigidly and fully regulated by *rule* that in the conception of inquiry it often allows the rules to displace their human users.» (Sigmund Koch: *Psychology as Science*, p. 8 in *Philosophy of Psychology* (ed. S.C. Brown), Macmillan Press 1979)

Det følger av denne diskusjonen at vi ikke betrakter teknologien som (verdi-)nøytral. Ulike teknologier – f.eks. datateknologien – har intenderte, innebygde egenskaper som fremmer bestemte typer anvendelser og som utelukker andre. Når teknologien samtidig får et prestisjepreg – som datateknologien har – ligger det en fare for at arbeidsoppgaver defineres ut fra hva teknologien kan bidra med, og ikke omvendt, at teknologien velges ut fra strukturen i arbeidsoppgaven. Teknologien medfører en avgrensning av arbeidsprosessen, ved å virke som et filter for hvordan problemer kan løses. Er teknologien kraftig øker sjansene for at den i tillegg vil virke som en generator: det tomrom den selv skaper, fyller den med teknologi-inspirerte problemstillinger. På denne måten skjer det en dobbelforskyvning i forhold til den opprinnelige problemformulering.

Ad 2. Rekonstruksjon og evaluering av forskningsprosessen.

Utgangspunktet for denne problemstillingen er en antagelse om at det hører til forskningens egenart at forskningsprosessen skal kunne kontrolleres, rekonstrueres og evalueres av forskeren. Ved anvendelse av edb-teknologi i den humanistiske forskningsprosessen får man ofte en arbeidsdeling mellom en problemorientert forsker og en metodeorientert forsker. De metoder og analyseteknikker som benyttes kan være så kompliserte at de representerer et eget spesialfelt som krever betydelig kompetanse. Man kan derfor havne i den paradoksale situasjon at den problemorienterte forskeren ikke er i stand til kritisk å vurdere anvendelsen av metoder og analyseteknikker i sitt eget forskningsprosjekt, samtidig som den metode-orienterte forskeren er ute av stand til å vurdere innholdet i det som skal analyseres.

Nå er ikke dette et problem som er unikt for denne type arbeid, problemet ligger latent i all type arbeidsdeling som er motivert ut fra en oppsplitting av arbeid i ulike kompetanseområder. Også innenfor vitenskapene er man for lengst kommet forbi den tid da det var mulig å

være å jour innenfor alle vitenskapelige områder. I dag får man heller være fornøyd hvis man klarer å følge med på sitt eget spesialfelt innenfor *en* vitenskap. Teamwork og tverrvitenskapelige prosjektgrupper har erstattet universalgeniet. Humaniora er vel det forskningsområdet hvor forskningen fortsatt drives som énmannsprosjekter (i en viss utstrekning), og hvor «prosesserings»-tilsnittet (Koch) kanskje ennå ikke er så fremtredende som i en del andre disipliner. Innføring av edb-teknologi i den humanistiske forskningsprosessen kan forandre på dette. Spørsmålet blir da om dette er en ønsket utvikling, og om hva som eventuelt må gjøres hvis man ønsker å forhindre at så skjer.

De innvendinger og problemer som er nevnt er ikke av en slik karakter at det skulle være logisk umulig å anvende edb-teknologi i den humanistiske forskningsprosessen. Praksis viser at *det* ville være en altfor rigorøs påstand. Problemet er hvordan humanioras egenart kan bevares samtidig som disse disiplinene tar i bruk nye forskningsverktøy som f.eks. edb.

Reflekterende tilnærming

Ovenfor har vi sett på hvilke muligheter og problemer anvendelse av edb og edb-faglig kunnskap innenfor humaniora fører med seg. Vi har kalt denne retningen for en «anvendelsesorientert» tilnærming for å indikere den kunnskaps-«approach» de edb-orienterte fagene bringer videre til det humanistiske forskningsparadigmet. I denne delen skal vi se litt nærmere på den motsatte kobling; overføring av humanvitenskapelige synsmåter og kunnskap til edb-fagene.

I denne tilnærmingen er forskningsområdet edb-faglig. Man studerer edb-teknologi, anvendelse av edb-teknologi, osv. Humaniora kan bidra med en perspektivutvidelse ved at dette forskningsparadigmet tilfører kategorier og begreper som kan gjøre det mulig å gjennomføre en eksplisitt beskrivelse og behandling av fenomen som edb-fagene ikke har begreppskjemaer for. Eksempler på områder hvor en slik kobling kan være interessant er:

- brukerorienterte systembeskrivelser (lingvistisk og psykologisk metodeutvikling)
- informasjonssystemers funksjon betraktet som handling (handlingsteori, filosofi)
- informasjonssystemer betraktet som kommunikasjonssystem (kommunikasjonsteori, lingvistik)
- representasjon av informasjon ved hjelp av tegn (semiotikk)
- hvordan forstår/oppfatter mennesker den informasjon som produseres av edb-baserte systemer (kognitiv teori, hermeneutikk)
- informasjonssystemets språk/forandring av språkbruk (grammatikk, lingvistik)

Når vi har valgt å kalle denne koblingen for en «reflekterende» tilnærming henger det sammen med vårt syn på humaniora og den humanvitenskapelige forskningstradisjon. Siden disse vitenskapene ikke er anvendelsesorienterte i sin natur er det litt vanskelig å tenke seg disse anvendt som redskaper eller teknikker i den edb-vitenskapelige forskningsprosessen på samme måte som synes mulig i den omvendte kombinasjonen. Dette indikerer at problemstillingen i forbindelse med

denne tilnærmingen ikke er faren for at humaniora skal «overta» edb-vitenskapene, men heller det motsatte: å definere en plattform som i det hele tatt slipper denne tradisjonen inn i edb-vitenskapene. Hvis en slik plattform lar seg utarbeide kan kanskje edb-fagene gjennomgå noe av den samme utvikling som organisasjonsvitenskapen har gjort:

«(...) contemporary writers are concerned rather less with «techniques» and rather more with knowledge; that they have moved away from an over-riding concern with prescription to an attempt to apply the insights of a wide range of disciplines and conceptual schemes in order to gain an understanding of organisational life; and that, consequently, they have asked a broader range of questions and attained more significant answers than was the case in earlier work.» (David Silverman: op.cit. s. 2)

Hva skulle en slik plattform bestå av? Ovenfor har vi antydnet noen delproblemer hvor humanvitenskapelig kunnskap og tilnærming vil kunne være *nyttig* for edb-vitenskapene. Vår antagelse er imidlertid at denne plattformen kan gis en dypere forankring enn bare utnyttelse av humanvitenskapelig produserte del-kunnskaper. Vi mener at en rekke humanvitenskapelige problemstillinger er «immanente» også innenfor de edb-orienterte fag, problemet er ikke hvorvidt edb-fagene skal forholde seg til disse eller ikke. Det gjør de uansett. Problemet er hvorvidt dette skal bringes på begrep, og betraktes som genuint edb-faglige problem. Vanligvis vil man hevde at det er en styrke for en vitenskap at den klarer å inn-reflektere flest mulige aspekter rundt forskningsobjektet, metodeanvendelsen osv. Riktignok er det legitimt å sette ting «i parentes»; altså bevisst bortse fra mulige innfallsvinkler med henblikk på å gjøre forskningsoppgavene mer overkommelige. På den annen side betraktes det som en svakhet ved et forskningsprosjekt hvis relevante innfallsvinkler er utelatt uten at dette er spesifikt presisert og begrunnet.

På hvilken måte berøres edb-fagene av (ufrakommelige) humanvitenskapelige problemstillinger? Dette kan best illustreres gjennom et eksempel. Da må vi ha i mente vår karakteristikkk av humaniora som en tradisjon som (vitenskapelig) forholder seg til mennesket og dets handlinger: Informasjonssystemer (særlig administrative) utvikles av og for mennesker og brukes av mennesker. Informasjonssystemene tjener en hensikt, nemlig som redskap for menneskets informasjonsbehandling. Når man utvikler et redskap er det rimelig at dette utformes ut fra de behov redskapsbrukeren har, og at man tar hensyn til egenskaper hos redskapsbrukeren. Hvilke egenskaper har denne redskapsbrukeren? Hvilke informasjonsbehov skal man søke å dekke gjennom systemet? Hva er informasjon? Å besvare slike spørsmål, og en rekke andre, involverer forsøk på kategorisering av mennesket og menneskelige handlinger. Dermed er humanvitenskapelige problemstillinger en del av grunnlaget for anvendelsesorienteringen overhodet. Betraktes mennesket f.eks. éndimensjonalt som et produksjonspotensial (i forbindelse med utvikling og drift av informasjonssystemer) får dette karakter av en fordom hvis forskeren ikke er bevisst (eller dokumenterer at han er bevisst) at dette perspektivet ikke er uttømmende for en beskrivelse av mennesket som helhetlig individ. En slik

ureflektert éndimensjonalitet har produsert vitenskapelige retninger som f.eks. «scientific management» (innenfor organisasjonsvitenskapen), en type vitenskap som de fleste i dag betrakter som inhuman og lite etterstrebellesverdige.

Nå kan man selvsagt stille spørsmålet om det er en forutsetning for edb-forskeren at han er (forskningskompetent) filosof, lingvist, historiker, osv. for at han skal kunne utføre meningsfull edb-faglig forskning. Selvsagt ville det være ønskelig med en slik universalkompetanse, men at denne fordringen skulle kunne imøtekommes er lite realistisk. Sannsynligvis heller ikke nødvendig. Foreløpig får vi greie oss med at disse problemstillingene overhodet blir reist innenfor det edb-faglige forskningsmiljøet, og samtidig håpe på en bevisstgjøring omkring disse forhold innenfor de ulike miljøene. Vi vil også påpeke de humanvitenskapelige miljøenes ansvar for at så skjer.

Muligheter for syntese?

Analysen av de to kombinasjonene av edb og humaniora er med vilje rendyrket. I en praktisk tillempling vil humaniora neppe fremstå som fullstendig likegyldig vis-à-vis anvendelsespotensialet til den kunnskapen den selv produserer. Samtidig innebærer anvendelsesorientert forskning (og dens parallell: praktisk systemutvikling) alltid et moment av refleksjon. Men det ligger likevel noe vesentlig i den grafiske beskrivelsen av de to tilnærmingene som to piler i *motsatt* retning. Vi tror ikke at denne motsetningen oppløses i det øyeblikket den konfronteres med den praktiske virkelighet. Ulikhetene mellom disse to forskningstradisjonene lar seg ikke oppheve så lett. Vi skal prøve å illustrere denne ulikheten ved hjelp av *Habermas* sin klassifisering av vitenskapene ut fra hvilke *kunnskapsinteresser* som virker konstituerende for disse. Dette må betraktes mer som en metaforisk enn som en teoretisk argumentasjon.

Vi kan assosiere den anvendelsesorienterte retningen med det som *Habermas* kaller for den *tekniske* kunnskapsinteresse. Denne får sin realisering gjennom de empirisk-analytiske vitenskapene:

«Empirisk-analytisk viten er (...) mulig prognostisk viten. *Meningen* i slike prognoser, nemlig deres tekniske anvendbarhet, gir seg da også først ut fra de regler vi bruker når vi anvender teorien på virkeligheten.» (J. *Habermas*: Erkjennelse og interesse, i *Vitenskap som ideologi*, s. 19, Gyldendal 1969)

Den reflekterende tilnærming kan i sin tur assosieres med det *Habermas* kaller for en *praktisk* kunnskapsinteresse. Denne får sin realisering gjennom den humanvitenskapelige tradisjon:

«*De historisk-hermeneutiske* vitenskaper vinner sin erkjennelse i en annen metodologisk ramme. Her konstituerer meningen i utsagnenes gyldighet seg ikke i det tekniske rådveldes referansesystem. (...) Forståelsen av mening og ikke iakttagelse baner veien til fakta. Den systematiske overprøving av antagelser om lovmessigheter tilsvarende her utlegning av tekster.» (*Habermas*: op.cit. s. 20)

Hos *Habermas* knyttes de ulike vitenskapelige tradisjoner sammen gjennom en filosofisk kunnskaps-antropologi. De ulike kunnskapsparadigmer reflekterer ulike måter som mennesket forholder seg til verden

på. Disse kunnskapsparadigmene er knyttet til ulike handlingskategorier: henholdsvis *arbeid* og *språk*. Både arbeid og språk er nødvendige kategorier for menneskelig eksistens. Ut fra denne betraktningen blir disse kategoriene komplementære. Men denne komplementariteten blir meningsfull bare hvis den sees i sammenheng med en tredje type kunnskapsinteresse: nemlig den emansipatoriske:

«Den metodologiske ramme som fastlegger meningen i gyldigheten av denne kategori (...), har begrepet *selvrefleksjon* som målestokk. Selvrefleksjonen løser subjektet fra avhengigheten av hypostaserte makter. Selvrefleksjon er bestemt av en frigjørende erkjennelsesinteresse.» (Habermas: op.cit. s. 310)

Gjennom den emansipatoriske kunnskapsinteressen får arbeidet og språket sin mening. Refleksjon utvides dermed til selv-refleksjon. Dette poenget har en klart praktisk konsekvens for hvordan humaniora må inn-reflekteres i edb-fagene: selv-refleksjonen er en individorientert kategori; man kan ikke bedrive selv-refleksjon på vegne av andre. Dermed oppheves ikke spenningen mellom anvendelse og refleksjon gjennom prosjektadministrative tiltak. Det humanvitenskapelige perspektivet må inn i den anvendelsesorienterte kunnskapsprosessen som en integrert del av den enkelte forskers forståelseshorisont. I tillegg til en utnyttelse av humanvitenskapelige delkunnskaper for å bedre den «tekniske» problemhåndteringen. Dermed kan man oppnå *en forandring gjennom forståelse* (Tore Nordenstam), noe den éndimensjonalt tekniske referanseramme vanskelig kan klare.

Sammenfatning

Det er den anvendelsesorienterte tilnærming som har dominert innføringen av edb-teknologien fra dens begynnelse. Denne tilnærmingen synes også å gjelde innenfor området: bruk av edb i den humanistiske forskningsprosessen. Denne tilnærmingen ligger selvsagt i edb-teknologiens natur: den er et redskap. Overfor den vitenskapbaserte teknologien har humaniora, og dermed refleksjonen, stadig blitt skjøvet lengre og lengre i bakgrunnen. Dermed er det veldig fristende å omdefinere humanvitenskapenes status og innretning, slik at disse også kan bli «nyttige», og gjennom sitt anvendelsespotensiale legitimere behovet for forskningsmidler. Men dette fremmer ikke nødvendigvis refleksjonen, snarere tvert imot, gjennom en slik prosess kan den miste et av sine siste fristeder.

Det er ikke en ren tilfeldighet at ordet «humanistisk» har to betydninger. Det kan assosieres til humaniora som vitenskapelig disiplin, men også til humanismen som livsholdning. Humanismen innebærer respekt for det enkelte mennesket og dets dannelse som en udelelig helhet. Den tar avstand fra den éndimensjonale oppfatningen av mennesker som arbeidskraft eller «human resources». Humanistisk forskning, når den er reflekterende, kan (men ikke nødvendigvis må) fremme slike humanistiske idealer ved at det i humaniora er *legitimt* å betrakte enkelte individer som udelelige helheter ut fra deres egne utgangspunkter (verdier, kunnskap, estetikk, osv.) i større utstrekning enn hva som er mulig innenfor andre disipliner. Dette skaper et berøringspunkt mellom de to betydningene av ordet «humanisme» som

pålegger de humanvitenskapelige disipliner forpliktelser på vegne av enkeltmennesket.

Litteratur

Habermas, Jürgen: Erkjennelse og interesse, i *Vitenskap som ideologi*, Gyldendal 1969.

Jacobsen, Per: Humaniora och teknik – kan de mötas? UHÅ-rapport: 1982-6.

Koch, Sigmund: Psychology as Science, in *Philosophy of Psychology* (ed. S.C. Brown), Macmillan Press 1979.

Nordenstam, Tore: Förståelse och förändring, i Göransson (red.): *Ideologi och systemutveckling*, Studentlitteratur 1978.

Silverman, David: *The Theory of Organisations*, Heineman – London 1970.

Opphavsrett og bruk av tekster i maskinlesbar form

Jon Bing

1. Utgangspunktet

Uttrykkene «åndsverk» eller «litterært verk» er kanskje nok så adstadiige, og vekker forestillinger om lødig litteratur og kalveskinnsbind med gullsnitt. I juridisk terminologi er det imidlertid betegnelsen for ethvert stykke tekst som er skapt på en selvstendig måte – brev, annonser, avisartikler, lyrikk, sakprosa osv.

Tekst er en av de vanligste måtene å kommunisere på. Tekster vil derfor ofte være kildemateriale for forskere. Ikke bare språkforskere, men også andre – f.eks. historikere – som utnytter tekstene qua kilder.

Det er to hovedgrunner til at forskere ønsker å gjøre eksisterende tekster tilgjengelig i maskinlesbar form. For det første kan det være en hensiktsmessig utgangspunkt for analyse av tekstene, f.eks. fremstilling av sorterte frekvenslister. For det andre kan datamaskinbaserte informasjonssystemer gi mulighet for effektivt å gjenfinne interessante tekster i en større tekstsamling.

I stigende utstrekning vil tekstene allerede være i maskinlesbar form i forbindelse med produksjonen, enten fordi forfatteren benytter et tekstbehandlingsanlegg for skriving og kommunikasjon, eller fordi teksten konverteres til maskinlesbar form i forbindelse med setting.

Ethvert åndsverk har sin opphavsmann, og opphavsretten setter grenser for hvordan åndsverk kan utnyttes – grenser som også gjelder for trivielle tekster. I denne lille artikkelen skal noen av reglene for utnyttelse av tekster i maskinlesbar form skisseres.

Det bør understrekes at datamaskinteknologien har satt gjeldende opphavsrett på prøve. Det finnes mange regler eller begreper som er problematiske i forhold til den nye teknologien, og det gjenstår ennå mye analyse før detaljene – også praktisk viktige detaljer – festner seg.* I denne artikkelen vil det imidlertid ikke gjøres noe forsøk på å forfine oversiktsbildet. Det er noen enkle og viktige hovedregler som vil bli behandlet.

Enkelte opphavsrettslige aspekter ved datamaskinbaserte systemer er under utredning ved Institutt for rettsinformatikk, Universitetet i Oslo. Dette instituttet vil kunne gi råd i forbindelse med rettslige aspekter ved datamaskinbaserte systemer, og er også interessert i å få eksempler fra praksis som korrektiv til mer teoretiske analyser.

2. Opphavsmannens enerett

Som nevnt er det små krav til at noe skal kvalifiseres som et åndsverk. Loven (ål prg 1) definerer åndsverk som «litterære, vitenskapelige eller kunstneriske verk av enhver art og uansett uttrykksmåte eller uttrykksform». I ordet «skape» leses det inn et krav om individuell, skapende

*En slags oversikt finner man i Jon Bing: «Opphavsrettslige aspekter ved datamaskinbaserte informasjonssystemer», publisert av Televerkets forskningsinstitutt som del av TELSAM-rapport nr. 6, *Jus, tele og data I* (1980).

innsats. En rutinemessig og arbeidskrevende sammenstilling av elementer – f.eks. en bibliotek katalog – vil normalt ikke ha den tilstrekkelige åndsverkkvalitet. Men det er likevel ikke stort som skal til. Og i vurderingen ligger det *ikke* noe krav om at verket skal ha kunstneriske eller estetiske kvaliteter. Ethvert brev, ethvert avisinnlegg, uansett hvor hjelpeløst i formulering og uansett hvor uinteressant i innhold, vil være et åndsverk. I praksis må man derfor regne enhver tekst som et åndsverk.

Dessuten finnes det i åndsverkloven (ål prg 43) en spesiell beskyttelse for «formular, katalog, tabell og lignende arbeid som sammenstiller et større antall opplysninger». Etter denne bestemmelsen nyter en del frembringelser som *ikke* er åndsverk et lignende vern. Den bibliotek katalog som ble nevnt ovenfor, vil sannsynligvis være vernet etter denne bestemmelsen. Og det må antas at en så triviell publikasjon som telefonkatalogen også er vernet etter denne «katalogregelen».

Utgangspunktet er at opphavsmannen, det vil altså si den som har skapt verket, har enerett til å råde over det «ved å fremstille eksemplarer av det og ved å gjøre det tilgjengelig for almenheten» (ål prg 2, 1. ledd). Dette er et strengt krav. Det betyr at ingen andre kan råde over et åndsverk på noen som helst måte med mindre minst én av to betingelser er oppfylt:

- (1) Vedkommende har fått rett til å råde over verket på den aktuelle måte i avtale med opphavsmannen eller en som representerer opphavsmannen (en organisasjon, en arving e.l.); eller
- (2) det kan påvises en lovbestemmelse som gir vedkommende rett til å råde over verket på den aktuelle måte.

Med mindre en av disse betingelsene er oppfylt, vil det altså være urettmessig å råde over verket – normalt også straffbart.

Et av de vanligste spørsmålene knytter seg til urettmessig bearbeidelse eller «plagiat». For at det skal bli spørsmål om dette, må man direkte ha utnyttet de formuleringer eller de ord opphavsmannen har brukt for å gi uttrykk for tankene. Hvis man f.eks. utgir en annens resonnement eller resultater som sine egne, men gir uttrykk for dem med egne ord og setninger, kan man nok beskyldes for uetisk bruk av kilder osv., men noen opphavsrettskrenkelse er det ikke. Man sier gjerne at Einstein har opphavsrett til den artikkelen hvor han fremsetter sin relativitetsteori, men ikke til relativitetsteorien.

I opphavsretten settes det altså en grense mellom form og innhold. Selv om grensen ikke er så enkel å trekke som antydnet her, viser den likevel at opphavsretten ikke skaper noen hindring for fri meningsutveksling og informasjonsformidling – men skaper nokså snevre grenser for mulighetene for å utnytte en annens form.

3. Datamaskinbaserte systemer

I opphavsrettslig sammenheng skaper i utgangspunktet datamaskinbaserte systemer ikke noe nytt. Det har vært en livlig diskusjon i utlandet, særlig i USA, omkring opphavsrett og datamaskinbaserte systemer. Dette springer imidlertid først og fremst ut av en lovgivning som er forskjellig fra vår egen, og som for øvrig ble endret såpass sent

som i 1974 for å avhjelpe noen av problemene.

Utgangspunktet er at en maskinlesbar kopi av verket er et *eksemplar* av verket i opphavsrettslig forstand, og derfor kommer inn under opphavsmannens enerett. Det er i prinsippet ikke noen forskjell på å fotokopiere en bok og å registrere boken i maskinlesbar form – enten ved inntasting på nytt eller optisk lesning. Dette gjøres helt klart av ål prg 2, 2. ledd:

«Som fremstilling av eksemplar regnes også overføring til innretning som kan gjengi verket».

Blant de «innretninger som kan gjengi» verk må datamaskiner utvilsomt regnes (selv om formuleringen opprinnelig først og fremst tok sikte på fonogrammer).

Eksemplar-begrepet er formelt. Nye eksemplarer fremstilles f.eks. regelmessig når man tar reservekopier (backup) av eksisterende materialer. Nye eksemplarer vil ofte oppstå når materiale kommuniseres fra én maskin til en annen. All eksemplarfremstilling er i utgangspunktet del av opphavsmannens enerett.

Dette gjelder også når det maskinlesbare eksemplar er fremstilt ved avtale med opphavsmannen, men den videre bruk mangler hjemmel i slik avtale.

Som eksempel kan man nevne Norsk tekstarkiv ved Universitetet i Bergen. De maskinlesbare eksemplar av norske bøker som finnes i arkivet, er opprinnelig blitt til i forbindelse med trykningen av bøkene. De er da konvertert til maskinlesbar form av trykkeriet, som senere har produsert offset-sats til boken ved hjelp av en datamaskinstyrt fotosetting. Trykkeriets maskinlesbare eksemplar er lovlig fremstilt på grunnlag av avtale med det aktuelle forlag, som igjen har en utgivelsesavtale med forfatteren.

Tekstarkivet mangler imidlertid slik avtale. Det er allerede urettmessig av trykkeri eller forlag å utlevere det maskinlesbare eksemplar av boken til arkivet uten avtale med forfatteren. Og straks arkivet tar en kopi for oversendelse til en interessert forsker, skjer en klar krenkelse av opphavsmannens enerett (ål prg 2, 1. ledd jf. prg 54), med mindre det foreligger avtale med opphavsmannen.

Eksempelet er valgt fordi det er kjent. Forhandlinger mellom Tekstarkivet og de skjønnlitterære skribentorganisasjonene er innledet.

Rettighetshaverne er generelt sett på vakt i forhold til urettmessig bruk av deres verk – antagelig skremmer erfaringer fra fotokopiering og videorammer tilstrekkelig til å gjøre de etter hvert forholdsvis sterke skribentorganisasjonene interesserte i å komme utviklingen i forkjøpet. Et eget samarbeidsutvalg mellom skjønn- og faglitterære skribentorganisasjoner, nettopp med sikte på å behandle spørsmål i tilknytning til datamaskinbaserte systemer, er allerede etablert. Det er også sannsynlig at Opphavsrettsutvalget, som for tiden arbeider med en generell revisjon av åndsverkloven, vil ta opp dette i sin neste utredning. Utvalget har samarbeid med tilsvarende utvalg i de andre nordiske land, og man vil kunne vente en nordisk koordinering. Denne kan kanskje få et enda bredere grunnlag i en rekommandasjon som i høst vil komme opp til behandling i eksekutivkomiteen i de internasjonale organisasjoner som administrerer de opphavsrettskonvensjoner Norge er medlem av

(World Intellectual Property Organization og UNESCO).

Ovenfor er opphavsmannens enerett understreket. Det er også ment å være en advarsel. Hvis man som forsker arbeider med tekster, har man en plikt til å orientere seg om grunnprinsippene for bruken av åndsverk.

Men det finnes enkelte unntak for praktisk viktige situasjoner. I denne artikkelen vil ikke alle disse unntakene nevnes, men bare to hovedpunkter som har særlig betydning nettopp for utnyttelse ved hjelp av edb.

4. Verk uten vern etter åndsverkloven

Utgangspunktet nevnt ovenfor var at enhver tekst måtte antas å være et åndsverk. Det er et holdbart utgangspunkt, men det gjelder tre praktisk viktige unntak.

For det første faller opphavsmannens enerett bort etter en bestemt *vernetid*. Etter ål prg 40 regnes denne til 50 år etter utløpet av opphavsmannens dødsår. Etter lov av 2.12.1955 nr. 1 er det gjort en praktisk viktig endring i denne regelen – vernetiden forlenges med 6 år for opphavsmenn som var avgått ved døden før lovens dato, men for hvis verker vernetiden ikke var utløpt (jf. lovens prg 1, 1. ledd). Denne noe besynderlige bestemmelse er kommet inn som en slags kompensasjon for krigsårene og gjør det litt vanskeligere å beregne utløpet av vernetiden. Men det kan være en nyttig påminnelse om at det ofte er problematisk å beregne vernetid, særlig hvis det er en utenlandsk opphavsmann inne i bildet. For et utenlandsk verk oversatt til norsk løper det f.eks. separate vernetider for originalen og den norske oversettelsen.

For det andre er «*lover, administrative forskrifter, rettsavgjørelser og andre offentlige saksdokumenter*» uten vern etter åndsverkloven (jf. ål prg 9). Dette er en praktisk viktig bestemmelse. Når Lovdata i vår tilbyr sin database av lover og forskrifter til landets jurister, tilbyr den altså en database med tekster som mangler opphavsrettslig beskyttelse. Imidlertid er det et par punkter som i denne forbindelse bør understrekes.

For det første kan *bearbeidelser* av de nevnte dokumentene være åndsverk. Et typisk eksempel er referat av dommer. Når dommer trykkes, gjengis det ofte et sammendrag innledningsvis. Dette er gjerne utarbeidet av en redaktør – og redaktøren har opphavsrett til dette sammendraget. Når Lovdata bringer sammendrag av høyesterettsavgjørelser som del av sin base, skjer det etter avtale med redaksjonen i Norsk Retstidende.

For det annet kan *samlinger* av de nevnte dokumentene ha vern qua samling. Enten fordi samlingen er så original at den regnes som et åndsverk i seg selv. Eller – kanskje mer praktisk – samlingen får et «katalogvern» som nevnt foran.

For det tredje er «*offentlige saksdokumenter*» ikke ethvert dokument som finnes i offentlige arkiver, eller som omfattes av offentlighetsloven. Et brev som en person har skrevet til et forvaltningsorgan er knapt et offentlig saksdokument i åndsverklovens forstand, selv om organets svar kanskje er det. Her er det usikre grenser som kanskje kan skape problemer i praksis – særlig fordi grensene sikkert må trekkes forskjellig etter åndsverk- og offentlighetsloven. Men som hovedregel må man

anse alle dokumenter som organet selv har utferdiget som «offentlige saksdokumenter», sikkert også slike som regnes som «interne» etter offentlighetsloven og derfor er unndradd offentlighet.

Endelig har man en særlig regel om «*forhandlinger i Stortinget, i kommunestyrene, i møter av andre valgte offentlige myndigheter eller i rettsaker og på offentlige møter som holdes for å behandle almene spørsmål*» og «*ordskifteprogram i Norsk Rikskringkasting*» (ål prg 22). Disse tekstene – enten de foreligger skrevet, eller man gjør et opptak som man får skrevet ned – kan enhver gjengi offentlig uten samtykke fra opphavsmannen. Verkene er altså ikke uten vern etter åndsverkloven, men det er gjort et formidabelt innhugg i vernet – opphavsmannens enerett til offentliggjøring er eliminert. Fremdeles har opphavsmannen igjen f.eks. sine ideelle interesser.

Så helt lett å identifisere hva slags tekster det her dreier seg om, er det kanskje ikke. Når er f.eks. et møte «offentlig», og når behandler det «almene spørsmål» – to betingelser som samtidig må være oppfylt for at unntaket blir aktuelt. Men selv om slike uklarheter må klebe ved bestemmelsen, er det mange klare og viktige tekster som faller inn under den – tekster som nettopp er viktige sett fra en forskers synsvinkel. Man ser også at bestemmelsen er et supplement til den som er nevnt foran om lover og offentlige saksdokumenter.

5. «Privat bruk»

En svært viktig bestemmelse for forskere, er ål prg 11:

«Når det ikke skjer i ervervsøyemed, kan enkelte eksemplar av et offentliggjort verk fremstilles til privat bruk. Slike eksemplar må ikke utnyttes i annet øyemed.»

Bestemmelsen har en uskyldig historie i ønsket om å la folk skrive dikt i minnebøker o.l. Men i moderne reprografis virkelighet er det f.eks. denne bestemmelsen som – hvis noen – gir rett til å ta fotokopier av artikler m.v.

Bestemmelsen er slett ikke lett å anvende eller beskrive. Utgangspunktet er imidlertid enkelt nok: Når det skjer til eget bruk, kan man fremstille eksemplar av verket – og det gjelder også maskinlesbare eksemplar. En forsker kan altså til privat bruk fremstille et eksemplar av et verk ved å taste det inn eller lese det optisk.

Hovedproblemet knytter seg til å tolke uttrykket *privat*. Dette omfatter naturligvis den rent personlige bruk f.eks. innen familiekretsen, men det har mindre interesse for forskeren. «Privat» omfatter imidlertid også yrkesmessig bruk. F.eks. kan språkforskeren fremstille et maskinlesbart eksemplar av et verk på universitetets sentrale anlegg, analysere og bearbeide det – gjerne lage et arkiv av slike verk for eget bruk.

Flere forhold vil imidlertid frarøve fremstillingen dets preg av å være privat.

Jeg vil anta at selve det forhold at eksemplarene fremstilles i *maskinlesbar form* kan være et argument mot at det foreligger privat bruk – særlig hvis man benytter fremmed hjelp for registreringer, legger verkene inn på et sentralt anlegg med dårlig sikkerhet m.v. Hvis man f.eks. bruker et eget kontonummer, men andre i nærmiljøet har

rutinemessig adgang til vedkommendes arbeidsområde på anlegget, vil det være et argument for at eksemplaret ikke er privat.

Antallet eksemplarer og antallet personer som får tilgang til eksemplaret eller eksemplarene spiller en viktig rolle. Det skal nok lite til for at professoren som gjør sine registrerte eksemplarer tilgjengelig for hovedfagstudenter anses for å gjøre annet enn «privat» bruk av dem. Og hvis f.eks. en annotert grunntekst mangfoldiggjøres systematisk (også i maskinlesbar form), vil det lett anses for å ha overskredet grensen for «privat» bruk.

Prosjektorganisasjonen vil spille en rolle. Hvis f.eks. eksemplarene fremstilles av en eller annen fellestjeneste (som f. eks. Tekstarkivet), så er man helt klart utenfor det private området. Men det samme gjelder hvis man fremstiller eksemplarer som en felles ressurs for et forskningsmiljø, der det er flere personer som samarbeider om et prosjekt o.l. Jo sterkere preg eksemplarfremstillingen får av systematisk og generell utnyttelse, jo lenger fjerner man seg fra privat bruk.

Endelig er det en meget viktig begrensning i retten til privat bruk. Den gjelder bare i forhold til *offentliggjort verk*. Etter ål prg 8 regnes åndsverk som offentliggjort når «det med samtykke av opphavsmannen er gjort tilgjengelig for almenheten». F.eks. vil et brev eller et manuskript *ikke* være offentliggjort – og da er det heller ikke adgang til å fremstille private eksemplarer av det, det er faktisk ikke lov til å skrive ned eller registrere biter av det – heller ikke å sitere det (jf. ål prg 13, 1. ledd).

Med sine begrensninger er retten til privat bruk den viktigste innkrenkningen i opphavsmannens enerett til eksemplarfremstilling. Dagens forskningsorganisasjon, kanskje særlig hvor edb blir tatt i bruk, er likevel slik at det ofte vil oppstå tvil om denne retten kan påberopes: Bildet av den ensomme forsker som på sin terminal taster inn et verk, for senere på sitt lukkede kontor å analysere eller benytte det, er vel forholdsvis urealistisk. Mer vanlig er det vel at forskeren inngår i et miljø hvor flere samarbeider, i alle fall om grunnlagsmateriale. Og i en slik sammenheng blir det lett kunstig å påberope seg retten til å fremstille private eksemplarer.

Det er også sannsynlig at bestemmelsen vil bli noe endret i den pågående revisjon av åndsverkloven, nettopp fordi den teknologiske utviklingen (da særlig av reprografi) har rokket ved forutsetningene for bestemmelsen.

6. Avslutning

Når man vurderer opphavsrettslige bestemmelsers anvendelse på datamaskinbaserte systemer, så er det de velkjente reglene med lang tradisjon i norsk åndsrett som man griper til. Det er kanskje i bunn og grunn ikke noe særlig nytt – forskere som arbeider med maskinlesbare tekster underlegges nøyaktig de samme regler som forskere som arbeider med skrivemaskin eller kulepenn.

Likevel vil det nok i praksis kunne være forskjeller. Og de forskjellene kommer av at en datamaskin – særlig et typisk universitetsanlegg – er forskjellig fra en skrivemaskin. Anlegget gjør det uhyre enkelt å produsere nye eksemplarer, enkelt å spre eksemplarer, bearbeide eksem-

plarer osv. Risikoen for at opphavsmannen skal miste kontrollen med utnyttelsen av sitt eget verk, blir større. Og denne større risikoen fører til at de generelle bestemmelsene kanskje blir påberopt hyppigere, at man har mer oppmerksomhet rettet mot den bruk som skjer ved hjelp av edb enn ved hjelp av kulepenn.

I praksis vil det derfor kanskje nettopp være bruk av edb som gjør at forskere oppdager og finner at man må ta hensyn til de opphavsrettslige reglene. Det kan kanskje føles som en hindring for fri forskning – men i opphavsretten balanseres flere interesser mot hverandre, og man er sikkert ikke uenig i at også opphavsmannens interesse i økonomisk vederlag for eget arbeid og vern av ideelle verdier i eget åndsverk, bør bli tatt vare på. Så får man heller ta opp diskusjonen om gjeldende rett har funnet det riktige balansepunktet.

EDB og kvinner – en statusrapport

Eva Møller

Når vi her bruker spalteplass på en artikkel som ikke direkte behandler de humanistiske fagområder, har det flere årsaker: Både edb-teknologi og kvinneproblematikken kan anses som tverrvitenskapelige områder med relevans også for de humanistiske fagområder – og det samme gjelder sammenhengen mellom dem.

Artikkelen må ses i sammenheng med i hvert fall 2 NAVF-prosjekter, som begge er av interesse for de humanistiske fagområder: prosjektet for humanistisk arbeidslivsforskning, som er under oppstartning, samt opprettelsen av det tverrfaglige sekretariat for kvinneforskning.

Videre er artikkelen utsprunget av den aktuelle debatt og forskning omkring betydningen av ny teknologi for kvinners situasjon, – et arbeid som bl.a. er fremmet gjennom *Handlingsplan for likestilling* og NAVFs sekretariat for kvinneforskning. Artikkelen står i stor gjeld til rapporter fra problemformuleringsseminar på Lysebu april 1982, utarbeidet av *Maren Britt Baadshaug*.

Artikkelens intensjon er å gi et overblikk over situasjonen: hvor står vi? Hvilke utviklingstendenser er i gang? Som naturlig er vil artikkelen på noen områder reise flere spørsmål enn de som kan besvares.

Den overordnede problemstilling utkrystalliseres i 2 hovedområder:

1. Problemkomplekset omkring edb-teknologiens konsekvenser for sysselsettingen. – Vil man oppleve reduksjon i antall arbeidsplasser på grunn av edb-teknologi, og vil denne i så fall ramme de typiske kvinneyrker spesielt hardt? Hvilke konsekvenser får edb-teknologien for jobbinnholdet? Dette er spørsmål som reises i mange sammenhenger, men her spør vi etter forhold som gjør seg gjeldende for kvinner spesielt.
2. Edb-arbeid som kvinneyrke. Vi vil se på både studieforhold og arbeidssituasjon for den gruppe som utfører såkalt kvalifisert edb-arbeid, d.v.s. kvinner med utdanning fra høyskole, universitet e.l.

Konsekvenser for sysselsettingen. – Arbeidsledighet?

I debatten om edb-teknologiens konsekvenser er et viktig tema risikoen for arbeidsledighet. På dette felt (som på de fleste andre i debatten) er der motstridende synspunkter, og forskningsrapporter og utredninger gir ikke noe entydig svar.

Fra noen hold vurderes risikoen for forhøyet arbeidsledighet p.g.a. edb-teknologi som liten. Dette standpunkt kommer f.eks. til uttrykk i utredningen «Økonomiske og sosiale virkninger av ny datateknologi», en rapport fra en tjenestemannsgruppe oppnevnt av Finansdepartementet (NOU 1981:14). Her konkluderer man: «At datateknologien ... skulle føre til omfattende massearbeidsløshet synes lite rimelig å påstå. Det er vanskelig å trekke noen klar konklusjon om datateknologiens innvirkning på den samlede sysselsetting.» (s. 14) Imidlertid tar man forbehold om at datateknologien nok vil medføre omstillinger og endringer i sysselsettingens fordeling mellom næringer og regioner, mellom ulike yrkeskategorier o.s.v., og at datateknologien nok vil

overflødiggjøre såvel som skape arbeidsplasser. For noen vil teknologi-
en gi nye muligheter, mens den for andre vil medføre utstøting og
utesting fra arbeidsmarkedet.

I utredningen stilles spørsmålet om kvinner er en spesielt utsatt
gruppe sysselsettingsmessig sett. Etter en vurdering av kvinners plasser-
ing innen ulike næringsgrupper svares det benektende på spørsmålet,
men det pekes på mulige innvendinger mot konklusjonen:

- dersom sysselsettingen totalt svikter vil kvinnene kunne være de
første som faller ut.
- problemet drøftes bare i forhold til de kvinner som allerede er i
arbeid, ikke i forhold til de som ønsker arbeid.

Tatt i betraktning at arbeidsledigheten generelt har vært voksende
siden rapporten ble avgitt (oktober 1980), samt at en omfattende skjult
arbeidsledighet blant kvinner er blitt dokumentert av de senere års
kvinneforskning, kan det være grunn til å tillegge forholdene noe vekt.

Når mange frykter at kvinner vil bli spesielt hardt rammet av
arbeidsledighet som følge av datateknologi, skyldes det en rekke
forhold som er med på å karakterisere kvinners forhold på arbeidsmark-
edet:

- mange kvinnearbeidsplasser finnes innen områder hvor det satses
sterkt på automatisering, f.eks. kontor, varehandel, bank og forsik-
ring.
- kvinner er typisk ansatt i bunnen av organisasjonshierarkiet.
- kvinner stiller utdanningsmessig dårligere enn menn.
- kvinner har oftere enn menn perioder utenfor arbeidsmarkedet, og
kvinner i arbeid har oftere deltidsarbeid.

En undersøkelse av datateknologi i banker viser hvorledes en rekke
av de ovennevnte forhold spiller inn innenfor denne næring (Halvorsen,
1982):

Knut Halvorsen stiller spørsmålet hvilke nye jobber som kommer til
innenfor bankvesenet, hvilke som forsvinner og hvilke følger en endring
i jobbstruktur og kvalifikasjonskrav vil få for ulike grupper av ansatte i
bankene. I 1982 er 53% av alle ansatte i bankene kvinner, av disse
jobber 30% på deltid mot 2% av mennene.

Opp gjennom 60- og 70-årene har innføring av datateknologi ikke
hatt negative konsekvenser for sysselsettingen i bank, hvor den totale
sysselsetting er blitt tredoblet på 20 år. Imidlertid peker Halvorsen på
faktorer som tyder på at en rekke av de arbeidsoppgaver kvinner har i
banker, både innen kundebetjening og administrative oppgaver, står i
fare for å forsvinne i neste fase av automatiseringen. Dette gjelder
f.eks. jobber som blir automatisert ved innføring av selvbetjeningssys-
temer (automatiske seddelkassetter, uttaksautomater) og integrert
tekst- og databehandling. De resterende jobber, samt tilveksten av nye,
blir mer kvalifiserte og vil kreve en bedre utdanningsbakgrunn eller
omskolering.

Om den utvikling som kan forventes oppsummerer Halvorsen:

«De enkle, rutinemessige oppgaver blir færre mens stabs- og
støttefunksjoner får økende betydning. (...) Med kvinners utgangs-
punkt har de dårligere muligheter enn menn for å tilpasse seg
endringsprosessene. Jobbmulighetene for folk uten særlig utdan-

ning blir generelt dårligere, mulighetene for å jobbe deltid blir i hvert fall ikke bedre. Inngangsporten i bank blir på et høyere nivå enn tidligere (...). På de laveste nivåene blir det stadig vanskeligere å få jobb og avvisningsprosenten øker». (Halvorsen, 1982, s. 30-31).

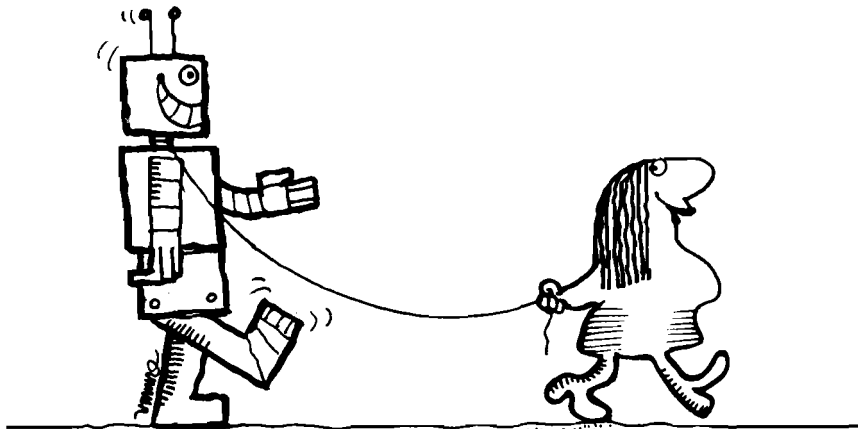
Denne såvel som andre undersøkelser peker på at en stor del kvinner stiller svakt sammen med andre ressursvake grupper. Kvinner, de unge og de eldre (og sikkert enda flere), stiller svakt når det gjelder den omstilling og omgruppering som de fleste er enige om blir en konsekvens av innføring av datateknologi.

Omsorgsfunksjonen som forsvinner

For å vurdere konsekvensene av datateknologien er det ikke tilstrekkelig å undersøke antallet av arbeidsplasser. Det er minst like viktig å spørre om kvaliteten på de arbeidsplasser som blir igjen etter en automatisering – er arbeidet blitt «bedre» eller «dårligere», mer eller mindre tilfredsstillende for dem som skal utføre det?

Det finnes neppe noe entydig svar på spørsmålet om automatisering bidrar til bedre eller dårligere jobber – det vil avhenge av hvilket arbeid det er tale om, hvorledes automatisering tas i bruk og hvilke krav man skal stille til et tilfredsstillende arbeid.

Det er gjort flere forsøk på å definere de krav som må stilles til et tilfredsstillende arbeid. Nesten klassisk i edb-litteraturen er Thørsruds og Emerys sett med psykologiske jobbkrav fra 1970, som legger vekt på behov for innhold og variasjon i jobben, behov for å lære noe, å treffe beslutninger og å oppnå en viss respekt for sitt arbeid. Videre behov for å se sammenheng mellom arbeidet og omverdenen, samt behov for å se at jobben er forenlig med en ønskverdig fremtid.



Kvinner og edb: slik...

(Tegning: Gunnar Johnsen, fra en brosjyre utgitt av Østfold distrikthøgskole).

Ovenstående og liknende sett med spesifikasjoner som i noen tilfeller brukes som arbeidsgrunnlag ved utformingen av edb-systemer er sjelden (eller aldri) kjønnsesifikke – og, vil man formodentlig med en viss rett hevde, gjelder både for menn og kvinner. Men kan man ikke tenke seg at de krav som stilles til et tilfredsstillende arbeid varierer med den ansattes kjønn? Muligens kunne man i tillegg til ovenstående (eller i stedet for noe av ovenstående) sette andre faktorer, som kvinner ville tillegge større vekt, og som kvinner ville prioritere å få tatt hensyn til i forbindelse med utformingen av et automatisert system.

Hos *Maja Lisa Perby* fremheves hvorledes kvinner ansatt i det danske televerk er bekymret for at den omsorgsfunksjon de av og til utøver overfor gamle, ensomme og ulykkelige vil forsvinne ved overgangen til et automatisert system:

«Deras sätt att uttrycka sina farhågor stämmer överens med hur jag hört kvinnliga handelsanställda prata om sitt arbete: de beskriver sin förlorade *sociala roll* med samma saknad och besvikelse som en manlig yrkesarbetare, t.ex. en typograf, kan tala om *yrkeskunskapar* som han inte längre har användning för» (Perby, 1981, s. 45-46).

Tilsvarende uttrykk for ansvar for omsorgsfunksjoner finnes også andre plasser, f.eks. hos kvinner ansatt i de svenske forsikringskasser. (Boman, 1981). (At de kvinnelige omsorgsfunksjoner i mange tilfelle ikke er erkjent som en del av selve arbeidet, og i hvert fall ikke er avlønnert i overensstemmelse med f.eks. en typografs faglige kvalifikasjoner, er en annen side av saken!)

At kvinner har et annet utgangspunkt enn menn og andre krav til ny teknikk tar man konsekvensen av i prosjektet «Kvinnors arbete, teknik och alternativ» (KATA). KATA-prosjektet er et forskningsprosjekt ved



eller slik...

(Tegning: Øystein Reigem).

Kungliga Tekniska Högskolan i Stockholm, hvor man vil finne alternative måter for å anvende teknikk. Man har her som utgangspunkt at kvinners spesifikke kunnskaper og erfaringer, som er knyttet til kvinners stilling i produksjon og reproduksjon, må tas vare på og anvendes som en verdifull ressurs når ny teknologi skal innføres.

Edb som kvinneyrke

Mens vi i foregående avsnitt mest har sett på situasjonen for kvinner som «ofre» for edb-anvendelse, eller som brukere av edb-systemer, skal vi i dette avsnitt vende blikket mot kvinner som til en viss grad har definert seg selv som «subjekter» i forhold til teknologien, de høytutdannede med edb som spesialer, som er med på å utvikle og drive edb-systemene.

Utdannelse

Muligheter for en videregående utdanning innen edb-området finnes det flere av: ved universitetene, ved flere distriktshøgskoler, ved ingeniørhøgskolene samt Norges Tekniske Høgskole.

I overensstemmelse med den generelle likestillingspolitiske målsetting har det gjennom ulike tiltak vært forsøkt å heve kvinneandelen ved edb-studiene. Noen lukkede studier gir kvinnelige søkere tilleggspoeng, og ellers har det blitt utarbeidet studieveiledningsmateriale som spesielt retter seg mot potensielle kvinnelige søkere.

Generelt har kvinneandelen ved de ulike edb-studier gjennom de siste årene vært stigende. Trondheim ingeniørhøgskole (TIH) hører til blant de som kan vise til den største stigning i kvinneandel, fra 18,8% i 1978 til 52% i 1982. – I 1982 er der for første gang tatt opp flere kvinner enn menn på edb-studiet, og edb-linjen tegner seg dermed sammen med kjemilinjens som «kvinnefag» ved TIH.

Arbeidssituasjonen

Vi mangler i dag en samlet oversikt over kvinners plassering innen edb-yrkene når det gjelder lønnsforhold, stillingskategorier, rekrutteringsbakgrunn, typer arbeidsoppgaver o.s.v.

Data fra Storbritannia, USA og Vest-Tyskland viser imidlertid:
– Kvinner er som regel i de lavest betalte yrkesgrupper i dataindustrien.
– Kvinner får lavere lønn selv når de er i samme stillingsgruppe som menn. (Baadshaug 1982 s. 21)

En mindre, norsk undersøkelse antyder noe tilsvarende, nemlig likhet mellom kvinnelige og mannlige systemfolk m.h.t. alder og ansiennitet, men kvinnene tjener betydelig mindre enn mennene (Baadshaug, 1982).

Nærmere undersøkelse av forholdene for kvinner i edb-yrkene er imidlertid i gang. I regi av en norsk dataforening har en del kvinner sluttet seg sammen, og en av målsetningene for denne gruppen er å undersøke forholdene for kvinner i edb-yrkene. Det har vært avholdt en konferanse om temaet, og på NordData 83 er en sesjon avsatt til det samme.

Kvinne-erfaringer – ressurs eller handikapp

Av og til argumenteres det for å få flere kvinner inn i tekniske yrker med den begrunnelse at kvinners erfaring utgjør en viktig ressurs som

kan tilføre området verdifulle egenskaper. Denne tankegang kommer f.eks. til uttrykk i følgende sitat av personalsjef O. Gjerde, IBM:

«... et sterkere innslag av kvinner (vil) trolig ha en positiv innflytelse på miljøet i bedriften. Men forutsetningen for å tilføre bedriftsorganisasjonen noe nytt er at kvinner i denne prosessen tar vare på sin egenart, og at bedriften har vilje – og evne – til å nyttiggjøre seg et sterkere innslag av såkalte «myke verdier» på en konstruktiv måte» (Gjerde, 1982).

Er det imidlertid tilfellet at kvinner representerer andre synspunkter og verdier enn menn innenfor de tekniske fag?

Hos *Elin Kvande* (1982) vises hvorledes kvinnelige studenter ved NTH til en viss grad kan ses å representere et oppgjør med dominerende holdninger blant teknologer. Hun finner f.eks. forskjeller mellom menn og kvinner når det gjelder forholdet til linjevalg, hvor kvinner i høyere grad enn menn velger linjer ut fra interesse, eller begrunner sine linjevalg med at de ønsker å gjøre noe samfunnsnyttig eller noe hvor de kan se mennesket i teknikken. Til forskjell velger menn oftere linje ut fra hensyn til prestisje eller fremtidig karrieremulighet. Kvande mener i sin undersøkelse å kunne se et motstandspotensiale mot noen av de verdier som tradisjonelt har inngått i sivilingeniørens yrke og utdanning, noe hun forklarer med den kvinnelige sosialisering til å ivareta omsorgsfunksjoner på samfunnsnivå såvel som på individnivå.

Samtidig må det fremheves at kvinner ikke nødvendigvis ønsker å skjære bort den tekniske dimensjonen, men snarere ønsker en integrering, slik som det kommer til uttrykk i følgende sitat fra en av de studentene Kvande har intervjuet: «Jag skulle gärne ha ändret på ämnenas innehåll och t.ex. få inn ämnen, som idag skils ut som icke-tekniska, som en naturlig del av de obligatoriska ämnena, det skulle förmänskliga dem» (Kvande, 1982 s. 45). Samme forhold fremheves hos Baadshaug (1982), der det stilles spørsmålsteget ved den måten kvinner tilsynelatende frivillig har glidd inn i oppgaver som angår helheten i et prosjekt eller oppgaver som angår systemets brukertilpasning innenfor større samarbeidsprosjekt, men hvor de like gjerne selv kunne ha tenkt seg å arbeide med mer matematisk orienterte deler av prosjektene. Baadshaug stiller spørsmålet: «I hvor stor grad skal vi forvente at kvinner fører med seg sin «ansvarsrasjonalitet» når de trer inn på nye arenaer? Er dette bare et eksempel på en «ny byrde» kvinner pålegger seg selv?» (s. 19)

Vi ser at kvinner innen edb-yrkene er underlagt en rekke til dels motsattrettede krav og forventninger, fra seg selv og fra omgivelsene. I en likestillingspolitisk sammenheng er det viktig å få klarlagt de krav og forventninger kvinner konfronteres med når de gir seg i kast med mannsdominerte yrker. I tillegg til de konkrete spørsmål vi tidligere pekte på vedrørende lønnsforhold, rekrutteringsbakgrunn og karrieremulighet, ville det f.eks. være interessant å få belyst om kvinner har spesielle roller innenfor edb-yrkene, og hvorfor og hvordan de tildeles disse roller. Videre om forventningen om å ivareta spesielle forpliktelser virker hemmende i en kærrieresammenheng, slik som man kan formode. Mye tyder på at det innenfor edb-yrket ennå gjenstår å finne en avbalansering mellom *likestilling*, i den forstand at kvinner forventes å

utføre tekniske jobber like så godt og på samme måte som menn, og *likeverd*, forstått slik at eventuelt spesielle «kvinnelige verdier» tillegges selvstendig verdi som kan være med på å prege fagets videre utvikling.

Litteratur

Baadshaug, Maren Brit (red.): Kvinner og ny teknologi. Rapport fra problemformuleringsseminar om kvinner og teknologi, Lysebu, 22.-24. april 1982.

Baadshaug, Maren Brit: Betydningen av ny teknologi for kvinners sysselsetting. Notat i tilknytning til midlene til forskning under Handlingsplan for likestilling. 1982.

Boman, Ann: «Omsorg og solidaritet - ohållbare argument?» i *Kvinnovetenskaplig tidsskrift* 1-2:1981.

Gjerde, Oddvin (IBM): «Kan bedriften ta spesielle kvinnehensyn? Hva kan de oppnå med dette?» Foredrag ved konferansen «EDB og kvinner», arrangert av Den Norske Dataforening, 11.11.1982.

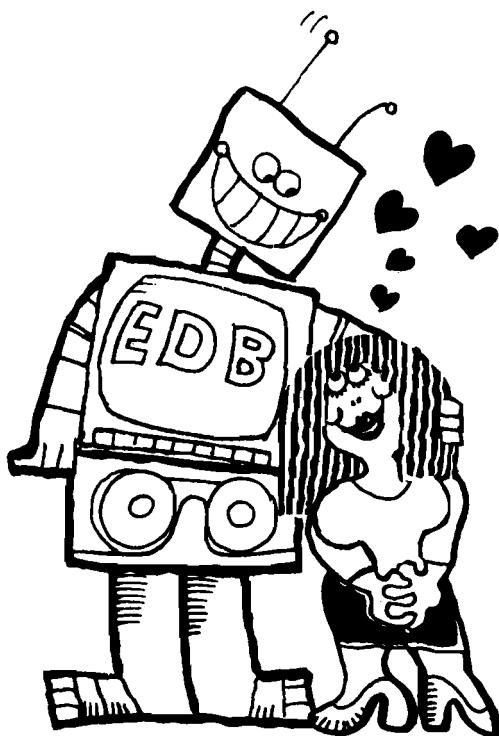
Halvorsen, Knut: Datateknologi i bank. I Baadshaug (red.), 1982.

Kvande, Elin: «Anpassning och protest» i *Kvinnovetenskaplig tidsskrift* 3:1982.

NOU 1981:14. Økonomiske og sosiale virkninger av ny datateknologi. Rapport fra en tjenestemannsgruppe oppnevnt av Finansdepartementet 29. november 1979.

Perby, Maja Lisa: «Kvinnor, datateknologi och arbete» i *Kvinnovetenskaplig tidsskrift* 1-2:1981.

Thorsrud, E. & Emery, F.: Mot en ny bedriftsorganisasjon. Oslo 1970.



(Tegning: Gunnar Johnsen).

LÆREBOKSPRÅK analysert ved hjelp av EDB

Anne Hvenekilde og Anne Golden

De fremmedspråklige elevene i grunnskolen har ofte store problemer med språket i de norske lærebøkene - også etter at de har lært et brukbart dagligspråk og er blitt integrert i norske klasser. Det er særlig lærebøkene i O-fag (orienteringsfag, dvs. naturfag, geografi, historie og samfunnsfag) som skaper problemer, og problemene skyldes først og fremst at elevene mangler bredden i ordforrådet. I prosjektet LÆREBOKSPRÅK ved Nordisk institutt, Universitetet i Oslo har vi analysert språket - og spesielt ordforrådet - innenfor fagområdene fysikk, geografi og historie for 4.-9. klasse og utarbeidet støttemateriell på grunnlag av analysene.

På forskjellige skoler brukes det forskjellige lærebøker, og vi ønsket derfor ikke å binde støttematerialet til en spesiell lærebok eller en spesiell serie. Vi valgte ut flere lærebøker eller deler av lærebøker for hvert av de seks klassetrinnene. Nærmere opplysninger om tekstgrunnlaget kan fås ved forespørsel.

Tekstene fikk vi skrevet inn på edb. Vi vurderte bruk av optisk lesing, men forkastet det fordi det både var mye bilder og varierende bokstavtyper og tekstformat i de enkelte bøkene.

Tekstene innenfor hvert fagområde ble slått sammen for hvert klassetrinn, slik at vi til sammen fikk 18 enheter. Fordelingen av ord innen de forskjellige enhetene var slik:

	FYSIKK	HISTORIE	GEOGRAFI
9 klasse	68.593	42.555	63.731
8 "	67.919	52.606	59.862
7 "	52.058	45.458	58.120
6 "	15.639	37.816	19.301
5 "	8.891	80.152	61.530
4 "	8.449	44.139	38.454
Tilsammen	221.549	302.726	300.998

Det samlede antall ord for materialet var 825.273.

Etter hvert som enhetene var registrert på edb, behandlet vi hver enhet på følgende måte:

- Ved hjelp av programmet OLISTX (fra Det historisk-filosofiske fakultets programpakke) fikk vi sortert grafordene (sekvenser av tegn som står mellom to mellomrom) alfabetisk, samtidig som vi fikk angitt frekvensen. Grafordene ble også nummerert fortløpende.

- b) Vi vurderte listen med graford for å finne fram til homografene. Disse skrev vi inn på en egen fil. I denne vurderingen tok vi hensyn til fagområdene, idet vi regnet enkelte av homografkomponentene for lite sannsynlige i et fagområde (f.eks. *pil* i fysikk), men aktuelle i et annet (f.eks. *pil* i geografi).
- c) Ved hjelp av programmet KWICTX (fra HF's programpakke) kjørte vi ut konkordansene til disse ordene. Dette programmet gav oss det aktuelle ordet midt på linjen og en bit av konteksten både foran og etter. Dette gav oss tilstrekkelig informasjon til å skille de forskjellige homografkomponentene fra hverandre.
- d) Vi ønsket å telle sammen både alle bøyingsformer og alle avledningsformer som hørte til samme rot, og alle disse formene samlet kalte vi et *rotlemma*. For å få ut totalfrekvensen for de enkelte rotlemmaene, kodet vi inn numrene på de grafordene som tilhørte samme rotlemma, eventuelt med utskilling av homografer med annen betydning.
- e) Et spesiallaget program LEMMA som Ivar Fønnes ved HF-data utarbeidet for dette prosjektet (nå innlemmet i HF's programpakke) foretok homografsepareringen og rotlemmatismen som vi hadde kodet inn.
- f) Rotlemmaene ble sortert etter fallende frekvens. Dette gav følgende type liste der vi har totalfrekvensen av rotlemmaet og det valgte oppslagsordet til venstre, og spesifiseringen av hvilke ord som inngår i rotlemmaet og deres frekvenser til høyre:

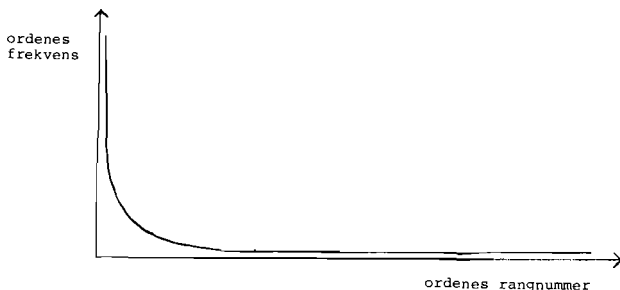
```

64 is      : is(15) isen(46) iset(2) ising(1)
44 kald    : kald(12) kalde(12) kaldere(2)
              kaldest(3) kaldeste(2) kaldt(13)
44 kanal   : kanal(7) kanalen(16) kanalene(9)
              kanaler(12)

```

- g) På grunnlag av de endelige frekvenslistene tegnet vi så frekvenskurver for hver liste.

I alle vanlige tekster er det slik at ordene stort sett faller i to grupper: de er enten svært vanlige eller svært sjeldne. Hvis vi tegner en kurve basert på en oversikt over ordene ut fra deres hyppighet, får kurven alltid omtrent samme form:



På den vertikale akse finner vi ordenes frekvens, og på den horisontale akse ordenes nummer rangert etter fallende frekvens. De ordene som forekommer hyppigst er funksjonsordene (grammatiske ord). Dette er ord som har lite betydningsinnhold i seg selv, men først og fremst markerer hvilket forhold det er mellom de andre ordene i teksten. De andre ordene, som har større betydningsinnhold, har langt lavere hyppighet. Disse finner vi i høyre del av kurven, og de består for det meste av substantiver, verb og adjektiver. Kurven fortsetter langt ut til høyre. I de fleste tekster forekommer omtrent halvparten av alle de forskjellige ordene bare én gang i teksten. Dette betyr at hvis fremmedspråklige elever skal lære alle ordene i en O-fagsbok, vil ordene i 50% av tilfellene aldri dukke opp igjen i den samme boka etter at de er lest og lært.

For hver av de 18 frekvenslistene tegnet vi en slik kurve i relativt stor målestokk. Hver kurve kuttet vi av der hvor den flatet seg ut. Den delen av ordlista som hadde frekvensverdier som falt innenfor den avskårne kurven, bearbeidet vi så videre. På denne delen av lista – som altså inneholdt de mest frekvente ordene – fjernet vi de vanligste funksjonsordene og i tillegg en del andre vanlige ord (f.eks. *gå, komme, stå* o.l.) som vi forutsetter må være kjente for at elevene i det hele tatt skal ha noen mulighet for å lære seg fag gjennom norsk. De ordene som ble igjen, var da vårt råstoff for utarbeidelsen av øvingshefter for fremmedspråklige elever. Dette var til sammen 1515 forskjellige rotlemmer i fysikk, 1459 i geografi og 1917 i historie.

Vi utarbeidet 6 øvingshefter, ett for barnetrinnet og ett for ungdomstrinnet for hvert av fagområdene. Heftene for barnetrinnet la vi opp på den måten at vi skrev forholdsvis korte tekster rundt de emnene som ble behandlet i lærebøkene, men vi holdt oss til de ordene som vi hadde kommet fram til gjennom analysen. Mellom tekstene laget vi oppgaver med øving av de samme ordene, f.eks. kryssord, utfyllingsoppgaver, synonymoppgaver, hyponymioppgaver, sortering av ord ut fra betydning, lagning av sammensetninger og avledninger osv. Heftene for ungdomstrinnet ble utformet litt annerledes. Her består hver leksjon av to deler: ordforklaringer og øvingsoppgaver. Ordene er ordnet slik at de fleste av ordene innenfor hver enhet samler seg omkring ett eller to temaer, men istedenfor lesetekster er det ordforklaringer og eksempler på bruk av ordene i forskjellige sammenhenger.

Gjennom arbeidet med prosjektet er vi blitt opptatt av en rekke problemstillinger som vi kan arbeide videre med ved hjelp av edb og det tekstkorpuset vi har liggende. Noe av dette er vi i gang med, andre ting håper vi å kunne gå løs på seinere:

1. *Sammensatte ord*

I analysen av ordforrådet slo vi sammen frekvensene til alle de ordene vi regnet til samme rotlemma, men innenfor den tidsrammen vi hadde, måtte vi kapitulere overfor forekomstene av disse ordene som del av sammensatte ord. Vi regnet sammen *arbeid*, en *arbeider*, har *arbeidet* osv., men fikk ikke med første del av *arbeidsliv* blant disse. Nå er *Tove Fjeldvig* og *Anne Golden* nær ved å avslutte et arbeid med automatisk rotlemmatisering og vil da som neste trinn arbeide med automatisk deling av sammensetninger.

2. *Lettleste utgaver*

Noen av tekstene i vårt korpus er også kommet i parallellutgaver bearbeidet for elever med lesevansker. Vi har også fått skrevet inn disse tekstene på edb og arbeider nå med en kvantitativ sammenliknende analyse mellom originaltekstene og de omarbeidede tekstene. Disse kvantitative analysene er utgangspunktet for kvalitative analyser på spesielle områder som de kvantitative analysene avdekker som interessante.

I arbeidet med de kvantitative analysene bruker vi standardprogrammer som gir forskjellige typer statistiske opplysninger som f.eks. antall forskjellige ord og antall løpende ord, gjennomsnittlig ord- og periode-lengde. I både de kvantitative og de kvalitative analysene har vi god nytte av programmer som tilrettelegger teksten på forskjellige måter.

3. *Lærebokspråket fra klassetrinn til klassetrinn*

Vi har utført kvantitative sammenlikninger av bøker av samme forfattere innenfor samme fag, men på forskjellige klassetrinn for å finne eventuelle forskjeller i språket fra klassetrinn til klassetrinn ut over forskjellene i ordforrådet. Denne analysen må følges opp av nærlæsning av de delene av tekstene som ifølge de kvantitative analysene ser interessante ut.

4. *Forholdet mellom «fagord» og andre ord*

I en separat undersøkelse har vi fått lærere i skolen til å krysse av de ordene på våre endelige lister (dvs. de ordene som kommer med i støttematerialet) som de vil regne som «fagord» innenfor det aktuelle området (fysikk, geografi eller historie), og som de derfor ville forklare for norske skoleelever. De ordene som ikke blir forklart av læreren, er det spesielt viktig å få godt representert i støttematerialet. De er forutsatt kjent av de norske elevene, men kan være like ukjente som fagordene for de fremmedspråklige elevene. Det ser ut til at fagordene utgjør en større andel i fysikktekstene enn i historietekstene, og at geografitekstene inntar en mellomstilling.

5. *Fast og løst sammensatte verb*

Fordi vår analyse bare hadde med enkeltord, måtte vi gi opp overfor et viktig og vanskelig område for utlendinger som skal lære norsk: løst sammensatte verb (f.eks. *ta på, gi opp, henge i* osv.). Siden det er et begrenset antall ord som brukes som verbalpartikler i slike verb (her: *på, opp, i*), lar det seg gjøre å få kartlagt slike verb ved å kjøpe konkordansprogrammer med utgangspunkt i mulige verbalpartikler. Det er et omfattende arbeid å gjennomgå konkordansene og skille ut de løst sammensatte verbene. Uten tilrettelegging ved hjelp av edb ville det være uoverkommelig å foreta en kartlegging av slike verb innenfor et korpus som vårt, men *med* edb håper vi det skal ligge innen rekkevidde.

6. *Kollokasjoner*

Som språklærere synes vi også det ville være interessant å undersøke kollokasjoner, dvs. ord som forekommer ved siden av hverandre eller nær hverandre i teksten. I arbeidet med konkordansene ble vi slått av hvor ofte ikke bare ord som er del av samme faste uttrykk (idiom) forekom sammen, men også hvor ofte f.eks. de samme to ordene var sideordnet eller et spesielt adjektiv var knyttet til et spesielt substantiv. Det ville være interessant å beskrive typer og frekvenser på dette feltet

og vurdere «ensidige» og «tosidige» kollokeringstendenser. La oss tenke oss at vi har to ord sideordnet med konjunksjonen *og*: *A og B*. Da kan det være slik at de fleste tilfellene av *A* forekommer i sekvensen *A og B*, men at de fleste tilfellene av *B* forekommer i andre kontekster. I så fall bør *A* læres som del av *A og B*, men *B* bør læres separat først og så eventuelt på et seinere stadium som del av *A og B*.

Arbeidet under punkt 4 nærmer seg avslutningen, punkt 1, 2 og 3 er godt i gang, punkt 5 har vi så vidt snust på og punkt 6 er på idéstadiet.

Prosjektet LÆREBOKSPRÅK

Prosjektledere: Anne Hvenekilde og Anne Golden

Tid: 1981-1982

Sted: Nordisk institutt, Universitetet i Oslo

Finansiering: Kirke- og undervisningsdepartementets skoleforskningsmidler (kap 252)

EDB i fornminneforvaltning ved et arkeologisk museum

Trond Løken

Som ett av de 5 arkeologiske regionmuseer i Norge er Arkeologisk museum i Stavanger tillagt ansvaret for forvaltning av kulturminner eldre enn reformasjonen (1537) i Rogaland fylke, med unntak av kirker, klostre, borger og byanlegg. Museet har også ansvar for registrering av slike kulturminner for det økonomiske kartverket og ivaretagelse av alle funn fra samme periode. Dette betyr at museet har et register med beskrivelse av over ca. 7000 fornminner som omfatter minst 25000 objekter samt 10000 museumsnummer, med flere hundre tusen gjenstander.

Som forvaltningsinstitusjon skal disse datamengdene brukes til å foreta så riktige avgjørelser som mulig i museets deltagelse i samfunnets planleggings- og utbyggingsprosess. Som forskningsinstitusjon skal de samme datamengder danne grunnlaget for kunnskapsproduksjon om forhistorien i Rogaland, og som formidlingsinstitusjon skal opplysninger om funn og fornminner og forhistorie gis til den interesserte lokalbefolkning.

Behovet for å skaffe seg EDB-registre for funn- og fornminner har vært følt lenge. De første forsøk ble gjort rundt 1970, men datidens teknologi var ikke tilfredsstillende idet informasjonen i liten grad egner seg for koding. Behov innen forvaltningsvirksomheten, især til effektiv framfinning av fornminnedata, bedre ajourføringsrutiner og lettere tilgjengelighet for eksterne brukere, førte til at når museet skulle gå inn i EDB-alderen var det fornminneregisteret som ble prioritert.

Problemene med et manuelt fornminneregister er 2-delt:

1. *Internt* eksisterer registeret for fornminner, registrert for økonomisk kartverk i manuelle trykte registre, ordnet etter kommune og gård samt avmerkninger på ØK-kartblad og flybilder. Dette gir bare 2 innganger til datamassen. Især savnes mulighet for oversikt over hvilke typer fornminner som finnes i bestemte områder. *Resultatet* er at avgjørelser om endringer/frigivelse i planområdet må tas på et dårligere grunnlag enn datamassen egentlig gir mulighet til.
2. *Eksternt*, hos planleggende organer, finnes trykte registre og avmerkning på kartblad, men uten mulighet for sikker kobling mellom kartblad og register. Dessuten finnes kart over verneverdige områder hvor fornminner er gruppert i 3 nivåer etter verneverdighet. For de eksterne brukere av våre data er disse manuelle registre vanskelig tilgjengelige, og den informasjon som finnes blir erfaringsmessig lite brukt i den tidlige planleggingsprosessen. *Resultatet* av dette er at mange konflikter oppstår når museet får planene til uttalelse. Det blir behov for endringer i planene eller utgravinger dersom endringer ikke er mulig å få til. Fornminner blir også ødelagt.

Eksempelvis behandlet AmS i 1980 208 offentlig utarbeidete planer, fig. 1. Mindre enn halvparten (100) angikk ikke fornminner. Av de øvrige 108 kunne bare 21 godkjennes uten videre, mens 87 (42%)

representerte problemer på én eller annen måte. 26 planer måtte omarbeides med det dette innebærer av ressursforbruk. Ved 24 planer måtte museet foreta undersøkelser. I de siste årene har slike undersøkelser kostet samfunnet ca. 3/4 mill. kroner bare i Rogaland. 20 planer, hovedsakelig nydyrkingsplaner, kunne ikke godtas. Ved 17 planer, også hovedsakelig nydyrkingsplaner, var fornminner ødelagt i forbindelse med at planene ble virkeliggjort fordi museet ikke, eller for sent, fikk tilsendt planene.

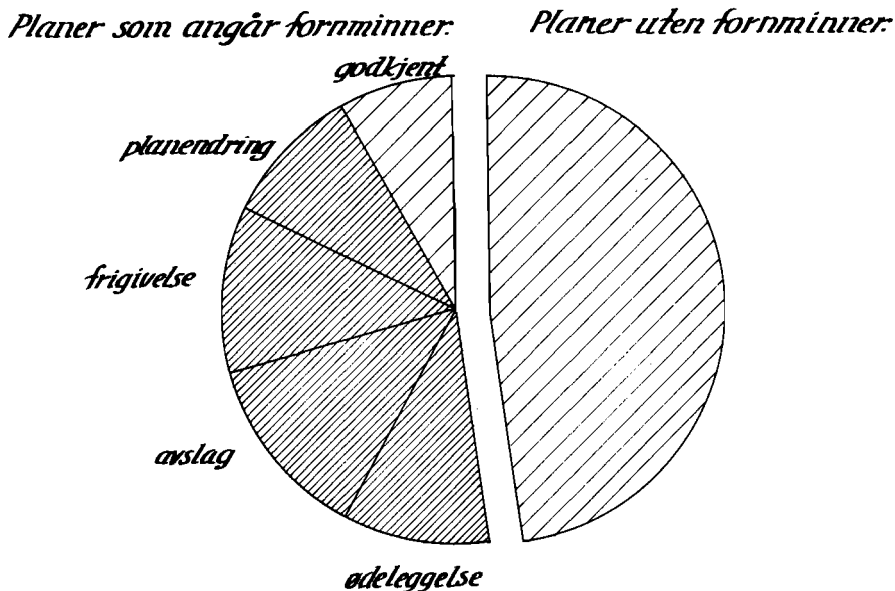


Fig. 1

Konklusjonen må bli at betydelige museale og andre ressurser skulle kunne spares dersom fornminnedata var lett tilgjengelig og ble brukt tidlig i planprosessen. Dessuten ville det kunne bevares en del flere fornminner for ettertiden.

I 1980 hørte vi på AmS om TEDAK (Temadelt datakataster: et EDB-system som kombinerer kartinformasjon og registerinformasjon), og dette ga støtet til at museet ba Rogalandsdata og Rogalandsforskning om hjelp til å utvikle et EDB-system som kunne effektivisere og forbedre vårt planarbeid og gjøre våre data mer tilgjengelige for fylket, kommuner og andre planleggingsinstanser. Så langt er et fornminneregister (Amis I) utviklet og operativt. Dette er bare en del av de data som museet bruker i planarbeidet. Men det er den viktigste del for planleggingsinstansene, idet kjente, synlige fornminner er det som disse uten videre har mulighet til å kunne ta hensyn til.

Før jeg går nærmere inn på Amis I vil jeg kort vise hvilke andre informasjoner som brukes i museets planarbeid og deres plass i

arbeidsprosessen, fig. 2.

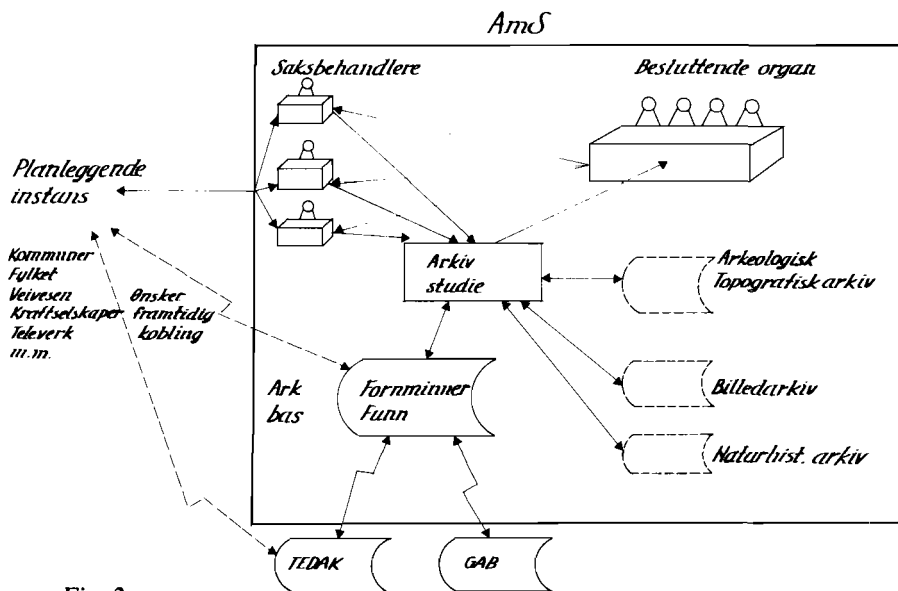


Fig. 2

En plan til museet går til saksbehandler som skaffer fram den informasjon som finnes. Denne er delt på 5 registre/arkiv hvor en må gjøre i alt 9 oppslag.

1. Fornminneregistret for ØK, flybilde og eldre fornminneregistre (3)
2. Funnregister (1)
3. Arkeologisk-topografisk arkiv og reguleringsplanarkiv (2)
4. Naturhistorisk arkiv (1)
5. Billedarkiv (2)

Det grunnlagsmateriale som er fremskaffet, vurderes i et fornminneutvalg. Deretter må saksbehandler i mer enn halvparten av tilfellene på befaring i planleggingsområdet, for å kontrollere eventuelle arkivopplysninger, lete etter fornminner og vurdere disse i terrenget. Deretter foretas en endelig vurdering/avgjørelse i fornminneutvalget og saken returneres til planleggende instans.

Med et totalvolum på 300 saker, vil en effektivisering av datafangsten bety mye for ressursforbruk i museets planarbeid.

Som antydnet på fig. 2 tenker vi oss et system hvor all denne informasjonsinnsamling skal kunne foregå ved hjelp av en terminal. *Fornminneregisteret* er laget og operativt for ca. halvparten av fylkets kommuner, det er koblet til GAB (register for opplysninger om Grunneiendom, Adresse med eieropplysninger og Bygning.) Mulighet for kobling til TEDAK er ivarettatt.

arbeidsprosessen, fig. 2.

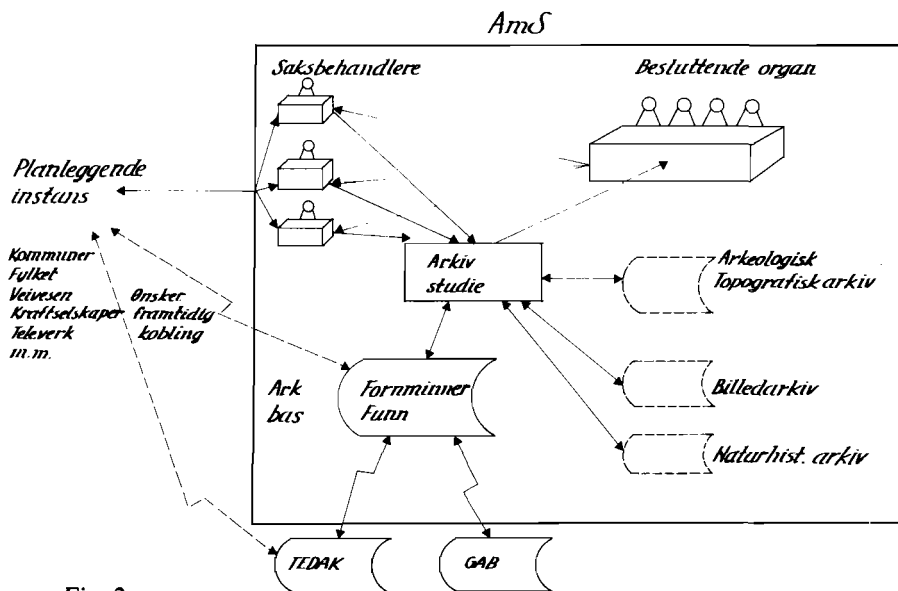


Fig. 2

En plan til museet går til saksbehandler som skaffer fram den informasjon som finnes. Denne er delt på 5 registre/arkiv hvor en må gjøre i alt 9 oppslag.

1. Fornminneregistret for ØK, flybilde og eldre fornminneregistre (3)
2. Funnregister (1)
3. Arkeologisk-topografisk arkiv og reguleringsplanarkiv (2)
4. Naturhistorisk arkiv (1)
5. Billedarkiv (2)

Det grunnlagsmateriale som er fremskaffet, vurderes i et fornminneutvalg. Deretter må saksbehandler i mer enn halvparten av tilfellene på befaring i planleggingsområdet, for å kontrollere eventuelle arkivopplysninger, lete etter fornminner og vurdere disse i terrenget. Deretter foretas en endelig vurdering/avgjørelse i fornminneutvalget og saken returneres til planleggende instans.

Med et totalvolum på 300 saker, vil en effektivisering av datafangsten bety mye for ressursforbruk i museets planarbeid.

Som antydnet på fig. 2 tenker vi oss et system hvor all denne informasjonsinnsamling skal kunne foregå ved hjelp av en terminal. *Fornminneregisteret* er laget og operativt for ca. halvparten av fylkets kommuner, det er koblet til GAB (register for opplysninger om Grunneiendom, Adresse med eieropplysninger og Bygning.) Mulighet for kobling til TEDAK er ivaretatt.

Funnregister: Brukerkravspesifikasjon er utarbeidet, slik at systemarbeidet forhåpentligvis kan ta til i løpet av vinteren.

Arkeologisk topografisk arkiv: Vi tenker oss avfotoografering av dokumentene og bruk av et framfinningssystem som Kodak har utviklet (ITM 150).

Naturhistorisk arkiv og Billedarkiv: Vi har ikke gått særlig mye inn i denne problematikken ennå, men disse arkivene bør være med i en helhetsløsning.

Når vi har våre EDB-registre for fornminner og funn operative, tenker vi oss at en naturlig saksgang i planarbeidet vil være at planleggende instans bruker disse registrene som grunnlagsmateriale på linje med annet grunnlagsmateriale som setter begrensninger for utnyttelse av et område. Andelen konflikter mellom fornminnevern – arealdisponering, burde kunne reduseres merkbart.

Vi skal så se litt nærmere på fornminneregisteret Amis I. For tiden er programmet og en datamengde på ca. 7 Mbyte knyttet til en ND 10S på Rogalandsdata. En multiplekser gir oss mulighet til å bruke 4 terminaler på 1 fast linje til Rogalandsdata. På sikt vil det være naturlig å anskaffe egen datamaskin, slik at vi kan få flere brukere/bruksområder ved museet. Men en fast linje til Rogalandsdata vil fortsatt være nødvendig av hensyn til sammenkobling av funnregisteret (Amis I) og TEDAK, og for å ha tilgang til den del av museets data som fortsatt vil bli lagret ved Rogalandsdata.

Amis I er et interaktivt system hvor vi kommuniserer med systemet i dialogform. En enkel grafisk framstilling, fig. 3, av systemet viser at det består i registrering, lagring, redigering og gjenfinning som hovedkomponenter. Dertil hentes informasjon fra GAB og det er gitt mulighet for tilknytning til en forvaltningsmessig kartdatabase (TEDAK).

Det etablerte systemet har 5 hovedfunksjoner, fig. 4.

Registrering av dokumenter (fig. 4:I) foregår direkte på terminal, på skjema (fig. 5) som hentes fram for hver ny registrering. Skjemaet er 2-delt. I øverste del legges data i forhåndsdefinerte felt. Her er det lagt inn viktige informasjoner som alle fornminner har. Kommune, gnr. og bnr. er valgt som stedsidentifikasjon fordi den er stabil og vel kjent. Gnr./bnr. er repeterbar for de tilfeller hvor *ett* fornminne ligger på flere gårder/bruk. Likeledes er fornminnets art, antall og datering repeterbar fordi et fornminneområde kan inneholde flere typer fornminner.

Skjemaets *nedre del* er basert på verbal beskrivelse i fritekst. Det er, i fast rekkefølge, lagt inn en rekke ledeord som normalt brukes ved fornminneregistrering. Disse kan uten videre byttes ut dersom andre typer kulturminner krever det. I denne delen er det ingen begrensninger på informasjonsmengden.

Registrering av data direkte til datamaskin eliminerer en rekke av de vanlige feilkilder som kan oppstå når datatilretteleggelsen må skje i flere faser.

Lasteprogrammet (fig. 4:II) sørger i tillegg for å søke etter feil i den fast formaterte delen, gi feilmelding og legge dokumentene som inneholder feil på egen feilfil. Disse rettes så opp i editor før de kan lastes til databasen.

Redigeringsprogrammet (fig. 4:III). Fornminnedata er dynamiske og

endres f.eks. ved utgravninger, skade ved bruk, eller ved at forskningen gir oss et annet syn på hva fornminner er. Databasen må derfor til enhver tid kunne oppjusteres ved hjelp av redigeringsprogrammet. I den faste, formaterte delen er det også mulig å endre eller tilføye oppslagsord. Det eneste som ikke kan endres er Flyfoto/reg.nr. som er hovedidentifikasjon i systemet.

Hentprogrammet (fig. 4:IV) henter opplysninger om gårdens navn, eier og adresse fra GAB, og disse data tas automatisk inn i utskriftprogrammet. Utskrift til eksternt bruk (fig. 6) er organisert med vekt på gårdsnavn, gnr./bnr. og eier. I denne utskriften kommer interne opplysninger ikke med. Det er planlagt å utvikle et program som med regelmessig mellomrom (2-4 ganger i året) undersøker om noen fornminner har fått nye eiere, slik at museet kan sende disse registreringspapirer, kart og informasjonsmateriale. Med den tiltakende mobilitet i samfunnet er dette eneste måte å sørge for at grunneierne har kjennskap til sine fornminner, slik at de ikke ødelegges av mangel på kunnskap.

Det kartmateriale som bør sendes ut i denne forbindelse, må, når koplingen til TEDAK forhåpentligvis er gjennomført, være eiendoms-kart med inntegnede fornminner, eventuelt kombinert med et oversikts-kart.

Utskriftsprogrammet (fig. 4:V) er utformet slik at informasjon fra databasen skal være lett å få tak i, ved at en i dialogform finner fram til bestemte fornminner eller fornminner på bestemte gårder.

Søkeprogrammer er til nå utarbeidet for hovedindikasjon, sted, fornminnets art og datering. Det vil for museets planarbeid og forskning på fornminnematerialet også være ønskelig med søkeprogrammer som kombinerer de nevnte søkekriterier.

Amis I er laget med tanke på kobling til TEDAK. Det har i en prøvekommune vært gjort forsøk for å se hvordan dette kan fungere. Fornminnene i halvdelen av Strand er digitalisert. Vi kan få tegnet ut fornminner i hele kommunen på et kartblad (fig. 7), og for en eiendom. Ved å peke på fornminner på skjermen, kan vi få listet ut de opplysninger som ligger i Amis I. Vi kan også oppgi fornminneidentifikasjon og få tegnet ut kart, i ønsket målestokk, over området omkring fornminnet. Likeledes kan vi få tegnet utbredelseskart over bestemte typer fornminner hvor vi selv har valgt den kartinformasjon som skal med.

En videre utvikling av Amis I med kobling til TEDAK, hvor fornminneregisteret plasseres i et totalt geodatasystem, vil etter AmS' oppfatning gjøre at informasjon om fornminner vil være lett tilgjengelig nær sagt uansett hvilken innfallsvinkel en har i planleggingsprosessen.

AMIS

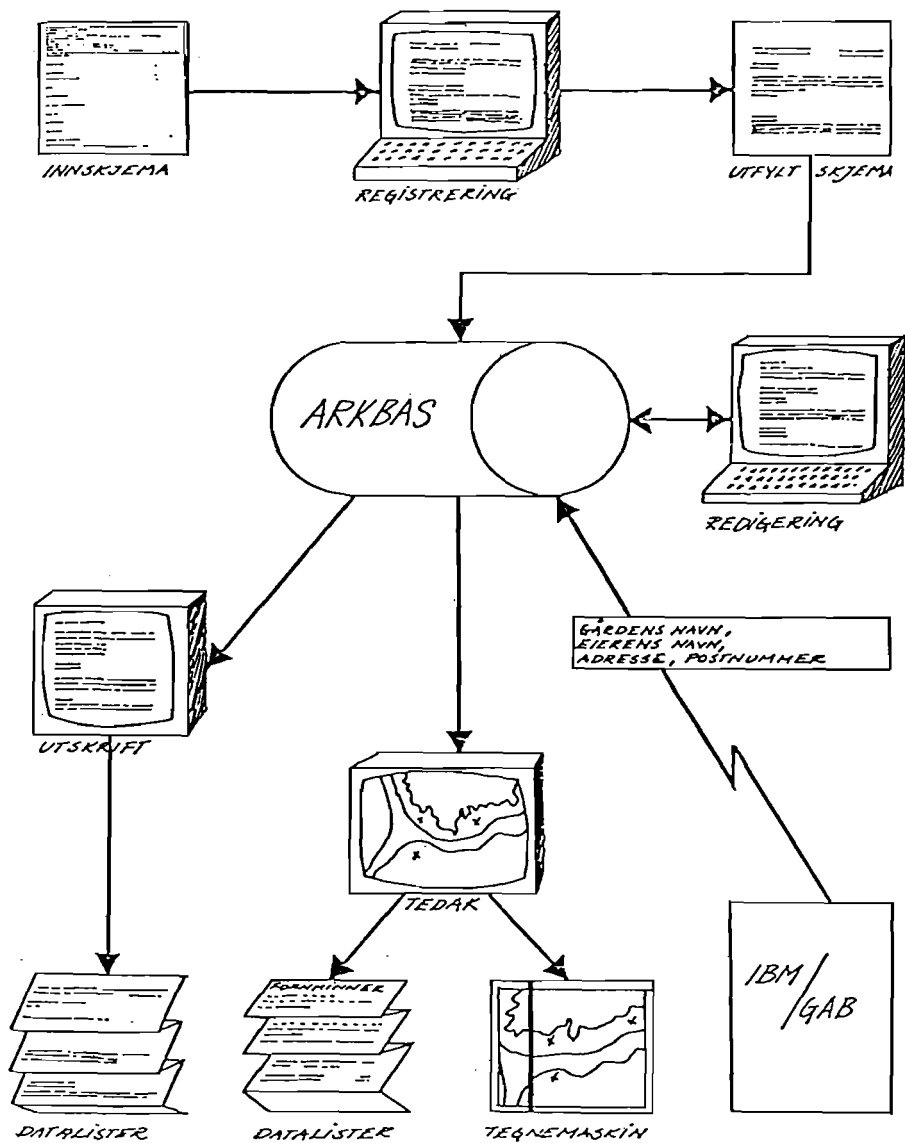


Fig. 3

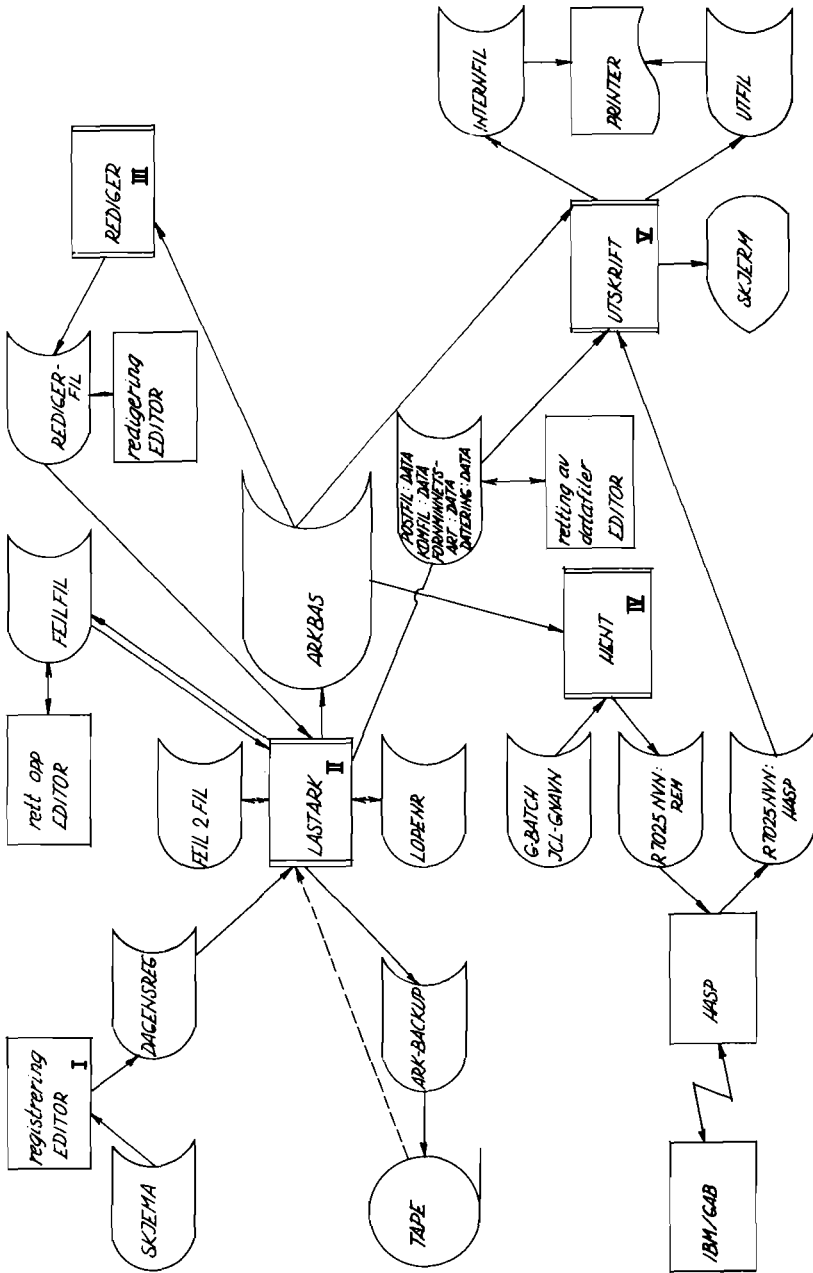


Fig. 4

.....
 FLYFOTO: _____ REGNR: _____
 KOMMUNE: _____
 GNR: _____ BNR: _____ BRUKER: _____ POSTNR: _____
 LOKALITETSHAVN: _____
 ØK-KART: _____
 REG. AV: _____ DATO: __/__/____
 KONTR AV: _____ DATO: __/__/____
 FORM. ART: _____ ANT: _____ DATERING: _____

 TERRENGETS ART:

ORIENTERING:

UTSTREKNING:

FORNMINNEBESKRIVELSE:

FUNN:

UNDERSØKELSER:

NAT.HIST. UNDERSØKELSER:

FOTO NR:

LITTERATUR:

TRADISJON:

INTERNE OPPLYSNINGER:

Fig. 5

GÅRDENS NAVN: Nag
GNR: 41 BNR: 11 Strand
GÅRDENS EIER: Nag Tor T
BRUKER :

FLYFOTO: 3720 F4 NR: R26
ØK: AN-024.5.1

REG. VED : O. Espedal DATO: ?/07-71
KONTROLL VED: DATO: / -

FORNMINNETS ART: Gravreys ANTALL : 2
DATERING :

FORNMINNETS ART: Helleristning ANTALL : 1
DATERING : Bronsealder

LOKALITETSNAVN:

TERRENGETS ART: SV-ventd berg mot fjorden. Noe småbjørk og einer.
Svaberg mot sjøen. Vid utsikt mot V og S.

ORIENTERING: Ca 200 m S for Solbakk-krysset, ca 130 m SV for Rv. 13.

UTSTREKNING: NV-SØ ca 60 x 30 m.

FORNMINNEBESKRIVELSE: Lengst i NV:

1) Helleristningsfelt. For det meste klart markert. Ligger på fjell som heller ca 30 grader mot N. Orientert Ø-V. Består av minst 29 skipsfigurer med kjøel, reling og stevner. 7 av skipene har mannskapsstreker. L. 0,48 - 1,6 m. Videre finnes 7 konsentriske sirkler med d. 15-20 cm, en spiral med d. ca 18 cm og ei skålgrop med d. ca 6 cm. Hoggesporene er 2-3 cm breie og inntil 1 cm djupe. Figurene er nå raudmåla. Feltet er ca 9 x 4 m stort.

Ca 13 m Ø for 1 og på fjell:

2) Rundreys. Klart markert, men mindre godt synleg. Langs kanten lyng, einer og bjørkekratt. Mye stein syner i midten. Delvis restaurert i 1971. D. ca 10 m, h. ca 0,7 m.

Ca 18 m SSØ for 2 og på fjell:

3) Rundreys. Delvis klart markert, men mindre godt synlig. Lyng, eine og bjørkekratt langs kanten. I SV-kant en stor stein ca 1,3 m h. Kratter i midten, d. ca 3 m, dybde ca 0,7 m. Røysas d. 10 m, h. ca 0,7 m.

LITTERATUR: AmS Top.ark. u/Nag: Innb. v/Jan Petersen, 6/10-23 og 5/11-28.

AmS Top.ark. u/Nag: Innb. v/Jone Johnsen, 29/8-69.

Eva og Per Fett: Sydvestnorske helleristninger, s. 25-26.

Fig. 6

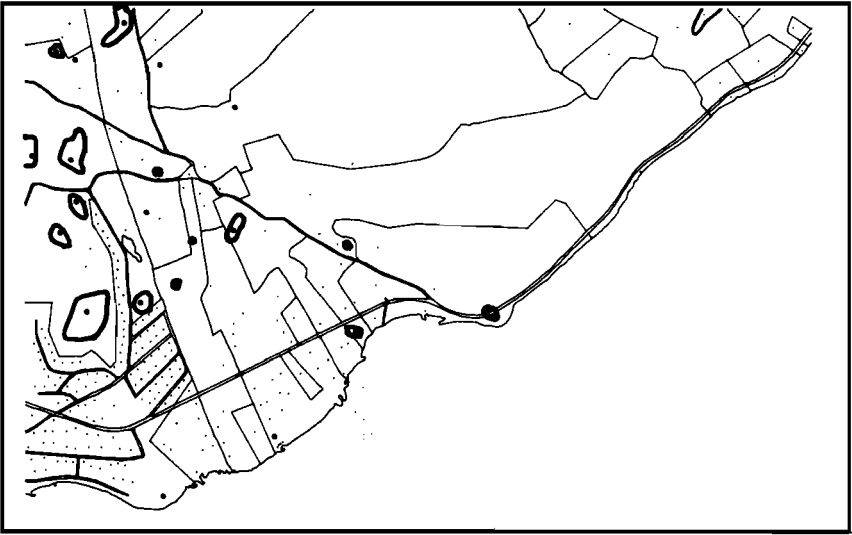


Fig. 7

Edb og ordbøker

Nokre prosjekt innan allmenn leksikografi og terminografi

Håvard Hjulstad

Det er mange sider ved leksikografisk arbeid (ordboksarbeid) som eignar seg godt til databehandling. Grunnlagsmaterialet er i regelen svært stort (t.d. «heile» den nyare norske litteraturen). Resultatet – den ferdige ordboka – skal gjerne ordnast alfabetisk, og det er ei vanskeleg oppgåve (stikkprøver syner at det finst mange alfabetiseringsfeil i mest alle ordbøker som er alfabetiserte manuelt). Enda viktigare er det kanskje at dei ulike leksikografiske presentasjonsformene alle har det felles at dei er strengt formaliserte både når det gjeld oppbygging av ordartiklane og når det gjeld detaljutforming av opplysningane.

Når tida er eg engasjert ved to av dei institusjonane her i landet som står sentralt i leksikografisk arbeid: Norsk leksikografisk institutt (NLI) ved Universitetet i Oslo arbeider med allmennspråkleg leksikografi. Norsk termbank (NoTe) ved Prosjekt for datamaskinell språkbehandling ved Nordisk institutt ved Universitetet i Bergen arbeider med fagspråkleg leksikografi – terminologi/terminografi.

Allmenn leksikografi og terminologi/terminografi

Det er ei vanleg – og naturleg – misoppfatning at terminologi og allmenn leksikografi er nett det same, berre at ordtilfanget ikkje er likt. Det er laga mange fagordbøker, både her i landet og i andre land, som byggjer på allmennleksikografiske og ikkje på terminologiske prinsipp. Dette skiljet har noko å seie for databehandlinga også, og eg skal seie nokre ord om det som særmerkjer dei to greinene.

Utgangspunktet for allmennordboka er «ordet»; utgangspunktet for den terminologiske ordboka er «omgrepet». Eit døme gjer kanskje dette litt klarare: Om ein slår opp i ei større allmennordbok på verbet «lodde», får ein greie på at det finst minst to slike verb. Det eine (I) finst i «lodde ut kaker» (1). Det andre verbet «lodde» (II) har minst to tydingar: som i «lodde djupna i vatnet» (2) og «lodde saman to metallstykke» (3). Det som knyter saman 2 og 3, og som held 2 og 3 skilde frå 1, er ordhistoria – opplysningar som er knytte til ordforma, ikkje til tydingsinnhaldet. I ei terminologisk ordbok måtte tankegangen vere slik: Tydingsinnhalda 1, 2 og 3 uttrykkjer (minst) tre skilde omgrep. På innhaldssida er det ikkje klare grunnar til å kople to av dei saman. Resultatet må bli tre skilde ordartiklar – eller «postar» i dataspråket.

Når ein hengjer på eit eller fleire språk til, ser ein verknaden av dette. På engelsk kan ein for «lodde» finne t.d. 1 «raffle», 2 «sound», 3 «solder». Omgrepa 2 og 3 er uttrykte med to heilt ulike termar. I ei ordbok med engelsk utgangspunkt ville ein aldri finne på å fortelje om 2 og 3 på same stad.

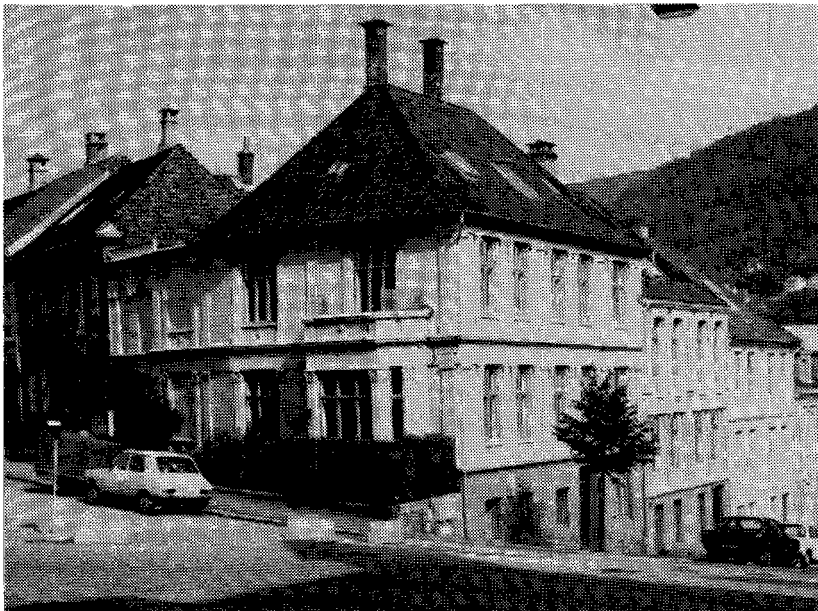
I fleirspråklege ordbøker talar ein om «utgangsspråk» og «målspråk». Ei terminologisk ordbok har ideelt sett ikkje noko «utgangsspråk», berre eit eller fleire «målspråk». Utgangspunktet er tydingsinnhaldet – omgrepet. Det kan vere uttrykt i definisjon(ar) på eit eller fleire språk. Så kan ein fortelje kva dette kan heite på ulike språk. Og under

same posten kan ein fortelje om alle dei termene på kvart språk som er nytta for omgrepet. Ordboka kan ein sortere på termene i eit av språka eller på anna vis (fagkode e.l.). Ein databehandlar blir fort glad i ei slik ordbok, for ho passar lett inn i eit «reint og pent» skjema.

Norsk termbank

Det er slike ordbøker vi arbeider med ved NoTe. Ikkje så langt frå sanninga er det å seie at det er *ei* slik ordbok vi arbeider med ved NoTe. Denne ordboka inneheld no om lag 40000 postar («ordartiklar») og om lag 135000 termar på ei rekkje språk, først og fremst norsk (bokmål og nynorsk), engelsk, tysk og fransk. Ordboka blir stadig oppdatert med at nytt materiale kjem til. Ho omfattar mange fagområde, først og fremst tekniske fag, men det aukar på med «mjukare» fagområde også. Vi blir mata med data frå Rådet for teknisk terminologi (RIT), Norsk språkråd, Norges Standardiseringsforbund, Universitetsforlaget og andre kjelder og samarbeidspartnarar.

Enno er ikkje denne ordboka kopla saman i ei eining i form av ein einskild fil eller ein einskild database. Ho skal nok koplast saman til ein database – tida skulle vere moden for det no. Men på utarbeidingsstadiet for einskildprosjekt er det mest tenleg med småfiler og enkle programsystem for eit enkelt format. Da kan ein greie seg langt i det daglege arbeidet utan hjelp av edb-ekspertar på «bitfiklingsnivå» og på databaseutviklingsnivå.



*Norsk Termbank holder til i Strømg. 53.
(Foto: Rune Midtvedt).*

Det er eit slikt enkelt opplegg vi har laga oss ved NoTe. Formatet har vi utvikla med sideblikk til andre termbankformat. Programsystemet har vakse fram av seg sjølv etter som ulike behov har meldt seg.

Formatet byggjer på prinsippet i den terminologiske ordboka. Termar på alle språk er på same nivå, utan tanke om «utgangsspråk» og «målspråk». Dei er alle knytte til omgrepet, som er verbalisert i ein (eller stundom fleire) definisjon(ar). Til kvar term er det mogleg å knyte ulike opplysningar, slik som ordklasse/kjønn, uttale osv. Til posten kan ein også knyte opplysningar om fagområde osv. Formatet er svært fritt, slik at det ikkje er vanskeleg å passe inn ordboksprosjekt av ulikt omfang og ulik karakter.

Programma har «vakse fram av seg sjølv» (stort sett med programmeringshjelp av meg sjølv), men vi har freista å passe oss for å skrive ad hoc-program («bruk-og-kast-program»). Dei fleste programma er så fleksible at dei kan brukast på alle dei ulike typane av data vi har. Dei er for det meste skrivne i programmeringsspråket Fortran-77 på Univac-1100 ved Universitetet i Bergen.

Det vil føre for langt å fortelje om «alt» vi kan gjere med data, men dei fleste programma kan grupperast slik: 1) Feltgenerering. Maskinell eller maskinstøtta generering av nye felt, t.d. sorteringsfelt, løpenummerfelt. 2) Feilfinning. Det er program som finn ulovlege feltkodar, ulovlege eller usannsynlege teiknsekvensar osv. 3) Sortering. Data kan sorterast på det feltet ein måtte ønskje. 4) Utskrift. Ein kan velje mellom fleire utskriftsformat. 5) Til trykk. Data kan trykkjast med hjelp av fotosats eller i offset frå original skriven ut med skjønnskriver. 6) Søking. Ein kan plukke ut postar etter ulike kriterium, eventuelt rette dei opp og leggje den retta versjonen tilbake. 7) Konkordans. Ein kan lage konkordansar over innhaldet i gitte felt, eventuelt samsortert med termene i eit gitt språk.

Sidan NoTe starta i 1979, har vi m.a. vore med på desse prosjekta som har resultert i trykte ordbøker:

Norsk markedsføringsordbok

Norsk dataordbok

Oljeordliste

RTT40, Ordbok for teleteknikk

RTT41, Ordbok for varsling og alarm

RTT42, Ordbok for VVS

RTT43, Ordbok for petroleumsvirksomhet

RTT44, Ordbok for bygningsglass

RTT45, Ordbok for vannturbiner

RTT, Ordbok for sanering og rehabilitering av VA-nett

Allmennordbøker

Det står att å finne ei form som ein kan støype alle ordbøker i. Jamvel om ein avgrensar det til allmennordbøker, kan det sjå rådlaust ut. Eit problem med allmennordbøkene er mangfaldet i utforminga av ordartiklane. Kvart einskilt system kan vere enkelt og greitt frå databehandlaren sin synsstad. Det er når ein skal lage eit generelt system som skal passe for alle allmennordbøker, problema melder seg.

I samarbeid med Norsk leksikografisk institutt og Det Norske

Samlaget har eg arbeidd med databehandling av allmennordbøker. Frå skoleordlister for grunnskolen og den vidaregåande skolen, via handordbøkene for nynorsk og bokmål, til Norsk ordbok som skal bli på kanskje 12 band. Det har vore naturleg å freiste å gi alle desse ordbøkene eit sams format, slik at dei same programma kan nyttast til alle. Resultatet er eit format med plass til både redaksjonell fridom og streng formalisme. Ei negativ side er plassen: Somme opplysningar står to gonger i same posten, ein gong som ordboksredaktøren vil ha det og ein gong som «maskina» vil ha det.

Alle ordbøker fortel korleis eit ord skal skrivast. Men ei vil skrive om «fotball», ei anna om «-ball» under artikkelen om «fot», ei tredje om «fot/ball», osv. Dei får lov til det som formatet er lagt opp. Men posten inneheld også opplysning om at det er «fotball» det er snakk om. Eit program (som må lagast for kvar einskild ordbok) hentar fram denne opplysninga frå ordartikkelen. Sameleis om ein fortel om bøyinga at det skal vere «-en -ar -ane» eller «m» eller «/en», kan desse opplysningane formaliserast til «M1». På dette viset er det t.d. enkelt å sortere saman fleire ordbøker.

Enno har eg ikkje freista å formalisere andre opplysningar. For tur står etymologi, uttale og tydingsinnhald (semantikk). Dei to første skulle ikkje by på så store datamaskinelle problem. Det er først og fremst eit spørsmål om «å gjere det». Formalisert semantikk er derimot langt frå problemløst. Men det skulle vere interessant å prøve...

Også den frie delen av posten er noko formalisert. Leksikografen har alltid tenkt i «felt», meir eller mindre medvite. Først har ein t.d. eit «oppslagsfelt», så eit «uttalefelt», så eit «etymologifelt», så eit «definisjonsfelt» osv. Denne feltinndelinga er det tatt vare på, slik at det skal vere mogleg å søkje etter ord innanfor t.d. definisjonane eller å dra ut alle etymologiane osv. Men innhaldet i desse felte er fritt, og rekkjefølgja er fri innanfor dei reglane leksikografen sjølv set opp.

Det allmennleksikografiske formatet liknar på NoTe-formatet på overflata. På den måten har eg kunna nytte fleire programdelar frå det eine på det andre. Fortran-77 var ikkje tilgjengeleg på DEC-10 ved Universitetet i Oslo, og dei fleste programma for «Allex-formatet» er skrivne i programmeringspråket Pascal.

Datamaskinbruken ved NLI er ikkje så «intensiv» som ved NoTe. Det er ikkje i same grad aktuelt å sortere data att og fram, dra delar frå den eine ordboka og slå saman med den andre, osv. Men maskina blir ein god del nytta til kontroll av at det redaktøren har gjort og det dataregistreraren har gjort, er i samsvar med reglane.

Norske handordbøker

Mellom dei viktigaste prosjekta ved NLI for tida er dei to Norsk handordbok for bokmål og nynorsk. Redigeringa er nett no (november 1982) slutført. Dataregistreringa er godt over halvvegs for nynorsk-utgåva sin del – noko mindre for bokmålsutgåva. Resultatet blir eit band med om lag 1000 sider for kvar av utgåvene. Dei kjem ut ein gong tidleg i 1984. Ordbøkene har med det ein ventar seg av slike ordbøker: rettskriving, bøyingsopplysningar, uttale (der det «trengst»), etymologi, tydingsforklaringar, bruksdøme.

Dette materialet er altså maskintilgjengeleg. Eit slikt materiale er ein «ønskedraum» for edb-leksikologen, og det bør gi grunnlag for ei rad prosjekt. Ordlaging, bøying og uttale av framandord, definisjonsspråk er nokre stikkord. Slike studiar kan vere interessante både med tanke på å forstå språket og med tanke på å skrive betre ordbøker.

Norsk-vietnamesisk ordbok

Eva Møller

Våren 1983 avsluttes arbeidet med en norsk-vietnamesisk ordbok, som er utført i Bergen. Edb har vært benyttet under det meste av arbeidet med tilrettelegging og korrekturlæsning av materialet. Alt er registrert via skjermterminal på universitetets Univac-anlegg, og kan overføres til trykkeri for fotosetting når ordboken er klar til utgivelse.

Arbeidet med ordboken ble påbegynt høsten 1979 som et rent vietnamesisk prosjekt, utført av en gruppe på 6 vietnamesere bosatt i Norge. Hovedoversetter har vært *Nguyen Quoc Khanh*. Det ble senere utvidet til et norsk-vietnamesisk samarbeidsprosjekt med deltakelse av Institutt for fonetikk og lingvistikk og Nordisk institutt, Universitetet i Bergen.

Flere nordmenn har deltatt i prosjektgruppen; amanuensis *Marianne Haslev Skånland* har vært leder for prosjektet de siste 2 år.

Edb-arbeid i tilknytning til prosjektet er utført ved Prosjekt for datamaskinell språkbehandling, Nordisk institutt, ved Matematisk institutt, ved HF's EDB-seksjon og ved Institutt for fonetikk og lingvistikk. Denne presentasjon av arbeidet bygger på samtaler med M. H. Skånland og programmerer *A-G. Strøm-Erichsen*.

Ordbokens omfang

Ved tilretteleggingen av ordboken har man forsøkt å tilpasse ordtilfanget til behovet hos innvandrere i Norge.

Ved en første utvelgelse av ord ble Norsk ordregistrant benyttet, en datamaskinelt lagret «ordliste» som bygger på en del eksisterende norske ordlister. Videre fikk norsklærerne ved UiB's Kontor for utenlandsstudenter lister og valgte ut hvilke ord de anså for viktige.

Dette utvalg er senere supplert med en rekke andre kilder, f.eks. varelister fra handlende og brosjyremateriale fra helsevesen, likningsvesen og andre offentlige kontorer. Ikke minst har oversettergruppen bearbeidet og systematisert ordutvalget.

Ordboken vil totalt inneholde mellom 16 og 17.000 oppslagsord som er oversatt til vietnamesisk. Av disse er ca 11.000 hovedoppslag, som har grammatiske opplysninger om ordklasse, bøyingsformer, uttale og evt. sideformer. De resterende 5-6.000 er underoppslag, som bare har opplysning om ordklasse.

I tilknytning til hovedoppslaget er gitt eksempler på vanlig bruk av ordet. I den grad det har vist seg nødvendig er også eksemplene oversatt til vietnamesisk. Dette gjelder f.eks. idiomatiske uttrykk, eller eksempler på vanlig bruk som er spesielt vanskelig å forstå når en har vietnamesisk bakgrunn. Videre er alle sammensatte ord oversatt, fordi det viste seg at sammensetninger var spesielt vanskelige å forstå uten eksplisitt forklaring.

Det vietnamesiske alfabet

Vietnamesisk er det eneste av østens språk som benytter alfabetisk skrift. Det bygger på det latinske alfabet, om enn med en lydrepresenta-

sjon som er noe annerledes enn den vi er vant med. Figur 1 viser et eksempel på vietnamesisk trykk.

Đi mua sắm,

Bà Giác : Mình ơi, mình vào tiệm này sắm một ít quần áo đi.
 Ông Giác : Thôi, anh không mua sắm gì nữa cả. Mình mới mua sắm tuần trước, bây giờ còn mua sắm gì nữa?
 Bà Giác : Nhưng em còn phải mua cho thằng Tuấn hai đôi tất, cho con Liên hai cái áo.
 Ông Giác : Em thì lúc nào cũng mua với sắm.
 Bà Giác : Còn mình thì không lúc nào mua sắm cả.

Fig. 1

Som det framgår er den største forskjell i forhold til vestens skriftspråk den omfattende bruk av diakritika. I mange tilfelle brukes 2 diakritika sammen med en bokstav.

Diakritika brukes for det første for å skille mellom vokaler, som det vises i følgende oppsett:

i	ư	u
e	ơ	o
ê	ă	ô
a		â

Fig. 2

For det andre brukes diakritika for markering av tonelag. Vietnamesisk har 6 tonemer, hvorav 5 er markerte. (Til sammenligning har norsk 2 tonelag, umarkerte). Figur 3 viser hvilke diakritiske tegn som benyttes for markering av de forskjellige tonemer. Her benyttes vokalen *a* som eksempel, men i realiteten forekommer alle vokalene fra figur 1 i kombinasjon med alle tonemer.

—	/	\	✓	✓ [?]	- 1 -
a	á	à	ǎ	ã	ạ

Fig. 3

Koding

For å registrere og lagre vietnamesisk skrift ved hjelp av vanlig tilgjengelig edb-utstyr har det vært nødvendig å definere et kodesystem. I tillegg til bruken av vanlige små og store bokstaver skal det være mulig å angi flere «lag» av diakritika for hvert tegn. På grunnlag av det tilgjengelige tegnrepertoar på datamaskin og terminaler har man definert entydige koder, hvor hvert diakritisk tegn er representert ved «S» pluss et annet tegn, f.eks.:

\$1 = /
\$3 = ^
\$, = v

Fig. 4

I sammensetninger til ord kan kodingen se slik ut:

<u>Norsk:</u>	<u>Vietnamesisk:</u>	<u>Kodet:</u>
"omtrent"	độ	\$-d\$3\$.o
"champagne"	ruợu sâm-banh	r\$,u\$,\$.ou s\$3am-banh

Fig. 5

Dvs. at alle vietnamesiske ord i ordboken er registrert og lagret i maskinen i den kodete form som er vist her.

Den kodete versjon er imidlertid vanskelig å lese og dermed umulig å arbeide med i korrekturfasen. Av hensyn til lesbarheten er kodingen konvertert til et tegnsett som til en viss grad ligner det vietnamesiske, og som allikevel benytter det tegnsett som finnes på en vanlig skriver. Et spesielt program for utskrivning gir mulighet til å forsyne alfabetet med diakritiske tegn både over og under linjen, dvs. at et tegn kan strekke seg over flere linjers høyde.

Når materialet er klart til utgivelse overføres det fra universitetets datamaskin via magnetbånd til en fotosetter som kan sette vietnamesisk skrift. Her omsettes den kodete versjon til de korrekte vietnamesiske skrifttegn, og ordboken kan trykkes uten ytterligere korrekturlesing.

Videre anvendelse av materialet

I en sammenfattende vurdering av edb-anvendelse ved et prosjekt av denne art framhever Marianne Haslev Skånland det tosidige aspekt: Selve registreringsarbeidet er svært tidkrevende, men når først materialet er registrert gir edb-anvendelse god hjelp ved korrekturlesing og oppretting.

Den omfattende registrering som er foretatt innbyr imidlertid til videre anvendelse av det samme materiale, anvendelser som til dels allerede er planlagt. Videre er det besluttet at materialet kan stilles til disposisjon for andre prosjekter. Interesserte kan henvende seg til styreren ved Nordisk institutt (for tiden prof. *Willy Dahl*) med forespørsler.

EDB OG HUMANIORA

Automatisk analyse av norsk

Intervju med Per-Kristian Halvorsen

Jostein H. Hauge

Per-Kristian Halvorsen har fra 1.1.82 arbeidet som forsker-NAVF ved NAVFs EDB-senter for humanistisk forskning. Han arbeider med automatisk syntaktisk og semantisk analyse av norsk språk. Virksomheten blir drevet i nært samarbeid med forskningsgrupper ved Xerox Corp. Palo Alto Research Center og Massachusetts Institute of Technology (MIT) i USA.

Datamaskinell behandling av naturlig språk forbinder mange med talende kassaapparater, instrumenter i biler som gir deg beskjed hvis du har glemt å lukke døren eller fylle på olje osv. Er ditt arbeid rettet mot slike anvendelser?

Mange vil nok oppfatte slike resultater som det mest bemerkelsesverdige ved det emnet vi nå beskjeftiger oss med. Men jeg er selv mest interessert i grunnforskingsdelen, selv om jeg også klart ser behovet for å omsette den kunnskapen som vinnes, i mer praktisk anvendbare sammenhenger.

Kan du si litt om din egen faglige bakgrunn?

Jeg har kanskje en litt ukonvensjonell faglig bakgrunn. Jeg begynte å studere lingvistikk i Oslo i 1972, men fant ut at det var begrensede muligheter for lingvistiske studier der. Følgelig valgte jeg å dra til University of Texas hvor jeg tok hovedfag og senere doktorgrad. En periode var jeg vitenskapelig assistent ved Lingvistisk institutt, Universitetet i Oslo, før jeg ble rekrutteringsstipendiat i NAVF. Jeg hadde også permisjon en stund for å arbeide ved University of Texas som «assistent professor». Det er i hovedtrekk min bakgrunn før jeg den 1.1.82 tiltrådte som forsker ved NAVFs EDB-senter for humanistisk forskning i Bergen.

Finnes det utviklingstrekk innenfor lingvistikk i dag som er spesielt interessante for deg som arbeider med faget innenfor en datamaskinell ramme?

Ja, jeg synes at den utvikling vi har hatt ved en del lingvistiske institutter i retning av større vekt på studiet av selve *språkprosessen* virker meget interessant og lovende. Det er i dag slik at en innenfor informatikk, lingvistikk og også deler av psykologi har et delvis sammenfallende syn på språkprosessen. Dette kommer til å gi oss en ny helhetsforståelse av virkemåten til det mest komplekse informasjonssystem som finnes, nemlig den menneskelige språkevne. I dag ser en nødvendigheten av å trekke flere fag inn i studiet av språkprosessen. Det har utviklet seg et nytt arbeidsområde som kalles kognitiv vitenskap eller på engelsk cognitive science.

Hva er ditt eget faglige spesialfelt innenfor dette området?

Først og fremst er jeg interessert i formell semantikk for naturlig språk. D.v.s. en semantikk som er eksplisitt og forståelig og som er bygget på bestemte modeller for å beskrive semantiske forhold.

For meg, som for mange andre lingvister, har det vært naturlig å ta utgangspunkt i modeller som bygger på matematisk logikk, f.eks. arbeider av *Richard Montague*.

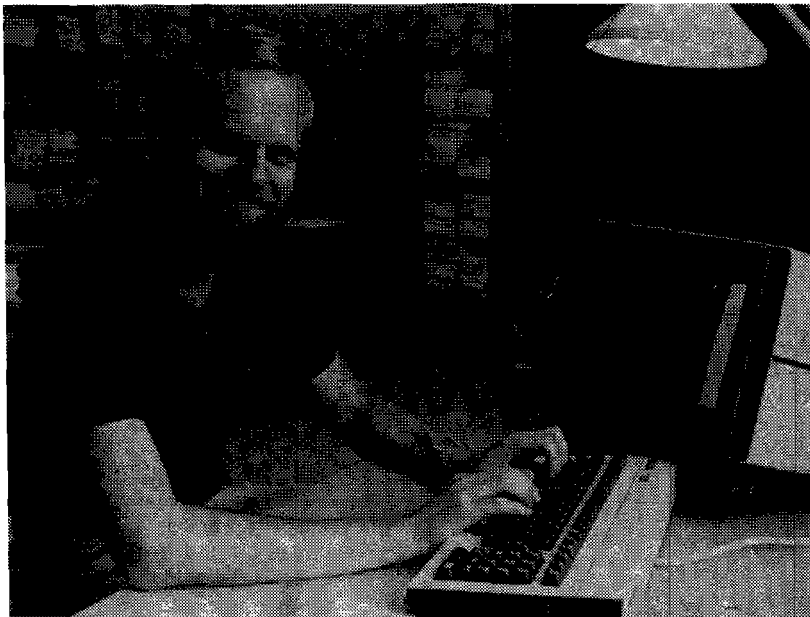
Når det gjelder min egen utvikling innenfor dette interessefeltet, kan jeg si at jeg – som så mange andre – begynte med datamaskinell syntaktisk analyse av språk. Dette er i dag et relativt veletablert felt. Det har vært drevet mye forskning på området, og det foreligger en rekke modeller som kan benyttes.

Blant syntaksforskere kan det naturlig nok også være uenighet om både teorier og de praktiske anvendelser av dem, f.eks. ved syntaktisk kategorisering. Men de fleste benytter seg i dag av en hierarkisk trestruktur som grunnlag for sin språkanalyse.

Men kan en ikke si at også semantikken har lange røtter innenfor lingvistisk forskning?

Jo, det er rett nok, men da taler en gjerne om semantiske studier som tar sitt utgangspunkt i studiet av enkeltstående ord og strukturer.

Når det gjelder formell semantisk analyse er det imidlertid i dag flere modeller for semantisk representasjon.



Per-Kristian Halvorsen.

(Foto: AV-avdelingen, Universitetet i Bergen).

En del forskere anvender semantiske nett, som gir informasjon om sammenhengen mellom de semantiske elementene. Et annet system er det som kalles «procedural semantics». Et eksempel på bruk av et slikt system er det spørre svar-systemet som kalles LUNAR. Det ble utviklet i forbindelse med det amerikanske Apollo-programmet. En tredje måte å representere betydning på er å gå veien via et formelt logisk språk. Dette utgangspunktet gjør det mulig å bestemme setningsinnhold på nye måter, f.eks. om setningen er sann eller falsk, gitt en modell av språket og en partiell modell av verden.

Slike systemer kan imidlertid bli uhåndterlige når det gjelder å implementere dem maskinelt. Selv har jeg konsentrert meg om nettopp dette problemet. Jeg har utviklet en semantisk teori for leksikalsk-funksjonelle grammatikker (LFG). De semantiske representasjonene er bevislig ekvivalente med tradisjonelle logiske representasjonsformer. De er like presise, men samtidig har semantikken for LFG trekk som gjør den meget bedre egnet til å være inndata for enkle, eksisterende tolkningsmoduler.

Er det ikke et kjempearbeid å anvende en så eksakt modell i praktisk arbeid med å beskrive språklig innhold f.eks. på setningsnivå?

Det tror jeg blir det samme for alle de tre modellene som jeg har nevnt. Det er åpenbart at også de semantiske nettverk har lett for å bli svært komplekse hvis de skal dekke ulike sider ved den språklige virkelighet. Men her kommer også det inn at en trenger ikke skissere like nøyaktig alle språklige forhold. Ofte vil det være slik at en kun beskriver en del av virkeligheten, f.eks. i forbindelse med spørre-svar systemer som bare dekker et avgrenset saksområde.

Synes du det er forskjell på forskningsvirksomheten i USA og i Norge på ditt felt?

I USA er det først og fremst en mer omfattende forskningsvirksomhet i gang både i bredde og dybde. Det er klart at vi i Norge ikke selvstendig kan utvikle paralleller til de utenlandske systemene som etter hvert kommer og som til dels vil være tilgjengelige i kommersiell form. Det vi kan håpe på er at utviklingen stadig sterkere går i retning av modulære systemer slik at vi lettere kan modifisere dem for norske behov.

Det vil også være nødvendig med selvstendig ekspertise i å beskrive språk så klart og formelt at resultatene kan nyttes i automatiserte systemer. Den forskningsretning som jeg arbeider innenfor (LFG), gir en lingvistisk teori som er tilfredsstillende for språkforskeren og som samtidig lett lar seg implementere på datamaskin. En kan bruke vanlige lingvistiske teknikker og beskrivelser i et hvilket som helst språk og samtidig få lagt grunnlaget for edb-baserte språkforståelsessystemer for disse språkene.

Du har også tilknytning til Massachusetts Institute of Technology (MIT). Dette er et navn som det står respekt av innenfor mange vitenskaper. Hva har MIT betydd for deg?

Det viktigste for meg er at oppholdene ved MIT har gitt meg innpass i et miljø som interesserer seg for å studere språkprosessen, og ikke bare språkssystemet som en statisk spørrelse. På 60-tallet var det slik at den lingvistiske teori var konsentrert om studiet av såkalt kompetansegram-

matikk, d.v.s. studium av hvilken språkkunnskap menneskene har. Det som ble forsømt, var studiet av språkbruken: hvordan vi kan utnytte vår språkkunnskap i ulike situasjoner, hvordan vi kan tolke språk, forklare hva som ligger under våre språkfeil o.s.v.

Ved MIT er det en stor gruppe vitenskapelige medarbeidere som samarbeider om å studere hele språkprosessen. Denne gruppen er sammensatt av både psykologer, lingvister, informatikere og ingeniører. Gruppen arbeider særlig med å bygge modeller som kan forklare de kognitive prosesser som kommer til uttrykk ved bruk av språk. Noe jeg også setter pris på ved MIT, er muligheten til å gjøre meg kjent med det utstyret som brukes for å studere forskjellige sider ved språkprosessen. Jeg tenker her på utstyr for taleanalyse og talesyntese og programmeringssystemer som egner seg til kognitiv modellbygging.

Den innsatsen som drives ved MIT, virker på mange måter imponerende. Et studium av språkprosessen vil innbefatte mange vitenskaper. Kreves det da ikke et miljø av MITs størrelse og mangesidighet for å kunne drive med denne typen forskning?

Nei, det er ikke slik at kvalitet nødvendigvis har med størrelse å gjøre. MIT er for så vidt en liten institusjon. Det lingvistiske instituttet har f.eks. bare 8 lærere.

Hva kommer det da av MIT på så mange områder er ledende?

Det har kanskje sammenheng med at MIT har et godt rykte og trekker til seg gode lærere og forskere og fremfor alt gode studenter. Det er også karakteristisk for MIT at der er en interdisiplinær ånd og arbeidsform. Som jeg tidligere var inne på er det en sterk interesse hos f.eks. ingeniører for å få kontakt med lingvister og psykologer. Dette har vist seg å resultere i fruktbart samarbeid. De materielle forhold ligger på mange måter ikke så langt over dem vi har i vårt land.

I Norge har det imidlertid vært en tendens til å spre de midlene en har til forskning, over en mengde institusjoner. En kunne oppnå en større konsentrasjon av forskningsvirksomheten og trolig bedre forskningsresultater dersom en opprettet sentra som arbeidet med ulike, spesialiserte forskningsoppgaver. Spredningseffekt kunne en oppnå ved å spre disse sentra heller enn forskerne innenfor hvert enkelt felt.

Er det ikke slik at også utviklingen innenfor datamaskinteknikk vil virke inn på forskningsarbeidet innenfor språklig databehandling?

Jo, det er riktig. Og særlig blir det fremtredende hvis vi utvider området til kunstig intelligens mer generelt. Her har de nye vitenskapelige datamaskiner betydd et gjennombrudd. Det finnes i dag det vi kunne kalle vitenskapelige personlige datamaskiner som har en meget stor adresseringskapasitet og svært gode grafiske egenskaper, f.eks. LISP-maskinene. Disse maskinene gir deg mulighet til å arbeide under full kontroll med store datamengder – i praksis vil du aldri slippe opp for lagerplass. Både innenfor språklig databehandling og innenfor kunstig intelligens er det av avgjørende betydning å ha nærmest ubegrensede adresseringsmuligheter. Slik sett blir maskinene mer slagkraftige enn store datamaskiner. Det systemet jeg arbeider med, tar f.eks. så stor plass at jeg ikke kan kjøre det innenfor det normale driftsopplegget på f.eks. en DEC 20 datamaskin. Det finnes selvsagt muligheter for å omgå slike kapasitetsproblemer, f.eks. ved å dele opp

erne, et klart medansvar for å bringe den teknologiske utvikling inn i de baner som gagnar menneskene best.

Jeg kan her vise til eksempler fra USA på at maskinleverandører trekker inn lingvister for å utvikle f.eks. spørre-svar systemer i naturlig språk. Dette gjelder f.eks. IBM og Hewlett Packard. Samme utvikling har funnet sted innen store forskningsinstitutter som Stanford Research Institute International (SRI), Bell Laboratories o.fl.

La meg til slutt nevne at vi her i Senteret har tatt initiativ til å etablere en nasjonal kontaktgruppe for dem som er interessert i datamaskinell analyse av naturlig språk. Vi håper denne kontaktgruppen kan være en møteplass for ulike forskningsprosjekter og gi grobunn for videre utvikling og samarbeid på feltet. Det siste møtet i gruppen ble holdt ved Regnesentret, Universitetet i Trondheim i februar, og der møtte det representanter fra flere forskningsmiljøer i Bergen, Trondheim og Oslo.

Fra mitt utsiktspunkt synes det klart at dette er et forskningsfelt som vil få stadig større oppmerksomhet i årene som kommer.

Se også rapporten fra seminaret om automatisk analyse av naturlig språk på s. 69.

Tekstbehandling i humanistisk forskning

Svein Skotheim

Med utgangspunkt i den økende bruk av mikrodatamaskiner innen forskning og høgre utdanning inviterte Humanistiske Data fire aktive forskere og lærere til en samtale. Innen humanistisk forskning er vel *tekstbehandling* det mest nærliggende bruksområde for mikro-maskiner, og samtalen begynte naturlig der, men også andre anvendelser blir berørt: kommunikasjon med stormaskiner, behandling av mindre datamengder, publiseringsteknikk, frihet i valg av forskningsmål osv.

Deltakerne i samtalen er *Ralph Jewell*, amanuensis ved Filosofisk institutt (UiB), redaktør *Johannes Kvestad*, Hardanger Folkemuseum, amanuensis *Jan Oldervoll*, Historisk institutt (UiB) og amanuensis *Helge Sandøy*, Nordisk institutt (UiB).

I den innledende runden skisserte Ralph Jewell sitt utgangspunkt for bruk av edb og tekstbehandling, nemlig en vitenskapsfilosofisk avhandling om Bergenskolen i meteorologi. Avhandlingen er skrevet inn på DOC-systemet på Univac, og Jewell understreker at bruk av elektronisk tekstbehandling ga nye perspektiver på stoffet i flere sammenhenger. For det første kunne nye ideer og forslag med utgangspunkt i kildematerialet lettere innarbeides etter hvert. Han noterte disse ideene i tilknytning til dokumentene og stoffet for øvrig, og på et senere tidspunkt viste det seg da å være lettere å samle ideene, og få forslagene gruppert og vurdert.

Det viste seg også å være en fordel å ha dokumentene i maskinutskrift i stedet for i originalutgavene fordi en lettere kunne sammenstille dem med annet kildemateriale på en systematisk måte.

Jewell nevner 4 stikkord som årsak til at tekstbehandling for ham føles som et stort fremskritt innen forskningsarbeid:

1. Det er lettere å ta vare på viktige sammenhenger som er bare svakt antydning i kildematerialet.
2. Det er mulig å effektivisere behandlingen av materialet på et foreløpig grunnlag.
3. Det er en mer effektiv og smidig måte å presentere stoffet på, uten unødig bruk av tid og krefter på de forskjellige stadier i forskningsarbeidet.
4. Det er lettere for andre å følge opp med videre forskningsarbeid med utgangspunkt i det foreliggende materialet.

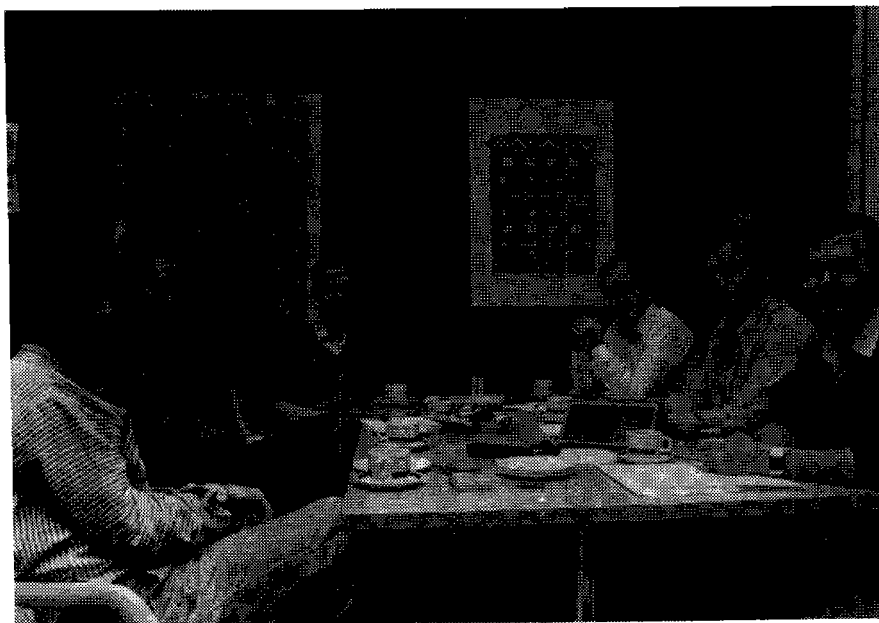
Helge Sandøys erfaringsgrunnlag bestod først og fremst i innskriving og redigering av en lærebok i dialektologi. I dette arbeidet var det store manusproblemer, blant annet komplisert lydskrift som krever stor presisjon. Ved vanlig maskinskriving må hele sider ofte skrives én gang til p.g.a. ny korrektur. Ved omigjenskriving kan nye feil oppstå, noe som i prinsippet er umulig på tekstbehandlingsmaskinen. På disse maskinene blir presisjonen bare bedre for hver ny korrektur. Ved utarbeiding av lærebøker er det stadig behov for endringer av tekst og pedagogisk fremstilling etter hvert som en vinner erfaring i forelesnings-sammenheng. Slike endringer og flyttinger gjøres svært enkelt på tekstbehandlingsmaskiner.

Teksten ble i dette tilfellet lagt inn på en ABC80/100 mikrodata-maskin. Den ble deretter overført til Univac-maskinen ved UiB og befinner seg nå på HF-fakultetets nye mikroanlegg METRIC. Det er et mål å overføre teksten videre til fotosetter.

I tillegg til dette arbeidet hadde Sandøy erfaring med den maskinelle bearbeiding av en dialektbibliografi som førsteamanuensis *Oddvar Nes* redigerer. Her er det snakk om et stort manusgrunnlag. Titlene i bibliografien blir skrevet inn på tekstbehandlingsmaskin, og korrektur/redigering blir foretatt der. Tekstmassen blir senere overført til Univac-maskinen ved UiB, og der blir titlene kodet og kryss-sortert. Senere blir teksten igjen overført til tekstbehandlingsmaskinen og klargjort for trykking. Databasen i Univac-maskinen kan ajourføres hvert år, og nyutgaver blir enklere å redigere.

Sandøy karakteriserte seg som sjøllært i denne sammenhengen og hadde stort sett funnet fram gjennom prøving og feiling ved terminalen.

Johannes Kvestad arbeider med lokalhistorie ved Hardanger Folke-museum og skriver bl.a. en bygdebok for Ulvik kommune. I forbindelse med dette arbeidet bruker han Hardanger Folkemuseums mikromaskin til å skrive inn manus etter hvert som arbeidet skrider fram. Den forrige bygdebokforfatter etterlot seg en mengde manuskripter i forskjellige utgaver som viser mindre endringer og omdisponeringer. Dette må ha vært et enormt arbeid ettersom de forskjellige manus finnes i opptil 2 og



*F.v.: Jan Oldervoll, Johannes Kvestad, Helge Sandøy, Ralph Jewell, Svein Skotheim, Jostein H. Hauge.
(Foto: AV-avdelingen, Universitetet i Bergen).*

3 maskinskrevne utgaver. I dag er arbeidet med redigering og retting sterkt forenklet med tekstbehandlingsmaskinen, og det er hele tiden lett å få utskrifter og utdrag som kan sendes til de forskjellige kilder for godkjenning.

Kvestads problem er først og fremst kapasiteten på de to diskettstasjonene som finnes på maskinen (400.000 tegn i hver). Straks et manus blir lengre enn 40 sider, føler han seg på gyngende grunn og legger inn kopier på nye disketter etter hvert. Med strømproblemer og usikre værforhold på Vestlandet, føler han seg ikke helt trygg på teknikken og har bl.a. opplevd at relativt store tekstmengder er forsvunnet. Med et kopisystem unngås dette, men det tar tid og føles tungvint.

Alt i alt opplever han imidlertid mikromaskinen som et verdifullt hjelpemiddel, og han arbeider for tiden med å utarbeide et rasjonelt opplegg for overføring av tekstmateriale til trykkerier for publisering. I dette arbeidet har han hatt mye nytte av kontakten med Senteret.

Jan Oldervoll er veteranen i denne sammenheng. Han har drevet med bruk av hullkort/edb/tekstbehandling siden 1970 og føler at teknikken, slik den arter seg i dag, er et fantastisk godt hjelpemiddel. De problemer som eksisterer, er stort sett mulig å løse. Han har imidlertid lært ikke å stille urimelige forventninger til teknikken. Han understreker at det er viktig å ha ekspertise i nærheten som kan løse enkle problemer under arbeidets gang, og han understreker også at *motivasjon* er et viktig begrep i forbindelse med denne teknikken. De som har lyst til å bruke tekstbehandling, lærer det fort og finner stort sett ut at det er et verdifullt hjelpemiddel.

Av de mange fordeler med tekstbehandling nevner han bl.a. at forfatteren lettere retter i manus fordi han vet at det ikke koster så mye arbeid. Han understreker også fordelene i forbindelse med publisering, særlig ved komplisert sats som er svært dyrt ved den tradisjonelle produksjonsmåten. Tekstbehandling gjør det mulig å kode materiale slik at det kan overføres direkte til fotosats.

Oldervoll har ellers som nevnt brukt teknikken i 13 år, og i en spirituell gjennomgang gir han oss innblikk i erfaringene på de ulike utviklingstrinn. Han summerer opp med å si at erfaringene er blitt bedre og bedre, og at teknikken i dag befinner seg på et brukbart nivå, lett tilgjengelig for alle.

Etter denne generelle åpningsrunden ble mer spesielle problemer i forbindelse med tekstbehandling og mikromaskiner tatt opp.

Ralph Jewell utdypet nærmere hvordan han følte at han fikk en større frihet intellektuelt og forskningsmessig ved at han kunne «bolte seg» med ideer og forslag knyttet til dokumenter og materialet for øvrig. Det var også mulig å eksperimentere med valg av metode i fremstillingen, f.eks. om det passet bedre med en historisk-biografisk metode eller en mer filosofisk fremstillingsmåte. Ved prøving og feiling kom han fram til at de to fremstillingsmåter kunne forenes, og denne utprøving ville ikke vært mulig med en tradisjonell arbeidsmetode fordi det ville blitt for tidkrevende. Jewell mente ellers at en fordel ved å arbeide mot en stormaskin, er at den lettere kan kombinere tradisjonell tekstbehandling med bruk av databaser.

Kvestad ønsket seg bedre kommunikasjon mellom mikromaskiner og

stormmaskiner, og dessuten bedre lagerkapasitet på mikromaskinene.

Oldervoll mente at dette ønsket snart ville bli innfridd, men at fordelene ved stormmaskiner neppe kunne oppnås fullt ut med mikromaskiner, først og fremst i forbindelse med prosesseringskapasitet. Å lete og redigere i en større tekstmengde vil i lang tid fremover ta nokså mye lengre tid på mikromaskiner enn på stormmaskiner. Det vil også være begrenset manipulasjonsmulighet m.h.t. søking osv.

Jewell berørte deretter et prinsipielt viktig spørsmål. I forbindelse med avhandlingen om Bergensskolen i meteorologi var det svært viktig å analysere *utviklingen av forskningsprosessen*. Mange av avhandlingene til de forskjellige forskerne i miljøet fantes i flere utgaver, f.eks. en 8-siders artikkel av Jack Bjerknes som ble regnet som svært sentral og fundamental innen meteorologi på den tiden. Dette arbeidet fins i mange utgaver, og en nærmere analyse av de forskjellige utgavene frem mot det endelige publiserte resultat ga mange viktige problemstillinger av betydning for Jewells arbeid. Han pekte på at ved bruk av tekstbehandling vil dette perspektivet forsvinne helt ettersom tidligere utgaver blir slettet i maskinene.

En diskusjon omkring problemet viste at dette var noe en bare måtte akseptere. Det samme gjaldt den gangen telefonen ble et vanlig instrument og reduserte brevskrivningen i sterk grad.

Oldervoll kom inn på muligheten av å la maskinen lese ordlister og konkordanser og dermed korrigere manuskripter videre. F.eks. kan maskinen som Det historisk-filosofiske fakultet ved Universitetet i Bergen har kjøpt (METRIC) lese korrektur mot en ordliste, f.eks. Tanums eller Helleviks som fra før foreligger på edb.

Oldervoll nevnte også muligheten til å føre inn synonymordlister, f.eks. kalle opp på skjermen de synonymer en har for et historisk begrep, og velge det beste. Kanskje vil denne teknikken bli spesielt aktuell i forbindelse med fremmedspråk.

Til slutt var det en prinsipiell debatt om bruk av edb og tekstbehandling kunne medføre endring av hele den humanistiske forskning. Sandøy pekte på at uansett hvilke maskiner vi kan få som hjelpemidler, må vi selv produsere ideer og problemstillinger. Fordelene med maskinene ligger først og fremst i søke- og sorteringsmuligheter. En fristelse vi står overfor, er å samle for mye data.

Jewell syntes problemstillingen var svært viktig og viste bl.a. til at en kjent amerikansk professor klaget over at studentene i dag manglet meteorologiske problemstillinger, men til gjengjeld hadde mye tall og numeriske analyser som de arbeidet med. Han mente derfor at alle burde være oppmerksom på faren for at forskning blir «databehandling» i snever forstand.

Oldervoll avviser også problemstillingen om at maskinene leder forskningsaktiviteten. Tvert imot er det blitt mulig å ta opp helt nye problemstillinger og å kunne gå inn i gamle problemstillinger på en ny måte. Som eksempel nevnte han holdninger til uekte fødsler i de gamle samfunn. Her var det sterke normer ettersom det var viktige faktorer i samfunnet som ble berørt. Hittil hadde forskningen stort sett benyttet «myke kilder», d.v.s. rettsdokument m.v. Databehandling hadde imidlertid gjort det mulig å ta opp problemet på en ny måte, f.eks. i forhold

til folketellingene. Dermed hadde forskningen ført til andre resultater enn tidligere, det viste seg nemlig at reaksjonene i lokalsamfunnene ikke var så sterkt negative som de myke kilder kunne tyde på.

RAPPORTER

Den andre nasjonale namnegranskarkonferansen, Universitetet i Oslo

Årets konferanse, som fant sted 19. november, hadde som arbeidstitel «Edb og namn». I år som i fjor var Institutt for namnegransking ved Universitetet i Oslo arrangør, og amanuensis *Botolv Helleland* stod for den praktiske gjennomføringen. Omkring 60 personer var til stede, blant dem deltakere fra Sverige og Danmark.

Konferansen inneholdt både en mer omfattende orientering om bruken av edb ved Institutt for namnegransking og kortere presentasjoner av prosjekter i inn- og utland. *Håvard Hjulstad* ga ei innføring i programsystemet FELTED som han har laget for instituttet. Det er et generelt system som tar seg av hele databehandlingsprosessen: innskriving, korrigering, utskrift, sortering og søking. I «generelt» ligger at systemet ikke er laget spesielt for ett bestemt navneprosjekt, men det er så fleksibelt at det er anvendelig for de fleste typer navnedata.* Fra instituttets side ble det uttrykt stor tilfredshet med systemet.

Den neste delen av programmet var viet omtale av en del prosjekter. Denne delen burde ha fått mer tid, fordi her var det en god del nyttig stoff og konkrete erfaringer å hente.

Bjørn Geirr Harsson og *Kristian Krogh*, Norges Geografiske Oppmåling, orienterte om koordinatfesting av navn og navn knyttet til hovedkartverket i M711-serien.

Diskusjonen etter disse innleggene var interessant sett fra edb-synspunkt. Det ble hevdet at ved å velge en felles form for ei gruppe av navn, ble det innført en standardisering. F.eks. kunne en ved å velge betegnelsen *fjell* som samleform for betegnelser på høydedrag (fjell, topp, nut, pigg, berg o.a.) gjøre voldsomt mot data bl.a. fordi denne betegnelsen har forskjellig mening i enkelte strøk av landet. Likeså vil valg av den mest frekvente navneforma av ei gruppe personnavn vekke motforestillinger, særlig hvis det er noe eldre materiale med avvikende skrivemåte i forhold til dagens (Erich, Niels). Slike «standardiseringer» er vi nødt til å gjøre for maskinelt å kunne slå sammen varianter av samme navn. Imidlertid ligger det ikke noe annet i dette enn at en gjør et valg mellom flere mulige betegnelser – i dataene ligger like fullt

* Se Hjulstads artikkel i *Humanistiske Data* 3/4-82. (Red. anm.)

variantene lagret. Vi har bare fått en nøkkel som åpner for sammenhengen mellom ulike varianter. Og hvilke varianter som skal høre i hop, er bestemt av navneforskeren alene. På bakgrunn av de motforestillinger som er nevnt ovenfor, kan det være et alternativ å velge mer nøytrale samleformer, f.eks. tallkoder. De vil ikke gi de uheldige (og egentlig uinteressante) assosiasjoner som en meningsbærende samleform vil gjøre.

Navnemateriale, både steds- og personnavn, er en datatype som egner seg svært godt for systematisering ved hjelp av edb. Ofte gjelder det store datamengder, dataene kan ha mange tilleggskoder og mye tilleggsinformasjon. Det kan være aktuelt å gruppere ordsamlingen etter bestemte nøkler, sortere etter bestemte ledd eller nøkler og skrive dataene ut på en bestemt form. Fra edb-synspunkt er problemene de samme som ved arkivarbeid generelt.

Jeg tror årets konferanse har vært med på å legge forholdene til rette for et fruktbart samarbeid mellom navnegranskere og edb-medarbeidere.

Eirik Lien

Red. anm.:

Botolv Helleland opplyser at Institutt for namnegransking har utarbeidet en rapport som inneholder de foredragene som ble holdt på konferansen samt et referat fra drøftingene mellom deltakerne. En omtale av rapporten vil bli gitt i et senere nummer av Humanistiske Data.

Helleland forteller også at Samarbeidsnemnda for namnegransking nedsatte ei nemnd 20.11.82 som skal utrede faglige problemer i forbindelse med databehandling av stedsnavn. Amanuensis Per Bjørn Pedersen, Rogaland Distrikthøgskole, er formann i nemnda.

Humaniora-uken ved Universitetet i Oslo 1983

Ved Det historisk-filosofiske fakultet, Universitetet i Oslo ble det i tiden 24.-29. januar 1983 holdt spesielle forelesninger, utstillinger, konserter m.m. under fellestittelen «Språk og Kultur – de humanistiske fag i 80-årene». Meningen med uken var å vise verden utenfor Blindern hva HF-fakultetet står for. Programmet var bygd opp om 5 hovedtemaer.

- 1) Humanistiske fag i en teknologisk tidsalder.
- 2) De humanistiske fag som formidlere av fremmede kulturer.
- 3) Den nasjonale arven.
- 4) Wagner og europeisk kultur.
- 5) Humanistisk mangfold.

Jeg skal i denne rapporten konsentrere meg om de arrangement som har en eller annen form for tilknytning til edb-bruk.

Det var laget 4 spesialutstillinger basert på ulike forskningsprosjekt, 3

fra historie og 1 fra norsk. Den største av disse utstillingene, «Fra stua full til tobarnskull», omhandler barnebegrensning i Norge fra slutten av forrige århundre og fremover, og bygger delvis på resultater fra prosjektet «Fruktbarhetsfallet i Norge rundt 1900», som ble omtalt i nr. 3/4 1982 av Humanistiske Data. Ingen steder i utstillingen ble det vist eksplisitt til hvordan datamaskinen var blitt nyttet i forsøkene på å finne svar på hovedproblemstillingen: Hvordan fallet i fruktbarhet artet seg, hvilke faktorer bestemte utviklingen.

De andre utstillingene anskueliggjorde mer direkte datamaskinens betydning. Både «Folketellingen 1875 Christiania» og «EDB og lokalhistorie» viste hvordan datamaskinen kunne lette tilgangen til og utnyttelsen av eldre kildemateriale. De sentrale historiske kilder frem til slutten av 1800-tallet er skrevet med en skrift (gotisk skrift) som vi i dag ikke umiddelbart kan tyde. Dessuten lider mange kilder av mangel på registre, noe som gjør det vanskelig å finne frem. Datamaskinen vil etter at kildene er skrevet inn, kunne sortere opplysningene etter de kriterier brukeren bestemmer. Man kan også få skrevet ut en fortløpende versjon av kilden. Som en service overfor publikum var det lagt ut alfabetiske og geografiske (dvs. personene kommer i den rekkefølgen de er skrevet inn i originallistene) lister, slik at interesserte kan slå opp for å finne oldeforeldre, besteforeldre eller andre.



Fra utstillingen «Fra stua full til tobarnskull».



Fra utstillingen «EDB og lokalhistorie». Kåre Andersen forklarer.

I «EDB og lokalhistorie» var det tatt utgangspunkt i et sogn (Hidra i Vest-Agder) og en kildetype (skifteprotokoller fra 1700-tallet) og vist hvordan datamaskinen letter bearbeidelsen av denne type kilder. Skiftene er ikke skrevet inn kildetro, men ved et spesialprogram der opplysningene skrives inn sammen med koder. Disse kodene kan siden brukes som sorteringskriteria for oversikter over f.eks. gjeldsforhold.



Ivar Fønnes gjennomgår datautskriften fra utstillingen «Innvandrere og lærebokspråk».

Den siste utstillingen var hentet fra et forskningsprosjekt om innvandrere og lærebokspråk.* Det har vist seg at fremmedspråklige elever som snakker et brukbart norsk, har problemer i enkelte fag fordi ordforrådet ikke strekker til. Utstillingen viser hvordan man ved analyse av tekstene i lærebøker i visse fag har kunnet finne frem til ord som er så hyppig forekommende at de er nødvendige for at elevene skal få med seg meningen i det som står i læreboka. Det har vært nødvendig med flere omganger med sorterte lister og kodete versjoner før man har kunnet avgrense dette ordforrådet. Til hjelp for fremmedspråklige elever utarbeides det så oppgavehefter som skal gi dem øvelse i å gjenkjenne ord, forklare ord, finne synonymer og lage avledninger. På denne måten kan datamaskinen brukes til å løse problemer for ulike grupper i samfunnet.

*Se artikkelen på s. 27 (Red. anm.)

I tillegg til utstillingene er det holdt flere forelesninger innenfor fagfeltene leksikografi, språkanalyse og musikk. *Ivar Fønnes* holdt et foredrag om HF's studietilbud «EDB for humanister» for et fullsatt auditorium. Han redegjorde for bakgrunnen til at dette studietilbudet har kommet i gang. Rent formelt skyldes det at endringer i studiereglementet åpnet adgang til ett-semesterens emner som del av cand.mag.-graden. Men den reelle begrunnelse var av mer faglig art. I samfunnet i dag møter vi edb på stadig flere områder, f.eks. hver gang vi er innom en bank. På markedet har det dessuten etter hvert kommet mange rimelige tilbud på mikromaskiner, slik at det om noen år ikke vil være uvanlig å finne data-anlegg i private hjem. Den stadig større bruk av edb vil etter hvert forme/endre vårt samfunnsliv. Vil vi forstå hva det er som skjer og selv være med på å forme denne utviklingen, er det viktig å få vite mer om hva edb innebærer. Kunnskaper i edb blir stadig etterlyst i stillingsannonser, og i skolen, som er humanistenes største arbeidsplass, er edb valgfag på de fleste studieretninger, og dessuten med i læreplanen for handels- og kontorfaglig og yrkesfaglig studieretning. En arbeidsgruppe i KUD arbeider med spørsmålet om hvordan edb skal innpasses i skoleverket, og et av de problem som da oppstår, er lærerutdanning. Foreløpig gir «EDB for humanister» ingen formell undervisningskompetanse, men det arbeides med saken. En viktig oppgave for humanistene vil være å utarbeide et meningsfylt opplæringstilbud som et alternativ til konstruksjon av dataspill. Som eksempel på et slikt opplæringstilbud ble nevnt det arbeid som Demografiske Databasen i Umeå driver.

Det er først og fremst innen forskningen at humanistene til nå har brukt edb. Det dreier seg om to typer av hovedutnyttelse:

- a) Informasjonsbehandling, dvs. søking i store datamasser, sorteringer og produksjon av tekster for publisering.
- b) Statistiske beregninger av forskjellig art og en grafisk presentasjon av resultatene.

De fagene som først og fremst har nyttet edb, er historie, språk- og litteraturforskning, musikkforskning og leksikografi. Disse fagfeltene er det som er grunnlaget for de eksempler som benyttes i «EDB for humanister».

Selv om undervisningen er knyttet til et bestemt programmeringspråk og et spesielt data-anlegg, vil de kunnskaper man erhverver seg innen programmering og det vide spektrum av edb-metoder man kommer borti, kunne brukes også i andre sammenhenger.

Fønnes gikk så igjennom selve studieopplegget for semesteremnet, før han avsluttet med å referere fra en undersøkelse blant studenter i de 2 første kullene (bare 1 kull har foreløpig rukket å ta eksamen) om hvilken bakgrunn de forskjellige har. I alt har 40 studenter svart på spørreskjema. Omtrent halvparten av studentene er i alderen 20-30 år, flere yngre menn enn yngre kvinner selv om forholdet mellom de to kjønn er ca. 50-50. De aller fleste har kun fag fra det historisk-filosofiske fakultet, men de mannlige studentene er mer spredt i sin faglige bakgrunn enn de kvinnelige studentene. Fordelingen av studenter etter hvor lang utdanning de hadde idet de begynte på semesteremnet, viser at ca. halvparten hadde oppnådd cand.mag.-graden eller mer. Over

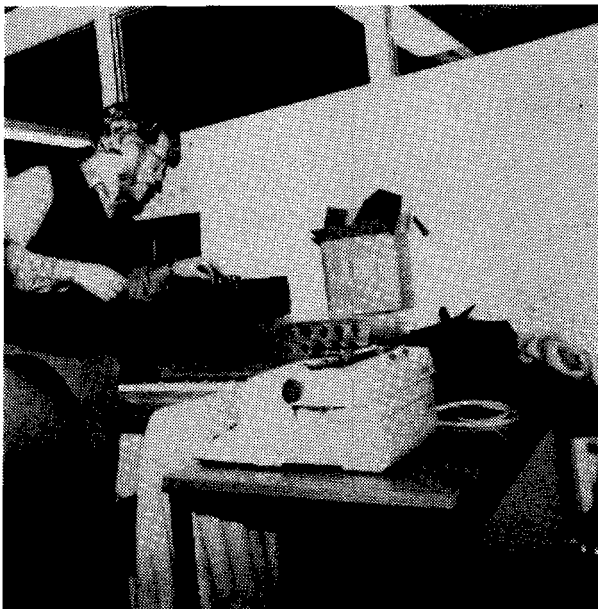
halvparten av studentene tar semesteremnet ved siden av jobb. Ganske få av dem har noen særlige forkunnskaper i edb fra før, ytterst få har vært borti programmering før.

Håvard Hjulstad og Dagfinn Worren viste i sitt foredrag «Datamaskinen som ordgranskar» hvordan datamaskinene er et nyttig redskap i ordboksarbeid. Mye hurtigere enn den menneskelige hjerne kan de sortere ord alfabetisk (forlengs og baklengs), telle opp frekvenser og angi ordforråd i en tekst. Men det er viktig å presisere at maskinen bare er et hjelpemiddel til å lette forskerens arbeid. Det fins problem maskinen ikke kan løse – den kan f.eks. ikke skille to ord som har samme skrivemåte, men ulike betydninger. Den har også problem med å slå sammen ulike former av samme ord. Dette problemet er forsøkt løst ved et eget taggingsprogram der forskeren for hvert ord angir grunnform og ordklasse.

Patrick Chaffey redegjorde i sitt foredrag «The Norwegian Term Bank Project» for et samarbeidsprosjekt for flerspråklig administrativ nomenklatur mellom Norsk Termbank i Bergen, Universitetene i Oslo og Surrey, Utenriksdepartementet og Agder DH. Det har etter hvert meldt seg et behov for å få standardisert oversettelsen fra norsk til fremmede språk av visse typer data som navn på institusjoner, fagforeninger og organisasjoner, titler på norske lover og bestemmelser og spesielle faguttrykk. Ulike oversettelser av samme institusjonsnavn kan skape forestillingen om at det dreier seg om flere ulike institusjoner. Norsk navngiving følger et klart mønster, derfor gjelder det å finne oversettelser som er bygget opp på samme måte. Prosjektet foreslår en oversettelse av f.eks. et institusjonsnavn og i over 90% av tilfellene blir forslaget god tatt. All informasjon blir så lagret i Termbanken i Bergen slik at den skal være lett tilgjengelig.

Ved Institutt for musikkvitenskap ble det først holdt en demonstrasjon av Musikusprosjektet (tidligere presentert i Humanistiske Data 1-1981). Datamaskinen er benyttet for å overføre musikk fra et spesialbygd orgel til maskinlesbar form. Datamaskinen kan så «spille av» stykket slik at man kan lytte etter feil, før et annet program settes i gang for å skrive ut noter. Problemet med å få registrert musikken direkte ved innspilling via orgelet lå i at datamaskinen har vanskelig for å skille ulike stemmer fra hverandre. Derfor må én og én stemme spilles inn om gangen. Et annet problem var rytmen/takten. Da man ikke holder tonene nede den matematisk riktige periode, vil maskinen ha problemer med å bestemme taktlengden. Dette er foreløpig forsøkt løst med en «trampe takten» pedal. Orgelet er også utrustet slik at man kan endre hastighet og transkribere musikken til et annet toneleie (man kan så få skrevet ut noter). Den andre delen av Musikusprosjektet besto i en analysemodell for stilanalyse av musikk. I paneldebatten etterpå ble datamaskinens nytteverdi diskutert. Datamaskinen er godt egnet til å gjøre «manuelt» arbeid som f.eks. notetegning (for komponister vil det bety stor besparelse) og opptelling av kvantitative størrelser (f.eks. telle akkorder). Men det fins også farer. Forskeren må alltid huske at datamaskinen skal tilpasses hans/hennes problemstillinger og ikke omvendt. Siden datamaskinen har sitt opphav i naturvitenskapen, er man også litt engstelig for å trekkes over til naturvitenskapelige metoder

slik at man lager stilanalyser som ikke er musikalsk relevante. Det manuelle arbeid ved stilanalyse har den fordel fremfor datamaskinbehandling at forskeren blir bedre kjent med materialet og kan justere metodene underveis.



Tor Sverre Lande demonstrerer innspillingsorgelet som benyttes i Musikusprosjektet.

Som en avslutning på edb-arrangementene ble det holdt en demonstrasjon av edb-anvendelse i humaniora ved HF-Data. Det ble vist tekstbehandling, dataregistrering, statistikkanalyse og grafisk presentasjon av resultater. På grunn av ukjent utstyr fikk man også demonstrert litt av de mange frustrasjonsmomenter man opplever i sitt møte med datamaskinen.

Arrangementene var godt besøkt, spesielt de som var av mer generell art. Til Ivar Fønnes' foredrag og demonstrasjon av edb-anvendelse i Humaniora kom det flere enn lokalene egentlig var beregnet på, slik at flere måtte sitte på gulvet.

Hege Brit Randsborg

Seminar om automatisk analyse av naturlig språk

Den 31. januar 1983 ble det holdt et seminar om automatisert analyse av naturlig språk ved Regnesenteret, Universitetet i Trondheim (RUNIT). Dette var en oppfølging av et tilsvarende seminar som ble arrangert ved NAVFs EDB-senter for humanistisk forskning i Bergen i februar 1982. Seminaret var organisert som et uformelt presentasjons- og diskusjonsforum for prosjekter innen automatisert analyse av naturlig språk og for utveksling av ideer og erfaringer. Det var 14 deltakere fra Trondheim, Bergen, Oslo, Stavanger og Kristiansand.

Per-Kristian Halvorsen, forsker/NAVF ved Senteret, presenterte et system for språkanalyse basert på leksikalsk-funksjonell grammatikk (LFG). Han startet med å understreke nytten av å bruke en modell fra lingvistikken i forbindelse med automatisert språkbehandling. Det er ingen liten oppgave å skrive en grammatikk fra grunnen av. Ved å nytte en lingvistisk modell kan en bygge på det arbeid som er gjort av andre. Det er stadig flere som bruker LFG som et beskrivelsessystem. Han satte opp de resultatene en ønsker av et automatisert språkanalyse-system.

1. De tematiske/semantiske roller som de forskjellige fraser har i setningen
2. Kontrollrelasjoner for verb i leddsetninger
3. Tolkning av kvantorer
4. Oppløsning av pronomener, både setningsbundne som refleksiver og ikke-setningsbundne pronomener.

LFG modellen har 3 strukturer som brukes til å nå disse resultater:

- 1) *konstituent struktur* med de enkelte fraser
- 2) *funksjonell struktur* som bruker grammatiske relasjoner som subjekt og objekt
- 3) *semantisk struktur* som inneholder semantisk predikat, argumenter, kvantorer o.l.

Strukturene 2 og 3 er implementert som asykliske, rettete grafer og fremkommer som løsning av et sett med ligninger som er tilknyttet frasestrukturreglene. Det er samme løsningsalgoritme som gir den funksjonelle og semantiske struktur og på denne måten har en f.eks. mulighet til å bruke informasjon fra den semantiske prosessen under syntaksanalysen eller omvendt.

Guri Verne, hovedfagstudent ved Institutt for Informatikk ved UiO presenterte arbeid som gjøres i forbindelse med to hovedfagsoppgaver der. Utgangspunktet er et grammatikkfragment i Montague grammatikk for engelsk som en tilpasser til norsk. En ønsker å undersøke hvor godt denne grammatikken virker for norsk, hvor fleksibelt systemet er f.eks. i forbindelse med utvidelser og om reglene er formalisert på en slik måte at disse kan implementeres direkte. Montagues grammatikkmodell er sentrert om semantikk og er spesielt egnet for behandling av tvetydigheter, forekommer av kvantorer og pronomenbinding. Det er laget programmer (i SIMULA) som leser inn en norsk setning, foretar syntaksanalyse, oversetter til intensjonallogikk, foretar forenklinger og gir ut en eller flere logiske uttrykk som representerer ulike meninger av setningen.

Tore Amblé fra RUNIT gjennomgikk systemet for et verbfritt norsk som brukes i prosjektet Mjuke System for kommunikasjon med datamaskin. Det er bare tillatt med verbene *ha* og *være*. Hele syklusen fra en stiller et spørsmål i naturlig språk til en får et svar tilbake gjennomløpes. De enkelte trinn i analysen består av følgende trinn:

Leksikalsk analyse der bøyning og feilstaving behandles

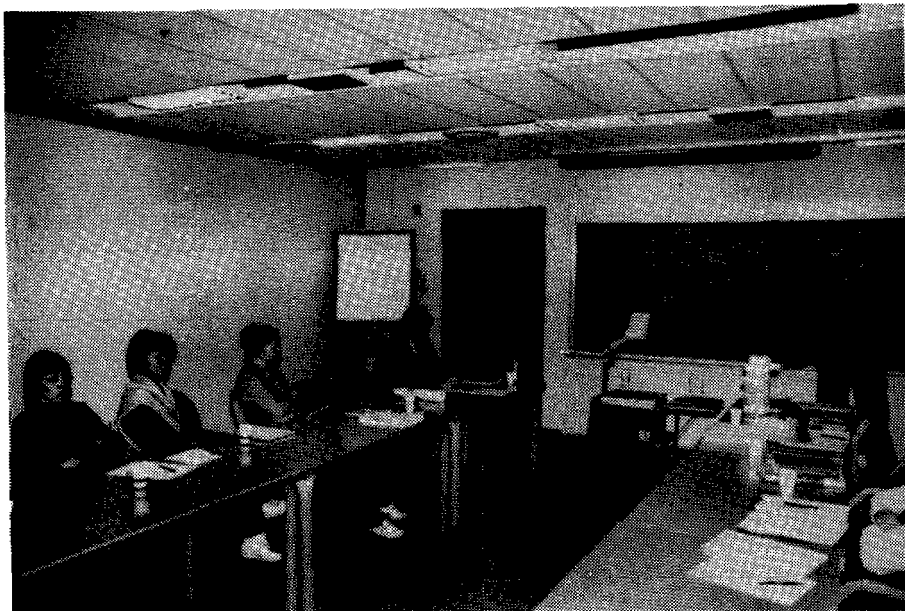
Syntaktisk analyse

Semantisk analyse

Logisk analyse

Faktisk analyse

Den syntaktiske, semantiske og logiske analysen gjøres i PROLOG, et programsystem basert på første ordens predikatlogikk. Grammatikken som brukes er ad hoc og skrevet i en utvidet BNF form. Amblé ønsket kontakt med lingvister for videre studier av dette subsettet av norsk f.eks. til å definere et større subsett ved hjelp av dette verbfrie subsettet. Han fortsatte med å gi en oversikt over PROLOG og kom til slutt med noen inntrykk fra den første konferanse om logisk programmering som ble holdt i Marseilles i september 1982. Bl.a. er japanerne opptatt med å ta i bruk et PROLOG-lignende språk i sitt 5. generasjons datamaskinkonsept. Det er da tale om systemer med en hastighet på 1000-10000 ganger det PROLOG har i dag.



Deltakere i seminaret. T.h. Per-Kristian Halvorsen.

Torhild Wessel fra RUNIT orienterte om de ideer en har for pronomenbehandling i Mjuka System. En ønsker at brukerne skal ha mindre skrivearbeid, slippe å gjenta tidligere informasjon og ha mindre kommunikasjon mot sluttssystem, dvs. at tidligere svar kan brukes om igjen. Kravene som stilles, er at det må være en naturlig bruk av pronomen, og tvetydigheter skal ikke forekomme. Hun gikk gjennom med eksempler det som kunne tenkes å bli referert til ved hjelp av pronomener som DEN, DET, DE, DEM, DENNE, DISSE, slik som alle tema i en gitt samtale, siste tema eller deler av en svarmengde som er synlig på terminalen i forhold til den som bruker må bla i. Det ville være naturlig å skille DISSE, DENNE fra DE, DEN, DEM for å skille mellom første og siste svarmengde og synlig/usynlig skjermbilde.

Knut Hofland fra Senteret orienterte om et prosjekt som gjelder grammatisk merking av det engelske LOB tekstkorpuset. Det brukes her en ordliste (ca. 7000 ord) og en suffiksliste (ca. 700) for å finne mulige ordklasser til et ord. For å entydiggjøre markeringen og å separere homografer brukes sannsynligheter for kombinasjon av to og to ordklasser.

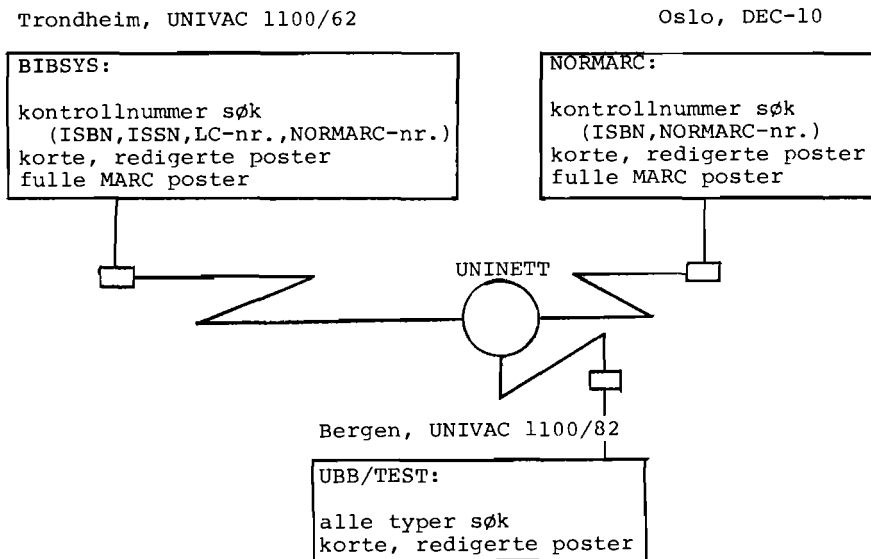
Deltakerne var enig om at et slikt seminar er et nyttig forum for utveksling av informasjon. Det ble avtalt å holde et nytt seminar i Oslo vinteren 1984.

Knut Hofland

BIBNETT/BIBNETT 2

BIBNETT

BIBNETT prosjektet ble kort beskrevet i Humanistiske Data 2-81. Prosjektet ble utvidet i 3 måneder, slik at det ble avsluttet 31.3.82. Mulighetene i BIBNETT ved prosjektets utløp:



BIBSYS kan motta og utføre søkinger, be om søking og sende og motta poster.

NORMARC kan motta og utføre søkinger, sende poster. Det kan også brukes som initierende system.

UBB/TEST kan motta og utføre søkinger, sende poster. Systemet kan ikke benyttes som initierende system.

Testperioden var fra 15.1.-31.3.82, men på grunn av ustabil nettværk har det ikke vært noen reell testperiode. Det ble allikevel utført tester daglig i denne perioden. De mest interessante data i BIBNETT er tiden brukt pr. sesjon. En sesjon består av følgende meldinger:

- start sesjon
- svar på start sesjon
- søking i et annet biblioteksystem
- svar på søkingen
- postoverføring, ber om å få overført en bestemt post
- svar på postoverføring, posten blir overført til ens eget system
- avslutt sesjon
- svar på avslutt sesjon

Den lengste tiden for terminalbrukeren viste seg å være å få startet sesjonen. Dette tok ofte mer enn ett minutt. En sesjon hvor ca. 1200 tegn ble overført tok gjennomsnittlig 5 minutter.

Hovedhensikten med BIBNETT var å avklare om program-til-program kommunikasjon mellom uavhengig utviklede programsystemer var teknisk og økonomisk realistisk i Norge.

BIBNETT prosjektet har vist at det er mulig å knytte sammen uavhengige utviklede biblioteksystem med dagens teknologi selv om biblioteksystemene bruker forskjellige typer databaser, har ulike brukerdialoger og kjøres på forskjellige maskiner. Det viste seg å være mulig å få overført hvert systems brukerdialog til standard meldinger og å overføre bibliografiske poster fra et systems database og lagre dem direkte i en annen database.

BIBNETT prosjektet viste at det var mulig å bruke en interaktiv tjeneste, med modifikasjon, i kommunikasjon mellom systemene, men også at slik tjeneste ikke var god nok for denne type bruk.

Program-til-program kommunikasjon mellom biblioteksystemene er avhengig av et stabilt nettværk og informative feilmeldinger når feil oppstår.

Antall kjøringer i testperioden var for få til å benyttes som evaluering om program-til-program kommunikasjon er økonomisk realistisk i Norge.

Ref.: *The BIBNETT Project; system to system communication using a computer network.* Liv A. Holm.

Application Protocol for Library and Information Systems; the network protocol from the BIBNETT project. Liv A. Holm.

BIBNETT 2

BRODD (tidligere Norsk dokumentdata) har fått bevilget penger fra Norges teknisk-naturvitenskapelige forskningsråd til et forskningsprosjekt «BIBNETT 2». Riksbibliotekstjenesten yder ressurser til prosjektet. De andre samarbeidspartnere er BIBSYS (RUNIT), Norsk Senter for

Informatikk (NSI) og Universitetsbibliotekene i Oslo og Bergen.

Biblioteksystemene som er med i BIBNETT 2 er: NORMARC og SAMKAT (katalogsystem), UB, Oslo. Disse kjøres på DEC-10 maskin.

BIBSYS, Trondheim, kjøres på UNIVAC 1100/62 og POLYDOC på mikromaskin (MYCRON)

Formålet med BIBNETT 2 er å fremme utnyttelse av og bidra til datanettforskning ved at man på grunnlag av resultatene i BIBNETT 1 og UNINETT skal kunne utføre følgende:

- 1) Utføre mer kompliserte dialoger mellom uavhengig utviklede informasjonssystemer på forskjellige stormaskiner. Dette skal gjøres ved at NORMARC og BIBSYS skal kunne utføre mer kompliserte dialoger.
- 2) Det må skapes tilsvarende dialogmuligheter mellom mikromaskin og stormaskin. For å få til dette må lavere lags nettverksprotokoller implementeres på MYCRON. POLYDOC skal utføre dialoger med SAMKAT eller NORMARC. Det vil være en stor fordel for alle POLYDOC-brukere om deres lokale systemer kan kommunisere med andre små og store systemer.
- 3) Informasjonssystem og katalogsystem må kunne oppdatere hverandres database. BIBSYS og POLYDOC skal uoppfordret sende bibliografiske poster til SAMKAT.
- 4) Multicast-problemer sett fra nett-brukers side må analyseres. Man må klargjøre de krav som multicasting setter til informasjonssystemene, og om informasjonssystemene må stille krav til en multicasting tjeneste.

Prosjektet er planlagt avsluttet i 1984.

Åse Støgård

Dokumentasjonssenter for kulturpolitikk og kulturforskning

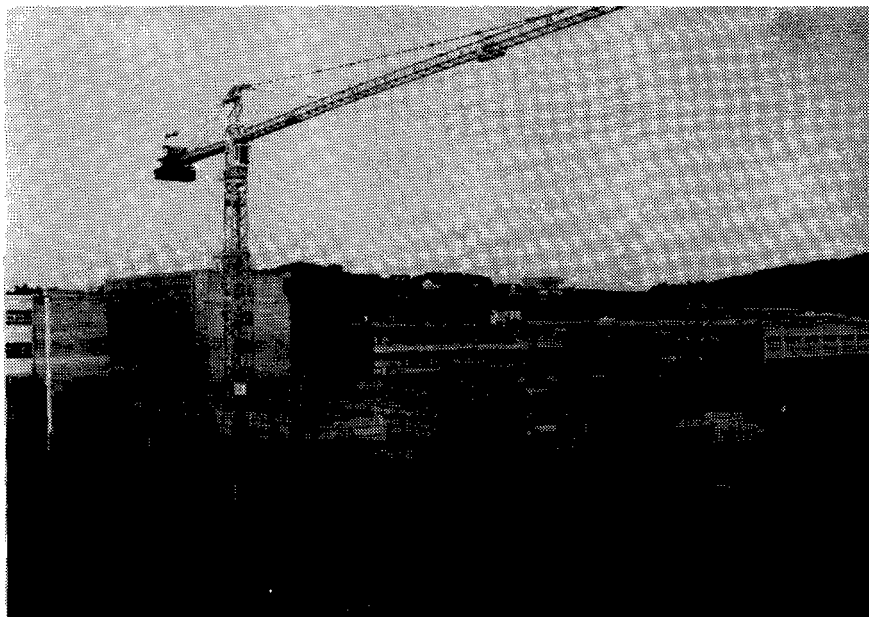
Rogaland distriktshøgskole (RDH) og Rogalandsforskning (RF) har i noen år arbeidet med planer om å få opprettet et dokumentasjonssenter for kulturspørsmål i Stavanger.* I disse dager legges fram en innstilling om dette sentret for Kultur- og Vitenskapsdepartementet (KVD). Det er den tredje innstilling som er laget om dette temaet. Denne siste innstillingen er et resultat av et prosjekt drevet i samarbeid mellom de tre ovennevnte institusjonene: RDH, RF og KVD. Prosjektet har vart i to år. Det har hatt en styringsgruppe med representanter for kulturforvaltningen (kultursjef *Roald Håland* fra Rogaland og første-konsulent *Asbjørn Langeland* fra Statens bibliotekstilsyn), fra forskningsrådene og edb-miljøet (direktør *Jostein H. Hauge* fra NAVFs

*Jf. meldingen i HD 1-82. (Red. anm.).

EDB-senter) og fra initiativtakerne (førsteamanuensis *Lennart Rosenlund* fra RDH/RF). Sistnevnte har sammen med to amanuenser og høgskolebibliotekaren ved RDH dannet en arbeidsgruppe i Stavanger. Denne arbeidsgruppa har i prosjektperioden drevet med innsamling og behandling av relevante dokumenter, foretatt studiereiser til europeiske dokumentasjonssentra, laget et edb-opplegg i samarbeid med NAVFs EDB-senter i Bergen, og skrevet innstillingen.

Faglig sett vil Senteret avgrense sin virksomhet til å gjelde kulturpolitikk i vid forstand: kulturens plass i samfunnet, kulturytringenes og kulturorganisasjonenes betydning for samfunnsutviklingen og for sosiale grupperingers og enkeltpersoners tilpasning i samfunnet o.l. Dette vil utelukke kulturytringene i seg selv, bl.a. fordi disse har egne institusjoner med lange faglige tradisjoner.

Målgruppen vil i første rekke være ansatte i kulturforvaltningen, kulturorganisasjonene, de frivillige organisasjonene i kulturlivet, kulturforskere og -studenter og kulturinstitusjonene (bibliotek, muséer, teatre osv.). Det vil bli lagt særlig vekt på å etablere et nært og godt samarbeid med andre dokumentasjonssentre i inn- og utland. Det tas sikte på å supplere eksisterende tilbud, ikke på å dublere dem.



*Reisingen av bygg VI ved Rogaland distriktshøgskole. Bygget blir ferdigstilt sommeren 1984.
(Foto: Rogaland DH).*

De tilbud Sentret vil konsentrere seg om, er disse:

1. **Forskning.** Sentret vil formidle kontakt mellom bruker-grupper og forskningsmiljøene. I noen grad vil Sentret kunne bidra med hjelp til forskningsarbeidet, som å utføre deler av prosjekter, hjelpe til med søknader om midler osv.
2. **Kontaktformidling** mellom brukergruppene innbyrdes og mellom dem og de ulike deler av kulturlivet vil stå særlig sentralt. Her vil kurs, seminarer, meldingsblad være viktige virkemidler.
3. **Dokumenttjenester.** I samarbeid med RDH-biblioteket vil Sentret bygge opp bibliotekstjenester innenfor feltet. Videre vil det bli bygget opp edb-tjenester i samarbeid med NAVFs EDB-senter for produksjon av bibliografier, emne-lister, SDI-tjenester (Selective Dissemination of Information) osv.
4. **Distribusjon av trykte produkter.** Det vil fortrinnsvis dreie seg om meldingsbladet, bibliografier og de rapporter og andre dokumenter som blir utarbeidet i Sentrets regi.

Styringsgruppa anbefaler opprettelse av Sentret i Stavanger i samarbeid med RDH og RF. Den anbefaler en grunnmodell med tre ansatte, og en finansiering over statsbudsjettet av denne grunnmodellen.

Endelig anbefales det at Sentrets framtidige styre arbeider videre med alternativ finansiering for de ulike aktiviteter Sentret kan tenkes å arbeide med. Dette gjelder i første rekke ekspansive tiltak, men også i noen grad for en hurtigere gjennomføring av Sentrets ordinære aktiviteter.

Innstillingen vil bli overlevert KVD våren 1983.

Knut Bjander

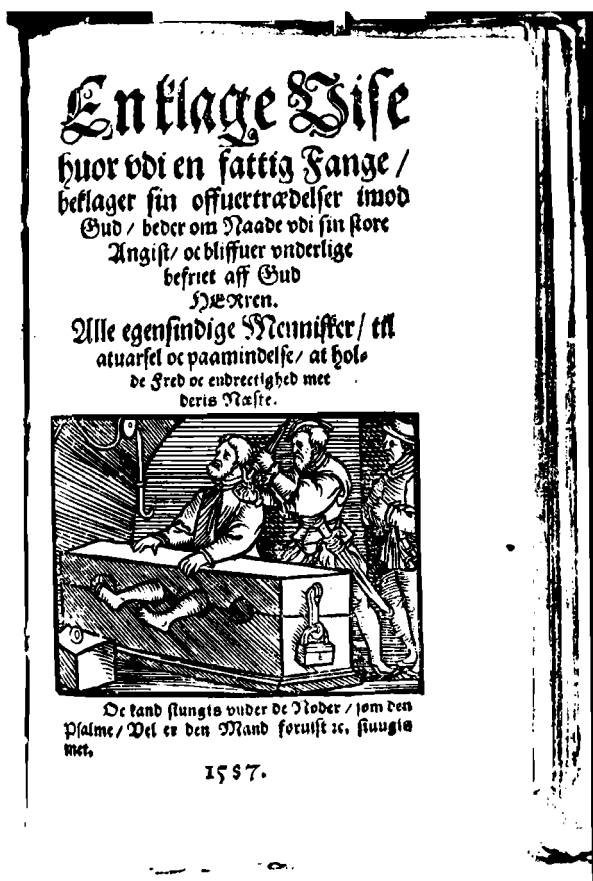
NTAI-prosjektet. Eit fellesnordisk eksperiment med edb-katalogisering av folkloristisk arkivmateriale

Etter den 2. arkiv- og dokumentasjonskonferansen for nordiske tradisjonsarkiv i Åbo 1978, vart det oppnemnt ei nordisk arbeidsgruppe som fekk i oppgåve å utarbeida og testa eit edb-indekseringssystem som skulle kunna nyttast ved dei nordiske tradisjonsarkiva. Med i gruppa var: arkivar *Carsten Bregenhøj*, Dansk Folkemindesamling (formann), arkivchef *Pekka Laaksonen*, Suomalaisen Kirjallisuuden Seuran Kansanrunousarkisto, antikvarie *Gunnar Ternhag*, Dalarnas museum og underskrivne. Prosjektet har gått under namnet NTAI (Nordiska traditionsarkiv och indexering). Nordiska Institutet för Folkdiktning i Åbo (NIF) har finansiert reise- og møteverksemd, medan NORDINFO og NOS-H har finansiert sjølve edb-arbeidet som har vore utført ved Datacentralen af 1959 i København.

Sjølve prosjektet var lagt opp som eit eksperiment der vi skulle bruka edb i katalogiseringa av folkloristisk kjeldemateriale, samstundes som vi skulle sikta mot løysingar som kunne fremja samarbeid mellom dei nordiske tradisjonsarkiva. Gruppa måtte såleis finna fram til system for

katalogisering som sikra ein felles katalogiseringsstandard for basisopplýsningar, samstundes som det tillet kvart einskilt arkiv å halde på sin særart.

Gruppa utarbeidde eit skjema i samarbeid med edb-konsulentar på Datacentralen. Deretter fekk ein tilsendt ulike typar arkivmateriale frå fleire nordiske arkiv for å prøva skjemaet på dette materialet. Ein assistent stod for analysen av materialet, i samarbeid med formannen for prosjektet, Carsten Bregenhøj. Sjølv punchinga vart utført ved Datacentralen. Alt materiale vart lagt inn i databasen på originalspråket.



Tittelside på et dansk skillingsvisetrykk fra 1587.
(Foto: Etno-folkloristisk institutt, Universitetet i Bergen).

Sigbjørn Århus.

Under våren 1983 kommer de till programpaketet hörande instruktions- och läroböckerna att färdigställas. I fullständigt skick kommer STAR att vara både ett programpaket och ett pedagogiskt hjälpmedel för användning av statistiska och kvantitativa metoder i arkeologisk forskning. Paketet består av sex delar:

- (1) Datamaskinprogram
- (2) «Introduction and STAR manual»
- (3) «Student textbook and STAR examples»
- (4) «Archaeology for statisticians»
- (5) «STAR algorithms»
- (6) Programdokumentation

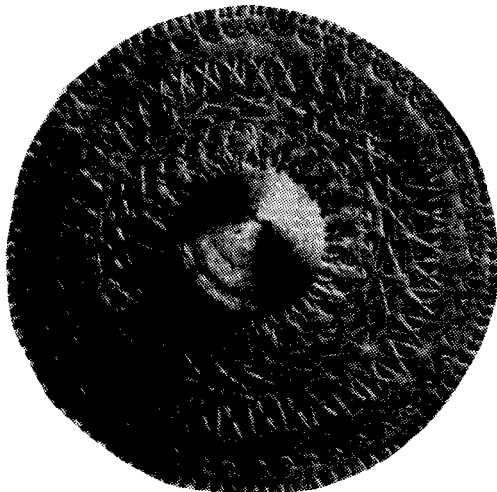
Delarna 2-4 är kompendier, dels tekniska brukarmanualer och algoritmsamling, dels teoretiska läroböcker.

I programpaketet ingår statistiska och kvantitativa metoder:

- (1) Enkel univariat och bivariat statistik
- (2) 4 seriationsmetoder
- (3) De vanligaste multivariata analysmetoderna
- (4) Discriminant analys

Programpaketet tillåter också en viss datamanipulation. Följande generella egenskaper hos programpaketet kan framhållas:

- (1) Programmen läser data i numerisk tabellform
- (2) De enskilda enheterna och variablerna tilldelas en intern löpande numrering, som gör att enskilda arkeologiska fynd kan identifieras i utskrifterna oberoende av datamanipuleringar
- (3) Analysresultaten presenteras grafiskt



*Gullsmykke fra Hauge («Krosshaugen»), Klepp i Rogaland.
(Foto: Ann-Mari Olsen, Historisk museum, Universitetet i Bergen).*

Sentret kan således citera programpaketets eget første meddelande till brukaren på terminalskärmen:

WELCOME TO STAR

Programpaketet kan implementeras på flertalet datamaskiner med tillgång till FORTRAN 77 compiler och NAG biblioteket. Tekniska upplysningar om programmen lämnas genom Sentret. Förfrågningar om paketets användning, pedagogiska sidor och övriga brukarfrågor riktas till Stig Welinder (Universitetets oldsaksamling, Oslo).

Stig Welinder

Avisinform-prosjektet

«Informasjonssystem for norske aviser» (Avisinform) er et prosjekt som blir utført ved Senter for medieforskning, et tverrfaglig forum ved Universitetet i Bergen. Prosjektgruppa består av dosent *Per Torsvik*, som er faglig leder, og to vitenskapelige assistenter, *Trine Syvertsen* og *Karl Knappskog*. Prosjektet, som er finansiert av Rådet for anvendt presseforskning og etterutdanning, kom i gang for alvor i 1980, og skal avsluttes i år.

Avisinform har sitt utspring i en del seminarer som ble holdt i 1975 om teknologiske framskritt og sosial endring. Per Torsvik foreslo da å sette i gang et prosjekt som skulle utrede hvordan datamaskiner og telekommunikasjoner vil få innvirkning på avisene og den journalistiske arbeidsmåten. Den raske utviklingen i medietilbudet vil føre til økt konkurranse for pressen. Avisene kan imidlertid satse mer enn i dag på bakgrunnsstoff, og behøver i så fall bedre arkiv- og informasjonssystemer.

En av prosjektets målsettinger er å greie ut den konkrete organiseringen av informasjon i et edb-basert arkiv for avismateriale. Det skal i denne forbindelsen gis forslag til konstruksjon av en tesaurus for slikt materiale. I forlengelsen av dette arbeidet skal de tekniske, økonomiske og juridiske sidene ved etablering av pressedatabanker belyses. Behovet for et automatisert arkivsystem og innstillingene til samarbeid mellom avisene inngår i denne utredningen.

En undersøkelse som prosjektet utførte i 1981, viste at bare 1/3 av avisene har organiserte klipparkiv. 3-4 aviser driver nå forsøk med edb-baserte arkiv.

Konsulent *Elin Solstrand* ved NAVFs EDB-senter har stått for de tekniske sidene ved opprettingen av to prøvedatabaser som inngår i prosjektet. Begge databasene er basert på ca. 700 artikler fra 4 aviser. Den ene inneholder den autentiske teksten, den andre referanser til artiklene. Ved bruk av søkesystemene NOVA*STATUS og NORDMARC skal det gjøres sammenligninger i gjenfinningseffektiviteten ved indeksering og i fulltekst. Mange mener det er vanskelig å basere gjenfinning på teksten alene, fordi avisspråk er inkonsekvent, lite presist og inneholder mange nyord. Dersom forsøkene viser at denne antagelsen er riktig, skal årsakene til gjenfinningssvikten undersøkes, sammen med måter å berike tekstene på, f.eks. tilføyelse av emneord eller sammendrag.

Disse prøvedatabasene er kun til testformål. Det opprinnelige målet var å lage et prototypesystem som kan legges inn i ulike redaksjonelle og maskinelle miljøer, men nå tar prosjektet utgangspunkt i at det eksisterer mange ulike søkesystemer på markedet.

Avisinform har utarbeidet en arkivmodell hvor den nyeste artikkelproduksjonen er lagret i fulltekst, men hvor en på et eller annet tidspunkt vurderer hvilket stoff skal lagres lenger. Da kan det foretas en manuell indeksering av artiklene, som deretter blir lagret på f.eks. microfiche.

Prosjektet foreslår samarbeidsmodeller for en informasjonssentral for aviser på ulike geografiske nivåer. Det kan f.eks. bygges opp regionale databaser som så kan kobles sammen til et nasjonalt system. Undersøkelser som Avisinform har utført, tyder på at særlig de mindre avisene er skeptiske til et slikt samarbeid, enten fordi de er så lokalt orientert at de ikke har behov for det, eller av økonomiske grunner. Et automatisert arkivsystem er imidlertid for ressurskrevende til å kunne tas i bruk av mindre og økonomisk svake aviser uten samarbeid. En slik informasjonssentral kan for øvrig også inneholde kunnskap om databanker og søkespråk, og søke i nasjonale og utenlandske databaser ved forespørsel fra medlemsavisene.



*Fra Bergens Tidendes klipparkiv.
(Foto: Birkhaug & Omdal, Bergens Tidende).*

Avisinform har dessuten gjort undersøkelser om journalistenes bruk av klipparkiv sammenlignet med andre kilder. Disse undersøkelsene viser bl.a. at nesten 90% av de spurte journalistene er interessert i å søke i et edb-basert arkivsystem via skjermterminal.

En av de planlagte arbeidsoppgavene for prosjektet er å foreta en kvantitativ og kvalitativ bruksregistrering i arkivene til en del utvalgte aviser. Spørsmål som blir rettet til arkivene over et lengre tidsrom skal registreres og klassifiseres etter type.

Mulige følger av innføringen av datateknologi for avisansatte skal også kartlegges. Journalister og arkivpersonale skal intervjues om sitt syn på sin egen yrkesrolle både nå og i framtida, om samarbeidet mellom arkiv og redaksjon, og om mulige forbedringer av arkivene.

Et informasjonssystem med basis i avisenes eget stoff vil være verdifullt også for andre enn avisene. Bibliotekvesenet, skoleverket, bedrifter og forskere er interessert i avisstoff. En pressedatabank, enten det er enkeltaviser som står for driften, eller den er et resultat av et samarbeid, vil kunne gjøre avisstoff mer tilgjengelig for offentligheten.

De pågående undersøkelsene i regi av Avisinform-prosjektet er også relevante i forbindelse med det nye studietilbudet i mediefag ved Universitetet i Bergen. Prosjektet kan bidra med nyttig informasjon til spesielt den delen av studiet som gjelder kildebruk og -kritikk.

Kristin Natvig

EXTEMPLO – et nordisk elektronisk tidsskrift

Behovet for rask nyhetsformidling innen IoD-sektoren (Informasjon og Dokumentasjon) har gjort det naturlig å prøve elektronisk publisering. Dette sier *Gunvor Svartz-Malmberg*, ansvarlig leder for Extemplo-prosjektet og kilden til opplysningene som blir gitt i denne rapporten. Ettersom terminalutrustningen ved forskningsbibliotekene og informasjonssentralene var svært god ved prosjektstart, kunne man henvende seg til en leserkrets som var vant til å arbeide med tekstinformasjon formidlet via telekommunikasjonsnett. Dermed kunne et forsøk med et elektronisk tidsskrift for IoD-sektoren også gi viktige erfaringer for framtidig elektronisk publisering av vitenskapelige rapporter og artikler.

Med støtte fra NORDINFO er Extemplo blitt gitt ut som et kostnadsfritt, sannnordisk nyhetsorgan i tidsrommet 1979-1982. Tidsskriftet har hovedsakelig inneholdt et kalendarium over kurs, konferanser, seminarer og forelesninger innen vitenskapelig IoD i Norden, samt informasjon fra de ulike nordiske online-brukergruppene. Redaksjonen har bestått av representanter fra Danmark, Finland, Norge og Sverige, som har redigert hver sin nasjonale spalte. Under prosjekttiden har bladet utkommet to ganger i måneden, men det har vært mulig å få inn notiser som «Siste nytt». For dette formålet har den elektroniske posten (MAIL-programmet) vært spesielt anvendelig.

Forsøkene med Extemplo fortsetter nå i tilknytning til prosjektet «Elektronisk nyhetsformidling». Begge prosjektene får økonomisk støtte fra den svenske Delegationen för vetenskaplig och teknisk

informationsforsörjning (DFI). «Elektronisk nyhetsförmedling» har til formål å studere lignende prosjekt i andre land og ved utgivelser skaffe erfaringer i metoder som kan anvendes for elektronisk publisering av tidsskrifter. Redaksjonsarbeidet foretas på hjemmeterminaler.

Extemplo blir gitt ut via DEC10-maskinen ved SCANNET-sentralen Medicindata i Göteborg. For at tidsskriftet skal være lettest mulig tilgjengelig, legges det nå stor vekt på å lenke sammen de ulike systemene leserne benytter seg av i sitt arbeid (se fig. 1).

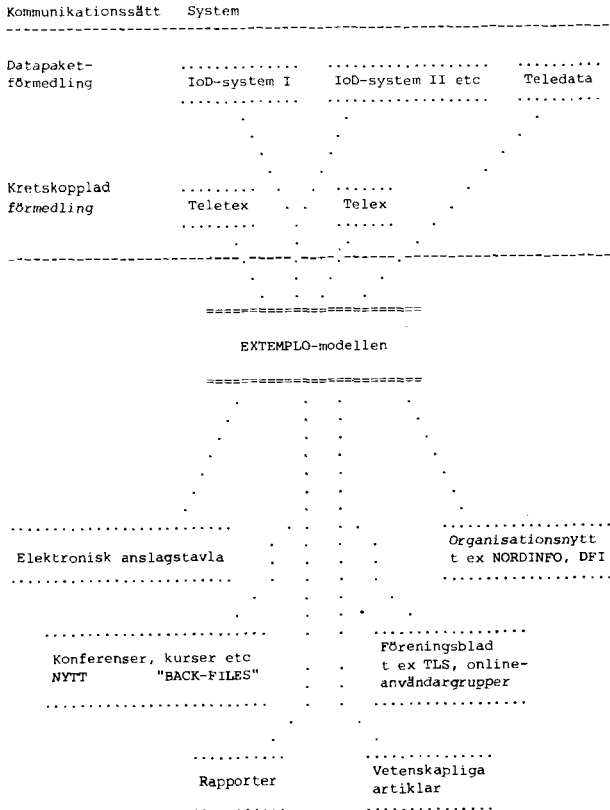


Fig. 1. System-system-kontakter og publikasjonstyper i en tänkt elektronisk nyhetsförmedling med EXTEMPLO som modell.

Siden høsten 1982 har Extemplo også vært tilgjengelig hos DataArkiv i Stockholm. Dette innebærer at en kan søke i tidsskriftet som en hvilken som helst database. To-veis kommunikasjon er også mulig nå - dvs. en kan sette opp meldinger om kurs, konferanser og andre nyheter

på den elektroniske oppslagstavlen via brev og telex til bladets redaksjon. Via DataArkiv får nye brukergrupper adgang til bladet: journalister, informasjonsmedarbeidere, markedsførere, datafolk, kommunebibliotekarer, skolepersonale osv. Extemplo er også blitt forberedt for Teledata.

Extemplos medarbeidere mener at bladet bør betraktes som en forgjenger til mer avanserte elektroniske tidsskrifter for publisering av vitenskapelige originalartikler. Elektronisk publisering vil gjøre det mulig for forfattere og lesere å føre en diskusjon via datamaskin om artiklens innhold, og gjøre disse debattinnleggene tilgjengelige for nye lesere.

Kristin Natvig



Nytt fra RHF/NAVF

Oppfølging og videreføring av «Humaniorautredningen»

Rådet for humanistisk forskning (RHF) gjennomførte i årene 1973-75 en større undersøkelse av virksomheten innen humanistisk forskning i Norge, og publiserte materialet i utredningen «Humanistisk forskning i Norge», («Humaniorautredningen») i 10 deler i 1975. Utredningen har ligget til grunn for RHF's forskningspolitiske planlegging i årene som fulgte. For øvrig fikk utredningen mest å si for den interne fagpolitiske diskusjon i forskningsmiljøene, og ble i mindre grad brukt som grunnlag for en mer utadrettet offentlig diskusjon om humanioras oppgaver og vilkår.

RHF har ønsket å gjennomføre en oppfølging av Humaniorautredningen 10 år etter, og arbeidet i 1981 og 1982 med planene for en ny utredning. Rådet behandlet på sitt møte 4. februar 1983 et forslag til opplegg for arbeidet. Etter denne behandlingen er det klart at Rådet denne gangen ikke bare ønsker å følge opp utredningen fra 10 år tilbake gjennom ajourføring av materialet og identifisering av utviklingslinjer, men også utarbeide dokumenter som kan inngå som deler av den offensiv som allerede er i gang fra det humanistiske forskningsmiljøet for å reise en diskusjon om de humanistiske fagenes legitimitet og muligheter i dagens samfunn.

Derfor vil arbeidet denne gangen omfatte to deler:

- en empirisk ajourføring av en del av datamaterialet om virksomheten innen humanistisk forskning i Norge. En stor del av dette arbeidet har NAVF's Utredningsinstitutt sagt seg villig til å utføre. Dette

materialet vil bli bearbeidet med henblikk på en noe mer utadrettet bruk enn hva tilfellet var for 10 år siden.

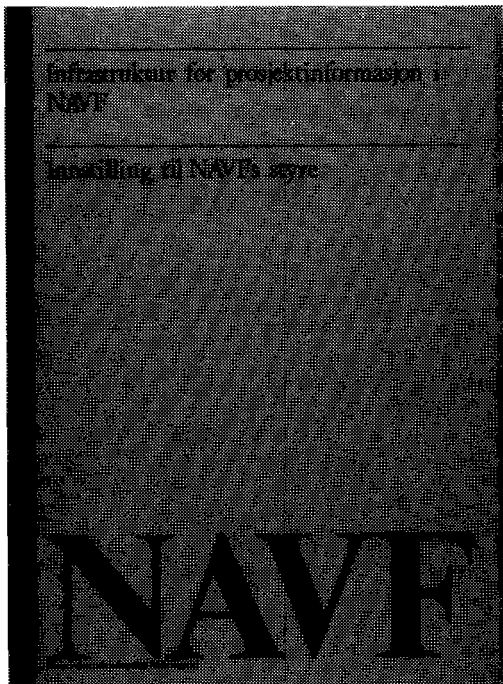
- en prinsipielt anlagt del som foreløpig er kalt «De humanistiske fagenes plass og betydning i vårt samfunn», der en teoretisk og historisk vil se nærmere på hvilke problemer som er framtreddende i forholdet mellom de humanistiske fag og det samfunn som kanskje i mindre og mindre grad etterspør den humanistiske viten og kompetanse.

Arbeidet vil bli ledet av en styringsgruppe/et utredningsutvalg bestående av 7 representanter for forskjellige fag og institusjoner, med Rådets ordfører, dosent *Ola Kai Ledang*, som formann.

Egil Kallerud

Dokumentasjon av pågående forskning

Etter oppdrag fra NAVFs styre har Senteret utarbeidet innstillingen «Infrastruktur for prosjektinformasjon i NAVF». Innstillingen var ferdig i november. Den er nå til behandling i de fem faglige råd i NAVF og vil bli ferdigbehandlet i løpet av våren.



Bakgrunnen for innstillingen er det arbeidet NAVF og de øvrige forskningsråd i den senere tid har utført for å bedre oversikten over igangværende forskningsaktiviteter.

I samarbeid med de respektive faglige råd i NAVF har NAVFs utredningsinstitutt og NAVFs EDB-senter for humanistisk forskning de siste årene utarbeidet oversikter over forskningsvirksomheten innenfor humanistisk forskning, helse- og sosialforskning og barneforskning. Også Norges landbruksvitenskapelige forskningsråd og Norges Teknisk-Naturvitenskapelige Forskningsråd har gitt ut flere oversikter over pågående forskning.

Oversikter over pågående forskning gir aktuell informasjon om forskningsaktiviteter til både forskersamfunnet, planleggende myndigheter og almenheten.

Forskningsrådenes samarbeidsutvalg har i overensstemmelse med anbefalingene i en utredning om prosjektinformasjon fra 1981 rådd til at «forskningsrådene fatter beslutning om å etablere et prosjektinformasjonssystem på sin sektor med sikte på samordning med de øvrige forskningsrådene gjennom et koordineringsutvalg».

I den nylig avgitte innstillingen til NAVF er vekten lagt på hvordan NAVF kan bygge opp et permanent dokumentasjonssystem for forskning innenfor Rådets faglige ansvarsområder. Det blir bl.a. foreslått at det trinnsvis bygges ut et system for prosjektinformasjon med basis i det humanistiske og det samfunnsvitenskapelige fagråd, som til nå – sammen med Rådet for forskning for samfunnsplanlegging (RFSP) – har vist størst interesse for denne type informasjon.

Innstillingen skiller mellom en datafaglig del og en dokumentasjonsfaglig del i slikt arbeid. Det blir derfor foreslått etablert et edb-faglig knutepunkt for prosjektinformasjon i NAVF som på sin side samarbeider nært med dokumentasjonstjenester knyttet til de enkelte fagområder. Informasjonsgrunnlaget for databaser, publikasjoner og spesielle temaoversikter vil dels bestå av informasjon fra den løpende søknadsbehandlingen i NAVF supplert med heldekkende eller temadefinerte kartlegginger som utføres av NAVF-organer eller etter ønske fra andre instanser.

Det blir hevdet at de forskningsutførende institusjonene, f.eks. universitetene, i dag ikke har interne rapporteringsordninger som kan danne basis for samvirke med et dokumentasjonsorgan i NAVF. Innstillingen legger vekt på nødvendigheten av å bedre dokumentasjonsstandarder ved slike institusjoner. Dette vil også lette formidlings- og planleggingsarbeidet i institusjonene.

Som resultat av edb-arbeid i Senteret i forbindelse med prosjektkataloger og senere utført utviklingsarbeid, er det utarbeidet en edb-teknisk systemløsning med tilhørende programvare som dekker alle arbeidsfaser fra prosjektregistrering til publisering av prosjektkataloger og interaktiv søking i databaser.

I en tid med større ønske om tverrfaglig forsknings- og utviklingsarbeid er det også viktig å kunne skaffe seg aktuell informasjon på tvers av de tradisjonelle faggrensene. Forskningsrådenes samarbeidsutvalg har derfor i 1982 nedsatt et utvalg som skal legge frem forslag til felles registreringsstandard for opplysninger om pågående forskning og faste

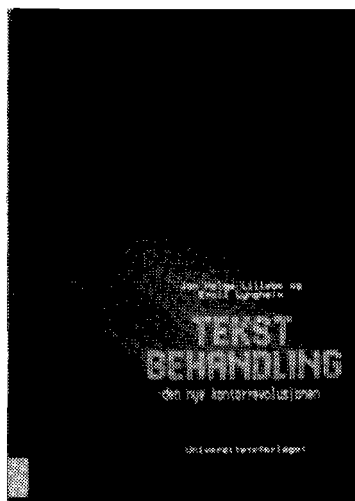
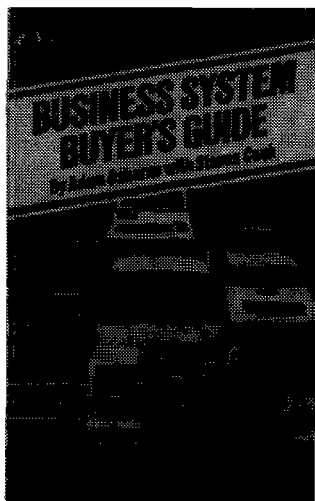
stor vekt på at kjøperen presist spesifiserer sitt behov.

Et eget kapittel er viet omtale av mikromaskinens oppbygning og virkemåte, inklusiv periferutstyr som skrivere, disketter o.l. Valg av programmeringsspråk behandles også og forskjellene blir greitt klargjort. Boken har klare råd når det gjelder billige løsninger som f.eks. bruk av eget TV til skjerm eller bruk av kassetter. Mikromaskinmarkedet er i rivende utvikling og løfter om nye produkter/forbedringer er en dominerende del av markedet. Osborne har her gode råd å gi. «When you are told that the program or hardware you are waiting for will be available in July (or any other month), you should ask: «Of which year?» og «He who buys on the cutting edge of technology shall be sacrificed upon it.». Mulige unntak fra disse gode prinsipper blir imidlertid også tatt med.

Boken har en sterk aksentuering av gode prinsipper som kjøperen bør være fortrolig med før han møter selgeren. Det synes jeg er nyttigere enn et rent markedsorienterende opplegg. Boken er på en måte holdningsskapende for kjøperen og ut fra mine erfaringer er dette viktig. Mikromaskinmarkedet er fremdeles helt spesielt, også i Norge må man forutsette at «competent vendors are a distinct minority». Det advares på den ene side mot å gå alt for sterkt etter lav pris («You get what you pay for.») og på den annen side imot for dyrt utstyr («Powerful performance characteristics are irrelevant if you do not need them.»).

Boken behandler vesentlig 8-bits maskiner. Siden den kom ut i 1981 har den ikke noe om den nyeste utvikling, f.eks. 16-bits maskiner og nettverk, og den omtaler vesentlig amerikansk utstyr. Bokens styrke slik jeg ser det er imidlertid at den går utover det rent tekniske og prismessige, og på en lettlest og fornøyeelig måte orienterer leseren om et nytt felt. Boken er fortsatt et godt kjøp!

Roald Skarsten



Lillebo, Jan Helge og Lyngheim, Eroll:
Tekstbehandling – Den nye kontorrevolusjonen. Universitetsforlaget, Oslo, 1982. 180 sider, kr 140.

Denne boka kom høsten 1982 og har fått relativt stor presseomtale m.v. i forbindelse med lanseringen. Ønsket fra forlaget er vel å gjøre den til standardverket om tekstbehandling, men en del svakheter kan gjøre dette vanskelig. Allerede prisen – kr 140,- – innebærer at folk som kun ønsker en første innføring kan slippe lettere fra det økonomisk ved å kjøpe mindre hefter o.l.

La det imidlertid være sagt med en gang at boka er god. Den er omfattende, og det meste av det som er tatt med synes å være relevant. Den er også pedagogisk godt lagt opp, bl.a. med et detaljert studieopplegg for brevkurs som en integrert del av boka. Den inneholder også en rekke kontrollspørsmål og introduserer hele tiden nye begreper på en fin måte.

Boka er delt inn i 3 deler. Den første delen definerer begrepet tekstbehandling og skisserer utviklingen fram til i dag. Den andre delen tar for seg den praktiske bruken av tekstbehandling og i siste delen ser en bl.a. på utviklingstendensene framover. Vi må med en gang si at del 2 synes å være for kortfattet, bl.a. er det kun avsatt knappe 20 sider til såkalt redigering av dokumenter. Dette er jo et svært sentralt punkt i all tekstbehandling og den vanlige bruker vil som regel ha en detaljert og grundig innføring i de mulighetene som finnes. I boka burde det vært laget flere eksempler, og hvert eksempel kunne vært kommentert mer detaljert. Det er tydelig at forfatterne har hatt et noe teoretisk forhold til tekstbehandling og lagt for stor vekt på å fremstille sammenhenger og forklaringer. Den enkelte elev og nybegynner vil på dette punkt og i andre sammenhenger finne stoffet for kompakt og dermed lite pedagogisk. Dette er ikke en gjennomgående feil, men gjelder først og fremst de mer praktisk orienterte kapitlene.

Når det gjelder del 1 som presenterer begrepet tekstbehandling og del 3 om utviklingstendensene m.v. er boka etter mitt syn uten lyte. En liste med ordforklaringer bak i boka er nyttig, men er vel på en del punkter unødvendig (ord som arbeidsgruppe, service, m.v. er vel unødvendig å ta med i en slik liste?)

Alt i alt er altså boka etter min mening en grei innføring i emnet. Den har sin store styrke i oversikt og framstilling av sammenhenger.

Svein Skotheim

Heggstad, Kolbjørn:
Norsk frekvensordbok. Universitetsforlaget, 1982. 160 sider, kr. 115.

Kolbjørn Heggstad bygger her på eit avismateriale på omlag 1 mill. ord henta frå Aftenposten, Dagbladet og Morgenbladet i 1968 (i alt 2/3 av materialet), NTB i 1971 og Stavanger Aftenblad, Adresseavisa og Drammens Tidende i 1972. Boka har 3 hovudlister, nemleg dei 10.000 hyppigaste orda ordna 1) etter rangnr., 2) alfabetisk og 3) ei liste der orda er ordna alfabetisk i grupper innom rangnr. 1-1000, 1001-2000,

osb. ned til rangnr. 5000. Rangnr. er gjeve etter at frekvensen er modifisert for spreinga (Juillands spreingsformel), slik at t.d. hyppige ord som er konsentrerte i delar av materialet får lågare rang enn ord som ikkje er nytta fullt så ofte, men har jamnare spreing. Alle personnamn er utelatne.

Hovudlistene gjev berre rangnr. og modifisert frekvens. Ein del lesarar ville vel ønska seg absolutt frekvens i tillegg. Tal og sortering gjeld ordformer, slik at ulike former av samme lemma (= oppslagsord) finst på ulike stader i hovudlistene. Når ulike ord er likt skrivne, er dei plasserte under eitt.

Hovudlistene er såleis trivielle, men godbitane finst på dei 20 første sidene: Her får vi døme på korleis ei lemmatisert liste (der ulike former av same lemma står under samla frekvens) kan ordnast. Vi får også ein smakebit på homografseparerte lister (som skil mellom likt skrive ord, som t.d. TRE: tal 87,3%, subst. 6,3%, verb 6,3%).

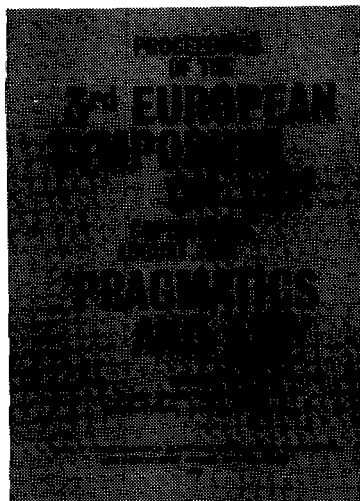
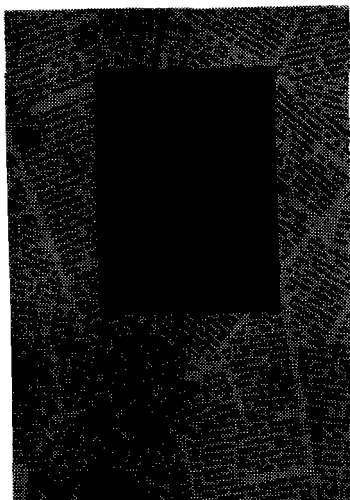
Dei mest brukte substantiv, verb og adjektiv finn vi i kvar si lemmatiserte liste, og ordklassfordelinga i heile materialet er berekna.

Forfatternar gjev vidare eit interessant oversyn over dei 30 vanlegaste forstavingar og endingar, han gjev nokre prøver på vanlege ordkombinasjonar, gjer greie for bokstavfrekvensar og ordlengder og gjev ei rangering av geografiske namn og personnamn.

Heggstad og medarbeidarane hans har her gjeve lingvistar, pedagogar og andre med interesse for kvantitative sider ved moderne norsk bokmål lettare tilgang til dette store materialet, som tidlegare berre har vore utgjeve i rapportform. Boka kan koma til nytte i svært ulike samanhengar, og bør også kunne inspirera til meir utførlege studiar innom fagfeltet.

Harald Solevåg, Gulbrand Alhaug, Anne Grete Strøm-Erichsen og Ivar Utne har vore med og arbeidt fram materialet.

Per Vestbøstad



**Høedt, Lundqvist, Picht og Qvistgaard (red):
Proceedings of the 3rd European Symposium on LSP. Erhvervsøkonomisk forlag, 1982. 643 s.**

Det 3. europeiske symposiet om LSP (Language for Specific Purposes), «Pragmatics and LSP», fant sted i København i august 1981. Flesteparten av foredragene som ble holdt på symposiet foreligger nå i bokform. De er samlet i følgende avdelinger:

tekstypologi

språkhandlinger og kommunikasjon

terminologi - leksikografi - oversettelse - tolking

didaktikk

Boka kan skaffes ved henvendelse til: *Fagsprogligt center, Handelshøjskolen i København, Fabrikvej 7, 2000 København F.*

K.N.

Linguistische Datenverarbeitungssoftware. Standard- und entwicklungsprogramme

Red.: Informationszentrum Sozialwissenschaften, Bonn/Institut für deutsche Sprache, Mannheim. Bonn/Mannheim 1982. Ca. 400 s., DM 78.

Første del av denne boka består av en dokumentasjon av 54 standardprogrammer/programsystemer til anvendelse på følgende områder: Indekser, konkordanser, frekvenslister, ordbøker, morfologi, fonologi, syntaks, innholdsanalyse m.m. I tillegg blir 20 eksperimentelle programmer/programsystemer dokumentert. Disse er til bruk ved maskinoversettelse, språkforståelse, simulering av språkatferd, tale-mønsterjerkjenning, pragmatikk og akustisk talegenerering.

Boka kan bestilles fra: *Institut für deutsche Sprache, Abt. Wissenschaftliche Dienst, Postfach 5409, D-6800 Mannheim 1, FRG.*

K. N.

Foredragssamling - Den 5. nordiske IoD-konferanse

Foredragene som ble holdt ved den 5. Nordiske IoD-konferansen i Trondheim i juni 1982 foreligger nå samlet i bokform. Motto for konferansen var «Menneske - System - Maskin», og de ca. 35 foredragene spenner over følgende emner: Edb-brukerens arbeidssituasjon, systemendringer som berører informasjonssøking, og utviklingen innen maskinvarer og elektroniske medier. Samlingen koster kr. 100 og kan skaffes ved henvendelse til: *Norsk dokumentasjonsgruppe, NSI, Postboks 350 - Blindern, Oslo 3.*

K. N.

MELDINGER

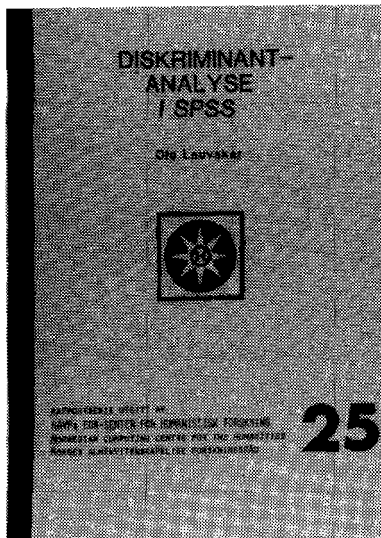
Senterrapport nr. 25

Ole Lauvskar: Diskriminantanalyse i SPSS

Denne rapporten har bakgrunn i eit samarbeid mellom *Ole Lauvskar*, statistikkonsulent ved Senteret, og mag.art. *Bergljot Solberg* i åra 1979-1981. Som utdanningsstipendiat-NAVF ved Historisk institutt, Universitetet i Bergen, var Solberg interessert i å få statistisk rettleiing ved analyse av eit spydmaterial frå yngre jernalder.

Då ein skulle velja metode for å klassifisera spyda, var diskriminantanalyse aktuell. Det er røynslene frå arbeidet med dette spydmaterial som ligg til grunn for framstillinga.

Rapporten startar med ei lett innføring i diskriminantanalysen. I siste delen vert bruken av SPSS i prosjektet gjennomgått. (SPSS er ein statistisk programpakke for samfunnsvitarar, men han kan godt brukast av andre.) SPSS-oppsettet og SPSS-utskrifta for diskriminantanalysen av spydmaterialet blir forklara og kommentert i detalj.



EDB-versjon av Nordic Archaeological Abstracts

Nordic Archaeological Abstracts har vært utgitt av arkeologiforskere hvert år siden 1974. Tidsskriftet, som er finansiert av de nordiske forskningsrådene, består av sammendrag av all litteratur som blir utgitt i Norden vedrørende arkeologisk forskning. Innholdet omfatter dessuten fellesnordiske bibliografier.

Tidsskriftet blir nå forsøksvis overført til edb for å gjøre materialet tilgjengelig for forskere utenfor Norden. Dette arbeidet blir finansiert av NORDINFO, og det er Norsk Dokumentdata som tar hånd om databasen. Prosjektets styringsgruppe har representanter fra bl.a. Arkeologisk museum i Stavanger og Riksantikvaren.



*Helleristning fra Ausevik, Flora i Sogn og Fjordane.
(Foto: Historisk museum, Universitetet i Bergen).*

Aktuelle tidsskrifter

SCOPE (Scholarly Communication: Online Publishing and Education) er et nytt amerikansk meldingsblad som utkommer annenhver måned. Bladet inneholder nyheter om mikroprosessor-basert teknologi: mikromaskiner, nettverk, programvare for undervisning, informasjonstjenester og databaser. Også nye publikasjoner og kommende konferanser blir meldt. Abonnementsprisen er \$47 pr. år (\$84 for to år). Bladets adresse: *SCOPE, Queens College/CUNY, Flushing, New York, 11367, USA.*

Det nye tverrfaglige tidsskriftet *Multilingua* blir utgitt med støtte fra EF-kommisjonen. Vi sakser fra tidsskriftets egen presentasjon:

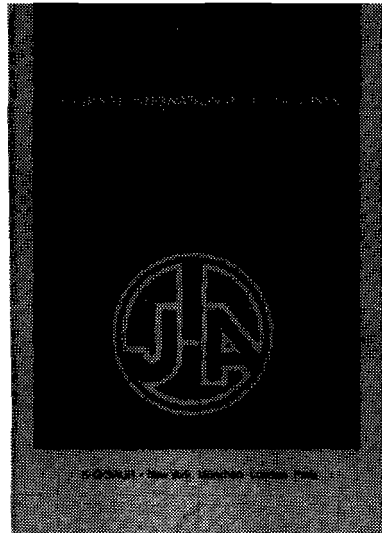
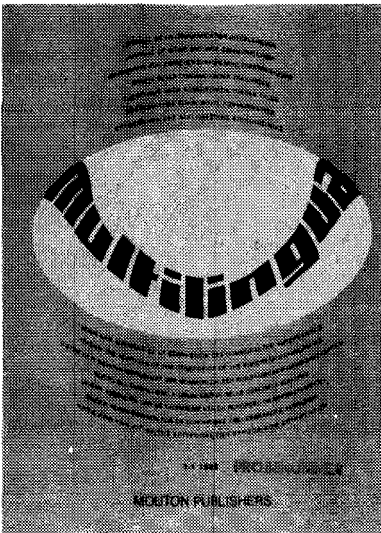
Multilingua ønsker å stimulere til diskusjon, ny forskning og aktiviteter som sikter på å overvinne språkbarrieren. Det skal fungere som et forum for kontakt og samarbeid blant de mange ulike yrkesgrupper involvert i mellomspråklig informasjonsarbeid, dokumentasjon og kommunikasjon. Det vil prøve å presentere, undersøke og øve innflytelse på språkbrukspolitikken i et lingvistisk variert samfunn.

Multilingua skal konsentrere seg om:

- anvendelsen av informasjonsteknologi i forbindelse med problemløsning innen mellomspråklig kommunikasjon
- moderne metoder for behandling og bruk av flerspråklige redskaper, slik som tesauri, automatiske ordbøker og maskinoversettelsessystemer
- lingvistisk forskning viet til en forbedring av forståelsen av naturlige språk og hjelp til overføring av informasjon mellom språk.

For å oppnå denne målsetningen, skal tidsskriftet publisere artikler, studier, forskningsnotater og -rapporter, anmeldelser og sammendrag på følgende områder: Språkpolitikk og -planlegging, flerlingvistiske aspekter i informasjons- og dokumentasjonsarbeid, spesielle språk, leksikografi, oversettelse, tolking, språkundervisning og terminologi.

Multilingua utgis i fire nummer á 64 A4-sider pr. år. Et abonnement koster DM 110 for institusjoner og DM 50 for privatpersoner, og kan tegnes hos: *Walter de Gruyter & Co., Postfach 110240, D-1000 Berlin 11, West Germany.*



The International Council on Archives utgir to tidsskrifter. *International Journal of Archives* utkommer to ganger i året og inneholder artikler med flerspråklige sammendrag om forskning i og om arkiver. Det gis dessuten informasjon om utgiverens virksomhet og kommende konferanser. Abonnement kan tegnes hos: *K. G. Saur Verlag KG, Pöissenbacherstrasse 26, D-8000 München 71, Federal Republic of Germany.*

ADPA er den årlige bulletinen til organisasjonens edb-komité. Bulletinen består av artikler om arkiver og edb, meldinger om komitéens aktiviteter, og bibliografier over relevant litteratur. Abonnements-adresse: *J. Pieyns, rue des Buissons, 65, B-4000 Liège, Belgique.*

Computers and Medieval Data Processing (Informatiques et Études Médiévales) er et engelsk/franskspråklig meldingsblad som utkommer to ganger i året. Bladet inneholder beskrivelser av nye edb-støttede prosjekter innen middelalderforskningen, referater fra seminarer o.l., bibliografier over relevant litteratur, og meldinger om kommende konferanser m.m. Abonnementet er gratis og kan tegnes hos: *Serge Lusignan, Institut d'études médiévales, Université de Montréal, C.P. 6128, Succ. «A», Montréal, Québec, Canada, H3C 3J7.*

Som vedlegg til høstnummeret av bladet er det hvert år et spørreskjema rettet mot alle som bruker edb i middelalderforskning. Utfylte spørreskjema danner grunnlaget for bladets fortløpende oversikter over prosjektbeskrivelser. Kopier av skjemaet kan fås ved henvendelse til Senteret.

Nytt

i

biblioteket



- Abrahamsson, Sixten (red.): *SCANNET: ABC for databasetilslutning.* Kirjastopalvelu, Helsingfors, 1980. 13 s.
- Bailey, Richard W. (ed.): *Computing in the Humanities. Papers from the fifth international conference on computing in the humanities, Michigan, May 1981.* North-Holland, Amsterdam, 1982. 191 s.
- Berrefjord, Per H. og Torp, Ernst B.: *Dataguiden: oppdateringsutgave 1982.* Ingeniørforlaget, Oslo, 1982. 124 s.
- Bræk, Rolf: *Håndbok i systemarbeid.* Tapir, Trondheim, 1982. 150 s.

- Bølviken, Erik: Preliminary lecture notes for the Nordic research course in multivariate analysis in archaeology. Univ. i Tromsø, 1981. 66 s.
- Fagerli, Mundal, Novakowski: NORMARC: veiledning i MARC-katalogisering og kort oversikt over arbeidsrutiner. Universitetsbiblioteket i Oslo, 1980. 56 s.
- Goetschalckx, J. and Rolling, L. (eds.): Lexicography in the electronic age. Proceedings of a Symposium held in Luxembourg, 7-9 July, 1981. North-Holland, Amsterdam, 1982. 276 s.
- Hajicová, Eva (ed.): COLING 1982. Abstracts. Univerzita Karlova, Praha, 1982. 308 s.
- Horecký, Ján (ed.): COLING 82. Proceedings of the Ninth International Conference on Computational Linguistics. North-Holland, Amsterdam, 1982. 432 s.
- Johnsen, Kjetil: Systemtekniske konsekvenser av persondatalovgivningen. Universitetsforlaget, Oslo, 1981. 202 s.
- Knuth, Donald E.: TEX and METAFONT: new directions in typesetting. Digital Press, Bedford, Mass., 1979.
- Lawson, Veronica (ed.): Practical experience of machine translation: proceedings of a conference, London 1981. North-Holland, Amsterdam, 1982. 199 s.
- Lie, Ulf og Skjoldmo, Sissel: Behovet for fremmedspråkkompetanse i næringslivet. Univ. i Bergen, 1982. 30 s.
- Ongstad, Per: Evaluering av søkesystemer. Norsk Dokumentdata, Oslo, 1982. 159 s.
- Salton, Gerard and McGill, Michael J.: Introduction to modern information retrieval. McGraw-Hill, New York, 1983. 448 s.
- Voje, Kirsten: Barneforskning: registrering av forskning om barns livsforhold og utvikling. NAVF, Oslo, 1982. 230 s.
- Zaks, Rodney: The CP/M handbook with MP/M. Sybex, Berkeley, Cal., 1980. 321 s.
- Aasen, Brattvåg, Eidhammer, Tøfte: Innføring i databehandling. Tanum-Nordli, Oslo, 1982. 375 s.

Infodial – Paris, 24.-27. mai 1983

Infodial, som er en internasjonal kongress og utstilling om databaser og databanker, vil innbefatte tekniske og faglige sesjoner om kunst og kultur. Opplysninger om programmet kan skaffes fra: *Infodial, 4 place de Valois, 75001 Paris, France.*

IDT83

IDT83 (Information, Documentation, Transfert des Connaissances), den 5. internasjonale kongress om informasjon og dokumentasjon, skal holdes i Grenoble 8.-10. juni 1983. Flere opplysninger kan fås fra: *Alpes Congrès, Avenue d'Innsbruck, 38029 Grenoble Cédex, France.*

6. skandinaviske forskningsseminar om systemering

Bergen 8.-11. august 1983

Seminarret ønsker å samle ca. 30 deltakere fra Danmark, Finland, Island, Sverige og Norge omkring temaet «Det menneskelige perspektiv i systemering». Det er et ønske at flest mulig deltakere presenterer egne forskningsprosjekter innen systemering.

All henvendelse vedrørende seminarret rettes til: *Professor Markku Nurminen, 6. Scandinavian Research Seminar on Systemeering, Institutt for informasjonsvitenskap, 5014 Bergen-Universitetet.*



IJCAI-83

Den 8. internasjonale konferanse om kunstig intelligens skal avholdes i Karlsruhe, Vest-Tyskland, 8.-12. august 1983. Målet for IJCAI-83 er å fremme den vitenskapelige utvekslingen innen og mellom alle greinene av kunstig intelligens blant verdens forskere.

Konferansen vil omfatte foredrag, paneldiskusjoner og programdemonstrasjoner. Det planlegges også en utstilling av maskin- og programsystemer til bruk i kunstig intelligens-forskning og -utvikling, bl.a. robot-teknologi og ekspert-systemer.

Kopier av de foreløpige konferanse- og utstillingsbrosjyrene kan fås ved henvendelse til Senteret.

Internasjonal konferanse i leksikografi

Exeter, september 1983

En internasjonal konferanse om ulike aspekter ved leksikografi skal finne sted ved University of Exeter 9.-12. september 1983. Det vil bli holdt noen forelesninger om generelle spørsmål innen ordboksproduksjon, men hovedvekten skal legges på seksjonsmøter om følgende mer spesialiserte emner: a) den historiske ordboka, b) den tospråklige

ordboka, c) ordbøker for elever, d) terminologistandardisering, e) maskinstøttet leksikografi.

Konferanseavgiften er \$90. Flere opplysninger kan skaffes fra: *Dr. R. K. Hartmann, The Language Center, University of Exeter, Exeter EX4 4QH, Devon, U.K.*

De nordiske datalingvistikkdagene 1983

De nordiske datalingvistikkdagene 1983 blir arrangert 3. og 4. oktober av Centrum för datalingvistik ved Uppsala universitet. Som tidligere år skal dagene fungere som et nordisk forum for presentasjon og diskusjon av datamaskinell metodikk og tillempling innen språkvitenskap. Tema for årets symposium blir automatisk morfologisk og syntaktisk analyse (parsing).

Foredragsholdere må levere sammendrag senest 1. juni. Endelig program for arrangementet vil foreligge i august.

I tilslutning til De nordiske datalingvistikkdagene skal det arrangeres et møte 5. og 6. oktober om datamaskinstøttet leksikografi og terminologi. Møtet blir en oppfølging av det som ble holdt i Sandbjerg i 1982 om samme tema. Norsk Termbank er arrangør av den faglige siden.

Påmeldingsblanketter til begge symposiene kan skaffes fra Senteret.



CHI '83 – foredragsinnbydelse

CHI '83-konferansen skal finne sted 12.-15. desember 1983 i Boston, Massachusetts. Sponsorer blir ACM SIGCHI og The Human Factors Society. Tema for konferansen er «menneskelige faktorer i edb-systemer», med fokus på presentasjonen av teori og praksis m.h.t. bruksverdien av interaktive systemer. Det ønskes foredrag som dreier seg om menneskelige faktorer innen programvare, maskinvare og dokumentasjon, spesielt innenfor følgende emner: metodologi, grense-snitt- og språkdesign, og evaluering av systemutførelse.

Fristen for innlevering av manuskripter er 1.5.83. Nærmere opplysninger kan fås ved henvendelse til Senteret.

SUMMARY

Om forholdet mellom EDB og humaniora

On the relationship between ADP and the Humanities

The background for this article by Professor Markku Nurminen and M.A. student Kjell Bjørn Rønning, the Department of Information Science, University of Bergen, is the fact that ADP technology is becoming an increasingly important tool in research in the Humanities.

The authors start the article by giving definitions of the Humanities and the ADP-orientated subjects respectively. The two fields of study seem to have a totally different basis, ways of approach to problems, and time horizons. Both an «application-orientated» approach and a «reflective» approach to the research object can be considered meaningful methods for combining the two.

The first approach implies using ADP technology and knowledge in research in the Humanities not only as a tool but as a means of development and management of information systems. Combining ADP and the Humanities in this way raises the question: To what degree can ADP technology «steer» the contents and process of research in the Humanities?

The «reflective» approach implies a transferral of humanistic views and knowledge to ADP subjects. The Humanities can make it possible to describe and treat phenomena that the ADP subjects lack concepts for. The problem here is to define a platform for incorporation of the humanistic tradition in these subjects.

In the authors' opinion a synthesis between these two research traditions is difficult, due to the fact that their differences cannot be easily resolved. This view is illustrated by Habermas' classification of the various sciences according to the «knowledge interests» that constitute them. Nurminen's and Rønning's conclusion is that the humanistic perspective must be integrated in the methodology of the computer sciences.

Opphavsrett og bruk av tekster i maskinlesbar form

Authors' rights and the use of texts in machine-readable form

University Lecturer Jon Bing, University of Oslo, starts this article by stating two reasons why researchers are interested in the availability of existing texts in machine-readable form: Analysis of texts can be facilitated, and computer-based information systems can make efficient retrieval of interesting texts in a collection possible.

As computer systems facilitate the production, spread and processing of texts, authors are paying more attention to the protection of their rights. However, many legal rules and terms are problematic in connection with new technology. Bing outlines the implications of authors' rights for machine-readable texts, both for public and private use. The Norwegian Research Center for Computers and Law, University of Oslo, is currently carrying out research in this field.

EDB og kvinner – en statusrapport

ADP and women

Consultant Eva Møller at the Centre raises two main issues in this article: the consequences of ADP technology for the employment of women in general and the contents of their work, and training and work conditions for highly qualified women who work with ADP.

Fears that women especially will suffer from unemployment as a result of data technology are based on the following circumstances: Many female occupations are found within sectors that are becoming highly automated; women are mainly employed at the bottom of the organisational hierarchy; they have less education, more frequent periods of non-employment and have part-time work more often than men. In this connection

Møller refers to a study of the consequences of data technology in Norwegian banks.

Women make special demands in regard to job satisfaction which are not taken into consideration when ADP technology is introduced. Scandinavian studies show that women employed at e.g. social security offices are worried about the loss of the *caring* aspects of their work caused by automated systems.

In accordance with goals for sexual equality in Norway efforts have been made to increase the number of female ADP students, with good results. However, European and American studies show that women have the lowest-paying positions within the data industry, and that they earn less than men in the same occupations.

Other studies show that female students are opposed to some of the attitudes which have formed a traditional part of the training and work of engineers. Women wish to integrate a human dimension with the technical. In Møller's opinion the demands and expectations women are confronted with in male-dominated occupations must be examined more closely.

LÆREBOKSPRÅK analysert ved hjelp av EDB Textbook language analysed with the aid of ADP

Foreign pupils in Norwegian elementary schools often have problems with textbook language due to a lack of variety in their vocabulary. University Lecturer Anne Hvenekilde and Assistant Teacher Anne Golden at the Department of Scandinavian Languages and Literature, the University of Oslo, have analysed the language in grade 4-9 textbooks for physics, geography and history, and prepared support materials on the basis of these analyses.

The texts within each subject were combined into units for each grade and registered via ADP. Using special ADP programs, graph words, homographs and «root lemmas» (consisting of conjugations and derivations of the same root) in each unit were sorted according to different criteria (e.g. frequency). The most frequent words (function words and other common words not included) constituted the basis for the preparation of exercise books for foreign pupils – one for grades 4-6 and one for grades 7-9 in each subject. The contents of these books are described.

Tasks that can be carried out with the aid of ADP and the text corpus at hand are automatic root lemmatisation and separation of compounds, quantitative and qualitative analysis of easy-to-read versions of the texts, analysis of textbook language from grade to grade, an investigation into the relationship between special subject words and other words, charting of different types of compound verbs, and descriptions of the various types and frequencies of collocations.

EDB i fornminneforvaltning ved et arkeologisk museum ADP in the administration of antiquities at an archaeological museum

According to Curator Trond Løken, an ADP register of the finds and antiquities administered by the Archaeological Museum in Stavanger has been desirable for many years. Manual registers pose a number of problems both internally and for planning authorities.

In 1980 the museum asked Rogalandsdata and the Rogaland Research Institute to help develop an ADP system which would make its data more easily accessible. Since then a system has been planned for collection via screen terminal of information found in registers of antiquities and finds and from the museum's archaeological-topographical archive. Natural history and picture archives will also be included in this system.

A register of antiquities (Amis I) is now operative. Amis I is an interactive system covering registration, correction, storage, editing, and retrieval. Løken describes these functions in detail, along with the possibilities of connecting the system to other relevant archives and data bases.

EDB og ordbøker ADP and dictionaries

This article by Consultant Håvard Hjulstad deals with some ADP projects in lexicography and terminography he is taking part in. Lexicographical work is well suited to data processing because it is based on extensive sources, alphabeticization is required, and lexicographical forms of presentation are strictly formalised.

After discussing the main differences between lexicography and terminology/terminography, Hjulstad describes the work taking place at the Norwegian Term Bank, University of Bergen, on a dictionary of mainly technical terms in Norwegian, English,

German, and French. The material is stored on small files which will be linked together in a data base. The program system which has been developed is simple but flexible enough to be used on all types of data. It includes possibilities for field generating, error finding, sorting, print-outs, printing via photo composition, searching, and concordances. The system has been used in the production of a number of technical dictionaries.

In co-operation with the Department of Norwegian Lexicography, University of Oslo, and a publishing company, Hjulstad is also at work on the data processing of general Norwegian dictionaries. These dictionaries have been given a common format so that the same programs can be used for all of them. The formalisation of various types of information included in the dictionaries is discussed.

Norsk-vietnamesisk ordbok Norwegian-Vietnamese dictionary

In this article Consultant Eva Møller at the Centre describes the production taking place in Bergen of a Norwegian-Vietnamese dictionary. ADP has been used in the preparation and proof-reading of the material.

The work on the dictionary was started in the autumn of 1979 by a group of Vietnamese living in Norway. Later the venture became a Norwegian-Vietnamese co-project involving several departments at the University of Bergen. Project leader the past two years has been University Lecturer Marianne Haslev Skånland of the Department of Phonetics and Linguistics.

The dictionary's vocabulary is adapted to the needs of Vietnamese immigrants in Norway. The selection of words is based on existing Norwegian dictionaries, brochures issued by public offices, and suggestions made by teachers of Norwegian for Vietnamese and the group of translators involved in the project. The complete dictionary will contain a total of 16-17.000 entries translated into Vietnamese. In connection with main entries examples of common usage are given and translated when deemed necessary.

Møller summarises both the principles upon which the Vietnamese alphabet is based and the code system that had to be defined in order to register and store Vietnamese written language with the University's ADP equipment. Details are also given of the conversion of the code system required for the process of proofreading and photocopying.

Automatisk analyse av norsk Automatic analysis of Norwegian

Director Jostein H. Hauge interviews Per-Kristian Halvorsen, researcher at the Centre since January 1982. Halvorsen's field is automatic syntactic and semantic analysis of the Norwegian language. His work is conducted in close co-operation with Xerox Corp. Palo Alto Research Center and MIT in the U.S.

In Halvorsen's view, the increasing emphasis on the study of the *language process* in the U.S. is highly valuable to the field of computational linguistics. Researchers in informatics, psychology and linguistics take part in these studies. The new scientific personal computers, which have a very large address space and extremely good graphic qualities, are also of great importance to this field.

Halvorsen is primarily engaged in the development of a formal semantics for natural languages, based on models for mathematical logics. Computer implementation of systems for representation of meaning via a logical language is difficult. Within this framework Halvorsen has developed a semantic theory for lexical-functional grammar (LFG).

The theory behind LFG enables using common linguistic techniques and descriptions in any language as the foundation for ADP-based language-understanding systems. According to Halvorsen, parallels to foreign systems cannot be developed independently in Norway, but one can hope for modular systems which can be modified to meet Norwegian requirements.

In order to strengthen developmental work and achieve better results in computational linguistics in Norway, a higher degree of co-operation between research institutes and contacts with the computer industry are essential. A step in this direction is the recent establishment of a national contact group for researchers involved in computational analysis of natural languages.

Tekstbehandling i humanistisk forskning Word processing in research in the Humanities

The use of microcomputers within research and higher education is on the increase.

This is the background for a debate, led by Information Consultant Svein Skotheim at the Centre, between University Lecturers Ralph Jewell (Philosophy), Jan Oldervoll (History), and Helge Sandøy (Scandinavian Languages and Literature) – all at the University of Bergen, and Editor Johannes Kvestad. Subjects touched upon included word processing, communication between microcomputers and main frames, techniques for publication, and how the use of computers can influence the choice of research strategies and topics.

In all the participants' experience microcomputers are valuable aids to research work. For example, electronic word processing enables new ideas and suggestions to be developed and incorporated more freely during the writing process. Word processing also simplifies writing, editing, and correction of textbooks and bibliographies.

There was general agreement that computers do not «steer» research – on the contrary, they stimulate innovative approaches and solutions. However, there is a danger that research can become mere data processing in a narrow sense.

Den andre nasjonale namnegranskar konferansen The second national conference for name researchers

Consultant Eirik Lien, University of Trondheim, reports from this conference, which was arranged by the Institute for Name Research, University of Oslo, in November. About 60 people took part.

The conference included a comprehensive review of the Institute's use of ADP, short presentations of both local and foreign projects, and a discussion. Mention is made of Håvard Hjulstad's program system FELTED, which he has created for the Institute (see HD 3/4-82).

In Lien's opinion, name material constitutes a type of data which is well suited to systemisation by the use of ADP. The problems involved are the same as those connected with archives in general.

The Institute for Name Research is to publish a report bringing the speeches of the conference along with a summary of the ensuing discussion. The report will be presented in a later issue of *Humanistiske Data*.

Humaniora-uken ved Universitetet i Oslo 1983 The Humanities week at the University of Oslo 1983

In January the Faculty of Arts, University of Oslo, arranged lectures, exhibitions, concerts etc. under the title «Language and Culture: the Humanities in the 80's». Research Assistant Hege Brit Randsborg reports that the programme was based on five main themes: Humanities in a technological age, the Humanities as mediators of foreign cultures, the national inheritance, Wagner and European culture, and humanistic plurality.

Four exhibitions drawing on ADP-based research projects were shown. These dealt with birth control in Norway from the end of the last century onwards (see *Humanistiske Data* 3/4-82), the 1875 Christiania census, ADP and local history, and problems immigrants encounter in connection with textbook language (see also p.).

Several lectures on ADP and lexicography, language analysis, and music were held, along with a presentation of the faculty's course on ADP and the Humanities. Demonstrations were given of a computer which transfers music from a specially devised organ to machine-readable form, and of the faculty's use of ADP in general. Randsborg gives detailed accounts of all of these activities.

BIBNETT/BIBNETT 2

The BIBNETT project, described in *Humanistiske Data* 2-81, was concluded in March 1982. According to Librarian Ase Støgård, the project showed that program-to-program communication between independently developed Norwegian library systems is technically possible. The systems can be linked together with the aid of current ADP technology in spite of the fact that they employ different types of data bases, user dialogues and computers.

The project BIBNETT 2 is to further the utilisation of, and contribute towards, computer network research on the basis of the results attained by BIBNETT 1 and the possibilities offered by UNINETT, the networking system operating between the universities. This project will be completed in 1984.

Dokumentasjonssenter for kulturpolitikk og kulturforskning Documentation centre for culture-studies and culture-politics

College Librarian Knut Bjander, Rogaland Research Institute, reports that a proposal for a documentation centre for cultural issues, located in Stavanger, will be submitted to the Department of Culture and Science this spring.

The centre will limit its activities to cultural politics in the broadest sense: the position of culture in society, and the importance of cultural expression and cultural organisations for the development of society.

The target groups will be employees in cultural administration; cultural organisations and institutions; and researchers and students of culture. Close co-operation with other documentation centres, both in Norway and abroad, is desirable.

The centre will offer the following: 1. Arrangement of contact between both different user-groups and these groups and cultural researchers/institutions etc. 2. Library services (in co-operation with the Rogaland Regional College), and ADP services (in co-operation with the Norwegian Computing Centre for the Humanities) for production of bibliographies, subject lists etc. 3. Distribution of its own newsletter and reports along with bibliographies.

NTAI-prosjektet The NTAI project

Curator Magne Velure, the Hardanger Folk Museum, reports that in 1978 a project group was appointed in order to develop and test an automatic indexing system to be used at the Nordic traditional archives. The NTAI project was to experiment with the use of ADP in cataloguing folkloristic source materials, keeping in mind the furthering of co-operation between the archives. The project group thus had to devise systems which would ensure joint cataloguing standards for basic information while at the same time allowing each archive to keep its distinctive qualities.

The group developed a form which was tested on various types of archive materials. Details are given on the types of information included in the form.

At the 3rd Nordic conference on traditional archives and documentation, held in Copenhagen in September last year, NTAI's test data base was presented. Members of the project group gave introductions to the use of ADP in cataloguing at traditional archives. The participants of the conference had the opportunity of practicing information retrieval from the data base via screen terminal.

In Velure's opinion the NTAI project will hardly lead to speedy adoption of cataloguing via ADP at the Nordic traditional archives. However, the project constitutes a basis for further development in this field, and shows that joint Nordic solutions are possible.

Programpaketen STAR The program package STAR

Research Fellow Stig Welinder presents the archaeological program package STAR, which he has developed together with members of the staff of the Centre. This spring the instruction and teaching manuals of the package, dealing with statistical and quantitative methods in archaeological research, will be completed.

The package consists of six parts: program, «Introduction and STAR manual», «Student textbook and STAR examples», «Archaeology for statisticians», «STAR algorithms», and program documentation. Additional information on STAR can be obtained from the Centre.

Avisinform-prosjektet «Information system for Norwegian newspapers»

This is a project being carried out by the Centre for Media Research at the University of Bergen, reports Kristin Natvig. The project was started in 1980 and will be completed this year.

One of the project's goals is to outline the concrete organisation of information in an ADP-based archive for newspaper materials. Suggestions are to be given for the construction of a thesaurus for this type of material. Furthermore, the technical, financial and legal aspects of the establishment of press data banks will be investigated. The need for an automated archive system and attitudes to co-operation between newspapers will also be accounted for.

Two test data bases have been established in connection with the project. One contains

700 newspaper articles, the other references to these articles. Retrieval efficiency in the two data bases will be compared. Reasons for retrieval failure will be examined along with methods of improving retrieval efficiency.

The project group has developed an archive model which involves both storing articles in full text and selection later on for further storage. The group has also suggested co-operation models for an information centre for newspapers on various geographical levels. Journalists' use of archives of cuttings compared to other sources has been investigated.

Planned tasks include a quantitative and qualitative registration of the use of newspaper archives, and an investigation into the consequences of computer technology for newspaper personnel.

EXTEMPLO - et nordisk elektronisk tidsskrift

EXTEMPLO - a Nordic electronic journal

Research Assistant Kristin Natvig at the Centre reports that according to project leader Gunvor Svartz-Malmberg, the electronic journal Extemplo is based on the need for fast news mediation within the sector of information and documentation. As information centres and research libraries are well-equipped with screen terminals, the journal can address readers used to working with text information despatched via telecommunications networks. Thus a journal of this type can form a useful basis for future electronic publication of scientific reports and articles.

Extemplo has been published bi-monthly and free of charge since 1979. The journal contains a calendar of courses, conferences, seminars, and lectures relevant to information and documentation services, along with information from the various Nordic online user groups. The editorial staff consists of representatives from the Nordic countries, each of whom edit a national column.

The publication of Extemplo is now connected with the Swedish project «Electronic news mediation». The goal of this project is to gain knowledge about methods applicable to electronic publication of journals.

Nytt fra RHF/NAVF

News from the Council for Research in the Humanities

In the years 1973-75 the Council carried out an investigation into research activities within the Humanities in Norway. The results were published in a report which is now to be followed up, reports Senior Consultant Egil Kallerud. The material will be brought up to date and developments identified. In addition, the relationship between the Humanities and society will be analysed historically and theoretically, as a contribution to the current debate about the Humanities' legitimacy and possibilities in modern-day society.

Director Jostein H. Hauge reports on the Centre's recent proposal concerning an infrastructure for documentation of research in progress in the Norwegian Research Council for Science and the Humanities. The background for the proposal is the work recently carried out by the five Norwegian research councils aimed at improving documentation activities. A committee has been appointed which is to suggest joint registration standards for information on research in progress and guidelines for efficient interdisciplinary exchange and mediation of information.

The Centre's proposal suggests that a co-ordinated system for each council be gradually built up on the basis of systems developed for the Humanities and Social Sciences. Required for a system of this kind is improvement of universities' documentation standards and a computer-based documentation centre. An ADP system which covers all phases of project registration has already been developed by the Centre.

Meldinger

News

The Centre has issued a new report in its series, written by Statistics Consultant Ole Lauvskar. It deals with discriminant analysis of Iron Age swords, based on the statistical program package SPSS.

On a trial basis, the journal Nordic Archaeological Abstracts is now available to foreign researchers via ADP.

The following publications are presented: Multilingua - a new journal of interlanguage communication; SCOPE - an American newsletter concerned with online publishing and education; International Journal of Archives and ADPA, both published by the International Council on Archives; and Computers and Medieval Data Processing.

Forthcoming conferences

Infodial - Paris, 24-27 May: international congress and exhibition on data bases and data banks, including technical and professional sessions on the liberal arts. For programme information write to: Infodial, 4 place de Valois, 75001 Paris, France.

IDT83 - Grenoble, 8-10 June: 5th International Congress on Information and Documentation. More information from: Alpes Congr s, Avenue d'Innsbruck, 38029 Grenoble C dex, France.

IJCAI-83 - Karlsruhe, 8-12 August: 8th International Joint Conference on Artificial Intelligence. General enquiries can be directed to: Saul Amarel, Computer Science Department, Hill Center/Busch Campus, New Brunswick, New Jersey, 08903, USA.

International conference on lexicography - Exeter, 9-12 September: more information can be obtained from: Dr. R.K. Hartmann, The Language Centre, University of Exeter, Exeter EX4 4QH, Devon, UK.

Nordic conference on computational linguistics - Uppsala, 3-4 October.

CHI '83 - Boston, 12-15 December: conference on human factors in computing systems. Papers must be submitted by May 1 to Richard W. Pew, Bolt Beranek and Newman, Inc., 10 Malton Street, Cambridge, Massachusetts, 02238, USA.

Forts. fra 2. omslagsside.

RAPPORT nr. 14. *NOVA*STATUS HÅNDBOK*

Del 1: Søking. Brukerveiledning. 3. opptrykk februar 1983. ISBN 82-7283-011-6 Pris kr. 20.

Del 2: Fil-beskrivelser. Systemdokumentasjon. Utsolgt.

Del 3: Generering og oppdatering av databaser. Pris kr. 20.

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

RAPPORT nr. 16. *Årsmelding 1979*. NAVFs EDB-senter for humanistisk forskning. ISBN 82-7283-013-2 Gratis.

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. *Statistiske metoder på arkeologisk materiale*. Rapport fra et seminar på Bryggens museum, Bergen 24.-26. november 1980. Mars 1981. ISBN 82-7283-017-5 Pris kr. 35.

RAPPORT nr. 20. *EDB-prosjekter i humanistiske fag 1980*. Juni 1981. 2. opptrykk oktober 1981. ISBN 82-7283-018-3 Pris kr. 45.

RAPPORT nr. 21. *Rune Johansen: Bruk av EDB i teatervitenskapelig forskning*. Mai 1981. ISBN 82-7283-019-1 Pris kr. 35.

RAPPORT nr. 22. *Årsmelding 1980*. NAVFs EDB-senter for humanistisk forskning. ISBN 82-7283-020-5 Gratis.

RAPPORT nr. 23. *Stig Welinder: A program package for archaeological use*. 1981. ISBN 82-7283-021-3 Pris kr. 45.

RAPPORT nr. 24. *Rapport fra seminar om bruk av edb innen teater og teatervitenskap*. Januar 1982. ISBN 82-7283-026-4 Pris kr. 50.

RAPPORT nr. 25. *Ole Lauvskar: Diskriminantanalyse i SPSS*. Desember 1982. ISBN 82-7283-028-0 Pris kr. 55.

RAPPORT nr. 26. *Stig Welinder: Paleodemography*. Oslo 1982. ISBN 82-7283-030-2 Pris kr. 55.

RAPPORT nr. 27. *Årsmelding 1981*. NAVFs EDB-senter for humanistisk forskning. ISBN 82-7283-029-9 Gratis.

Vi takker
NAVFs EDB senter
for oppdragene
med:

- WORD FREQUENCIES IN
BRITISH AND AMERICAN ENGLISH
- HELSE- OG SOSIALPOLITIKK
- HUMANISTISK FORSKNING

Bøkene er på tilsammen over 1200 trykte sider.
Teksten er overført fra datamaskinen til
Universitetet i Bergen, via telefon til vår fotosetter.

bergmann

BERGMANNS BOKTRYKKERI A/S
Vestre Torggate 22, 5000 Bergen
Telefon 31 65 20