

h

UMANISTISKE

DATA



**NAVFs edb-senter
for humanistisk forskning**
The Norwegian Computing Centre
for the Humanities

1-87



SENTERETS RAPPORTSERIE

RAPPORTER UTGITT F.O.M. 1980

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.

HUMANISTISKE DATA

1-87

***NAVF's EDB-SENTER
FOR HUMANISTISK FORSKNING***

***The Norwegian Computing Centre
for the Humanities***

HUMANISTISKE DATA

utgis av NAVFs edb-senter for humanistisk forskning.

REDAKSJON

Einar-Arne Drivenes, Jostein H. Hauge, Kristin Natvig (red.).

ADRESSE

Harald Hårfagresgt. 31, Boks 53 – Universitetet, 5027 Bergen. Tlf. 05–212954/55/56.

ABONNEMENT

Gratis for enkeltpersoner, kr. 70,- for institusjoner (3 nr. pr. år).

Artikler, rapporter og meldinger mottas gjerne.

Medarbeidere fra Senteret i dette nummer:

Einar-Arne Drivenes, Jostein H. Hauge, Claus Huitfeldt, Kristin Natvig, Espen S. Ore, Øystein Reigem, Per Vestbøstad.

Redaksjonen avsluttet 10. juni.

Forsidebildet: Depottunn fra yngre bronsealder, med forskjellige kvinnesmykker av bronse: beltesmykker, halsring og bøylespenne. Funnet i Fredrikstad, Østfold. (Foto: Universitetets Oldsaksamling, Universitetet i Oslo).

NAVFs EDB-SENTER FOR HUMANISTISK FORSKNING

ble opprettet av Norges allmennvitenskapelige forskningsråd i 1972. Senteret skal arbeide på nasjonal basis for bruk av edb i forskningsarbeidet i de humanistiske fagene.

Sentrale oppgaver er edb-tjenester (veiledning og betalte oppdrag) og program- og metodeutvikling. Senteret holder kurs, seminar og konferanser om bruk av edb i humanistiske fag. Foruten utgivelse av Humanistiske Data omfatter informasjonstjenestene en rapportserie, årsmelding og elektronisk informasjonsformidling.

Senteret er sekretariat og operativt edb-organ for International Computer Archive of Modern English (ICAME), og utgir tidskriftet ICAME Journal. Senteret er ansvarlig for administrasjon og drift av Norsk tekstarkiv og har det administrative ansvar for NAVFs sentral for informasjon om forskningsprosjekter (SIF) og Fagtjenesten for informasjon om humanistiske forskningsprosjekter (SIF-H).

Humanistiske Data is published by The Norwegian Computing Centre for the Humanities. Editorial group: Einar-Arne Drivenes, Jostein H. Hauge, Kristin Natvig (ed.).

The journal can be ordered free of charge from the address above. Contributions are welcome. On request the Centre can supply the addresses of contributors to the journal.

GRAFISK FORMGIVNING: Svart på Hvitt.

SATS OG TRYKK: A/S John Grieg.

INNHold

ARTIKLER

På kryss og tvers. Edb-registrering for tverrfaglige analyser av arkeologiske funn av menneskeskjeletter. <i>Berit J. Sellevold og Jenny-Rita Næss</i>	4
Bruken av edb ved de arkeologiske museene. <i>Egil Mikkelsen</i>	12
Utgravninger i middelalderbyene. <i>Petter B. Molaug</i>	17
Naturlig språk i kunnskapssystemer. <i>Ivar Utne</i>	21
Grunnlagsmateriale for informasjonssøking ved Norsk termbank. <i>Ivar Utne</i>	39
Forfatterens Høvelbenk – Intelligent tekstbehandling? <i>Per Vestbøstad</i>	54
Edb i språkfag. <i>Per M. Mathisen</i>	60
Erfaringer fra skriving av linjeorienterte undervisningsprogrammer i språk. <i>Signe Marie Sanne</i>	65
Senterets videoplateprosjekt. <i>Roger Erlandsen, Claus Huitfeldt og Øystein Reigem</i>	72
Humanistiske bibliografiske databaser. <i>Annema Hasund Langballe</i> ..	85

RAPPORTER

Edb for humanister. Erfaringer fra et nystartet grunnfag ved Det historisk-filosofiske fakultet, Universitetet i Oslo. <i>Asbjørn Brændeland og Jon Lanestedt</i>	88
Communication Theory and Semantics – A Progress Report. <i>Andrew J.I. Jones</i>	95
Tre residente programmer for MS-DOS maskiner. <i>Espen S. Ore</i>	99
Nordisk konferanse om tekstforståelse og informasjonssøking. <i>Jo-stein H. Hauge</i>	106
Informasjonsmøte om Senterets videoplateprosjekt. <i>Einar-Arne Drivenes</i>	109

MELDINGER	116
------------------------	-----

SUMMARY	139
----------------------	-----

PÅ KRYSS OG TVERS

EDB-REGISTRERING FOR TVERRFAGLIGE ANALYSER AV ARKEOLOGISKE FUNN AV MENNESKESKJELETTER

Berit J. Sellevold og Jenny-Rita Næss

Prosjektet «Jernalderens Mennesker» er et tverrfaglig prosjekt, støttet av Rådet for humanistisk forskning under NAVF. Det begynte i 1984, med et pilot-prosjekt på et halvt års varighet for å fastslå hvorvidt de arkeologiske funn av menneskeskjeletter som er gjort i Norge egnede seg som basismateriale i analyser av mennesker og levekår i jernalderen. Det viste seg at det var mye skjelettmateriale, og at det for en stor del var ganske velbevart. Men for at det skulle være brukbart i et forskningsprosjekt som dreiet seg om levekår og sosiale forhold måtte skjelettene kunne dokumenteres utvetydig med hensyn til funnkontekst. Det ble allerede på dette tidspunkt klart at det med den store mengden av funn ville være ønskelig med edb-registrering av materialet, også med hensyn til dokumentasjon og sortering, og med hensyn til data-analysene.

NAVF bevilget midler til et tre-årig prosjekt. Dette hovedprosjekt begynte i 1985, og utføres av en fysisk antropolog, Berit J. Sellevold (BJS) og en arkeolog, Jenny-Rita Næss (JRN). Prosjektet er todelt. Den ene delen går på registrering av samtlige gravfunn i norske samlinger, og vil bli lagt frem i en monografi. Den andre delen er en analyse av de deler av materialet som er best egnet for kulturhistoriske slutninger. Se Sellevold og Næss 1987 for nærmere beskrivelse av prosjektet.

MATERIALET

Prosjektets basismateriale består av menneskelige skjeletter og skjelettdeler som er funnet ved arkeologiske utgravninger eller annen virksomhet innenfor Norges grenser, samt alt dokumentasjonsmateriale forbundet med slike funn, som f.eks. gjenstander, opplysninger om gravens form og størrelse, og alt skriftlig materiale som måtte finnes.

Grunnleggende i prosjektet er de antropologiske undersøkelsene av skjelettfunnene, med målinger av knokler og tenner, og registreringer av genetiske trekk, patologiske forhold og anomalier, både på ben og på

tenner. Men for at disse observasjonene, dvs. disse antropologiske beskrivelsene, skal kunne tolkes, må funnene være plassert i sine kulturhistoriske sammenhenger. Vi må derfor utvetydig dokumentere skjelett-funnene med hensyn til funnkontekst. Dette betyr at de «kulturhistoriske» data må registreres for hvert enkelt skjelettfunn. Det sier seg selv at det blir en betydelig datamengde det er nødvendig å registrere for hver funn-enhet.

Med hensyn til analysene av det således registrerte materiale har vi behov for å kunne bevege oss på kryss og tvers innenfor data-massen. Basis-data består av både kvantitative og kvalitative informasjoner: De antropologiske registreringene er for en stor dels vedkommende kvantitative, og består eksempelvis av mål på gitte dimensjoner (f.eks. kraniets lengde og bredde, angitt i millimeter, eller lårbenets største lengde, også angitt i millimeter; men også vinkler og korder, f.eks. lårhalsens vinkel med lårbenet, pannebenets korde). De arkeologiske opplysningene vil typisk være av kvalitativ karakter, som opplysninger om en gravs beliggenhet i forhold til andre kulturminner, beskrivelse av gjenstander og disses plassering i forhold til både skjelett og gravkonstruksjon, osv.

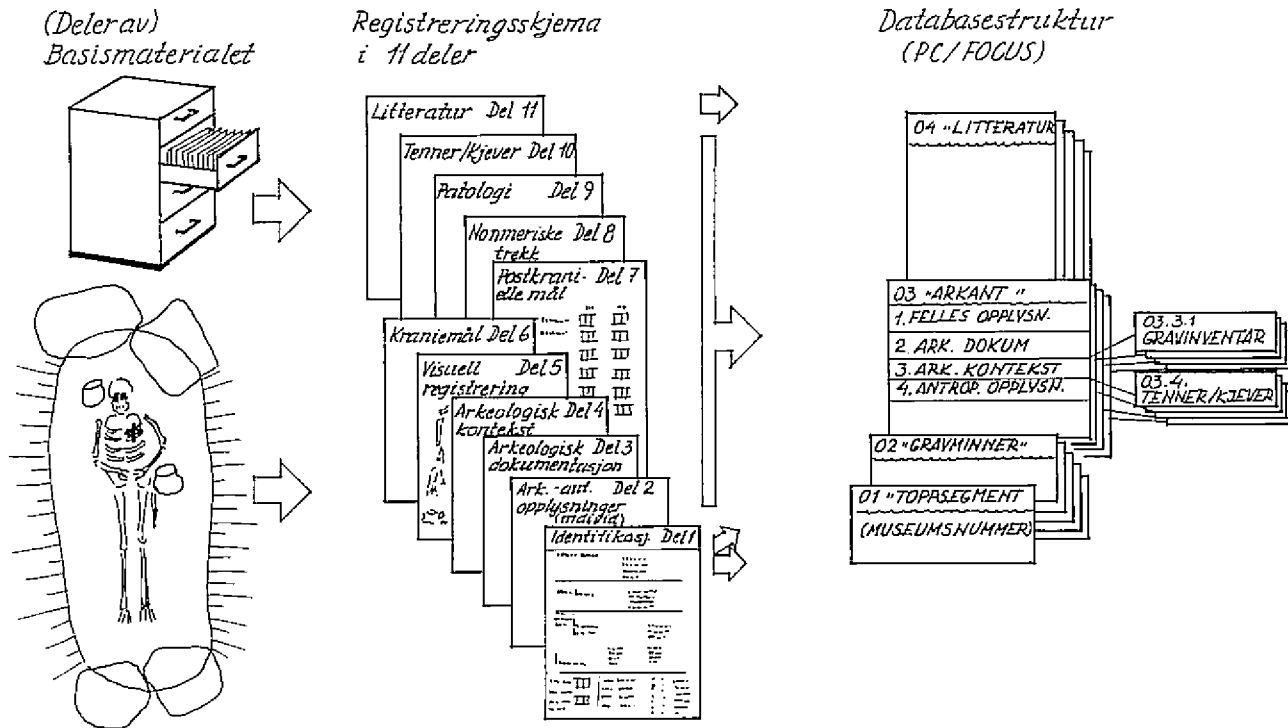
Vi hadde altså behov for å registrere både kvantitative og kvalitative data, og for å bruke disse data i fri kombinasjon i våre analyser.

For å løse dette «problemet» var det nødvendig å utvikle både felt-registreringsskjemaer til data-innsamlingen og en struktur for databasen. Arbeidet med disse to system-utviklingene har pågått parallelt.

DATABASE-STRUKTUREN

Det grunnleggende arbeidet med utformingen av en struktur ble utført av Aud Simonsen ved Arkeologisk museum i Stavanger (AmS), og tilrettelagt for dBase II på AmS' Osborne-maskin. dBase II holdt til forprosjektet, men det viste seg imidlertid meget snart at dBase II og Osborne ikke hadde kapasitet nok til å klare hovedprosjektets datamengde og våre analyse-behov. Vi søkte og fikk bevilget fra NAVF-RHF midler til innkjøp av to NCR PC 4i-mikromaskiner, med kapasitet på 640 K og med innebygget 10 mb harddisk. Denne maskinen har stor nok kapasitet for våre behov.

For å kunne behandle våre basisdata har vi, i samråd med edb-konsulent Morten Sørreime, anskaffet et PC/FOCUS database-program. Morten Sørreime har nå videreutviklet den opprinnelige struktur og har bearbeidet den for PC/FOCUS. Han har overført til PC/FOCUS alle data som under forprosjektet var lagt inn i dBase II. PC/FOCUS kan arbeide både hierarkisk og relasjonelt, og vi anser det derfor som det beste PC-



Figur 1.

programmet på markedet til våre behov for bevegelser på kryss og tvers, og «opp og ned» i materialet med hensyn til diakrone og synkrone analyser.

På høyre side i Figur 1 sees prinsippet i oppbyggingen av vår databasestruktur. Strukturen har fire hovedfelter: et *museumsnummer* (1), som består av en eller flere *graver* (2), som inneholder oldsaker og skjelettmateriale, som kan omfatte en eller flere gjenstander og et eller flere individer. For hvert *individ* (3) og *dettes funnkontekst* er, eller kan det være *litteraturhenvisninger* (4), som utgjør siste nivå i strukturen.

Innenfor hvert hovedfelt kan det være flere underfelter. I PC/FOCUS-databasen kalles hovedfeltene «segmenter». Disse er i utgangspunktet ordnet hierarkisk. Det første segmentet, inngangen til databasen, heter «TOPPSEG», og inneholder data som går på det enkelte museumsnummer (evt. samlingsnummer eller aksjesjonsnummer, hvis funnet ennå ikke er gitt museumsnummer). Data fra felt-registreringsskjemaets del 1 legges inn i «TOPPSEG».

Segmentet under «TOPPSEG» heter «GRAVNUM». Under hvert museumsnummer kan det være en eller flere graver. Hver grav blir registrert for seg, slik at det kan være flere «GRAVNUM»-segmenter under «TOPPSEG».

Segmentet under «GRAVNUM» kalles «ARKANT» (forkortelse for «arkeologisk-antropologiske data»). En grav kan inneholde et eller flere individer. Dette segmentet inneholder arkeologiske og antropologiske data om hvert enkelt individ. Vi har delt datamassen i dette segmentet inn i 4 seksjoner som hver inneholder logisk sammenhørende data. Den første dreier seg om felles arkeologisk-antropologiske opplysninger, med data fra feltskjemaets del 2; den andre seksjonen inneholder data vedrørende den arkeologiske dokumentasjonen (feltskjemaets del 3); i den tredje seksjonen lagres data vedrørende den arkeologiske konteksten (feltskjemaets del 4). Av tekniske grunner legges data fra del 4 vedrørende gravinventaret inn i et eget segment, men sorterer logisk inn i den tredje seksjonen under «ARKANT»-segmentet. Den fjerde seksjonen av «ARKANT»-segmentet lagrer antropologiske data fra feltskjemaets deler 6, 7, 8 og 9. Av tekniske grunner er data fra feltskjemaets del 10 (tenner og kjever) lagt inn i et eget segment, men sorterer logisk inn i «ARKANT»-segmentet.

Databasens siste segment, som ligger under «ARKANT», kalles «LITTERAT». Her registreres alle litteraturhenvisningene som går på alle deler av et funn. Disse registreres i fritekst, og segmentet har ubegrenset størrelse.

Figur 1 anskueliggjør databasens struktur, med referanse til felt-registreringsskjemaets forskjellige deler.

REGISTRERINGSSKJEMAENE

Data som skal legges inn i databasen, blir samlet inn på museer og i samlinger der arkeologisk skjelettmateriale blir oppbevart, samt i felt ved pågående utgravninger. I mangel av en transportabel datamaskin, og for tilfeller der det er helt upraktisk å benytte elektronisk datautstyr (f.eks. på små utgravninger ute i felt) har vi utviklet et registrerings skjema på A4-format papir. Organiseringen av dette skjemaet og databasestrukturen ble utviklet samtidig for å passe sammen – være «kompatible» – slik at det å legge inn data i maskinen fra registreringene på skjemadelene skal være så ukomplisert som mulig. Samtidig er skjemadelene utformet slik at det skal være logisk å arbeide med dem ved innsamling av data. Feltskjemaet vil bli publisert om kort tid i serien *AmS-Varia* fra Arkeologisk museum i Stavanger, med detaljert beskrivelse av delene og forklaring på anvendelsen.

Prinsipielt kan skjemaene beskrives som følger: For hvert enkelt individ er det 11 skjemadel: 1. identifikasjon, 2. arkeologisk-antropologiske opplysninger (individet), 3. arkeologisk dokumentasjon, 4. arkeologisk kontekst, 5. visuell registrering, 6. kranie mål, 7. postkranielle mål, 8. nonmetriske trekk, 9. patologi og anomalier, 10. tenner og kjever, og 11. litteraturhenvisninger.

Del 1, **Identifikasjon**, går på museumsnummeret, dvs. det sluttede funnet. Her registreres alle funnets museums- og/eller samlingsnumre, arbeidsnumre osv. Her plasseres funnet i sted og tid, og likbehandlingen (brent/ubrent) angis. Informasjonen i denne skjema-delen går inn i databasen i «TOPPSEG». I denne skjema-delen registreres også *gravnummer*. Hvert gravnummer går inn i segmentet «GRAVNUM» i databasen.

Del 2, **Arkeologisk-antropologiske opplysninger (individet)**, gjelder det enkelte skjelett. I denne skjema-delen registreres opplysninger om skjelettets bevaringstilstand i dag, kjønns- og aldersbestemmelser med angivelse av grunnlaget for disse bestemmelsene, hvor funnet befinner seg ved skjelettundersøkelsen, samt dets permanente oppbevaringssted. Her registreres også opplysninger om tapte skjeletter, og hvor det enkelte funn har vært før det kom til sitt nåværende magasin. Her noteres hva som skal skje med funnet (oppbevares/returneres el.l.). Disse opplysningene legges inn i database-segmentet «ARKANT», i den første seksjonen, «felles opplysn.»

Del 3, **Arkeologisk dokumentasjon**, omfatter opplysninger om funnomstendigheter, undersøkelser i felt, behandling og merking av funnet samt detaljerte opplysninger om hva som finnes av dokumentasjonsmateriale i arkivene, dvs. innberetninger, fotografier, planer osv. Opplysningene i

denne skjemadelen går inn i database-segmetet «ARKANT», i annen seksjon, «ark.dokum.»

Del 4, Arkeologisk kontekst. Opplysninger om funnmiljøet registreres her, sammen med beskrivelser av ytre og indre gravutforming og -materiale. Likbehandlingen og gravinnredningen registreres. Registreringen av grav-inventaret er så plasskrevende at disse data blir lagt ut i et eget segment i databasen (se ovenfor).

Del 5, Visuell registrering. Denne består av en skjematisk tegning av et skjelett. Her markeres hva som er til stede av skjelettet ved undersøkelsen, og til dels angis også bevaringstilstanden i og med at fragmenterte ben eller deler av ben blir skravert, mens det som er intakt blir malt sort med spritpenn. I dette prosjektet har vi ikke mulighet for å legge inn bilder i databasen, men når det senere blir aktuelt vil denne skjemadelen egne seg bra til å legges inn, og vil utgjøre en rask og grei dokumentasjon for hva som faktisk finnes av et gitt skjelett.

Del 6, Kraniemål. Her registreres 36 mål tatt på kraniet, og det er plass til utregnede resultater for 7 indekser (dvs. forholds-tall mellom to mål). Disse registreringene utgjør den kvantitative beskrivelsen av et gitt kranium eller deler av kranium.

Del 7, Postkranielle mål. På denne skjemadelen finnes rubrikker for registrering av målte dimensjoner på de 10 store rørknoklene: for hver overarm 5 mål, for hver av de to underarmsbenene 2 mål, hvert lårben 10 mål, og hvert skinneben 4 mål, i alt for et intakt skjelett 46 mål på det postkranielle skjelettet. Dertil er det rubrikker for 16 utregnede indekser, 2 for hvert overarmsben, lårben og skinneben, og 1 for hvert spoleben og albueben. I tillegg er det plass til opplysning om utregnet kroppshøyde basert på lårbenets, skinnebenets og/eller overarmsbenets maksimale lengde.

Del 8, Nonmetriske trekk. Her registreres i alfanumerisk kode forekomst eller fravær av små, ikkemålelige morfologiske variasjoner på knoklene. Slike trekk regnes i det store og hele for å være genetisk betingede, men miljøet kan også spille en rolle i utviklingen av noen av trekkene. I alt er det plass til registrering av 32 slike trekk. Det finnes et stort antall genetisk betingede trekk på skjelettet. Vi har valgt ut disse 32 fordi de er enkle å registrere og relativt utvetydige, slik at det skulle være mulig å reprodusere undersøkelsesresultatene. Av de 32 trekkene finnes 26 på kraniet og 6 på det postkranielle skjelettet. 19 av kranie-trekkene er bilaterale, 7 er unilaterale, mens alle 6 postkranielle trekk er bilaterale.

Del 9, Patologi og anomalier. Alle knokler og ben-fragmenter blir nøye undersøkt for forhold som avviker fra det normale, dvs. forekomst av

patologiske eller anomale forandringer på skjelettet. Noen av de vanlig forekommende forandringene registreres i alfanumerisk kode, men mange data i denne delen registreres også i fritekst. Med PC/FOCUS er det mulig å søke i fritekst og benytte dette i analyser.

Del 10, **Tenner og kjever**. Her registreres dentale observasjoner, mål og forekomst av genetiske variasjoner. For det komplette permanente tannsett er det 32 x 20 parametre som registreres, for melketannsettet 20 x 20, mens det for kjevne, utover patologiske variasjoner, også registreres tre parametre. Dessuten blir tenner og kjever vurdert med henblikk på alders- og kjønnsbestemmelse som en sjekk på kjønns- og aldersbestemmelse ut fra skjelettet som helhet. I tillegg registreres en del ikke-kodede observasjoner i fritekst. De odontologiske registreringer er så omfattende at de krever sitt eget segment i databasen, likesom det er tilfelle med gravinventar-registreringene (se ovenfor).

Del 11, **Litteratur**. I denne siste skjemadelen listes all relevant litteratur og skriftlig dokumentasjonsmateriale. Dette feltet er naturligvis basert på fritekst, og i databasen har det en ubegrenset størrelse.

KONKLUSJON

I det registreringssystemet som er beskrevet her er det tatt hensyn til alle arkeologiske funn av graver, både med og uten skjelettdeler, både brann-graver og ubrente graver. Registreringene er i form av 1) kontinuerlige variabler (mål på knokler og tenner), 2) alfanumeriske koder, og 3) fritekst. Registreringsskjemaet i 11 deler er konstruert slik at de enkelte skjemaer kan benyttes stort sett uavhengig av hverandre. Hvis det f.eks. er behov for registrering og analyser av utelukkende biologiske forhold kan f.eks. delene 1, 2, 5, 6, 7, 8, 9 og 10 trekkes ut. Likeledes kan skjemaet være brukbart også ved registrering av graver uten skjeletter. Skjemaets, og systemets, siktemål er å fungere i en tverrfaglig sammenheng, slik at alle arkeologiske funn av menneskeskjeletter skal inngå i sine kulturhistoriske kontekster. Dette er det overordnede prinsippet i utviklingen av registreringssystemet og database-strukturen.

LITTERATUR

Sellevoid, B.J. og J.-R. Næss 1987 «Iron Age people in Norway. Preliminary report on a research project.» *Norw. Archaeol. Review*, 20:1. Universitetsforlaget, Oslo.

Berit J. Sellevold er forsker-NAVF ved Riksantikvarens Utgravningskontor for Oslo. Jenny-Rita Næss er førstekonservator ved Arkeologisk museum i Stavanger.

BRUKEN AV EDB VED DE ARKEOLOGISKE MUSEENE

Egil Mikkelsen

Arkeologene i Norge har i mange år arbeidet med elektronisk databehandling i ulike forskningssammenheng, i første rekke for å bearbeide bestemte funngrupper. Arkeologmiljøet i Tromsø har her vært ledende. Jeg skal ikke forsøke å lage en fullstendig oversikt over denne bruken av edb, men konsentrere oppmerksomheten om tiltak som de fem arkeologiske museene har samarbeidet om for å få fram mer generelle registre for faste fornminner og arkeologisk funnmateriale.

FELLES BEHOV – FELLES LØSNINGER

Høsten 1980 nedsatte Den arkeologiske interimskommisjon (DAIK), som er et fellesorgan for de arkeologiske museene og Riksantikvaren, en edb-gruppe for å kartlegge behovet og oppgavene som kunne være aktuelle for edb-behandling. En rapport fra denne gruppen forelå i januar 1982. Siden da har edb-gruppen fortsatt sitt arbeide for å utvikle systemer for edb-registrering av faste fornminner og arkeologisk funnmateriale. I løpet av 1987 vil det foreligge en rapport om dette. Gruppen samarbeider med andre arkeologiske institusjoner i Sverige og Danmark, og det er avholdt to felles konferanser.

Gruppens hovedmålsetning har vært å få utarbeidet felles landsomfattende systemer i form av registrerings skjemaer som inneholder bestemte typer informasjon i klart definerte former, som skal inngå i lokale databaser. Minstekravet er at en gjensidig, mellom museene, skal kunne søke i hverandres databaser. Hovedproblemet har vært at en ikke har hatt ressurser til å kunne samordne innkjøp av maskinutstyr og programmer ved de enkelte museer, og at en heller ikke har hatt tilstrekkelig bistand fra edb-ekspertise.

STORE FUNN KREVER AVANSERTE PROGRAMMER OG MASKINUTSTYR

Datamengden og strukturen på dataene når det gjelder faste fornminner og funn er svært omfattende. Det er snakk om mange tusen faste

forminner, med en temmelig omfattende verbalbeskrivelse av det enkelte objekt. Det arkeologiske funnmaterialet ved de fem museene (i Oslo, Stavanger, Bergen, Trondheim og Tromsø) omfatter nærmere 100.000 katalognummer; mange av dem inneholder store samlede funn på flere titalls eller hundretalls gjenstander, som må registreres i sammenheng. Tar en med funnene fra gravningene i våre middelalderbyer, blir antallet ennå større. Edb-registrering av dette materialet krever store og avanserte programmer og maskinutstyr, som både kan håndtere slike programmer og har tilstrekkelig lagringskapasitet.

Lengst er en til nå kommet med edb-registrering av de faste forninnene. Her er det utarbeidet forslag til registreringsskjema, med spesifikasjoner av enkelte datafelt. Ved *Arkeologisk museum i Stavanger* er det laget et særskilt program, AMIS, basert på databaseprogrammet SIBAS II. På en *NORD-100 Compact* maskin er nå alle faste fornminner i Rogaland som er registrert i forbindelse med det økonomiske kartverket, lagt inn. Tidligere var registeret lagret ved Rogalandsdata, men dette viste seg å bli en for dyr løsning.



*En samling leirkar, glassbegre og drikkehorn fra eldre jernalder.
(Foto: Universitetets Oldsaksamling).*

SKJEMA FOR ARKEOLOGISKE FUNN

(01)MUSEUMSNR: :::::::::::::::::::: (02)AKSESJONSNR: ::::::::::
(03)HØRER SAMMEN : ::::::::::::::::::::
(04)FYLKE: :::::::::::::::::::: (05)KOMMUNE: ::::::::::::::::::::
(06)PGD: :::::::::::::::::::: (07)SOGN: ::::::::::::::::::::
(08)GNR: :::: (09)GÅRD: ::::::::::::::::::::
(10)BNR: :::::: (11)BRUK: ::::::::::::::::::::
(12)LOKALITET: (13)FORMINNENR: ::::::::::
(14) M/11-KART: :::::: (15)UTM-KOR: :::::::::::::::::::: (16)H.O.H.: ...
(17)ØK-KART: :::::: (18)NGO-KOR: ::::::::::::::::::::
(19)FUNNKATEGORI: ::::::::::::::::::::
(20)FUNNOMSTENDIGHET: ::::::::::::::::::::
(21)FUNNMILJØ: ::::::::::::::::::::
(22)PERIODE: ::::::::::::::::::::
(23)FUNNET AV:
(24)ÅR: :::::: (25)GAVE/KJØP: ::
(26)INNLEVERT AV:
(27)KATALOGISERT AV:
(28)ANDRE FUNNOPPLYSNINGER:
(29)LITTERATUR:

(30)NR: :::::::::::::::::::: (31)DATO: (32)DATERING: ::::::::::
(33)GJENSTAND: :::::::::::::::::::: (34)GJENSTANDSDEL: ::::::::::
(35)FORM: :::::::::::::::::::: (36)VARIANT: ::::::::::::::::::::
(37)MATERIALE: :::::::::::::::::::: (38)ANTALL: :::: (39)FRAGM:
(40)STMÅL(MM): :::: (41)MÅL(MM):
(42)VEKT(G): :::::: (43)FARGE: (44)DEI
(45)LOK-1: (46)LOK-2:
(47)LOK-3: (48)LOK-4:
(49)X: (50)Y: (51)Z: (52)KAT. AV:
(53)OPPBEVARINGSTED: (54)NÅR: ..
(55)TEGNING:
(56)FOTO: ::::::::::::::::::::
.....
(57)KONSERVERING:
(58)NAT. VIT. ANALYSER:
(59)ANDRE GJENSTANDSOPPL:
.....

Forslag til skjema for edb-registrering av arkeologisk funnmateriale. Skjemaset er foreløpig, og det vil bli foretatt noen mindre endringer i det. Den øvre delen er felles for ett samlet funn, f.eks. et gravfunn, mens den nedre delen skal kunne gjentas for hver enkelt gjenstand i funnet.

Universitetets Oldsaksamling i Oslo har tatt i bruk et standardprogram for dataregistrering av faste fornminner, NOTIS-IR, som er et fritekst-basert arkivsystem. Også her har en en en *NORD-100 Compact* (oppgradert til *CX*-versjon), med en internhukommelse på 1 Mb og en lagringskapasitet på 90 Mb. I 2 år har en nå lastet inn fornminnebeskrivelsene fra registreringene for økonomisk kartverk, for en stor del som sysselsettingsprosjekt. En del av de erfaringene en har høstet gjennom dette arbeidet kan summeres som følger:

- Datamengdene fyller mye hurtigere opp maskinen enn en er klar over på forhånd
- Det er problematisk å bruke et standardprogram, som ikke oppfyller alle de krav en stiller, bl.a. til repeterbarhet av felt og visse søkeoperasjoner
- Dataregistrering som sysselsettingsprosjekt er til hjelp for å få registrert store datamengder, men korrekturen blir svært omfattende og tidkrevende
- En må regne med en lang registrerings- og korrekturfase før databasen blir tilfredsstillende operativ for søking og bruk

På noe sikt vil det trolig være nødvendig å overføre fornminnedataene til et kraftigere program, f.eks. *SIFT*, som er utviklet av Rasjonaliseringsdirektoratet. Det vil imidlertid kreve kraftigere maskinkapasitet enn det *Oldsaksamlingen* disponerer i dag. Foreløpig ser vi det som viktigst å få våre data over i en elektronisk form, så får forbedringene komme etter hvert som søkebehovene blir større.

Tromsø Museum har også lastet inn en del av sine fornminneregistreringer på datamaskin, mens de arkeologiske museene i Bergen og Trondheim ennå ikke har begynt. Det er å håpe at en etter hvert kan få harmonisert ytterligere de løsninger som nå er i bruk eller planlegges ved de arkeologiske museene. Her er det imidlertid behov for en sterkere støtte fra myndighetenes side, både med hjelp fra edb-ekspertise og til mer samordnede anskaffelser av maskinutstyr og programmer.

ARBEIDET FRAMOVER

Den nasjonale edb-gruppens arbeide med et felles opplegg for edb-registrering av arkeologisk funnmateriale vil bli avsluttet våren 1987. Det er utarbeidet forslag til registrerings skjema og en beskrivelse av de enkelte datafelt, samt en del krav til dataprogrammet, som må velges senere. Skjemaet er laget som to deler, i den ene er det samlet fellesopplysninger om det samlede funnet, i den andre delen er det avsatt plass til beskrivel-

se av de enkelte gjenstander. Denne siste delen må kunne repeteres så mange ganger som det er antall gjenstander i funnet. Edb-gruppen har forsøkt å definere aktuelle begreper og fastlagt formen de ulike data skal registreres på, men det er fortsatt behov for utvikling av nomenklaturen, spesielt når det gjelder gjenstandsbetegnelser. Her er det imidlertid gjort en del arbeide, både i Norge og i Sverige, men det er behov for en samordning av systemene.

De databasene som skal omfatte det arkeologiske funnmaterialet blir svært store. En ønsker å vurdere bruken av programmet SIFT også i denne sammenheng, og det er nå behov for edb-ekspertise for å kunne komme videre. Lagringsbehovet må løses, og vi følger utviklingen innenfor optiske lagringsmedier, som vi mener er aktuelt i vår sammenheng. Registreringsfasen for disse store funnmengdene vil bli langvarig, og arbeidet vil kreve arkeologisk ekspertise som kan klassifisere, rette opp og bearbeide til dels svært gamle beskrivelser av funn og funnforhold.

Til slutt skal nevnes at stadig større deler av det arkeologiske miljøet i Norge nå tar i bruk elektronisk tekstbehandling, både i arbeidet med forvaltningssaker og i forskningssammenheng. Det er innenfor dette feltet en mest umiddelbart merker fordelene ved bruken av edb, og tekstbehandlingens tjener som et viktig middel til å bryte ned de barrierer som finnes for å ta i bruk denne nye teknologien. Når databasene for faste fornminner og arkeologisk funnmateriale en gang er etablert, vil dette representere en veldig rasjonaliseringsgevinst og innby til å ta opp nye problemfelt som våre manuelle arkiver har tjent som sperre for.

Egil Mikkelsen er førstekonservator ved Universitetets Oldsaksamling, Oslo.

Artikkelen ble publisert for første gang i *Museumsnytt* 4-86.

UTGRAVNINGER I MIDDELALDERBYENE

Petter B. Molaug

EDB-REGISTRERING AV GJENSTANDER

Edb ble tidlig tatt i bruk i norsk arkeologi, men dette var først og fremst innen området statistikk, som ledd i den vitenskapelige bearbeidelsen av et arkeologisk materiale. Her er det etterhvert gjort store fremskritt, særlig innen bruken av multivariate metoder. Bl.a. har NAVFs edb-senter laget en egen programpakke, kalt STAR. Som del i en rutinemessig registrering av oldsaksmateriale er edb-bruk langt nyere ved de arkeologiske institusjonene, selv om forsøk ble gjort tidlig enkelte steder.

Det gjøres nå forsøk på å samordne registreringsprosedyrene ved de 5 arkeologiske museene og Riksantikvaren, slik at det kan brukes det samme katalogoppsettet og de samme søkeordene (se artikkelen av Egil Mikkelson i dette nr. av HD). I denne oversikten vil jeg beskrive enkelte tidligere forsøk på edb-registrering innenfor området middelalderens gjenstandsmateriale, og også peke på noen mulige tendenser fremover.

Arkeologisk gjenstandsregistrering er tredelt.

- Del 1. *Funnomstendigheter.* Her kommer foruten stedsangivelse (stedsnavn/gårdsnavn, kommune osv.) nøyaktig lokalisering av gjenstander funnet ved arkeologiske utgravninger i horisontalplan, vertikalplan og i forhold til lag og konstruksjoner.
- Del 2. *Gjenstandsbeskrivelse.*
- Del 3. *Behandlings- og oppbevaringsdata.*


Ved utgravningen på Bryggen i Bergen (Historisk Museum) ble det alt i 1960-årene begynt med koding av feltopplysningene. Det ble utarbeidet et skjema med faste plasser for utfylling med bokstaver og tall, og ett slikt skjema ble fylt ut for alle gjenstandene som var funnet samme sted. Individuelle funn-nummer ble brukt som identifikasjon. Disse opplysningene ble senere kodet inn på hullkort og foreligger nå på data med muligheter for bruk av flere generelle og spesialtilpassete programmer.

I forbindelse med Gamlebyprosjektet for bearbeidelse av arkeologisk materiale fra de arkeologiske utgravningene i Gamlebyen, Oslo (Riksantikvaren og Universitetets Oldsaksamling), ble det besluttet å overføre

gjenstandsdataene fra gravningene i 1973-76 til edb. Fra 1977 ble gjenstandskortene renskrevet på skrivemaskin med OCR-B typer på trykte skjemaer med rubrikktekster og linjemarkering i rødt. I alt ca. 17.000 funnkort ble så optisk lest og overført til UNIVAC-maskinen ved Universitetet i Bergen i løpet av 1978 og 1979. Her ble dataene lagt på editorfiler og på filer beregnet for søking med NOVA*STATUS fritekst søkesystem. NAVFs edb-senter i Bergen stod for den praktiske tilretteleggingen av dette arbeidet. Senere er en del av dataene kodet for statistisk behandling med programpakken SPSS.

Kontakten med datamaskinen har skjedd ved oppringt samband fra Oslo via universitetenes datanett (UNINETT), ved oppdrag utført av personale ved NAVFs edb-senter, bl.a. utskrifter og ved egne besøk i Bergen. Prosjektet må kunne karakteriseres som vellykket så langt, idet det er produsert en rekke kataloger, ordlister, statistiske tabeller og beregninger.

Det kortet som ble benyttet bestod av 35 rubrikker fordelt på kortdelene som er nevnt ovenfor (se fig. 1). Hensikten med utformingen var



GAMLEBYEN, OSLO

01 Gjenstand	02 Gjenstand	03 Gjenstand	04 Gjenstand	05 Gjenstand	06 Gjenstand	07 Gjenstand	08 Gjenstand
09 Gjenstand	10 Gjenstand	11 Gjenstand	12 Gjenstand	13 Gjenstand	14 Gjenstand	15 Gjenstand	16 Gjenstand
17 Gjenstand	18 Gjenstand	19 Gjenstand	20 Gjenstand	21 Gjenstand	22 Gjenstand	23 Gjenstand	24 Gjenstand
25 Gjenstand	26 Gjenstand	27 Gjenstand	28 Gjenstand	29 Gjenstand	30 Gjenstand	31 Gjenstand	32 Gjenstand
33 Gjenstand	34 Gjenstand	35 Gjenstand	36 Gjenstand	37 Gjenstand	38 Gjenstand	39 Gjenstand	40 Gjenstand
41 Gjenstand	42 Gjenstand	43 Gjenstand	44 Gjenstand	45 Gjenstand	46 Gjenstand	47 Gjenstand	48 Gjenstand
49 Gjenstand	50 Gjenstand	51 Gjenstand	52 Gjenstand	53 Gjenstand	54 Gjenstand	55 Gjenstand	56 Gjenstand
57 Gjenstand	58 Gjenstand	59 Gjenstand	60 Gjenstand	61 Gjenstand	62 Gjenstand	63 Gjenstand	64 Gjenstand
65 Gjenstand	66 Gjenstand	67 Gjenstand	68 Gjenstand	69 Gjenstand	70 Gjenstand	71 Gjenstand	72 Gjenstand
73 Gjenstand	74 Gjenstand	75 Gjenstand	76 Gjenstand	77 Gjenstand	78 Gjenstand	79 Gjenstand	80 Gjenstand
81 Gjenstand	82 Gjenstand	83 Gjenstand	84 Gjenstand	85 Gjenstand	86 Gjenstand	87 Gjenstand	88 Gjenstand
89 Gjenstand	90 Gjenstand	91 Gjenstand	92 Gjenstand	93 Gjenstand	94 Gjenstand	95 Gjenstand	96 Gjenstand
97 Gjenstand	98 Gjenstand	99 Gjenstand	100 Gjenstand	101 Gjenstand	102 Gjenstand	103 Gjenstand	104 Gjenstand
105 Gjenstand	106 Gjenstand	107 Gjenstand	108 Gjenstand	109 Gjenstand	110 Gjenstand	111 Gjenstand	112 Gjenstand
113 Gjenstand	114 Gjenstand	115 Gjenstand	116 Gjenstand	117 Gjenstand	118 Gjenstand	119 Gjenstand	120 Gjenstand
121 Gjenstand	122 Gjenstand	123 Gjenstand	124 Gjenstand	125 Gjenstand	126 Gjenstand	127 Gjenstand	128 Gjenstand
129 Gjenstand	130 Gjenstand	131 Gjenstand	132 Gjenstand	133 Gjenstand	134 Gjenstand	135 Gjenstand	136 Gjenstand
137 Gjenstand	138 Gjenstand	139 Gjenstand	140 Gjenstand	141 Gjenstand	142 Gjenstand	143 Gjenstand	144 Gjenstand
145 Gjenstand	146 Gjenstand	147 Gjenstand	148 Gjenstand	149 Gjenstand	150 Gjenstand	151 Gjenstand	152 Gjenstand
153 Gjenstand	154 Gjenstand	155 Gjenstand	156 Gjenstand	157 Gjenstand	158 Gjenstand	159 Gjenstand	160 Gjenstand
161 Gjenstand	162 Gjenstand	163 Gjenstand	164 Gjenstand	165 Gjenstand	166 Gjenstand	167 Gjenstand	168 Gjenstand
169 Gjenstand	170 Gjenstand	171 Gjenstand	172 Gjenstand	173 Gjenstand	174 Gjenstand	175 Gjenstand	176 Gjenstand
177 Gjenstand	178 Gjenstand	179 Gjenstand	180 Gjenstand	181 Gjenstand	182 Gjenstand	183 Gjenstand	184 Gjenstand
185 Gjenstand	186 Gjenstand	187 Gjenstand	188 Gjenstand	189 Gjenstand	190 Gjenstand	191 Gjenstand	192 Gjenstand
193 Gjenstand	194 Gjenstand	195 Gjenstand	196 Gjenstand	197 Gjenstand	198 Gjenstand	199 Gjenstand	200 Gjenstand

Fig. 1. Utgravningene i Gamlebyen, Oslo. Skjema brukt ved utgravningene 1973-76 for katalogisering av gjenstander.

13.000 funn-nummer og flere mindre gravninger. Til dette er brukt programmet dBase II. Med assistanse fra NAVFs edb-senter er det laget underprogrammer for innlasting, oppdatering, søking og utskrivning.

Det ble laget et nytt skjema med 64 rubrikker, fordelt på 3 kortvarianter. Delen for funnomstendigheter tilsvarer omtrent det eldre kortet for optisk lesing, mens gjenstandsbeskrivelsen er forskjellig. De tre variantene er en for keramikk, en for lær og en for annet materiale. Det har vist seg å være tidsbesparende å kunne skrive dataene direkte inn på skjermen. Imidlertid fordrer effektiv søking at dataene er konsekvent skrevet, og her står det igjen noe korrekturarbeid. Det er ennå ikke laget statistikker og oversikter med grunnlag i dette materialet, men det er foretatt opprettinger og søking, og det er gjort katalogutskrifter for det hele. Sannsynligvis vil oppdatering av dataene på den lettvinde måten som nå er mulig gi den største gevinsten i forhold til et manuelt system. Det vil til enhver tid være mulig å produsere ajourførte utskrifter og statistikker.

I det gjenstandsskjemaet som DAIK's edb-komite har anbefalt, er det en del rubrikker som ikke er nødvendige å fylle ut på hvert kort når det er mange gjenstander fra en gravning. Dette gjelder særlig funnstedet. Omvendt er det ofte ønskelig med flere rubrikker for lokaliseringsdata og for gjenstandsbeskrivelse. Et felles skjema ønskes velkommen ut fra behovet for å kunne søke i databaser ved forskjellige arkeologiske institusjoner, men hensynet til dokumentasjon og forskning gjør at dette i mange tilfeller bør utvides med flere rubrikker eller sekundære databaser, og at rubrikker der teksten vil bli den samme på alle dokumentene må kunne sløyfes på det enkelte dokument. De store gravningene fra middelalder og nyere tid kommer ikke i noen prinsipielt annen stilling enn store utgravninger fra forhistorisk tid, f.eks. steinalderboplasser.

En viktig oppgave de nærmeste år vil være å kombinere tilgjengelighet med fleksibilitet og spesialtilpasning til problemløsning. Spesielt utfordrende vil dette være fordi det sannsynligvis vil bli introdusert automatiske registreringssystemer både for feltdokumentasjon (skrift og tegning) og gjenstandsbeskrivelse basert på digitalisering av former og integrerte programpakker. Flere prøveprosjekter har allerede vært gjennomført, og de tekniske løsningene er i prinsippet kjent.

Petter B. Molaug er antikvar ved Riksantikvarens utgravningskontor.

Artikkelen ble trykket for første gang i *Museumsnytt* 4-86.

NATURLIG SPRÅK I KUNNSKAPSSYSTEMER

Ivar Utne

1. INNLEDNING

Artikkelen omhandler bruk av naturlig språk i kunnskapssystemer. Naturlig språk vil her si språk som ikke uttrykkes i formaliser eller stikkord. Et kunnskapssystem er et edb-system med relativt avanserte og fleksible rutiner for databehandling og kontakt med bruker. Kunnskapssystemer særpreges av at de yter brukerne ekstra hjelp til å besvare kompliserte spørsmål, og da særlig innenfor emner som en tidligere ikke har brukt datamaskiner til å behandle. Det dreier seg i stor grad om vurderinger og tilnærmet kvalitativ eller myk kunnskap til forskjell fra kvantitativ kunnskap eller harddata.

Jeg åpner med en generell oversikt over kunnskapssystemer og emner som relaterer dem til naturlig språk. I den sammenhengen kommer jeg kort inn på anvendelse av samme teknologi i forbindelse med *registrering og gjenfinning av informasjon, dokumentasjon og maskinoversettelse*.

Størstedelen av artikkelen er en presentasjon av hvordan naturlig språk behandles i regelverk og representasjon. Et kunnskapssystem som kommuniserer med brukerne gjennom naturlig språk, består av hovedkomponentene *analyse av språk, representasjon av meningsinnhold, slutningsprosesser og produksjon av språk*. Slutningsprosesser blir utelatt i den spesielle behandlingen av regelverk og representasjon. *Dialog* som vil være en del av de naturlige språktekstene, blir heller ikke behandlet spesielt.

Hovedvekten legges på *språkets innholdsside*. Andre relevante temaer som har vært omtalt tidligere i *Humanistiske Data (HD)* eller som fins relativt lett tilgjengelig i nordiske publikasjoner, vil få liten vekt. Det vil si analyse av morfologi (Fjeldvig og Golden 1983, Fjeldvig og Golden 1984, Koskenniemi 1983, Karlsson 1984), syntaks (Lie 1980, Dyvik og Hofland 1983, Dyvik 1987), kunstig tale (Utne 1985, Slethei 1986) og den generelle kunnskapssiden ved ekspertsystem (som ofte brukes synonymt med kunnskapssystem) (Fjornes 1985).

Framstillingen bygger hovedsakelig på et utvalg av foredragene om innholdsanalyse og språkproduksjon fra Coling-konferansen 1986 i Bonn (Coling 86), men er betydelig forenklet.

2. GENERELL OVERSIKT

2.1. NATURLIG SPRÅK

Naturlig språk i kommunikasjon med datamaskinelle systemer vil si bruk av skriftlig eller muntlig språk med språklige uttrykksmåter slik det vil være naturlig for mennesker å bruke dem. I praksis vil det si fullstendige setninger, evt. også tekster. Det skiller seg fra bruk av kommandoer og menyer, inn- og utdata i tabellform, og representasjon i logiske eller matematiske formalismer.

2.2. KUNNSKAPSSYSTEMER

Kunnskapssystemer skiller seg fra andre datamaskinelle systemer, bl.a.

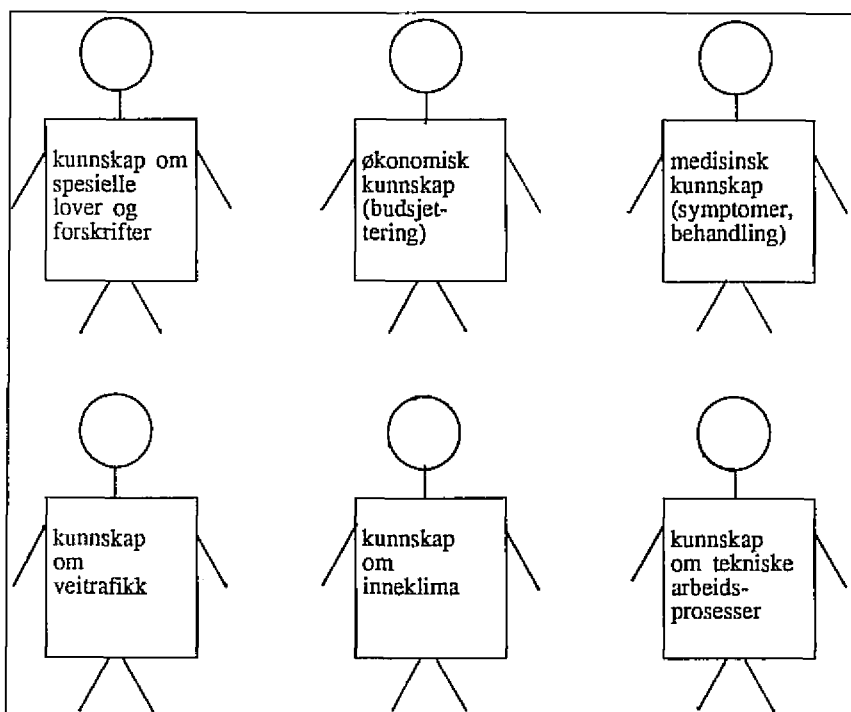


Fig. 1. Et kunnskapssystem behersker et utsnitt av kunnskapen innen avgrensede fagområder. En måte å se det på er at de har et «firkantet» forhold til virkeligheten.

databasesystemer, ved at de kan trekke relativt avanserte slutninger ut fra kompleks kunnskap og ved at de benytter avanserte teknikker for inn- og utmating av data.

Evnen til å *dra slutninger* baserer seg på at programmene håndterer programmerte slutningsrekker som utnytter kunnskap lagret i komplekse datastrukturer. I større grad enn tidligere har en å gjøre med data der samspillet mellom mange variabler avgjør svaret. Dessuten spiller tematisk og logisk strukturering en stor rolle, dvs. at bl.a. semantikk og logikk er viktige hjelpevitenskaper.

Kunnskapssystemer har ofte avansert teknikk for *inn- og utmating* av data. Det innebærer f.eks. bruk av naturlig språk både skriftlig og muntlig, avlesing av dataverdier fra måleapparater (lys- eller varmemålere), styring av utenforstående instrumenter (roboter), avlesing av grafiske bilder (hente ut informasjon fra tekniske tegninger) eller bruk av videokamera (analysere trafikkutvikling).

2.3. KUNNSKAPSSYSTEMER OG NATURLIG SPRÅK

Naturlig språk kan komme til å spille en viktig rolle for *innmating* av data, intern *representasjon* av data, *slutningsprosesser* (hvor også logikk inngår) og *utmating* av data. Status er at det fins utviklingssystemer som håndterer slike aspekt innenfor mindre kunnskapsområder og innenfor et avgrenset sett med språktrekk.

2.3.1. INNMATING AV DATA

Innmating av data i form av naturlig språk kan skje enten som skriftlig *skrevet tekst* fra datafiler og tastatur eller som *tale*. *Skriftlige inndata* må enten foreligge dataleselig eller bli gjort dataleselig ved hjelp av optiske lesere. Dataene må deretter analyseres slik at meningsinnholdet kan trekkes ut og kan gjøres om til en form som det er mulig å passe inn i systemets datastrukturer. Forskningen har kommet relativt langt innen syntaktisk analyse av enkeltsetninger, men kort innen innholdsanalyse av sammenhengende tekst.

Muntlige inndata må, før de går gjennom de samme prosesser som skriftlige inndata, omkodes fra talestrøm til skriftspråk. Dette krever samspill mellom avansert fonetisk kunnskap og språkanalyseteknikker som også brukes for skriftspråk. Foreløpig kan slike systemer bare identifisere enkeltord uttalt hver for seg.

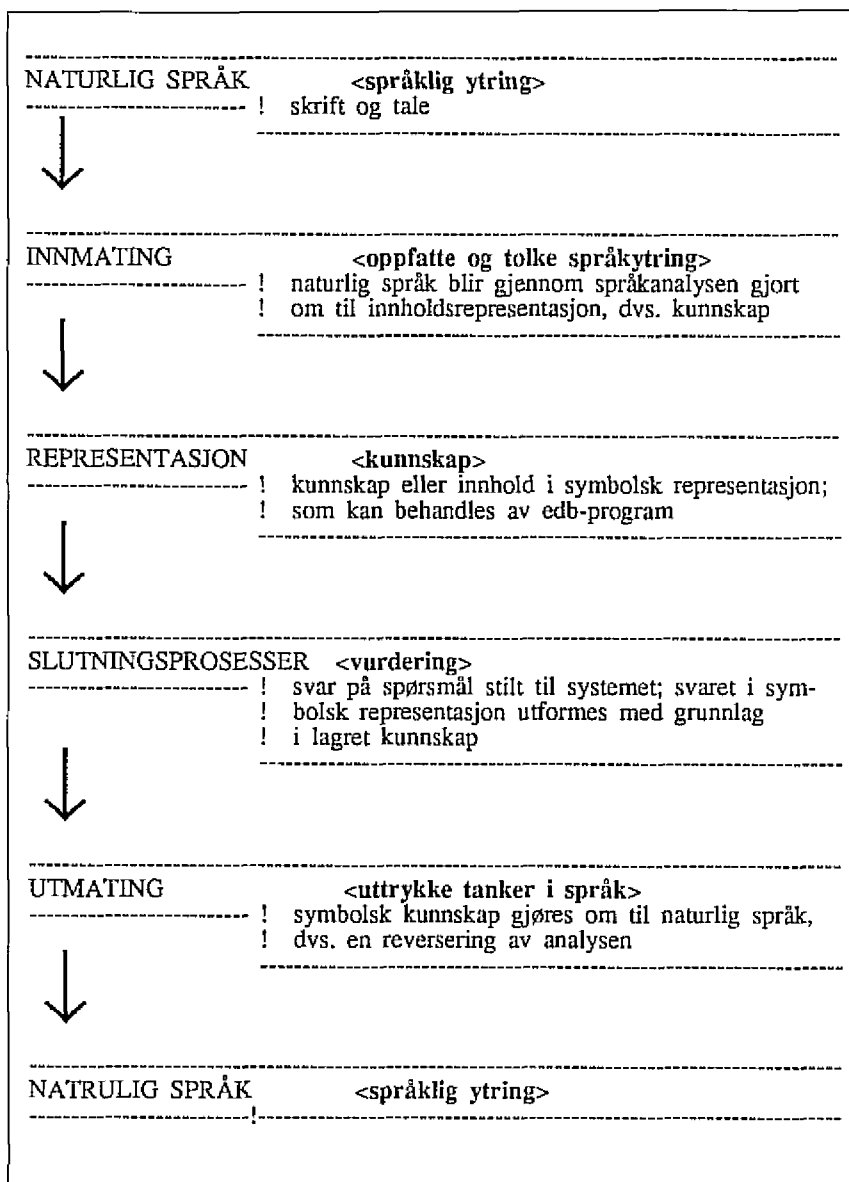


Fig. 2. Fra språk til kunnskap og tilbake til språk gjennom de ulike stadiene.

2.3.2. REPRESENTASJON AV DATA

Data representeres i datastrukturer, som i praksis enten vil si tabeller eller formaliserte utsagn. Disse gjør det mulig å representere og organisere innholdet slik at en tar vare på *betydningsrelasjoner mellom begreper*, og *innholdsstrukturer i tekst* (evt. setninger). Det første dreier seg om te-saurusstrukturer, det andre om tekstlingvistikk (eller tekstsemantikk) og formell logikk.

2.3.3. SLUTNINGSPROSESSER

Et kunnskapsbasert system skal svare på spørsmål. Til dette trengs prosedyrer som med grunnlag i eksisterende data (dvs. fakta og relasjoner) kan kombinere enhetene logisk. Dette utføres bl.a. av *logiske slutningsregler* som handlerer formell semantisk kunnskap og spørsmål. Jf Fjornes 1985.

2.3.4. UTMATING AV DATA

Utdata kan gå til instrumenter som utnytter informasjonen videre, eller de kan presenteres for brukere som skreven tekst, tale, grafikk eller tabeller. Presentasjon som *tekst* krever at formaliserte data omkodes til informasjonsenheter som igjen kan omformes til naturlig språk – overflattisk sett en reversering av språkanalysen. Utdata som *tale* krever fonetisk kunnskap, programvare og teknisk utstyr som gjør omforming fra tekst til tale mulig.

2.4. ANVENDELSE AV NATURLIG SPRÅK I ANDRE SYSTEMTYPER

2.4.1. REGISTRERING OG GJENFINNING AV INFORMASJON

Registrering og gjenfinning av informasjon har utviklet seg til å bli en svært sentral databehandlingsoppgave. Utvikling av databaseverktøy er et viktig bidrag til effektivisering av informasjonssystemer, dvs. systemer for lagring av store datamengder i strukturert form. Kunnskapssystemer skiller seg i prinsippet fra informasjonssystemer ved at de har mer fleksible rutiner for behandling og kombinerer av data, slik at de kan trekke slutninger. Bruk av naturlig språk i inndata, utdata og i dialogen med informasjons- og kunnskapssystemer er en viktig utfordring for å oppnå brukervennlighet.

2.4.2. DOKUMENTASJON

Kunnskapssystemer ordner data i en logisk struktur slik at innholdsmessig beslektete enheter knyttes sammen. En slik strukturering gjen-speiles i måten systemet analyserer og tilrettelegger inndata for intern lagring. Og en logisk struktur er en forutsetning for gjenfinning av informasjon.

En slik systematisk organisering gjør det ikke lettere bare for maskiner å finne fram, men også for mennesker. Det vil være et visst sammenfall med metoder en bruker for ordning av informasjon eller dokumentasjon i ikke-maskinelle arkiver. Således trekker en ofte på kunnskap fra arkiv- og biblioteksarbeid når en skal lage strukturer for maskinell organisering av informasjon. Bruk av tesauruser og klassifikasjonssystemer er eksempler. På den annen side vil den videreutviklingen som må til for kunnskapssystemer også kunne tilbakeføres til tradisjonelle dokumentasjonsvirksomheter. Og det kan føre til en strengere systematikk. Nettopp dette kan føre til krav om nye måter å sette opp formularer og arkivkort på – kanskje en overgang til nye formelle språkregler der en før brukte naturlig språk.

2.4.3. MASKINOVERSETTELSE

En viktig forutsetning for gode maskinoversettelser er oversettelse mellom naturlige språk gjennom en innholdsrepresentasjon, kombinert med analyse og produksjon av naturlig språk. Jf Utne 1987. Teknikkene for behandling av naturlig språk i kunnskapssystemer fokuserer sterkt på slike egenskaper, og de to fagfeltene har således et visst sammenfall i utviklingsbehov.

3. ANALYSE, REPRESENTASJON OG PRODUKSJON

I resten av framstillingen vil jeg presentere regelverk og representasjon knyttet til naturlig språk. Programvaren som tar seg av den språklige kommunikasjonen med kunnskapssystem, består av regler som sier noe om *innhold og språklige uttrykksmåter* i tekst.

Representasjon av *innhold* eller kunnskap krever et kunnskapsmønster som er en formalisert framstilling av begrep og innholdsstrukturer. Denne representasjonen inngår som del av *regelverk for analysen og regelverk for produksjon av tekst*. Den er dessuten *mal for datastrukturene* hvor det

analyserte innholdet lagres. Dette kunnskapsmønsteret er strukturert for å kunne gjenspeile de innholdsstrukturer som fins i en tekst – ideelt sett i alle typer tekster. Et slikt kunnskapsmønster vil naturligvis ikke i noe tilfelle være fullt dekkende, men kun en kunnskapsmodell.

Framstillingen av dette er delt i tre: *representasjon*, *analyse* og *produksjon*. Slutningsprosesser blir utelatt her. Jeg konsentrerer meg om tekst generelt, og tar ikke opp *dialog* spesielt, selv om dette står sentralt i brukernes kontakt med systemet.

4. REPRESENTASJON

Representasjon av innhold innebærer som tidligere nevnt en formalisering av innhold i tekst. Det som formaliseres, er innholdsstrukturer i tekst, også kalt *tekststrukturer*. Tekststrukturer omfatter *organisatoriske strukturer*, som oftest er synlige gjennom språklige eller grafiske uttrykk, og *innholdsstrukturer* som går fram av semantiske relasjoner i teksten.

4.1. ORGANISATORISKE STRUKTURER I TEKSTER

Organisatoriske strukturer omfatter konkrete grep som gjerne fungerer som et skjelett for teksten. Ofte kommer dette til uttrykk gjennom *overflateuttrykk*, som tegn eller funksjonsord (*derfor*, *i motsetning til*), eller gjennom uttrykk i teksten som *refererer til samme emne*.

4.1.1. OVERFLATEUTTRYKK

Overflateuttrykk vil oftest være *tegnbruk*, *tekstformat*, *bindeord* og *faste uttrykksmåter* som fungerer som sammenbindingsledd mellom setninger og evt. større tekstelementer. Jf Tucker et al 1986, Nishida et al 1986.

Tegnbruk har strukturerende funksjon ved bruk av parentes, komma eller tankestrek for å omkranse utfyllende kommentarer, synonyme uttrykksmåter og presiseringer. Dessuten har punktum, komma, utrops-tegn og spørsmålsteget sine hovedfunksjoner.

Tekstformat bidrar til å markere samhørende innholdsbolker som avsnitt, samt kategorier i f.eks. tabeller og figurer.

Bindeord er bl.a. *deretter* for tidsfølge, *hvis – så* for kondisjonale forhold, *derfor* og *fordi* for årsaksforhold, *som* for avgrensing, *eller* for alternativ, og for sidestilling.

Faste uttrykksmåter er f.eks. *inngår i* som generaliserende, *i motsetning til* for kontrast, *det vil si* for parafrase og adverb som uttrykker tidsrelasjoner som *samtidighet (samtidig, på samme tid)* og *tidssekvens (så, deretter)*.

4.1.2. REFERANSE

Tekster er vanligvis lenket sammen med samme begrep eller med semantisk beslektete begreper som kan følges fra setning til setning.

Semantiske rekker kan uttrykke at to påfølgende setninger henger sammen. Men de kan også vise at setninger et stykke fra hverandre henger sammen, og at setninger innimellom er et sidespor i form av utvidelse, generalisering, kontrast, digresjon o.a. Det vil altså si at rekkene viser når teksten kommer tilbake til et tidligere hovedtema. Dette blir utnyttet bl.a. av Nishida et al 1986.

Lenking gjennom *samme begrep* kan dels komme til uttrykk gjennom samme språklige uttrykk, dels også gjennom synonymer, kortformer eller pronominalisering.

Lenking gjennom *semantisk beslektete begrep* innebærer at teksten veksler mellom begrep som står i betydningsmessig slektskap (jf omtale av generiske og ontologiske relasjoner lenger nede). Slike relasjoner kommer til uttrykk som tekstbindingsmiddel, slik at når en *datamaskin* omtales i en setning, så vil den *prosessoren* som omtales like etter, med stor sannsynlighet tilhøre den tidligere nevnte datamaskinen. En *plugg* og en *ledning* som omtales i nærheten av hverandre, kan også implisere sammenheng.

4.2. INNHOLDSSTRUKTURER I TEKSTER – RAMMER

Innholdsstrukturer kommer til uttrykk gjennom innholdet i sammenhengende tekst, f.eks. mellom kjerneledd og beskriverledd i et syntagme, mellom setningsledd, mellom setninger og mellom tekstdeler. Setninger bindes internt sammen rundt en verbhandling. Til denne verbhandlingen knyttes vanligvis en aktiv utøver (oftest subjekt). Og mellom setninger kan det f.eks. være årsak-følge-forhold. En teksts innholdsstruktur kan gjenspeile forløp (eller prosess) eller være beskrivende (eller statisk). *Ramme* (eng. *frame*) brukes som fellesbetegnelse på alle slike tekstmodeller, mens *skjema* (eng. *schema*) brukes på undergruppen tekstmodeller for forløp.

En analyse som skal identifisere innholdsstrukturer krever spesifisering av rammer som grunnlag. I prinsippet kan slike systemer tolke hvilken som helst tekst, men det aktuelle programmet kan på et gitt tidspunkt bare forstå tekst som det har lagret kunnskapsrammer for. Eksempler på slike kunnskapsrammer er: kunnskap om billettkjøp, bevegelser i et trafikkrøss, en beskrivelse av datamaskiner eller avisartikler om terrorisme.

Rammene representerer prototypiske tekststrukturer og fungerer som maler for *analyse* og *representasjon*. For *analyse* brukes de som regelverk, dvs. en konstellasjon av begreper som til sammen kan danne en meningsfull sammenheng. For *representasjon* brukes de som tabellgrunnlag, dvs. som kategorier hvor en registrerer aktuelle forekomster av begrep.

Rammer kan enten modellere en statisk tekst, dvs. at rammene har en innholdsstruktur som tilsvarer en typisk ordening av temaet i en *beskrivelse*. Eller de kan fungere som rammer for *handlingsforløp*. Et system som til en viss grad danner skole, men som ikke vil bli gjennomgått her, er det amerikanske KL-ONE med forskjellige videreutviklinger, jf Brachman og Schmolze (1985).

Flere av disse modellene utnytter i stor grad kunnskap om *makrostrukturer* i tekster. Det vil si struktureringsprinsipper en gjerne bruker bevisst for å organisere eller disponere en tekst, som f.eks. slutningsrekker og sidesprang.

Følgende klassifiseringskriterier går delvis igjen i flere framstillinger (bl.a. hos Tucker et al 1986 og Nishida et al 1986).

1. *Tidsfølge*: tidsrekke, evt. samtidighet
2. *Betingelse*: betingelse med tilhørende konsekvens
3. *Utvidelse*: spesifisering
4. *Generalisering*: sammenfatning til generell form, f.eks. som definisjon eller term
5. *Kontrast*: motsetningsforhold, kontraster
6. *Parafrase*: tilnærmet likt innhold, men ikke helt identisk
7. *Ekvivalens*: samme innhold med forskjellig uttrykk
8. *Digresjon*: sidesprang vekk fra den logiske strukturen
9. *Alternativ*: alternativer, f.eks. uttrykt med *eller* mellom
10. *Sidestilling*: likeverdige ledd sidestilt med *og*, eller avsnitt på samme nivå

Organisatoriske grep som dette brukes for å binde sammen mindre tekstenheter til større. Det kan være deler av beskrivelser som bindes sammen til en større helhet som f.eks. en bruksanvisning eller annen teknisk dokumentasjon. Eller det kan være enkelthandlinger eller kortere

forløp som bindes sammen til lengre forløp eller kombineres med sidesprang til utfyllende kommentarer. All slik sammenbinding skjer med organisatoriske strukturer eller med innholdsstrukturer.

4.2.1. BESKRIVELSER – STATISKE TEKSTER

I en ramme for *beskrivelse* skal den generelle strukturen være ordnet slik det vil være naturlig å dokumentere typer og undertyper, gjenstander og deler av gjenstander, gjenstanders egenskaper mm. Langt på vei er det en anvendelse av såkalte *generiske* og *ontologiske* relasjoner.

Generiske relasjoner vil si at begrep står i over/underordningsforhold, dvs. er hhv. mer generelle eller spesielle i forhold til hverandre. Vi kan tenke oss veksling mellom *mikrodatamaskin* og det overordnede *datamaskin*, eller mellom *lagringsmedium* og det underordnede *diskett*.

Ontologiske relasjoner er alle andre relasjoner der begrep står i et innholdsmessig (semantisk) forhold til hverandre. Eksempler:

– *del-helhet*:

‘prossessor’ er en del av en ‘datamaskin’ (alle del-helhetsrelasjoner til en teknisk innretning kan til sammen betraktes som innretningens oppbygning; men helst bør måten disse er organisert på gå fram, dvs. inni, utenpå, over, under o.a.)

– *sammenheng*: en ‘plugg’ henger sammen med en ‘ledning’

– *sekvens*: ‘montering’ kommer før ‘bruk’

(*årsak-følge* er en type sekvens-relasjon som uttrykker forløp: ‘strømbrudd’ kan være årsak til ‘systemkrasj’)

– *funksjon*: ‘transport’ er en funksjon til ‘bil’

– *egenskap*: ‘flytende’ er egenskap til ‘kork’

En anvendelse av ontologiske relasjoner, som noen muligens vil kalle egenskaper, finner vi hos japanerne Nishida et al (1986). De arbeider med et kunnskapssystem som primært identifiserer trekkene eller kategoriene FUNKSJON (eng. FUNCTION), OPPBYGNING (eng. COMPOSITION) og EGENSKAP (eng. PROPERTY) i tekniske systemer eller gjenstander. Dahlgren og McDowell (1986) presenterer en mer detaljert modell.

Hahn og Reimer (1986) benytter dels generiske og dels ontologisk-

pregete nettverk for å representere kunnskap om datamaskiner. Til hver maskintype knytter de kategoriene MODELLNAVN, PRODUSENT, PROSESSORTYPE, STØRRELSE PÅ PRIMÆRHUKOMMELSEN, OPERATIVSYSTEM, DISKETTSTØRRELSE og TYPE PORTER. Flere av disse kategoriene kan f.eks. betraktes som del-helhetsrelasjoner med ytterligere spesifikasjon i form av egenskapsrelasjoner.

For å kunne fylle denne strukturen med data, må den underliggende rammen ha tabeller med mulige verdier som kan fylles i disse variablene, slik at f.eks. mulige verdier for OPERATIVSYSTEM er satt til *CP/M*, *PC-DOS*, *MS-DOS*, *XENIX* og *UNIX*. Om det presenteres andre operativsystem, blir denne variabelen stående uten verdi.

4.2.2. FORLØP – PROSESSUELLE TEKSTER

Et tekstforløp består noe forenklet sagt av *handlinger* som til sammen gir *forløp* – også kalt episoder.

4.2.2.1. HANDLING

Innenfor setninger snakker en gjerne om *handling* (eng. *event*) som er prosessuell, og innholdsmessig organisert rundt ett verb. En slik synsmåte er inspirert dels av Fillmores (1968) kasusgrammatikk, og dels av modeller utviklet av bl.a. psykologene Rummelhart, Lindsey og Norman (Rummelhart 1975).

I henhold til disse modellene sier en gjerne at det til verbet *gi* normalt knytter seg en aktiv giver – benevnt med den såkalte semantiske rollen UTØVER (eng. AGENT), noe som gis – benevnt som OBJEKT, og en mottaker – benevnt som MOTTAKER (eng. RECIPIENT). Utvalget av slike semantiske roller avhenger av type handling. Handlinger vil normalt inngå i tid og sted, uttrykt med rollene TID (eng. TIME) og STED (eng. PLACE). Og de kan knytte til seg KILDE (eng. SOURCE), som vil si råvare eller utgangspunkt, MÅL (eng. GOAL), RESULTAT (eng. RESULT) eller PRODUKT (eng. PRODUCT), som f.eks. i setningen:

UTØVER	VERBHANDLING	PRODUKT	KILDE
<i>Bakeren</i>	<i>baker</i>	<i>brød av mel, gjær, salt</i>	<i>og vann</i>

Det spesifiseres gjerne hvorvidt visse verb kan ha flere alternative oppsett av roller, og om visse roller er valgfrie. Videre stilles det krav til inn-

holdet i rollene, slik at en f.eks. kan kreve at verbet *kjøpe* må ha ‘menneske’ som UTØVER og ‘ting’ eller ‘tjeneste’ som OBJEKT, eller slik at verbene *sovne* og *våkne* må ha krav om at tid spesifiseres med et kort intervall evt. et tidspunkt, mens *sove* og *våke* må spesifiseres med lengre tidsrom. Dette kommer til nytte når en skal avgjøre hvilke begrep eller uttrykk i teksten som skal fylle den aktuelle rollen.

4.2.2.2. FORLØP – EPISODER

Hver handling inngår gjerne i kjeder av handlinger, *forløp*. Slike forløp er gjerne gjenkjennelige mønstre eller stereotyper som uttrykker *prosesser*, dvs. handlingsforløp, som f.eks. kjøp av billetter, bestilling av hotellrom og besøk på en restaurant.

Disse formelt representerte forløpene kalles gjerne *skjemaer*. Slike skjemaer inneholder spesifikasjon av handlinger som inngår i handlingsforløpet og hvorvidt aktuelle handlinger og tilhørende roller må være med (eller om de kan utelates), om de kan opptre i alternative rekkefølger, og om visse handlinger eller roller står i avhengighetsforhold til hverandre.

4.3. SAMMENFATTENDE OM REPRESENTASJON

Representasjon av innhold forutsetter regelverk som utnytter *organisatoriske* trekk, som i stor grad vil si overflateuttrykk, og *innholdsstrukturer*, som vil si modeller av kunnskapsmønster. Det innholdet som representeres lagres i strukturer, eller tabeller, som gjenspeiler slike kunnskapsmønster. *Analyse* og *produksjon*, som vi nå kommer til, utnytter alle de nevnte typer regelverk for å formidle mellom naturlig språk og innholdsrepresentasjon.

5. ANALYSE AV TEKST

Innholdsanalyse av tekst gjennomføres med *referanse til lagrete skjemaer* og med *støtte av regler som håndterer språklige uttrykk*. Så lenge teksten følger skjemaet med hensyn til obligatorisk innhold, kan en slå fast at analysen går etter rett spor. Hvis ikke, må en prøve andre skjemaer eller slå fast at det ikke fins skjema som gjør analysen mulig.

Analysen gjøres gjerne så eksplisitt som mulig. Det vil si at den for

hver handling fyller i et visst utvalg roller til tross for at de ikke er eksplisitt uttrykt. Informasjonen utledes da gjerne fra konteksten. Det kan være eksplisitering av referanse (bl.a. pronominalisering), klargjøring av felles innhold i forskjellige uttrykk (f.eks. *Reagan* og *presidenten*), samt identifisering av tid og sted.

Den språklige analysen ved innhenting av data vil enten være basert på *fullstendig parsing* eller på *identifisering av innholdsord* i teksten.

Fullstendig parsing innebærer morfologisk, syntaktisk og semantisk analyse, først av setninger og eventuelt videre av større enheter. Målet er en fullstendig innholdsanalyse som går veien om bl.a. syntaksanalyse. Jf Dyvik og Hofland 1983, og Dyvik 1987.

Identifisering av innholdsord vil si at programmet finner fram til ord som representerer begrep og relasjoner mellom begrep. Begrepene identifiseres gjerne av innholdsord som samsvarer med en underliggende ordliste i programmet, mens relasjonene blant annet identifiseres gjennom bindeord som også fins i en ordliste.

Uansett grad av parsing kombinerer analysen seg i de fleste tilfeller *overflateuttrykk* (bindeord, grafisk struktur), *referanse* (pronominalisering, synonymi og semantisk slektskap) og *innholdsstruktur*.

6. PRODUKSJON AV TEKST

Utmating av data dreier seg om å organisere representerte data eller data som er resultat av slutningsprosessen i passende informasjonsbiter slik at de danner naturlige *handlinger* som semantiske enheter. Hver av disse *handlingene* skal omdannes til setninger. Hver *handling* vil før utmating være forholdsvis eksplisitt med alle semantiske roller, i det minste de obligatoriske, fylt med innhold. Ved overgang til naturlig språk må en del redundans fjernes, slik at f.eks. tid og sted ikke gjentas for hver setning, og slik at alle begrep ikke uttrykkes med sine fulle former, men reduseres til pronomener, evt. utelates. Generering av språk krever regler på alle språknivå som motsvarer dem som benyttes ved analyse ved hjelp av fullstendig parsing.

Når kunnskap representert i innholdsstrukturer omdannes til overflatespråk, kan oppsettet nedenfor vise *mulige stadier*. Jf Kittredge et al 1986, Nishida et al 1986, Laubsch et al 1984, Hanakata et al 1986, Novak 1986, Danlos 1987a, Danlos 1987b.

1. Organisering av innholdet, slik at det faller i *naturlige innholdsblokker* i en naturlig og leservennlig progresjon eller orden.

I et kanadisk system (Kittredge et al (1986)) for automatisk generering av værmeldinger prioriterer en f.eks. hovedbolker primært etter viktighet, sekundært etter årsakskjeder, slik: *vind, skydekke, nedbør, tåke og sikt*. På neste nivå kommer først *generell* og deretter *lokal* informasjon. Og på nivået under dette ordnes informasjonen *kronologisk*.

2. Utplukking av én og én *handling*. Dette danner grunnlag for én og én setning rundt ett verb-begrep med tilhørende begreper for tilknyttete roller.
3. Valg av aktuelt *overflateuttrykk* med grunnlag i kontekstuelle trekk – enten ord eller lengre fraser.

Hos Laubsch et al (1984) kan begrepet ‘accident’ avhengig av kontekst uttrykkes med *accident, bug, error* eller *failure*.

Et konstruert eksempel for norsk bokmål er ‘spise’ som kan uttrykkes med *spise* eller *ete*, avhengig av UTØVEREN.

4. *Linearisering*, dvs. overgang fra semantisk representasjon til syntaktisk overflatestruktur. Det omfatter sideordning, underordning og rekkefølge.

Eksempler fra Laubsch et al (1984):

- fravær av UTØVER gir f.eks. passivkonstruksjon med logisk OBJEKT eller MOTTAKER på syntaktisk subjekts plass:

Maten blir spist.

- fravær av UTØVER og OBJEKT gir nominalisering:

Det ble gjort et vedtak.

5. *Redundans* reduseres eller fjernes, dvs. (eksemplene og kategoriene er konstruerte her)

- substantiv

- forkortes

Gro Harlem Brundtland -> Brundtland

- generaliseres

Gro Harlem Brundtland -> statsministeren

- pronominaliseres

Gro Harlem Brundtland -> hun

- utelates
Gabriel strikket og Gabriel så på TV
Gabriel strikket og så på TV
- tids- og stedsuttrykk (presisjon avtar med økende avstand)
 - forkortes – ofte kontekstavhengig
16. april 1987 kl. 14.45 -> kl. 14.45
 - gjøres mindre presist
fredag kl. 20.30 -> fredag kveld,
ved Oslo rådhus -> i Oslo
 - gjøres om til adverb eller generelt steds- eller tidsuttrykk
kl. 12 -> deretter
 når forrige handling foregikk sekvensielt før dette,
ethvert tidspunkt den 17. mars 1987 -> hele dagen,
ved Oslo rådhus -> der
 - utelates når det går fram av konteksten, f.eks. tidligere setninger/ytringer; i dypstrukturen vil tid og sted være representert for alle *handlinger*

Under denne siste poleringen av de språklige uttrykkene vil en del systemer også til en viss grad kunne motvirke monoton språkføring ved tilfeldig eller kontekstavhengig variasjon i uttrykksmåte.

Slik det her er beskrevet vil et kunnskapssystem i mer eller mindre naturlig språkdrakt kunne uttrykke lagret kunnskap. Det kan være svar på spørsmål fra brukerne eller referat av sammenhengende kunnskap. Forutsetningene for vellykket språkproduksjon er språkregler som er gode til å simulere språkatferd.

Arbeid med regler for språkproduksjon vil gi innsikt i skriveteknikk for oss selv også – stikkord er: entydighet, struktur, logikk, konsekvens og variasjon.

7. AVSLUTNING

Denne framstillingen har vært et forsøk på å gå veien fra og til naturlig

språk via representasjon av kunnskap, og samtidig presentere et viktig felt der vi som filologer kan bidra til utvikling innen feltet kunstig intelligens.

Mye av dette er kjente temaer som faller inn under filologiske fag, særlig for språk, litteratur og logikk, uten at det samles innen noe enkeltfag. Dertil kommer kunnskap om edb. Arbeid med dette emnet krever tverrfaglighet og samarbeid.

Det vil koste mange årsverk og mye penger å komme langt, og det vil være vanskelig å få støtte til ett stort prosjekt som skal gå over lang tid. Informasjonsteknologisk prosjektarbeid vil med stor sannsynlighet måtte preges av nitid og målrettet arbeid med mindre prosjekter, der en jevnlig må vise fram resultater som har praktisk anvendelse. En må sette seg delmål der det er overlappende delmål mellom behov for kunnskapsystemer, deriblant naturlig språk, på den ene siden, og maskinoversettelse, informasjonslagring og -gjenfinning inkludert dokumentadministrasjon på den andre siden.

Vi har å gjøre med et fagfelt der filologene på linje med andre fagmiljøer i tillegg til grunnforskning også må orientere seg mot praktisk anvendelige resultater og se seg om etter støtte fra næringslivet og de sektorer i det offentlige som primært ser etter resultater de kan bruke.

LITTERATUR

- Brachman, R.J. og J. Schmolze (1985): «An overview of the KL-ONE knowledge representation system». *Cognitive Science*, vol. 9, No. 2, s. 171-216
- Coling '84: *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics. Proceedings of Coling '84*. Stanford
- Coling '86: *11th International Conference on Computational Linguistics. Proceedings of Coling '86*. Bonn
- Dahlgren, K. og J. McDowell (1986): «Kind Types in Knowledge Representation». Coling '86, s. 216-221
- Danlos, Laurence (1987a): *The Linguistic Basis of Text Generation*. Cambridge
- Danlos, Laurence (1987b): «The Linguistic Basis for Text Generation». Foredrag på *Association for Computational Linguistics European Chapter 3rd Conference*. Trykkes i konferanserapporten.
- Dyvik, Helge J.J. (1986): «Unification-based Chart Parsing: Some Modifications and Extensions of D-Patr». *Nordisk seminar om maskinoversættelse*, s. 141-160. København

- Dyvik, Helge J.J. og Knut Hofland (1984): «Parsing basert på LFG: Et MIT/Xerox-system basert på norsk». Hein, Anna Sågvall (red.): *Føredrag ved De nordiska datalingvistik dagarna 1983*. Centrum för datorlingvistik, Uppsala Universitet
- Fjeldvig, Tove og Anne Golden (1983): «Automatisk rotlematisering». *HD* nr. 3, 1983, s. 22-35
- Fjeldvig, Tove og Anne Golden (1984): *Automatisk rotlematisering – et lingvistisk hjelpemiddel for tekstsøking*. CompLex nr. 9/84. Institutt for rettsinformatikk, Universitetet i Oslo
- Fjornes, Erik (1985): «Humanistiske fag og kunnskapsbaserte systemer». *HD* nr. 3, 1985, s. 4-10
- Hahn, Udo og Ulrich Reimer (1986): «TOPIC Essentials». Coling '86, s. 497-503
- Hanakata, K., A. Lesniewski og S. Yokoyama (1986): «Semantic based generation of Japanese German translation system». Coling '86, s. 560-562
- Karlsson, Fred (1984): «Finländsk datalingvistik». *HD* nr. 3, 1984, s. 14-20
- Kittredge, R., A. Polguere og E. Goldberg (1986): «Synthesizing Weather Forecasts from Formatted Data». Coling '86, s. 563-565
- Koskenniemi, Kimmo: *Two-level Morphology: A General Computational Model for Word-Form Recognition and Production*. Publications No. 11, 1983, Dept. of General Linguistics, University of Helsinki
- Laubsch, J., D. Roesner, K. Hanakata og A. Lesniewski (1984): «Language Generation from Conceptual Structure: Synthesis of German in a Japanese/German MT Project». Coling '84, s. 491-494
- Lie, Svein (1980): *Automatisk syntaktisk analyse. Del 1. Grammatikken*. Rapport nr. 17. NAVFs EDB-senter, Bergen
- Nishida, Fujio, Shinobu Takamatsu, Tadaaki Tani og Hirojii Kusaka (1986): «Text Analysis and Knowledge Extraction». Coling '86, s. 241-243
- Novak, Hans-Joachim (1986): «Generating a Coherent Text Describing a Traffic Scene». Coling '86, s. 570-575
- Rummelhart, D.E. (1975): «Notes on a Schema for Stories». Bobrow, D.G. og A. Collins (red.): *Representation and Understanding*, s. 211-236. New York
- Slethei, Kolbjørn (1986): «Tekst-til-tale-syntetisering». *HD* nr. 1, 1986, s. 34-40
- Tucker, Allen B., Sergei Nirenburg og Victor Raskin (1986): «Discourse and Cohesion in Expository Text». Coling '86, s. 181-183
- Utne, Ivar (1985): «Kunstig tale på norsk». *HD* nr. 3, 1985, s. 38-49

Utne, Ivar (1987): «Maskinoversettelse. Hvor står vi – hvor går vi?». *Språklig Samling*, nr. 1, 1987. Oslo

Ivar Utne er amanuensis ved Avd. for norsk leksikologi, Nordisk institutt, Universitetet i Bergen.

GRUNNLAGSMATERIALE FOR INFORMASJONSSØKING VED NORSK TERMBANK

Ivar Utne

SAMMENDRAG

Artikkelen tar opp hvordan fagterminologisk materiale ved Norsk termbank kan utnyttes som datagrunnlag i informasjonssøkesystemer. Det dreier seg om lister med termer/uttrykk som har varierende skrivemåter, forkortelser, synonymer innen samme språk, og emnesorterte ordlister (tesauruser).

NORSK TERMBANK OG AVDELING FOR NORSK LEKSIKOLOGI

Norsk termbank er en språklig oppdragsavdeling ved Universitetet i Bergen. I 1986 har antall ansatte ligget omkring 18-19. Prosjektene ved Termbanken er eksternt finansierte og sikter mot praktiske løsninger. Hovedbeskjeftigelsen er utvikling av teknisk fagterminologi, som til nå hovedsaklig har vært relatert til oljeindustrien. Virksomheten utnytter i stor grad datamaskinelle metoder for lagring og gjenfinning av ordboksmateriale, for språkkontroll og for datastøttet oversettelse. Edb-løsningene er primært støtte for terminologisk arbeid, men inngår også i planer for utviklingsarbeid i stor bredde innenfor feltet informasjonsteknologi, deriblant kunnskapsbaserte systemer og automatisert oversettelse.

Termbanken er knyttet til Avdeling for norsk leksikologi (ANL), Nordisk institutt. Der fins det store mengder allmennspråklige ordlister i maskinleselig form, bl.a. korrekturordlister (stavingslister), ordlister med orddelingspunkt, ordlister med oppslagsform og tilhørende bøyingsformer, ordlister med ordklasseinformasjon, tospråklige ordlister (norsk mot engelsk, tysk og fransk) og ordlister med eksempelfraser. Deler av dette materialet inngår i samarbeidsprosjekter med andre deler av Universitetet og NAVFs edb-senter for humanistisk forskning.

Termbanken og ANL utgir skriftserien *Norske språkdata* hvor bl.a. prosjektrapporter utkommer. Det fagterminologiske materialet gjøres tilgjengelig som ordlister etter avtale med oppdragsgiverne.

TERMINOLOGISK VARIASJON

Sentralt i prosjektarbeidet ved Termbanken står registrering og utvikling av norsk fagterminologi, samt registrering av engelske tilsvarende. Det opparbeidede materialet er grunnlag for oversettelse og utforming av brukerdokumentasjon innen oljeindustrien.

I denne artikkelen vil jeg ta utgangspunkt i disse dataene og peke på informasjonsteknologiske bruksområder. Jeg vil legge forholdsvis mye vekt på de språklige sidene. Hensikten er å kaste lys over språkvitenskapelige forutsetninger for informasjonssøking.

Mange måter å uttrykke det samme på

Under registreringen av terminologi har en kommet over betydelige mengder data der samme term er skrevet på ulike måter. Det gjelder både engelsk og norsk, og både stavemåter og oppdeling av ord. Registrering av skrivevarianter er ikke utført for alt materiale ved Termbanken. Omfanget av dette har tiltatt gradvis etter hvert som en har fått øynene opp for nytten.

Det er dessuten svært vanlig at samme gjenstand eller fenomen går under flere forskjellige navn, dvs. det eksisterer synonymi.

En slik uorden som her er antydning, gjør at arbeidet med å utarbeide fagordbøker for norsk stiller store krav til organisering og gjenfinning av termer som er skrivevarianter og synonymmer. Dette får særlig følger for datamaskinell behandling: alfabetisering, søk og konsekvenssjekk.

Konsekvent språk er ikke bare et praktisk problem under terminologiarbeid og oversetting. Det er også nødvendig for brukerne av dokumentasjonen at de vet hvordan ord og uttrykk blir skrevet. De skal kanskje også søke datamaskinelt i tekstmasser der de samme uttrykkene inngår. Et aktuelt eksempel er brukerdokumentasjonen for Gullfaksfeltet som blir lagt inn på CD-ROM for å muliggjøre maskinell gjenfinning.

TERMINOLOGI SOM DEL AV SØKESYSTEMER

Søk i terminologimateriale eller i tekster skjer normalt ved hjelp av informasjonssøkesystem. Slike søk foretas enten i fritekst eller på emneord.

Fritekstsøk innebærer at brukeren får fram tekstutdrag hvor søkekriteriet fins. En enklere variant av fritekstsøk finnes i tekstbehandlingssystemer.

mer med funksjoner som FINN eller SØK. Men bruksområdene er videre enn det. Fritekstsøk brukes i store tekstmengder og med mer nyanserte valgmuligheter, f.eks. med søk på flerordkombinasjoner, og søk på begynnelsen eller slutten av ord (høyre- og venstretrunkering). Aktuelle bruksområder er administrative, juridiske tekster og oppslagsverk. Blant annet fins leksikonet *Groliers Academic American Encyclopedia* (ca. 10.000 sider) på CD-ROM med program for fritekstsøk.

Søk på emneord forutsetter at tekstmengden/dokumentene er klassifisert etter emneord (ett eller flere for hvert dokument). Brukerne kan senere søke blant emneordene og få opplysning om hvilke tekster/dokumenter som er klassifisert under dem. Slike søk er særlig aktuelle for bibliotek og dokumententre i næringslivet. Tesauruser, dvs. emnesorterte ordlister, er aktuelle hjelpemidler for emnesøk.

For videre omtale av tekstsøkesystemer vises det til Fjeldvig 1986.

Utnytting av terminologimateriale i søkesystemer

Termbankens materiale består, som nevnt, av ordlister med skrivevarianter, synonymer og semantiske nettverk. Oversikt over skrivevariantene kan særlig komme til nytte ved fritekstsøk, synonymene vil kunne utnyttes både til fritekstsøk og emneordsøk, mens de semantiske nettverkene er særlig nyttige for søk på emneord.

Tabelldrevet program sjekker data

I praksis kan ordlistene bygges inn i et søkesystem, slik at søkeprogrammet har tilgang til informasjonen i en tabell som enten er lest inn i minnet eller fins på fil eller i en base. Programmet starter søkeprosessen med å undersøke om det fins alternative skrivemåter for søkekriteriet, og slår dem opp – helst slik at fullstendig tilslag kommer først. Slike lister er hovedtema for denne artikkelen.

Produktive regler

De etablerte filene eller tabellene kan bygges ut med tillegg av klart produktive regler for elementer som danner uttrykk (termer). Slike elementer kan være røtter som: AIR, DATA og DOWN og endelser som -OUR, -OR, -ING og -NING.

De produktive elementene utnyttes i forbindelse med flerleddete uttrykk som ikke forekommer i tabellen. Vi kan tenke oss søk på DATA PROCESSING, og at uttrykket (eller termen) ikke fins i tabellen over uttrykk. Men programmet finner DATA i tabellen over aktuelle ordlagingsselement. Dette gir grunnlag for å anta at termen kan skrives på opp til tre måter. I tillegg til særskrivning – DATA PROCESSING – kan en anta at det er mulig å skrive det som ett ord eller som et ord med bindestrek mellom røttene – DATA-PROCESSING. Tilsvarende vil programmet ved søk på SHUTDOWN finne DOWN som ordlagingsselement og utvide søket til SHUT-DOWN og SHUT DOWN.

På samme måte vil søk på ord som inneholder aktuelle endelser, også føre til søk på tilsvarende ord der den andre endelsesvarianten inngår.

Denne strategien med røtter og endelser fører naturligvis til mange gale maskinelle forslag til søk. Men den åpner samtidig muligheten for mer uttømmende søk. Produktive regler blir ikke fulgt opp senere i artikkelen.

Søk på flere språk samtidig

I en del av den tekstmassen Termbanken har hatt tilgang til, har samme tekst eller dokument både norske og engelske setninger og tekstavsnitt. Som nevnt ovenfor forekommer også engelske ord i norsk tekst og norske ord i engelsk tekst. Disse blandingsvariantene kommer vel dels av at nordmenn skriver engelsk med utilstrekkelige engelskkunnskaper og norsk med manglende kjennskap til norsk terminologi. Dessuten bunner vel noe i forglemmelser og i det faktum at det ikke fins kjente norske uttryksmåter.

Skal en ha muligheter for å søke i slike tekster, må søkesystemet kunne behandle denne tospråklige situasjonen. Det betyr at når en bruker vil søke etter MAGNETVENTIL i en norsk tekst, må en også finne de stedene der denne er kalt SOLENOID VALVE og SOLENOID VENTIL. Og når en søker etter SEKSKANTMUTRER, må en også finne oppføringer med HEXAGON NUTS.

I forlengelsen av dette kan det også være aktuelt å la brukerne søke med norske søkeord i engelsk tekst og med engelske søkeord i norsk tekst.

Slike tospråklige og flerspråklige søkemuligheter vil uten tvil føre til mange gale tilslag fordi det ikke alltid er 1:1-forhold mellom språkene. Flerspråklig søk blir ikke fulgt opp i denne artikkelen.

SKRIVEVARIANTER

Merknad: I det følgende vil den formen som er standardisert til bruk ved Termbanken, være uthevet.

Uklare normer

Ofte er det ikke en ensartet praksis for skrivemåter verken i det norske eller de engelskspråklige språksamfunnene. Det kan være uttrykk for mangel på norm, slik det ser ut til å være for sær- eller sammenskriving av SHUT DOWN/SHUTDOWN/SHUT-DOWN, SUB SEA/SUBSEA/SUB-SEA og DATA BASE/DATABASE.

Bruken av skilletegn mellom ordlagingselement er ikke entydig standardisert for norsk, f.eks.:

VANN/OLJE-BLANDING VANN-OLJE-BLANDING
OLJE/VANN-ANALYSATOR OLJE-VANN-ANALYSATOR

I Vinjes *Skriveregler. Bokmål.*:32 aksepteres dette i teknisk språk, men frarådes ellers fordi skråstreken heller betyr ELLER enn OG.

Valgfrihet

På norsk er det ofte valgfrihet, som f.eks. for **KOBLE/KOPLE** og **MUTTERE/MUTRER**, som alle er tillatte former etter norsk bokmålsrettskriving. **ADAPTOR/ADAPTER** er parallelle eksempler innenfor det engelske språkområdet.

Feil som går igjen

TINGLYSNING er galt innenfor gjeldende bokmålsrettskriving, mens **TINGLYSING** er norm. Førstnevnte form er eksempel på former som ikke er tillatt i bruk, men som ofte brukes i aviser og bøker og har uoffisiell aksept. Andre tilfeller der utbredt bruk står i motsetning til norm, er f.eks. disse som ender på **-ING/-NING**:

UTBLÅSING/UTBLÅSNING og
INFORMASJONSSØKING/INFORMASJONSSØKNING,

og disse med eller uten sammenføyings-S:

TILLUFTVIFTE/TILLUFTSVIFTE, FRISKLUFTMASKE/FRISKLUFTSMASKE.

Særskrivning av sammensatte substantivuttrykk som VARM RETT og POTET LØV regnes som engelsk påvirkning på norsk ordlagning. Det er en gal, men hyppig brukt skrivemåte.

Sok i ikke-standardiserte tekster

Under arbeidet med standardisering av skrivemåter ved Termbanken har en laget lister som inneholder standardiserte former og til dels andre registrerte ikke-anbefalte skrivevarianter. De fagterminologiske ordlistene (dvs. de tospråklige listene) er primært ordnet etter standardiserte skrivemåter. På den måten har en et forholdsvis godt grunnlag for å finne fram i materialet.

Lister med skrivevariasjon er et godt grunnlag for å finne fram i tekster som ikke er standardisert. I klartekst vil det si at edb-leselige lister som dokumenterer faktisk forekommende skrivevariasjon, kan styre søk på en skrivevariant inn mot samme uttrykk skrevet på andre måter også.

FORKORTELSER OG KORTFORMER

Mange forkortelser i fagtekster

Forkortelser fins i alle typer tekster, men tettheten øker jo viktigere det er å få mye informasjon inn på liten plass. Utbredelsen av forkortelser er betydelig i fagspråk, ikke minst innen teknisk fagspråk, der det er vanlig med termer som består av flere ord, og som det derfor er tungvint å gjenta. Forkortelser er en praktisk løsning.

Når behovet for å komprimere tekst øker, øker også mulighetene for at en tar i bruk sjeldne forkortelser eller skaper nye. Det er ikke uvanlig at samme forkortelse betyr mange forskjellige ting. I Amkreuz: *Abbreviations of Data Processing Terms*, hvor det inngår forkortelser for engelske, franske og tyske uttrykk, har BS 58 betydninger, f.eks. BACKSPACE, BLANK SPACE, BLOCK SIZE, BLOCK STRUCTURE, BRITISH STANDARD og BYTE SIZE. Ovennevnte verk har ca. 55.000 forskjellige forkortelser, dvs. forkortede termer. Ikke bare standardiserte forkortelser, men også andre skrivevarianter forekommer:

- På norsk er NUMMER standardisert til NR. eller NR., men NO. er også i bruk.

- På norsk er MD. standardisert for MÅNED og MÅNEDER men MND. og MNDR. (flertall) er vel mer brukt.

Løland og Thoresens *Norsk forkortingsordbok* har oversikt over standardiserte norske forkortelser.

Ad hoc-forkortelser – kortformer

Praktiske behov gjør at en går utenom etablerte forkortelser og tar i bruk andre avkortede former. Ofte er de laget ved at en har kuttet vekk slutten eller en del av ordet, f.eks. TETTST for TETTSTED. Andre ganger dreier det seg om ord som det ikke fins noen etablert forkortelse for, f.eks. V eller VN for VANN. Disse kaller jeg i fortsettelsen for kortformer. Valg av kortform er avhengig av hvor mye plass det er til disposisjon. Om en ikke har plass til teksten SALGSGASSKOMPRESSOR, kan en f.eks. skrive SALGSG.KOMPRESSOR eller SG.KOMPR.

Slike kortformer er svært utbredt i tekniske tekster. De er svært lite standardisert, og bruken av dem er svært impulsiv. Av dette følger at de er vanskelige å forstå for brukerne. Det hender også at den som har tatt i bruk en kortform, senere ikke selv kan huske hva som er ment. En- og tobokstavsformer er særlig vanskelige.

I det materialet Termbanken har arbeidd med, støter en ofte på dette fenomenet, særlig i skilt- og dataskjermtekster der plassen er svært begrenset. Mye informasjon skal inn på liten plass.

Eksempler på kortformer (flere eksempler fins i Simonsen 1986):

VÆSKEUTSKILLER	- VÆSKEUTSK
PUMPE	- PMP
TETNINGSOLJE	- TETO
POWER SUPPLY SELECTION BOARD	- PWR SUP SEL BD

Dette anarkiet av forkortelser og kortformer åpner for mye uklarhet, særlig på grunn av store muligheter for homonymi (samme skrivemåte med forskjellige betydninger). Om vi tenker oss dette lagt inn i en tabell for et søkeprogram, kan det for det første føre til at søk på en forkortelse vil føre fram til forskjellige termer som alle har dette som en mulig forkortelse eller kortform. For det andre vil søk på en term ikke bare føre fram til alle tekststeder med en eventuell forkortelse eller kortform for denne termen, men også alle forekomster av den samme forkortelsen/kortformen brukt for andre termer (dvs. homonymi).

Disse feilkildene gjør det neppe rimelig å bruke lister med kortformer

som grunnlag for søk rent generelt. Om kortformer skal benyttes, må det være tekstavhengig. Det kan altså være meningsfullt å bruke lister med kortformer til støtte for søk i tekster som har stor likhet med de tekstene som kortformene er hentet fra.

Forkorting av lengre fraser og setninger

Nedkorting av språklige uttrykk skjer også på frase- eller setningsplan. Eks.:

- VENTIL LUKKET (ER utelatt)
- LUFTINTAK, VARMESLØYFE 01-GX20 (FOR erstattet av komma)

Et søkesystem bør kunne utnytte dette, og utføre søkene også uten ord og endelser som ofte faller bort under forkorting av lengre uttrykk. Det kan være preposisjoner, hjelpeverb, modalverb og mulige bøyingsendelser som -EN, -ET og -ENE på norsk. Nyttan av dette vil naturligvis være avhengig av teksttype.

SYNONYMER

Variert synonymi innen nye fagområder

Under gjennomgang av teknisk dokumentasjon og annet skriftlig materiale for oljeindustrien blir en vitne til stor variasjon når det gjelder valg av termer.

Grunnlaget for variasjonen ligger bl.a. i at de tilsatte har ulik skole- og erfaringsbakgrunn, og at vareleveranser skjer fra leverandører som bruker hver sine termer for samme ting. Samme gjenstand kalles PINNE av noen og STIFT av andre. Eksempler på blanding mellom norsk og engelsk er nevnt tidligere. Dessuten brukes termer ofte galt. Dette kan illustreres med et eksempel fra allmennspråket. Der brukes termen STIKKONTAKT ofte feilaktig om PLUGG. STIKKONTAKT er imidlertid bare den delen av en elektrisk koplning som står fastmontert i veggen.

Eksempelene er ikke alltid like trivielle. Ofte er de så vanskelige at verken terminologene eller ingeniørene som rådspørres, kan gjøre rede for om det er samme sak.

Under gjennomgangen av en lagerliste kom terminologene ved Termbanken over følgende tre engelske termer som viste seg å være det sam-

me: WELLBORE WIPER, CABLE TYPE WALL CLEANER og HAMMER-LOK TURBOBONDER. Den norske termen er HULLSKRAPE MED VAIERSLØYFER, som er betegnelse på en type utstyr som brukes for å rense borehull. Begrepet ble i dette tilfellet oppklart ved hjelp av varekataloger fra leverandører til oljeindustrien. De tre termene representerte tilsvarende produkter fra tre ulike firmaer, jf. figur 1. Tilsvarende terminologiske temaer tas opp i Simonsen 1987.

Synonymien registreres og normeres

Terminologimaterialet ved Termbanken er så langt det er praktisk gjennomførlig, ordnet begrepsvis i begrepsposter. Se figur 2 som viser en slik post med engelsk og norsk hovedterm samt eksempler på synonymer og frarådte termer. I vår sammenheng vil hovedterm si den termen som er tilrådd for bruk, mens synonym er akseptabelt men ikke tilrådd fordi det vil føre til inkonsekvent termbruk. En term klassifiseres som frarådd når den er uheldig av en eller annen grunn. Den kan vekke gale forestillinger eller, som i figur 2, være pleonastisk – JOINT og CONNECTOR sier det samme. Også her har vi å gjøre med flere leverandører, og term-slektskapet ble avklart på samme måte som ovenfor.

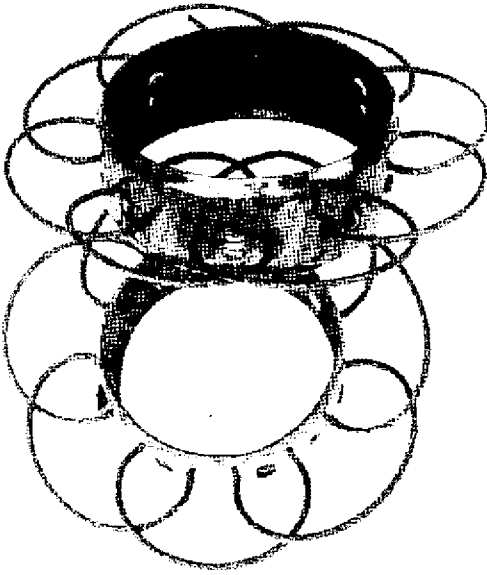
engelsk hovedterm	AUTOMATIC LOCKING CONNECTOR
engelsk synonym	SQUUNCH JOINT
engelsk synonym	STAB CONNECTOR
engelsk frarådd	SQUUNCH JOINT CONNECTOR
norsk hovedterm	LÅSERINGSKOPLING

Figur 2. Termpost med hovedtermer, synonymer og frarådte termer.

Termene i begrepspostene er nyttige både for fritekstsøk og søk på emneord. I fritekstsøk vil det være en realistisk utvidelse å få systemet til å søke etter og presentere tekstutdrag der også synonymer til søkekriteriet fins. I emneordsøk vil dette være et supplement som kan hjelpe brukeren med å spore opp eksisterende emneord. Materialet vil også være et nyttig grunnlag om en vil utvide en thesaurus med flere *brukt for*-relasjoner, dvs. flere synonymer til eksisterende begrep eller emneord.

TESAURUSER

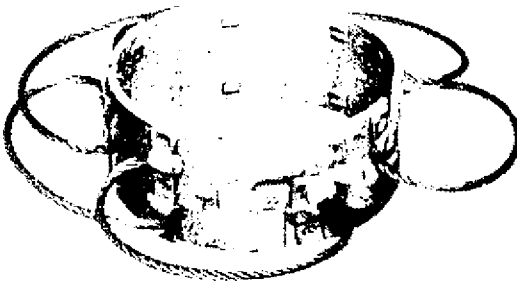
En thesaurus er en emnesortert ordliste. Den tjener som hjelp til å ord-



Wellbore Wipers



Cable Type Wall Cleaner



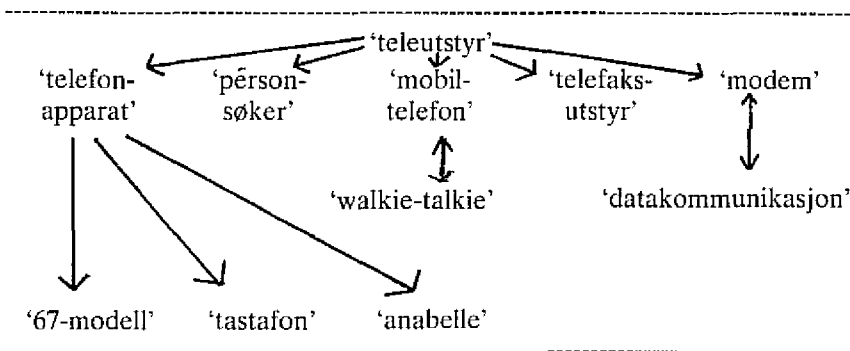
Hammer-Lok Turbobonder

ne og lette søkeveiene inn i arkiv og bibliotek. I praksis er informasjon-innholdet i dokumenter klassifisert med emneord. Disse emneordene fins i tesaurusen i hierarkisk strukturert form, også kalt tesaurusens semantiske nettverk. Ved søk kan brukeren ta seg fram til aktuelle dokumenter ved først å lete seg fram til aktuelle emneord i tesaurusens semantiske nettverk. Det kan være fra generelle til mer spesielle ord, eller det kan være fra et ord til et innholdsmessig beslektet ord.

En begrepsstruktur

Begrepsinnhold er symbolisert med ord i enkle anførselstegn. Når vi her skriver 'telefonapparat', mener vi altså begrepet eller betydningen. TELEFONAPPARAT med store bokstaver symboliserer det språklige uttrykket – termen. Skillet mellom term og begrep er sentralt i en tesaurus. Selve det semantiske nettverket eller emnegrupperingen er basert på begrepsinnhold. I vårt tilfelle struktureres det med begrepsrelasjonene *overordnet*, *underordnet* og *se også* (også kalt *assosiert*).

Figur 3 viser at 'teleutstyr' er samlebegrep for 'telefonapparat', 'mobiltelefon', 'modem' mm. 'Teleutstyr' er altså overbegrep og de tre andre er underbegrep. 'Telefonapparat' er igjen overbegrep for 'anabelle', '67-modellen' og 'tastafon'. 'Tastafon' er altså en type 'telefonapparat'. Pile-ne går fra overbegrep til underbegrep.



Figur 3. Illustrasjon av begrepsnettverk. Omtalt i teksten.

Det kan være en *se også*-relasjon mellom begreper som ikke kan ordnes verken hierarkisk oppadstigende eller som sideordnet (som følge av at de har felles overbegrep), men som det er nyttig å bli minnet om i søk,

dvs. som har et assosiert forhold til hverandre. Et slikt forhold vises mellom 'modem' og 'datakommunikasjon'. I figur 3 er det uttrykt med toveis-pil.

Begreps- og termrelasjoner

Det er knyttet en tilrådd term til hvert begrep. Dessuten kan det til et begrep også være knyttet en eller flere frarådte termer. I vårt eksempel knytter vi den tilrådte termen TELEUTSTYR til begrepet 'teleutstyr'. Dessuten knytter vi også den frarådte termen TELEKOMMUNIKASJONSUTSTYR til det samme begrepet. (Merknad: Denne klassifiseringen av TELEUTSTYR og TELEKOMMUNIKASJONSUTSTYR er kun gjort for vår illustrasjon, og bunner ikke i noen offisiell norm.) Vi sier at det er en *bruk*-relasjon fra den eller de frarådte termene til den tilrådte, og en *brukt for*-relasjon motsatt vei. I denne sammenhengen er alle termer som er knyttet til samme begrep, synonymer. Bare ett av dem er tilrådd som emneord.

Nettverket kan skrives som lister

Figur 4 viser en alfabetisk liste over termene i vårt vesle semantiske nettverk. Her er det termene og ikke begrepene som er vist. Det vil si at når en term, f.eks. DATAKOMMUNIKASJON, står med MODEM som *se også*-relasjon, vil det si at vi har å gjøre med de to underliggende begrepene 'datakommunikasjon' og 'modem' som det er en *se også*-relasjon mellom. I listen representeres begrepene med den tilrådte termen eller det språklige uttrykket. Listen viser også frarådte termer, f.eks. TELEKOMMUNIKASJONSUTSTYR. Disse er ikke ført opp med tilhørende semantisk relaterte termer, men kun med henvisning til tilrådd term.

Begrepsstrukturen som er vist i figur 4, er gjengitt i et begrepshierarki i figur 5. Dette hierarkiet er ordnet som den begrepsstrukturen vi refererer til. De språklige uttrykkene er her hovedtermene til de respektive begrepene.

I figur 6 er alle termer alfabetisert med henvisninger til begrepshierarkiet i figur 5. I de tilfellene termene er frarådte vises det til den tilrådte termen og den tilhørende plassering i begrepshierarkiet, jf TELEKOMMUNIKASJONSUTSTYR og TELEUTSTYR i figur 4 og 6.

67-MODELLEN	TELEFAKSUTSTYR
OT TELEFONAPPARAT	OT TELEUTSTYR
ANABELLE	TELEKOMMUNIKASJONSUTSTYR
OT TELEFONAPPARAT	BRUK TELEUTSTYR
DATAKOMMUNIKASJON	TELEUTSTYR
SO MODEM	UT PERSONSØKER
MOBILTELEFON	TELEFAKSUTSTYR
OT TELEUTSTYR	MODEM
SO WALKIE-TALKIE	TELEFONAPPARAT
MODEM	MOBILTELEFON
OT TELEUTSTYR	BF TELEKOMMUNIKASJONSUTSTYR
SO DATAKOMMUNIK.	WALKIE-TALKIE
TASTAFON	SO MOBILTELEFON
OT TELEFONAPPARAT	
TELEFONAPPARAT	
OT TELEUTSTYR	
UT 67-MODELLEN	
ANABELLE	
TASTAFON	

Figur 4. Alfabetisk liste av tesaurus. *SO* = se også, *OT* = overterm, *UT* = underterm og *BF* = brukt for.

-
- 1 TELEUTSTYR
 - 1.1 – MOBILTELEFON
 - 1.2 – MODEM
 - 1.3 – PERSONSØKER
 - 1.4 – TELEFAKSUTSTYR
 - 1.5 – TELEFONAPPARAT
 - 1.5.1 -- 67-MODELLEN
 - 1.5.2 -- ANABELLE
 - 1.5.3 -- TASTAFON
-

Figur 5. En liten hierarkisk struktur.

67-MODELLEN	1.5.1	
ANABELLE	1.5.2	
MOBILTELEFON	1.1	
MODEM	1.2	
PERSONSØKER	1.3	
TASTAFON	1.5.3	
TELEFAKSUTSTYR	1.4	
TELEFONAPPARAT	1.5	
TELEUTSTYR	1	
TELEKOMMUNIKASJONSUTSTYR	1	BRUK: TELEUTSTYR

Figur 6. En alfabetisert liste over termer med henvisning til plass i begrepshierarkiet.

Inngår som del av maskinelt dokumentsøk

Slike semantisk ordnete lister er særlig godt egnet som del av et system for maskinell gjenfinning, primært basert på emneordsøk. Vi kan tenke oss dette utnyttet som menyvalg, slik at en kan vandre fra et begrep til dets underordnete, sideordnete, overordnete eller assosierte begrep. På denne måten kan en utvide eller innskrenke søkeområdet, eller gå over til et semantisk beslektet område. Finner en ikke det en søker under MODEM, går en over til det assosierte emneordet DATAKOMMUNIKASJON.

Tesaurus-prosjektet ved Termbanken

Ved Termbanken ble det ved utgangen av 1986 avsluttet et prosjekt for videreutvikling av en norsk petroleumstesaurus, PETRUS. Den inneholder ca. 5500 begrep og 6500 termer. PETRUS har ti hovedhierarkier: 'egenskap', 'fenomen', 'jord/rom/tid/sted', 'organisme', 'prosess og informasjon', 'stoff/materiale', 'utstyr', 'vitenskap', 'økonomi og administrasjon', og 'generelt'. Alle begrepene er ordnet i en trestruktur under disse ti toppbegrepene. Resultatene av Termbankens arbeid skal overføres til Statoils arkiv og bibliotek.

Verktøyet Termbanken har utviklet for tesaurusbygging, er strukturert slik at det kan bygges ut til et fullstendig informasjonssøkesystem. Strukturen er nærmere omtalt i Pedersen 1986 (som har utviklet strukturen) og Utne 1986.

PETRUS-prosjektet er mer utførlig omtalt i Reinton 1986 og Reinton 1987.

LITTERATUR

Fjeldvig, Tove: *Tekstsøking – teori, metoder og systemer*. 2. utg. Universitetsforlaget. Oslo 1986.

Løland, Ståle og Toresen, Arnold: *Norsk forkortingsordbok. Norsk språkråd skrifter 14*. Oslo 1976.

Pedersen, Per-Bjørn: «Terminologiske oppdrag frå oljeindustrien til Universitetet i Bergen». Trykt i rapport fra *Symposium om datorstödd terminologi och lexicografi*. Centralen för Teknisk Terminologi. Helsinki 1986.

Simonsen, Gunnar: *Oppsummering av et fornorskningssprosjekt. Skilt og datasystemer på Gullfaks A*. I *Norske språkdata*. Nr. 12. Bergen 1986.

Simonsen, Gunnar: «En sentrifuge er en sentrifuge er en separator ???». Utkommer i *Nordisk tidsskrift for fagspråk og terminologi* 1/1987. Bergen.

Reinton, Jens: «PETRUS-prosjektet (En norsk petroleumstesaurus)» i rapport fra *ANLA-symposiet 1986: Fagspråk og oversettelse*. ANLA. Bergen 1986 (stensil).

Reinton, Jens: *PETRUS-prosjektet 1985-86. Sluttrapport for revisjonen av en petroleumstesaurus*. *Norske språkdata*. Nr. 15. Bergen 1987.

Utne, Ivar: «Computerized Lexicography at the Norwegian Term Bank». Utkommer i rapport fra konferansen *Standardization in Computerized Lexicography* (IAI, Univ. Saarbrücken og European Science Foundation). Foreligger som fotokopi. Termbanken 1986.

Utne, Ivar: «Datastøttet oversettelse av korttekster». I *Nordisk seminar om maskinoversettelse*. Foredrag udgivet af Annette Hartnack og Hanne Ruus. EUROTRA-DK, København. 1986.

Vinje, Finn-Erik: *Skriveregler*. Bokmål. 4. utg. 1984.

De omtalte listene distribueres ikke fritt.

Grunnlagsmaterialet for store deler av denne artikkelen er hentet fra flere terminologiprojekter ved Norsk termbank. De fleste terminologene på alle prosjekter arbeider med termposter og synonymi. Kari E. Haugland arbeider spesielt med sær- og sammenskriving. For øvrig vises det til litteraturlista.

Ivar Utne er amanuensis ved Avd. for norsk leksikologi, Nordisk institutt, Universitetet i Bergen.

FORFATTARENS HØVELBENK: INTELLIGENT TEKSTBEHANDLING?

Per Vestbøstad

1.1 KVA ER HØVELBENKEN?

Forfattarens Høvelbenk er eit knippe språklege analyseprogram som kan hjelpa folk å kontrollera målbruk og innhald i tekstane dei skriv. Eit vilkår for å kunna dra nytte av slike program er at teksten finst «på data» – og det gjer jo veldig mykje av det som kjem på trykk i dag.

Det er laga ein prototype av ein norskspråkleg høvelbenk i eit samarbeid mellom Avdeling for norsk leksikologi ved Nordisk institutt (UiB) og NAVFs edb-senter for humanistisk forskning i Bergen. Programmeringsarbeidet er gjort i TurboPascal av Jarle Ebeling og Anita Simon (ANL) og Espen Ore (NAVF).

Programma kan gjera følgjande på ein datalagra tekst:

Vanleg korrekturlesing

Analyserer ein del stilelement

Gjev komprimert utskrift av komposisjonen i teksten

Gjev nokre statistiske utrekningar på ord- og setningsnivå

Vurderer kva klassesteg teksten høver for

Gjev høve til å analysera ordtilfanget i teksten (konkordans og ordlister)

1.2 BAKGRUNNEN FOR DENNE PROGRAMTYPEN

Høvelbenk-programma er laga etter ide frå den amerikanske Writer's Workbench (WW), som tok form ved Bell Laboratories kring 1980 (1). Formålet med WW var å forbetra den språklege sida ved rapportar og tekniske handbøker som ingeniørane der skreiv. WW vart populært som arbeidsverktøy for tekniske forfattarar, og det vart etter kvart utbygt til å femna meir enn tredve delprogram.

Seinare vart WW prøvd ved amerikanske universitet og i high schools

i ulike variantar. Programmet synte seg å vera eigna både til å studera eksisterande tekstar og som verktøy i skriveprosessen.

1.3 KVEN KAN DRA NYTTE AV PROGRAMMA?

Utanom dei gruppene som alt er nemnde, nemleg tekniske forfattarar og studentar, ser me at denne typen programvare kan verta nyttig for ulike brukarar:

Rundskriv og forskrifter frå offentlege instansar er ikkje alltid like greie å skjønna. Kommentrarar og råd frå eit tekstanalyseverktøy ville truleg hjelpa forfattaren til å forenkla ein tung og vidløftig stil.

Ikkje alle forskarar er like vanhjelpne til å formidla forskingsproblem og -resultat, men mange kunne nok ha nytte av ei språkleg vurdering før rapportane kjem på trykk.

Journalistar arbeider ofte under eit sterkt tidspress, og burde kunna ha god hjelp av ei snøgg maskinell bearbeiding av artiklane før dei går til setjing. Institutt for Journalistikk (Fredrikstad) har såleis eit ferdig programverktøy, Analix (!), som gjer språklege analysar på artiklar skrivne i tekstbehandlaren Nortext 100.

I skuleverket har lærarar sett at Høvelbenk-liknande programvare kan vera til hjelp både i framandspråk og norsk. Dette gjeld nok for elevar flest, men særleg for dei med skrivevanskar bør slike program kunna gje betre høve til spesialtrening på aktuelle felt.

Men alle desse brukargruppene har litt ulike behov, og dei individuelle behova vil variera endå meir. Det er difor viktig at Høvelbenken er bygd opp av frittstående programmodular, og at systemet er ope for tilpassingar til ulike brukarar. Kvar brukar skal kunna velja sine egne kontrollar ut frå dei vanskane han strir med.

2. EINSKILDE PROGRAM I HØVELBENKEN

Etter som Høvelbenken ikkje skjønar tekstane, vil han berre kunna peika på *moglege problem*. Programma skal ikkje verka som ein raud blyant som finn feil i tekstane, men dei skal gje venlege råd om kva brukaren bør vurdere å omarbeida.

2.1 KORREKTURPROGRAMMET:

Her får ein kontrollert stavemåten mot ei ordliste, som førebels berre

tel 10.000 ordformer. Programmet strekar under ordformer som ikkje finst i ordlista. Samstundes vert desse orda samla i ei eiga liste, der brukaren sjølv kan merkja av dei orda han vil ha inn i hovudordlista. Programmet gjer også ein enkel teiknsetjingskontroll.

2.2 STILISTIKK:

Dette programmet peikar på uheldige formuleringar, overdriven bruk av moteord og gjentakingar både av einskilde ord og lengre uttrykk. Også her kan brukarane sjølv styra kva for ord og uttrykk programmet skal reagere på.

2.3 KOMPOSISJONEN:

Programmet lagar ein samantrengt versjon av teksten som berre inneheld overskriftene og første og siste periode i kvart avsnitt. Dette gjer det lettare å sjå korleis momenta i teksten er knytt i hop, og det kan ofte avdekkja om der er logiske sprang innom eit avsnitt.

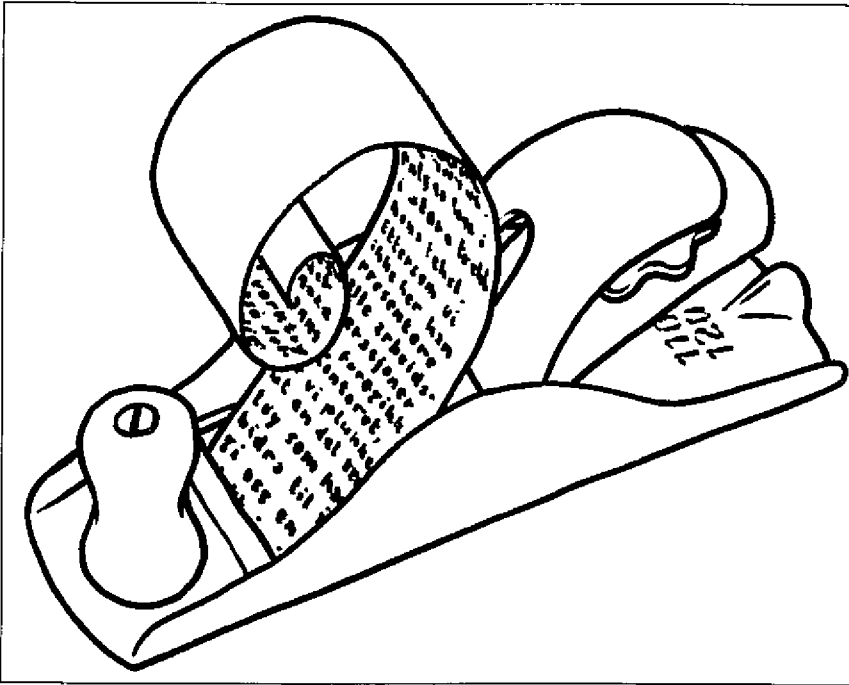
2.4 STATISTIKK:

Programmet gjev ei rekkje oppteljingar og utrekningar: Talet på ord, setningar og avsnitt, gjennomsnittslengdene, talet på lange periodar (> 40 ord) og talet på leddsetningar i kvar periode. Me får også vita kor mange upersonlege subjekt teksten har, og me får skrive ut alle ord som har meir enn 15 bokstavar. Programmet reknar ut eit LIX-tal, som gjev eit mål for kor lettlesen teksten er. I skulesamanheng kan LIX-talet gje ein peikepinn på kva klassesteg teksten høver for.

2.5 ORDLISTE- OG KONKORDANSPROGRAMMET:

Ordlistene og konkordansen er særleg berekna på bruk i undervisning der ein studerer tekstar språkleg og litterært. Me kan her få alfabetisk ordliste, frekvensordliste sortert etter fallande eller stigande frekvens og baklengsordliste (= finalalfabetisert liste). Den siste samlar alle ord med like endingar.

Konkordansen er ei alfabetisk ordliste, der ordet står oppført med kon-



Tegning: Øystein Reigem.

tekst – ein gong for kvar gong det er brukt i teksten. Slike lister er nyttige om ein vil studera bruken av einskilde ord og uttrykk t.d. i ein litterær tekst.

3. VURDERINGAR

3.1 FÅRE FOR EINSRETTING?

Ei viktig innvending er fåren for einsretting av skriftspråket dersom normerande programvare av denne typen vert teken i bruk, og mange vil meina at denne fåren vil vera spesielt stor i skulen.

Dersom Høvelbenken var laga ut frå ein filosofi om at der er fasitsvar heile vegen i skriveprosessen, ville desse innvendingane vera aktuelle. Men desse programma har ein heilt annan innfallsvinkel i det dei berre freistar å fanga opp språktrekk som velrøynde folk veit gjev lite effektiv

kommunikasjon. Her er altså lite aktiv normering. For det meste får brukaren varsame oppmodingar om å vurdera einiskilde ting på nytt. Det er sjølvstøtt viktig at toleransegrensene er vide og tilpassa brukargruppe/klassesteg. Erfarne pedagogar må inn i arbeidet med å leggja dette til rette.

3.2 UPERSONLEG UNDERVISNING?

I skulen vil ikkje denne typen programvare vera meint som erstatning for språklæraren, men som eit utvida tilbod der elevane kan få arbeida individuelt med ting dei treng øva ekstra på. Eleven treng ikkje køyra alle kontrollane, men kan velja ut dei mest aktuelle. Det ein tapar ved at dataprogram er upersonlege, kan ein vinna ved større høve til individualisering.

3.3 RØYNSLER FRÅ AMERIKANSKE UNIVERSITET:

Studentar ved Colorado State University (2) meinte at Writer's Workbench (WW) gav meir konkret og konsekvent tilbakemelding enn lærarane makt. Kritikken dei fekk var meir presis og gav bearbeidingsfasen meir konkret innhald. Denne fasen vart ikkje lenger berre mental skuggeboksing med eigne fantasilause formuleringar.

Lærarane trong på si side ikkje bruka så mykje tid til terping på elementære feil og fekk såleis tid til meir djuptgåande undervisning. Resultatet vart betre skriftlege arbeid frå studentane.

Lærarar ved University of California at Santa Cruz rapporterer at kommentarane frå programmet skjerpar sansen for grammatiske fenomen og stilistiske verkemiddel og innbyr til å eksperimentera med ulike uttrykksmåtar, fordi studentane får fleire alternativ «på bordet».

3.4 WW I ARBEIDSLIVET:

Folk som har brukt WW i jobben har lagt vekt på at programmet finn veike punkt dei sjølv ville ha oversett, at dei skriv betre med hjelp frå WW, og at dei har vorte meir medvitne om skriveprosessen. Eit lite moment, som kan vera viktig for mange, er at det er lettare å ta i mot kritikk frå eit edb-program enn frå sjefar og kollegar.

4. VIDARE PLANAR FOR EIN NORSK-SPRÅKLEG HØVELBENK

Ut frå vurderingane i kap. 3 skulle det vera gode grunnar for å satsa på ein norsk versjon av denne typen programvare.

Ved NAVFs edb-senter for humanistisk forskning planlegg vi eit seminar i september for å samla alle miljøa som har arbeidd med Høvelbenk-liknande programvare for norsk. Målet må vera å få til ein samla innsats for å få fram eit godt verktøy til hjelp for dei som skriv norsk – med variantar for kvar målform, der bokmål og nynorsk krev ulike løysingar.

Ut over dei funksjonane som er skisserte i demo-versjonen (kap.2) bør standardversjonen også kunna vurdere og kommentera:

1. frekvensen av passivsetningar og verbalsubstantiv
2. kompleksiteten i periodane (for mange undersetningar gjev kronglete språk, medan for få kan gje for enkel framstilling)
3. bruken av framandord/vanskelege ord/abstrakte ord

Høvelbenken bør også kunna gå gjennom eit sett med mønstertekstar slik at toleransegrensene for ein del av kontrollane kan tilpassast ulike teksttypar.

Det finst etter kvart fleire program i denne gata, som kan gje idéar til andre kontrollar og hjelpefunksjonar for skrivande folk. Dei såkalla idébehandlarane («idea processors»), som gjer det ekstra lett å omstrukturera kapitel og underavsnitt i ein tekst, er her særleg interessante.

NOTER

- (1) Cherry and Fox etc.: Computer aids for text analysis. Artikkel i «Bell Laboratories Record», May/June 1983
- (2) Charles R. Smith and Kathleen Kiefer: Using the Writer's Workbench Programs at Colorado State University. Artikkel i «International Conference on Computers and the Humanities» (ISBN 0-914894-96-X).

EDB I SPRÅKFAG

Per M. Mathisen

INNLEDNING

Adolf Øiens skole er forsøksskole under KUD/Datasekretariatets handlingsplan for utprøving av edb i undervisningen i grunnskole og videregående skole. Vi har vært i gang siden høsten 1985, men vi fikk maskin- og programvare først tidlig i november samme år, slik at vi har vært i full virksomhet i vel ett år. Undertegnede er ansvarlig for forsøksvirksomheten innenfor språkfagene og psykologi. Så langt har vi en del praktiske erfaringer å bygge på, men like viktig er kanskje de refleksjonene og erfaringene som har meldt seg i møtet mellom ideer og virkelighet. Disse vil bli forsøkt presentert nedenfor.

BAKGRUNN

Begrunnelsen for å ta inn edb i skolen kan være mangfoldig, men et aspekt ved dette vil jeg nevne her. Det synes å eksistere to ytterpunkter med hensyn til vilkår for å bruke edb i undervisningen:

1. Det en datamaskin kan brukes til skal ikke en lærer gjøre.
2. Det en datamaskin ikke kan gjøre bedre, skal en bruke lærere til.

Mitt inntrykk er at man f.eks. i USA synes å helle i retning av alternativ 1, mens man i Norge nokså klart tenderer mot 2, altså at hensikten med å bruke datamaskiner først og fremst er å gjøre undervisningen bedre og/eller mer effektiv.

Valget av alternativ 2 stiller egentlig store krav til datamaskin og programvare. Situasjonen er vel også i dag at det generelt sett er langt mellom de virkelig overbevisende pedagogiske anvendelsene av datamaskinen i undervisningen – men de finnes. Innen språkfagene er denne tendensen enda klarere: det finnes få program på markedet, og de fleste er innsiktet på lavere nivå i skoleverket, det vil ofte si enkle drilløvelser. Årsaken til at det finnes lite programvare for språkfagene *kan* være at det rett og slett ikke finnes fornuftige anvendelser her, men det kan også meget vel være at språkfagene står perifert i forhold til edb-utviklingen,

blant annet på grunn av sein kompetanseoppbygging hos språklærerne og liten forståelse blant edb-eksperterne for språkfagenes behov.

PROGRAMVARE

Det har vært vanlig å skille mellom undervisningsprogram og verktøyprogram til bruk i skolen. I vår situasjon har vi noen få brukbare anvendelser med hensyn til undervisningsprogram. Disse har gjerne effekt like mye i kraft av et middel til variasjon, som i kraft av pedagogisk velegnethet. (Denne situasjonen kan selvsagt endre seg, jeg vet om interessante ting som er under utvikling, men det norske markedet er trolig for lite til at vi kan håpe på noe særlig stort tilbud av undervisningsprogram.)

Med hensyn til verktøyprogrammene er det tekstbehandlingsprogram som er mest aktuelle for språkfaga. Det er vel også her man har gjort seg de største forhåpningene med hensyn til anvendelse av datamaskinen, og mange fagre vyer har vært presentert i ulike fagfora. Imidlertid er det svært få som har konkrete erfaringer å vise til. Det er egentlig overraskende, når så mange har vært så opptatt av ideen.

Skal man forsøke å finne årsaker til de mangelfulle erfaringene, kan man tenke seg følgende: 1) tekstbehandling krever minst én maskin pr. 2 elever, og svært få skoler har en slik maskinpark, 2) elevene må beherske tekstbehandlingsprogrammet og tastaturet for at utstyret skal kunne brukes på en meningsfylt måte. Dette siste er av *fundamental* betydning. Å la elever som har store problemer med å finne bokstavene på tastaturet skrive f.eks. stil på datamaskin, vil ta en evighet, og selv om man foretar mange rettinger og revisjoner av manuskriptet, vil enhver rasjonaliserings- eller tidsgevinst være helt illusorisk. Det samme kan gjøres med papir og blyant på mye kortere tid.

Hvor fører så dette oss – finnes det ingen fornuftige anvendelser for tekstbehandlingsutstyr? Vi har etter min mening to mulige veier å gå for å kunne utnytte tekstbehandling i språkfagene:

1. Dersom elevene *ikke* har funksjonell ferdighet på tastaturet, må en se bort ifra at elevene produserer egne tekster. Man må basere seg på skriving av enkeltord og bruk av pil- og funksjonstaster. Dette forutsetter at man har ferdig innskrevet tekst som utgangspunkt. Her tenker jeg meg at man kan bruke tekster fra pensum i språkfaga (eller med tilknytning til pensum), og oppgaven kan være av følgende type:

- 1) Skriv teksten om fra presens til preteritum, eller omvendt. (Øving i bruk av tempus – ofte inkonsekvens hos elevene i egne tekster. Hva skjer med teksten ved omskrivingen?)

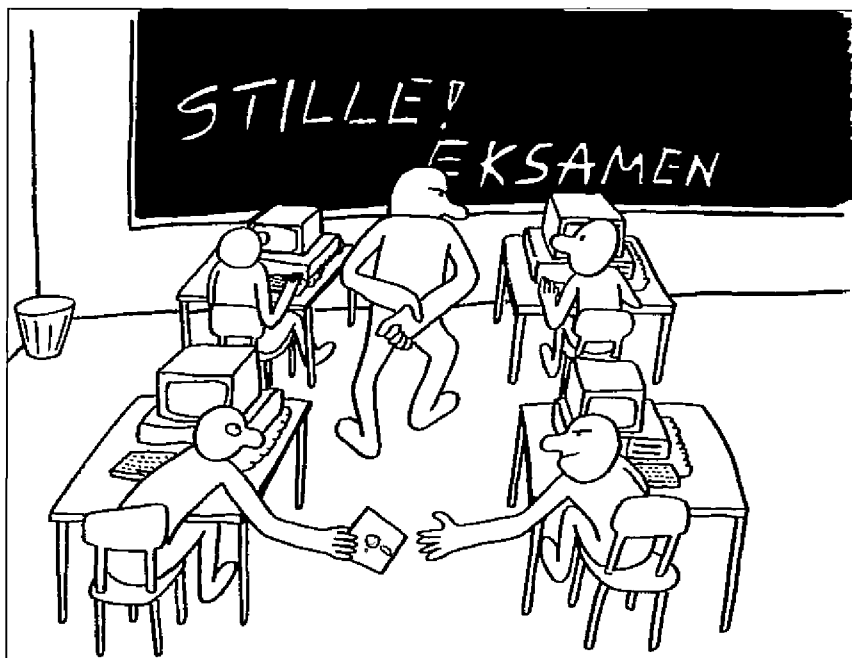
- 2) Skriv om fra muntlig stil til mer formell stil. (Et problem i for eksempel engelsk er at uformelle uttrykk/slang brukes ukritisk.)
- 3) Skriv om fra «fattig» og ensformig språk til rikt og variert språk, eller fra innfløkt språk til dagligdags språk. (Ordvalg/idiomatikk bør arbeides med i alle sammenhenger.)
- 4) Skriv om tekst med språkfeil til korrekt språk.
- 5) Klassiske eller modernistiske dikt kan presenteres som prosa, og elevene får i oppgave å gi teksten en plausibel lyrisk form, som så kan sammenlignes med originalen. (Kan gi økt innsikt i betydningen av rim, rytme, strofemønster, og lignende.)
- 6) Kvantitativ analyse av innskrevet tekst. (Setningslengde, lix, ordfrekvenser og så videre – kan være relevant for språkbruksanalyse/stilistikk.)

Listen med eksempel kan lett gjøres lengre, og tilpasses det enkelte språk og kurs. Teknikken baserer seg på at 2 elever arbeider sammen og diskuterer seg fram til «riktige» løsninger – altså brukes mer tid til vurdering og lesing enn til skrivning. Det endelige resultatet vil bli ryddig og pent, og vil i mange tilfeller kunne tas inn i et evalueringsprogram, for å gi tilbakemelding til elevene – straks og uten belastning for læreren. En annen mulighet er at man kan legge inn hjelpetekster som elevene får se dersom de gjør noe feil (f.eks. ord, uttrykk eller grammatikkregler).

Det kreative aspektet ved skrivningen må man i stor grad gi slipp på, selv om det godt kan tenkes framgangsmåter som gir mer av dette enn eksemplene ovenfor viser (for eksempel å fjerne og/eller skrive til avsnitt i en ferdig tekst, eller endre rekkefølgen på avsnitt eller innholdsmomenter, osv.).

2. Elevene *har* funksjonell ferdighet på tastatur. Dette forutsetter at elevene får eller har fått opplæring i maskinskriving eller tekstbehandling etter touch-metoden. I en slik situasjon kan det være aktuelt å se nærmere på hvilke fordeler tekstbehandling gir i forhold til vanlig tekstproduksjon med papir og blyant/penn.

Den største fordelen med maskinell tekstbehandling er selvsagt at man slipper å skrive om hele teksten når større endringer skal foretas, og man slipper å føre inn den ferdige kladden. Dette gjør det – blir det hevdet – mere meningsfylt og lystbetont å bruke tid på å forbedre innholdet. Reduksjon av skrivehemninger har også vært nevnt som et positivt argument, og som et nytt, spennende medium skaper det motivasjon hos elevene (så lenge det varer!). Dessuten blir alle utskrifter ryddige og lett-leste (fordel også for læreren), og det er mulig å oppnå full anonymitet – om det skulle være ønskelig. Endelig lærer elevene en arbeidsmåte som ser ut til å bli framtidens måte å produsere tekst på.



Tegning: Øystein Reigem.

De betraktninger som er gjort ovenfor baserer seg dels på andres vurderinger, dels på egne vurderinger, dels på egne *erfaringer* med tekstbehandling. Erfaringene er så langt så gode at vi har valgt å satse på tekstbehandling som et hovedprosjekt ved vår skole. Dette prosjektet vil jeg si litt mer om:

PROSESSORIENTERT SKRIVEOPPLÆRING OG EDB

I den senere tid har en «ny» form for opplæring i norsk skriftlig vunnet innpass i norske skoler. Denne innebærer at lærerne går inn i selve skriveprosessen og gir opplæringa der, i stedet for at kommentarer og retting gis etter at det ferdige produkt er levert inn. Et vesentlig trekk ved metoden er at eleven bearbejder teksten i flere omganger, før den endelige versjonen leveres inn. I en slik sammenheng skulle tekstbehandling være særlig velegnet. Vi har derfor opprettet en førsteklasse ved skolen, der alle elevene har valgt maskinskriving med tekstbehandling som valgfag.

Disse elevene vil få skriveopplæring etter de nye prinsippene, utnyttet i norsk og engelsk. Integreringen av disse to teknikkene vil skje etter hvert som tastaturferdigheten blir god nok.

Vi har også drevet forsøk med de andre mulighetene som er nevnt foran, dvs. å prøve ut tilgjengelige undervisningsprogram og å gjøre forsøk med oppgaver til ferdig innskrevet tekst. Det meste av dette vil jeg unnlate å komme inn på her.

LITT OM «DRILL-OG-ØVELSESPROGRAM»

Av de øvrige tingene vi arbeider med, vil jeg særlig nevne øvingsprogrammet *DRILL*, som en av edb-lærerne ved skolen har utviklet i samarbeid med språklærerne. Dette er et «forfatterprogram», som går ut på at elevene skal fylle ut «luker» i en tekst som læreren har preparert på forhånd. Programmet kan brukes i ulike språk og på mange ulike måter, og det kan legges inn svært varierte responser til eleven. Det verdifulle ved slike øvinger er at elevene får respons/hjelp med en gang når de gjør feil, og at de slipper forklaringer på ting de mestrer. Språklærerne har vært udelt positive til dette programmet, hvilket skulle antyde at «drill og øvelse» ikke nødvendigvis trenger å være skjellsord innenfor edb-metodikken.

KONKLUSJON

Det er i dag for tidlig å formulere noen endelige konklusjoner om hvorvidt bruk av datamaskiner egentlig er hensiktsmessig med det utstyret vi disponerer i dag; slike konklusjoner må utstå til forsøksperioden er over, høsten 1987. Men en viktig forutsetning vil jeg tillate meg å gjenta:

Skal effektiv utnyttelse av datamaskinen kunne oppnås (innen språkfagene, men trolig også ellers), er det en absolutt forutsetning at elevene har et minimum av tastaturferdighet. Så viktig og vanlig som datamaskinen ser ut til å bli i framtida, må det være en viktig oppgave for skolen å lære elevene å utnytte maskinen mest mulig effektivt. Det mest grunnleggende her er ikke å ha innsikt i programmering eller maskinens virkemåte, men faktisk å beherske tastaturet.

Per M. Mathisen er lektor ved Adolf Øiens skole.

ERFARINGER FRA SKRIVING AV INTERAKTIVE LINJEORIENTERTE UNDERVISNINGSPROGRAMMER I SPRÅK

Signe Marie Sanne

LINJEORIENTERT PROGRAM

Jeg har i en del måneder vært i gang med programmering (i FORTRAN) av språklige øvelser for våre italienskstudenter. Siden det virker som om ikke særlig mange på universitetshold har kastet seg ut i dette ennå, kunne jeg ha lyst til å fortelle om de erfaringene jeg som nybegynner i faget «programmering» har høstet. Jeg har hele tiden arbeidet på Sperry UNIVAC-anlegget ved Universitetet i Bergen, og det dreier seg om interaktiv kjøring av linjeorienterte programmer. Linjene (dvs. de som er programmert inn og som skrives ut på skjermen til studenten etter hvert) vil «rulle» over skjermen inntil studenten skal svare, og det innebærer at han alltid må skrive inn sitt svar nederst på skjermen hvor skrivehake og markør automatisk stiller seg. Studenten kan altså *ikke* bevege seg med markøren rundt på skjermen og skrive svaret rett inn i den setningen hvor de(t) manglende ord er erstattet med f.eks. prikker eller tomrom. Men dette vil ikke si at svaret aldri vil opptre i sin naturlige kontekst, for ved hjelp av *konkatenering* (= sammenkjeding) av setning og svar sørger vi for at studenten i neste øyeblikk har en «prikkfri» og korrekt setning foran seg på skjermen. Det vi samtidig sørger for, er at hvis alle svar avgis med *små* bokstaver (hvilket er mest praktisk, da studenten slipper å bruke SHIFT-tasten) og det manglende ordet opptrer først i setningen, har vi lagt inn en *konverteringsprosedyre*, slik at den ferdige setningen får stor forbokstav. Hvis vi i programmet har lagt inn våre data med små bokstaver, og studenten mot formodning svarer med store, eller bruker stor forbokstav, tar en ny konvertering seg av å først gjøre om de store bokstavene til små før testingen av svaret mot dataene starter.

REDAKSJONELLE PROBLEMER

Siden setningene «ruller» over skjermen, er det viktig at oppsettet av setningene på skjermen er klarest mulig og minst mulig trettende å se på for studenten. Av den grunn er det en del momenter jeg har prøvd å vurdere ved skrivingen av programmene, men det er først ved utprøving av forskjellige former redigering at disse spørsmålene kan besvares:

Skal setningene/instruksene være med små/store bokstaver?

Hvor stor avstand bør det være mellom linjene?

Kan setningen strekke seg over hele skjermen, eller bør bare venstre side, evt. midtpartiet benyttes?

Er det svært uheldig at mange setninger «løper» over skjermen før teksten står stille? Skal vi lage mange kunstige stopp hvor studenten får tid til å lese relativt få linjer ad gangen, før han trykker en tast for å lese videre?

Må setningen/uttrykket gjentas flere ganger hvis det opprinnelige spørsmålet/setningen er rullet ut av skjermbildet?

MASKINTEKNISKE PROBLEMER

Aksenter og apostrofer er på våre moderne maskiner ofte et problem. På UTS-terminalene som studentene foreløpig arbeider ved, finnes det bare to taster med tegn som kan minne om ovennevnte. Apostrofen heller i samme retning som lukket aksent, slik at jeg har valgt å gi studentene beskjed om å bruke apostrof både for apostrof og lukket aksent. På skjermen ser resultatet da brukbart ut (men ikke på en papirutskrift). En annen ulempe i denne forbindelse er at maskinen ikke godtar to tegn i samme posisjon, slik at når vi på italiensk skal ha små eller store bokstaver med åpen eller lukket aksent over, må vi på Sperry skrive aksent i neste posisjon, altså slik: a^{\cdot} eller A^{\cdot} .

På italiensk finnes en del forekomster av to ord skrevet fortløpende med apostrof imellom, f.eks. *all'ora*. Programmeringsspråket jeg har benyttet er FORTRAN77, og her skrives konstanter mellom to apostrofer. Hvis vi da har en øvelse som går ut på å fylle ut med preposisjon + bestemt artikkel, er det to svaralternativer, enten en form som skal forbindes med substantivet med en apostrof imellom (som ovenfor), eller en frittstående form med mellomrom foran substantivet, f.eks. *al bar*. Hvis vi nå skal foreta en test for å se om svaret ender med apostrof, får vi problemer. Når vi søker etter f.eks. en *a*, gjøres dette slik:

IF (SVAR(3:3).EQ.'a'),

men hvis vi søker etter en apostrof, vil det se slik ut:

IF (SVAR(3:3).EQ."')

Maskinen prøver å tolke det som står mellom de to første apostrofene, og der står det jo ingenting. Løsningen på dette problemet er forsåvidt enkel, idet man må huske å skrive inn apostrofen i datalinjene, og så lese den inn i en variabel, som vi kan kalle APOS. Da kan vi sette opp testen: IF(SVAR(3:3).EQ.APOS), som er en normal test.

Hvis vi ikke sitter ved en UTS-terminal, men tvertimot arbeider ved en PC som er tilkoblet Sperry via lokalnettet vårt, dukker nye problemer opp: Instruksene er skrevet på norsk, og alle norske bokstaver kommer nydelig ut på skjermen på UTS-terminalene (og på en eventuell papirutskrift), men forflytter vi oss til en PC og oppkaller det samme programmet via lokalnettet, vil vi på skjermen få rare klammer og streker istedenfor de norske bokstavene æ, ø og å. Når studentene en gang i fremtiden (forhåpentligvis i en nær) får benytte seg av en PC-lab som er knyttet til lokalnettet, er dette en ting som virker sjenerende. Så det er å håpe at en av universitetets edb-eksperter i mellomtiden har kommet med et program som omformer de rare tegnene på PC'en til de norske bokstavene.

TYPER PROGRAMMER

Jeg har foreløpig holdt meg til programmering av to hovedtyper øvelser:

- 1) En øvelsestype som går ut på å oppgi en norsk setning som skal oversettes til italiensk, og så en italiensk setning hvor manglende ordform(er) er erstattet av prikker, altså en slags utfyllingsoppgave av riktige grammatiske former.

Etter at studenten har svart og har sett den riktige setningen på skjermen, kan vi «blanke» ut skjermen og gi dem den norske setningen å oversette (denne gang uten noen oppgitte italienske ord).

- 2) En øvelsestype hvor studentene får oppgitt italienske setninger/uttrykk (+ evt. norsk oversettelse), og hvor de skal foreta en grammatisk analyse. Dette kan f.eks. være å vurdere et uttrykk og fortelle hva slags syntagme det er og hva slags grammatisk funksjon det har i setningen, eller vurdere en adverbial bisetning og fortelle hva slags type det dreier seg om.

Jeg har også laget en del øvelser som er en kombinasjon av disse to

NESTE EKSEMPEL:

Ero talmente stanca *che non mi reggevo in piedi*.

1) HVA SLAGS SETNING ER HER ISOLERT MELLOM STJERNENE?

tid	= TIDSetning	års	= ÅRSAKSetning
ben	= HENSIKTsetning	fel	= FØLGEsetning
inn	= INNROMMELSESetning	sam	= SAMMENLIGNINGSetning
bet	= BETINGELSESetning	adv	= ADVERSATIV setning

SVAR MED EN AV FØRORTELSENE:

nam
GALT – Setningen er en FØLGESETNING.

2) HVA ER KONJUNKSIONALET?

che
RIKTIG – KONJUNKSIONALET I SETNINGEN ER "che".

3) HVILKET LEDD I HOVEDSETNINGEN ER OGSÅ AV BETYDNING I SETNINGEN:

Ero talmente stanca *che non mi reggevo in piedi*.

talmente

RIKTIG

"talmente" og "che" viser at det er en FØLGESETNING.

Vil du ha mer informasjon om denne type setning? [j/n]

j

FØLGESETNINGER

FØLGEkonjunksjoner er:

{[ante/'cosi'/'tale'/talmente...]} CHE
IN/DI MODO CHE
{[CO]SICCHE'
-["'stlik[e]/s'd'sdan[ne]/s'dno...]} at'

{[troppo...]} PERCHE' = {[ah]for meget/nye...]} TIL AT

Modus: Hvis følgen fremsettes som en ØNSKET følge, skal det være KONJUNKTIV, hvis følgen har funnet sted INDIKATIV. {[troppo...]} PERCHE' skal ha KONJUNKTIV.

Trykk XMIT for å gå videre!

ETTER 5 SETNINGER HAR DU SVART RIKTIG PÅ 7 AV 11 SPØRSMÅL.

VIL DU FORTSETTE ØVELSEN [j/n]

n

TAKK FOR I DAG

Papiruskrift av en edb-øvelse i italiensk språk.

typer: Først utfylling med én av et visst antall oppgitte beslektede former, f.eks. CHE, CHI eller CUI, som alle kan tilhøre to forskjellige ordklasser og innlede to forskjellige typer bisetninger, og deretter en grammatisk analyse. Oppgaven blir da tredelt: Først fyller ut med riktig form, så fortelle hvilken ordklasse den tilhører, og så hva slags type bisetning den innleder.

ØVELSER PÅ FORSKJELLIGE NIVÅER

Jeg mener at det er fullt mulig å lage øvelser som egner seg både for begynnere og for viderekomne. Det er bare opp til den som lager øvelsen å ha fantasi nok til å lage forskjellige interessante typer. Den eneste begrensningen i denne type programmer er at spørsmålene må være entydige. Dette vil ikke si at det bare er ett svar til hvert spørsmål, for det er fullt mulig å legge inn flere alternative svar, men siden maskinen ikke er intelligent, kan den ikke gi seg til å analysere setningen, den bare tester det avgitte svaret mot svaralternativer vi allerede har lagt inn i programmet.

I tillegg er det fullt mulig å legge inn så mange kommentarer vi bare ønsker, og dette kan være en nyttig repetisjon av regler på ethvert nivå.

ALTERNATIVE SVAR

Hvis to eller flere alternative svar kan være riktig i setningen, legger man i datalinjene inn opplysninger om at det er flere mulige løsninger. Hvis studenten har svart med et av alternativene, svarer maskinen RIKTIG, og de alternative setningene oppgis da som en tilleggsopplysning. I en av øvelsene skal studentene f.eks. oversette med riktig endelse på partisippet: «Er dere kommet alt?». Siden «dere» kan være enten av hankjønn eller av blandet kjønn som skal ha hankjønnsbøyning av partisippet, eller bare bestå av kvinner, som skal ha hunkjønnsbøyning, må dette legges inn som alternative svarmuligheter.

ØYEBLIKKELIG FEED-BACK

I motsetning til ved skriftlige øvelser, hvor studenten først arbeider med øvelsen, kanskje leverer den inn noen dager etter, og får den korrigerte øvelsen tilbake en uke senere, får han ved maskinen øyeblikkelig

reaksjon på om han har svart riktig eller galt.

POENGBEREGNING

I programmene mine har jeg lagt inn at maskinen skal oppgi hvor mange riktige svar studenten har avgitt etter hver 5. setning. Dette er et helt vilkårlig tall, jeg har valgt ikke å gi det oftere for at studenten ikke skal være for opphengt i poengtall, men heller ikke sjeldnere, for at det skal kunne virke som en spore for studenten til å skjerpe seg.

FEIL AVGITT SVAR

I de øvelser jeg hittil har laget, har jeg ikke lagt inn en ny mulighet for studenten til å svare, hvis han har svart feil første gang. Etter avgitt svar sier maskinen GALT, og oppgir den riktige setningen straks. Jeg har vurdert som så at en slik prosedyre kan bidra til at studenten skjerper seg neste gang han gjør øvelsen, nettopp for å forbedre sitt poengtall. En annen mulighet jeg foreløpig ikke har inkorporert, er å legge inn alle spørsmål med galt svar i et nytt felt, og så når øvelsen er over, gi studenten anledning til å prøve seg på nytt på de feilbesvarte spørsmålene.

FORDEL/ULEMPE VED EDB-ØVELSER

Som ved skriftlige øvelser er det muligens en fare for at studenten etter at svaret er avgitt, hopper for fort over til neste eksempel. Hvis han desuten synes det er spennende med poengberegning (hvilket de fleste synes), kan dette kanskje medføre at han forter seg gjennom øvelsen. Men den store fordelen edb-øvelser her har, er at studenten har anledning til å bruke all den tid han bare ønsker til å studere inngående den korrigerte/riktige setningen. I en del øvelser har jeg lagt inn en litt fyldig grammatisk kommentar i tillegg, og som regel har jeg også norsk oversettelse av det italienske eksemplet, slik at det er mulig å studere både ordforråd og grammatikk samtidig. Og hvis studenten første gang har gått raskt gjennom øvelsen, er det temmelig sikkert at han neste gang vil konsentrere seg mere for å oppnå høyere poengtall.

NYTTIG BIEFFEKT

Maskinen er det mest «firkantete» vesen som eksisterer (i dobbelt forstand). Den tolker ingenting i beste mening, så hvis studenten ikke svarer med 100% nøyaktighet, blir svaret automatisk GALT. Dette gjelder f. eks. hvis studenten gjør en skrivefeil, eller glemmer en aksent eller en apostrof, da får han meldingen GALT. Altfor ofte tror studentene at de har nøyaktige kunnskaper om noe, men det vil maskinen raskt korrigere, kanskje med noe frustrerende virkning på studentene, men på lengre sikt er det en gevinst å hente at studentene blir tvunget til å være mer presise.

Signe Marie Sanne er universitetslektor ved Romansk institutt, Avd. B - Italiensk, Universitetet i Bergen.

SENTERETS VIDEOPLATEPROSJEKT

SLUTTRAPPORT

*Roger Erlandsen, Claus Huitfeldt og
Øystein Reigem*

Følgende artikkel består av utdrag fra *Videoplateteknologi i humanistiske fag*, som skal utgis i rapportserien til NAVFs edb-senter for humanistisk forskning. I tillegg til en utdypning av aspektene som er omtalt her, skal hovedrapporten inneholde kapitler om administrasjon og organisering, tilrettelegging av bildemateriale på videoplate, presentasjoner av prosjektet og økonomi (i alt 33 sider), samt åtte vedlegg som består av bl.a. referat fra prosjektets avsluttende seminar. (Red.)

BAKGRUNN

Mange humanistiske institusjoner ser et behov for kombinasjonen bil-
delagring og edb. Noen har databaser med søkbare opplysninger om
objekter (gjenstander, fotos, diverse arkivalia), men ingen mulighet for
en besiktigelse uten å gå til manuelle kataloger med eventuelle avbild-
ninger eller til magasinene hvor objektene står. Optiske lagringsmedier
gjør det mulig å edb-lagre «nye» typer informasjon som bilder, levende
film og lyd. Det er derfor et sterkt ønske i mange humanistiske fag- og
forskningsmiljøer om en nærmere kartlegging av de optiske lagringsme-
diens egnethet for lagring av humanistisk kildemateriale, og mulighete-
ne for kobling mellom edb og bildemateriale.

NAVFs edb-senter for humanistisk forskning (EDBH) og Sekretariatet
for fotoregistrering (SFFR) fant i lys av ovenstående tiden moden for en
innledende utprøving av optisk platteteknologi i humanistisk regi. Av for-
skjellige grunner (se nedenfor) fant en at videoplate var det mest egnede
medium for slik utprøving. En slik utprøving burde innebære en utprø-
ving av platens gjengivelses kvalitet. Dette forutsatte produksjon av en
egen videoplate. Produksjon av en egen videoplate var dessuten ønskelig
fordi det ikke fantes ferdig produserte plater med materiale av den type

det her var tale om, og fordi man gjennom produksjonsprosessen ville komme nærmere teknologien.

På grunn av de store innbyrdes forskjeller i arbeidsmåte og materialtype mellom institusjonene, og på grunn av de betydelige ressurser som kreves for videoplateproduksjon, anså man et samarbeid mellom flere humanistiske fag- og forskningsinstitusjoner både som det mest hensiktsmessige og som det eneste realistiske.

MÅLSETTING

Videoplateprosjektets målsetting har vært:

- å tilrettelegge et allsidig og representativt bildemateriale fra humanistiske fag- og forskningsinstitusjoner, eksperimentere med forskjellige opptaksteknikker etc., og legge dette materialet inn på en videoplate.
- å eksperimentere med edb-basert kobling av bildematerialet og referansedata. Dette innebærer tilrettelegging av referansedata på edb og utarbeiding av programvare. Hovedvekten har ligget på bildedatabaser og gjenfinning.
- å bygge opp kompetanse på feltet i de deltagende miljøer og å følge med i utviklingen av og spre kunnskap om optiske lagringsmedia, for dermed å sette de humanistiske miljøer i stand til å forberede en tidlig utnyttning av teknologien.

Prosjektet er altså et prøveprosjekt og har ikke hatt som målsetting å resultere i et umiddelbart anvendelig produkt.

FORBEREDENDE ARBEID

Valg av teknologi

EDBH og SFFR foretok i 1984 en innledende kartlegging av situasjonen på feltet.

Det finnes i dag en rekke forskjellige optiske medier, både analoge (dvs. videoplater) og digitale. På det tidspunktet prosjektet ble startet, var det bare én teknologi som var aktuell, nemlig LaserVision videoplate. Andre medier var enten ikke modne for ordinær bruk eller ville være økonomisk uoverkommelige.

Andre fordeler med å velge LaserVision videoplate var at utstyret er relativt rimelig, at en benytter seg av tildels kjent teknologi (TV), og at

utstyret også kan brukes uten datamaskin. Alle disse argumentene er viktige i humanistiske miljøer, hvor det som regel er økonomiske begrensninger og manglende ressurser til innkjøp av utstyr og satsing på ny teknologi.

LaserVision videoplate er en read-only-plate, dvs. at informasjonen på platen ikke kan endres. Alt materialet må legges tilrette på forhånd, og det presses et antall kopier av platen.

Bildekvaliteten er som godt TV, men på grunn av den begrensede oppløsningen et videosignal har, vil visse typer informasjon være lite egnet for videoplate (skrift, detaljerte tegninger).

Uten spesielt utstyr kan et videobilde ikke manipuleres videre i datamaskinen. F.eks. vil stående bilder fotografert liggende ikke kunne snus. Det er ikke mulig å ha bilder som er større enn skjermen. Det er ikke mulig å zoome inn på detaljer. Zooming og panorering *kan* simuleres ved at en fotograferer hvert objekt flere ganger.

Det finnes teknikker for også å legge inn digital informasjon (store mengder datalagret tekst, programmer og lyd) på videoplater. Heller ikke dette ble ansett for å være moden nok teknologi eller økonomisk forsvarlig. Når det gjelder referansedata taler dessuten mange grunner for å ha disse separat på vanlige magnetplater (disk, diskett).

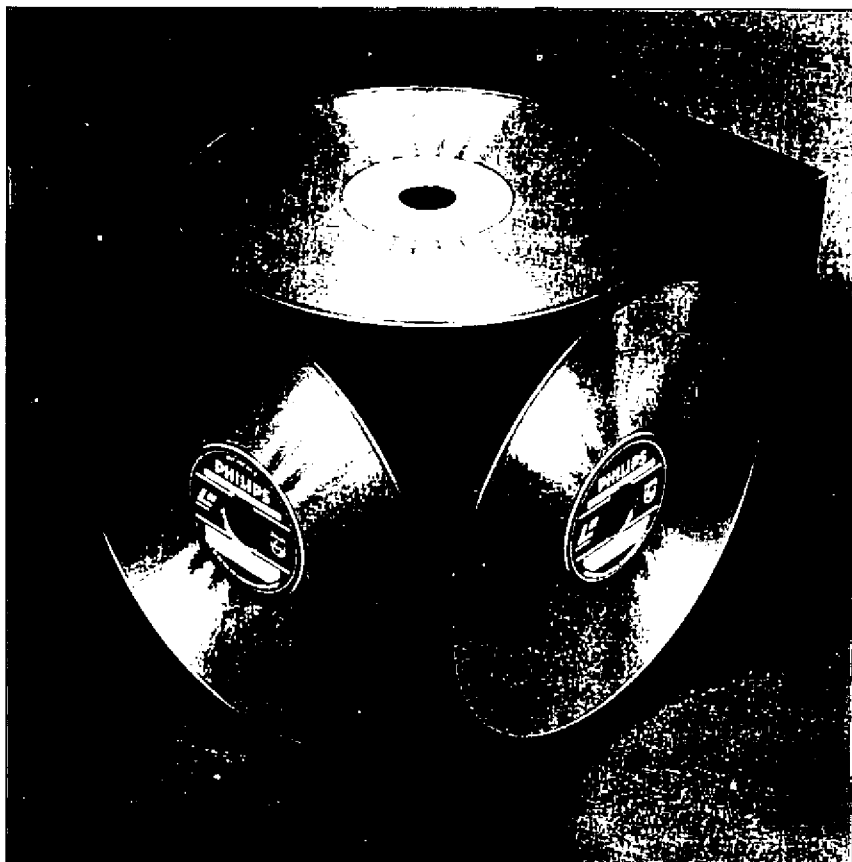
I mange anvendelser av videoplateteknologi ønsker man å legge tekst og grafikk fra PC-en over bildene fra videoplaten. Slikt utstyr finnes, men heller ikke dette ble kjøpt inn i prosjektet. Dette dels pga. at en løsning med bilde og referansedata på hver sin skjerm ble sett på som fornuftig for anvendelsene i prosjektet, dels pga. økonomien og dels pga. dårlig utvalg av slikt utstyr da spilleren ble kjøpt. For likevel å ha en mulighet til å legge tekst på videobildet, ble det valgt en videoplatespiller med teletekst. Ved å sende tekst kodet i et spesielt format til spilleren, kan en få tekst-TV-bokstaver (og enkel grafikk) på videobildet.

Valg av produksjonsmetode

De fleste aktuelle institusjoner har dårlig økonomi. Det var nødvendig å velge produksjonsmetode ut fra dette. En bestemte seg for å ta inn stillbilder via 16mm film, for senere å «scanne» filmen over på 1" videobånd (som er den formen mastering-fabrikk krever). Innsamling og avfotografering av materiale måtte i stor grad utføres av personale ved de deltagende institusjonene.

16mm-utstyr lot seg leie rimelig og var lett å betjene. De fleste sakkyn-dige mente at 16mm ga tilstrekkelig kvalitet.

Mastering og pressing av platen ble foretatt av et profesjonelt firma.



Valg av anvendelsesområde

En samling stillbilder som en får tilgang til via en database med opplysninger om bildene, vil vi her kalle en bildedatabase. Bildene kan også være sekvenser med levende video. Selve databasen kan være alt fra en konvensjonell database med strukturert informasjon til en tekstdatabase hvor en kan søke på fritekstopplysninger.

Bildedatabaser er bare en av flere mulige anvendelser av videoplate-teknologi i humaniora. Undervisning er det største «profesjonelle» anvendelsesområdet. I prosjektet ville en likevel konsentrere seg om bildedatabaser, både fordi anvendelsen er viktig og fordi den gikk naturlig inn

i forlengelsen av annet arbeid ved EDBH og SFFR. En mente også at en begrensning måtte til for å gjøre prosjektet overkommelig med de ressurser som stod til rådighet.

En ønsket i prosjektet å legge inn en rekke forskjellige typer materiale på platen, selv om en var klar over at mye ville ha dårlig kvalitet eller på annen måte være lite egnet for videoplate. Det var viktig å utforske grensene for den teknologien og produksjonsmetoden som var valgt. Selv om det på forhånd var klart at den aktuelle teknologien har visse begrensninger, ville prosjektet kunne gi erfaringer som var relevante for all edb-basert bildelagring.

Annet forberedende arbeid

Høsten 1984 kjøpte EDBH inn en Philips VP835B videoplatespiller fra Sverige. (Nå føres spillere i Norge også.) Dette er en såkalt profesjonell modell, noe raskere og mer robust enn modeller beregnet på hjemmemarkedet. B-en i navnet betyr at spilleren har en modul for teletekst, dvs. at den gir muligheten for å legge tekst-TV-aktig skrift og primitiv grafikk over videobildet. Spilleren har en vanlig serieinngang (RS232C), og den kan dermed kobles til en serieutgang på en vanlig mikromaskin. Spilleren styres fra en datamaskin ved å la et program sende enkle tall- og bokstavsekvenser til serieutgangen.

Det ble umiddelbart skrevet noen enkle programmer for å prøve ut styringen av spilleren. Det ble prøvd flere programmeringsspråk og flere maskiner.

TILRETTELEGGING AV REFERANSEDATA FOR BILDEMATERIALET

Platen inneholder ca. 16.000 stillbilder og ca. 10 min. levende video. For enkelte materialer forelå allerede referansedata i maskinleselig form før prosjektets begynnelse. Noen av institusjonene har foretatt dataregistrering selv i prosjektperioden, EDBH har sørget for resten. Som et resultat av dette og andre forhold har det i tilknytning til registreringsarbeidet også vært utført et stort omformateringsarbeid.

For å kunne koble referansedata og bilder på en videoplate, må det legges inn i referansedata opplysninger om hvor på platen de enkelte bildene ligger (hvilke «framenummer» de har). Arbeidet med dette utgjorde en betydelig del av tilretteleggingen av referansedata.

PROGRAMVARE

All programvare utviklet i videoplateprosjektet er skrevet for IBM-kompatible PC-er. Programvaren er beregnet på utstyr med to skjermer - PC-ens egen skjerm for tekst (databaser) og en TV-skjerm/videomonitor for bildemateriale fra videoplatespilleren.

Gregvid, Fregvid og Kunstvid

GREGVID, FREGVID og KUNSTVID er programmer for registrering og søking i databaser på Norske Kunst- og Kulturhistoriske Museers (NKKM) format. Programmene kobler databaser på disse formatene og tilhørende bilder på platen. GREGVID er laget for gjenstandsmateriale, FREGVID for fotos og KUNSTVID for kunstgjenstander. Alle bygger på eksisterende programmer uten video (GREG, FREG og KUNST).

Alle programmene er skrevet i fjerdegenerasjonsspråket DataFlex og går på IBM-kompatible PC-er med harddisk. Alle programmene virker på følgende måte:

- Søking: Ved oppstart kommer et tomt skjema fram på skjermen. (Hvert skjema går egentlig over flere skjermesider, men en kan lettvisnt bla mellom disse.) Noen av feltene i skjemaet er søkbare. Søking foretas ved å flytte cursor til et søkbart felt og skrive inn søkekriteriet. Når de framsøkte opplysningene kommer fram på skjermen, kan en kalle fram det tilhørende bildet fra videoplaten. I tillegg til å søke, kan en bla gjennom materialet ved å trykke Page Down og Page Up i et søkbart felt. Bladingen vil da gå alfabetisk på dette feltet. Til ett skjema kan det være knyttet flere bilder, og disse kan en bla gjennom på samme måten. Programmene kan ta både stillbilder og levende video, men styringen av levende video er noe primitiv.
- Registrering: GREGVID, FREGVID og KUNSTVID brukes også til registrering av opplysninger. I databasene ligger videoinformasjonen separat fra de ordinære opplysningene om objektene. En må bruke egne programmer (VIDREG, FVIDREG og VIDKUNST) til registrering av videoinformasjonen.

Ringvid og Resvid

RINGVID er et generelt program for kobling mellom videoplate og eksterne datakilder via oppringt linje. RINGVID gjør følgende fire ting:

- gjør PC-en til en terminal mot en oppringt maskin.
- plukker opp referanser til stillbilder som blir sendt fra den oppringte maskinen. Referansene må ha form som frame-nummer merket med et eget tegn foran, og kan f.eks. ligge ispedd en tekst som omtaler bildene, eller i et eget felt i et skjema med opplysninger om bildene.
- gir brukeren en mulighet til å bla i disse bildene.
- gjør frame-numrene om til kommandoer som sendes til videoplatespilleren.

I prosjektet er RINGVID hovedsakelig brukt mot databaser på SPERRY-anlegget til Universitetet i Bergen. Basene er realisert vha. gjenfinningssystemet SIFT. SIFT lagrer opplysninger i form av «dokumenter» som dels inneholder faste felt og dels fri tekst. Hvert dokument inneholder opplysninger om ett eller flere bilder. SIFT gir adskillig bedre søkemuligheter enn DataFlex.

For dokumentdatabaser er det naturlig at brukeren tillates å bla i bildene til det sist viste dokumentet. Den nåværende versjonen av RINGVID



Fylkeskonservatoren i Hordaland, Nils Georg Brekke, prøver ut Senterets videoplate på det avsluttende seminaret for prosjektdeltakerne i desember.

er tilpasset dette. Den er nemlig laget slik at den «glemmer» alle gamle frame-numre hver gang overskriften til et nytt dokument kommer på skjermen. Denne versjonen av RINGVID vil også kunne brukes på en løpende tekst som er delt i passende kapitler, såsant hvert kapittel er merket med en entydig tekststreng. Tegnstrengen som får RINGVID til å glemme, kan settes av brukeren under kjøringen av programmet.

RINGVID gir også brukeren en mulighet til å bla fritt på platen, uavhengig av frame-numrene som kommer fra det oppringte programmet.

RINGVID er ment som en prototyp som det så kan skrives versjoner av for spesielle formål. F.eks. kan det tenkes at en anvendelse krever levende videosekvenser i tillegg til stillbilder, eller at programmet må kunne huske frames lenger tilbake enn siste dokument/kapittel.

Mens RINGVID brukes for kobling til eksterne datakilder, brukes RESVID til kobling mot datakilder som finnes lokalt på PC-en – som for RINGVID gjerne i form av databaser eller tekstfiler.

RESVID er et «resident» program, dvs. at det lastes inn i maskinens hukommelse og ligger i bakgrunnen, klar til å snappe opp bildereferanser fra datastrømmen til skjermen. Den nåværende versjonen av RESVID er tilpasset en bestemt database, nemlig en base med dokumentmateriale fra Bergen Byarkiv. F.eks. har hvert skjema i Bergen Byarkiv-databasen to serier bilder tilknyttet, og RESVID gir muligheten for å velge mellom disse. Som RINGVID anses imidlertid RESVID som en prototyp det kan lages spesialversjoner av etter behov.

Vandring

VANDRING er et menybasert program for «vandring» i et stillbilde-materiale. For hvert bilde får en opp en liste med valgmuligheter – en såkalt meny. Hvilket nytt bilde som blir vist er avhengig av hvilket valg en gjør. Til hvert bilde kan en også ha en tekst. Denne blir vist over menyen.

Et slikt program kan f.eks. brukes til å simulere en fysisk vandring, og nettopp dette er gjort i prosjektet. Utgangspunktet er et fotomateriale fra Fylkeskonservatoren i Hordaland, nærmere bestemt en serie eksteriør- og interiørbilder fra Bogatunet, et verneverdig tun på Radøy i Hordaland.

VANDRING er skrevet i TurboPascal. Som for RINGVID og RESVID er det tenkt at det skal kunne skrives spesialversjoner. F.eks. kan det være aktuelt med en versjon som tar levende video.

VURDERINGER

Valg av teknologi

Da prosjektet ble startet opp, vurderte man det slik at read-only analoge optiske plater var den eneste typen optisk medium som både var økonomisk overkommelig og moden for ordinær bruk i forbindelse med de anvendelser som var aktuelle i prosjektet. Denne vurderingen ser ut til å ha vært riktig. Først idag ser en starten på en utvikling som på noe sikt kan gjøre andre, også skrivbare, digitale optiske platemedier aktuelle.

Avgjørelsen om ikke å benytte seg av mulighetene til å legge digital informasjon inn på platen ser også ut til å ha vært riktig. For et eksperimentprosjekt som dette, der det har vært nødvendig og ønskelig å kunne revidere referansedata og programvare underveis, har det vært gunstig nettopp å ha alle referansedata og programvare på ordinære overskrivbare magnetiske media.

Muligheten til å legge datamaskingenerert grafikk og tekst på videobildet har ikke vært noe savn i de anvendelser man har gjort av videoplate-teknikk i dette prosjektet. Selv ikke den noe begrensede muligheten man faktisk har hatt til å gjøre dette ved hjelp av teletekst har vært benyttet i særlig grad. Dersom man hadde tatt sikte på en annen type anvendelse, f.eks. undervisningsprogram eller presentasjonsopplegg, ville dette vært mer aktuelt. I forbindelse med det som har vært prosjektets primære anvendelsesområde, nemlig bildedatabaser, har en tvert om sett det som en betydelig fordel at en har basert seg på en to-skjerms løsning: Den funksjonelle delingen mellom datamaskinlagrede referansedata etc. på én skjerm og tilknyttede videobilder på en annen gjør det ofte lettere for brukeren å holde oversikt og å «vite hvor han er» enn om alt ble presentert på én skjerm.

Bildekvaliteten på prosjektets plate er som man forutså meget varierende. Det meste er akseptabelt, men noe åpenbart ubrukelig.

Deler av materialet er imidlertid av dårlig kvalitet i utgangspunktet, så ikke alt som på platen ser ut til å være av dårlig kvalitet skyldes dårlig gjengivelse.

M.h.t. deler av materialet skyldes utilfredsstillende gjengivelse feil og mangler i avfotograferingen til 16mm film eller på veien fra film til masterbånd.

Derimot er det også ganske klart at utilfredsstillende gjengivelse av noen av de materialene som finnes på platen faktisk skyldes begrensninger i selve platemediet. Dette gjelder særlig håndskrevet og maskinskrevet tekst, store og detaljrike kart og tegninger. Video har simpelthen ikke god nok oppløsning til å gjengi mange detaljer.

Også dette var forutsett, og forsåvidt ikke overraskende. Det var derfor eksperimentert med oppruting av håndskrevet materiale, kart, tegninger osv. i småbilder. Dette viste seg å være en god løsning i enkelte tilfelle.

For enkelte tekster viste det seg tilstrekkelig å dele opp hver enkelt tekstsider i en avbildning av sidens øvre og nedre halvdel. Materialet kunne dermed leses noenlunde uanstrengt, om enn halvsidevis.

For tekstmateriale later det imidlertid til å gå en grense ved det punkt der man under avbildning blir nødt til å dele tekstsidene vertikalt på tvers av linjene i teksten – dette kan i og for seg gjøre gjengivelsen god nok til at materialet blir leselig, men det å veksle mellom enkeltbilder under lesningen av hver enkelt linje er likevel uholdbart fra brukers synspunkt.

Når det gjelder kartmateriale, tegninger etc. skal vi her nøye oss med å notere at det å splitte et slikt materiale opp i småbiter for de fleste anvendelser vil kreve at man har tilgang til oversiktsbilder som gjør det mulig for brukeren å orientere seg i forhold til helheten på en hensiktsmessig måte. Dette krever for det første en meget nøyaktig planlegging før innfotografering av slikt materiale. For det annet bør en være oppmerksom på at det for enkelte materialer, f.eks. store og detaljerte strektegninger, kan være meget vanskelig å produsere et bilde av hele objektet som tilfredsstillende selv de mindre krav til gjengivelse man vanligvis stiller til et slikt oversiktsbilde.

Generelt kan konklusjonen på våre erfaringer sies å være at videoplateteknologi egner seg godt for levende film eller video og for stillbildemateriale der man ikke stiller spesielle krav til oppløsning og manipulerbarhet. Der slike krav stilles – noe som ofte vil være tilfelle m.h.t. f.eks. tekst, kart og tegninger – vil man i hvert enkelt tilfelle måtte vurdere om løsninger som kan tilbys innenfor rammene av denne teknologien vil kunne tilfredsstillende brukerkravene.

Valg av anvendelsesområde

Prosjektet kan trygt sies å ha dokumentert at denne type anvendelse av videoplater egner seg meget godt for den bruk av bildedatabaser som er aktuell og ønskelig for mange av de deltagende institusjoners vedkommende.

I ettertid kan det synes ønskelig at en hadde lagt mer arbeid ned i å finne fram til og vurdere materialer for annet enn rene bildedatabaser. En kunne også hatt demo-opplegg for andre typer anvendelser hvis en hadde tatt med annet materiale.

Prosjektets programvarearbeid kan sies å ha hatt hovedsakelig ett sikttemål: Til de fleste bildematerialene eksisterte allerede databaser med referansedata til bildene. Oppgaven var å koble enkeltposter i disse databasene til «sine» respektive enkeltbilder, bildeserier eller videosekvenser. Dette skulle gjøres på en slik måte at brukeren for en hvilken som helst fremsøkt post i databasen kan hente fram de tilsvarende bilder og til en viss grad styre fremvisningen av disse og tilknyttede bilder ved egne kommandoer.

Opgaven er løst dels ved at en har omarbeidet allerede eksisterende databaseprogrammer slik at videoplatespilleren kan styres fra disse programmene direkte (Gregvid, Fregvid, Kunstvid), dels ved at en har skrevet programmer som bare forutsetter at det legges inn ekstra data i allerede eksisterende databaser (Ringvid og Resvid). Den siste typen programmer ser ut til å gjøre den første typen overflødig i mange tilfelle – noe som vil bety en betydelig arbeidsbesparelse på programsiden, etter som disse er så generelle at de meget lett kan omskrives for forskjellige typer databaser.

Man har i all programvare utviklet ved dette prosjektet benyttet konvensjonelle 4.generasjons verktøy som Dataflex, dBase og Enable samt høynivåspråk som Basic og Pascal. Fordi hovedsikttemålet har vært utvikling av gjenfinningssystemer har såkalte forfatterspråk eller forfattersystemer ikke vært benyttet.

Når det har vist seg såvidt enkelt å tilrettelegge de aktuelle databasene for videoplate, kan dette sees som et resultat av at den type interaksjon det her er tale om i forbindelse med fremhenting av og søk i bilder i seg selv er meget enkel hva bildedelen angår. Også datastrukturen er stort sett meget enkel: hovedsakelig dreier det seg om et en-til-en forhold databasepost-enkeltbilde.

Et system for mer komplekse datastrukturer og en noe annen type interaksjon er også utarbeidet (Vandring). Heller ikke dette bød på spesielle programmeringstekniske vansker, men generelt krever slike oppgaver at det legges langt større vekt på spesifikasjons- og designfasen. Det er dessuten grunn til å tro at en her med fordel kunne benytte ferdige forfatterspråk i stedet for tradisjonelt programmeringsverktøy.

Det er først i utvikling av mer kompliserte systemer for bruk av interaktiv video for demonstrasjons- og undervisningsformål at bruk av forfatterspråk vil kunne oppvise fordeler i forhold til de enklere løsninger som ble valgt i dette prosjektet.

TILRÅDINGER OG KONKLUSJONER

Videoplatens fordeler som lagringsmedium for de formål som har vært aktuelle i vår forbindelse kan sies å være følgende: Utstyr for avspilling er billig – alt man trenger er en videoplatespiller med fjernsyn koblet til en personlig datamaskin. Utstyr for produksjon av kopier fra videobildet er også forholdsvis rimelig. Platen egner seg godt til produksjon i store opplag. Den gir mulighet for kompakt lagring av såvel stillbilder, levende video og lyd som digital informasjon i form av referansedata og programvare. Intet kjent lagringsmedium med tilnærmedesvis samme kapasitet, allsidighet og fleksibilitet er tilgjengelig i dag, selv om digitale plater snart vil komme med levende video, det eneste digitale plater har manglet. Den lave aksesstiden gjør rask søking i store bildemengder mulig.

Ulempene kan sies å være følgende: Platen er ikke overskrivbar. De siste leddene i plateproduksjonen er forholdsvis dyre og må settes bort til et profesjonelt produksjonsfirma. Tilrettelegging av bilder og data er et meget tidkrevende og omfattende arbeid som krever nøyaktig planlegging og i en viss grad profesjonell hjelp. Oppløsning, fargegjengivelse og mulighetene til å manipulere videobildet er ikke for alle formål tilstrekkelig.

Videoplate er følgelig et velegnet og økonomisk rimelig bildelagringsmedium for bruk til formål i forbindelse med enkle bildedatabasessystemer der det ikke stilles spesielt strenge krav til bildenes fargegjengivelse, oppløsning og manipulerbarhet.

Dette betyr i praksis at den i de fleste tilfelle egner seg godt som lagringsmedium for den type bildedatabaser som har vært videoplateprosjektets viktigste anvendelsesområde, hvor det er tale om bilder av kunst- og kulturhistorisk materiale i form av kunstverker, fotos og gjenstander. Den egner seg stort sett *ikke* for bilder av tekstmateriale, dokumenter, kart og tegninger.

Siden den første typen materiale utgjør hovedtyngden av museenes materialer og den siste typen utgjør hovedtyngden av arkivenes materialer, antar vi derfor at videoplateteknologi er et velegnet lagringsmedium for museene, men ikke for arkivene.

De fleste museer står likevel ikke i dag i posisjon til å benytte seg av denne teknologien – dels fordi de ikke råder over de nødvendige økonomiske midler, dels fordi de ikke har maskinlesbare referansedata til sitt bildemateriale, dels fordi de ikke har bildematerialet tilrettelagt i en form som egner seg for innscanning på video. Bl.a. av disse grunner avventer mange institusjoner videre utvikling av digitale optiske lagringsmedier.

I denne situasjonen vil vi gi følgende råd til museene: De bør (1) tilrettelegge referansedata til sine bildesamlinger i maskinleselig form, og

(2) avfotograferer gjenstander og bildemateriale til et høyoppløselig analogt medium av enhetlig format – fortrinnsvis perforert 35mm «kinofilm», som også kan benyttes som mikrofilm, til mikrokort o.l.

Dette er et tidkrevende og omfattende arbeid som det under enhver omstendighet er viktig å komme i gang med så snart som mulig, da det er en forutsetning for å kunne benytte seg av såvel digitale som analoge optiske lagringsmedier. Når dette arbeidet er utført, vil man stå i posisjon til å legge sitt materiale inn på videoplate. Samtidig har filmen en oppløsning og gjengivelses kvalitet forøvrig som vil danne grunnlag godt nok for overføring til alle forutsebare fremtidige optiske lagringsmedier, det være seg analoge eller digitale. Også ut fra almene bevaringshensyn vil dette derfor være en anbefalelsesverdig strategi.

ARBEID I FORLENGELSE AV PROSJEKTET

Videoplateprosjektet har, foruten de mer konkrete resultater som det er redegjort for i det ovenstående, resultert i en generell kompetanseheving og kunnskapsformidling om optisk lagringsteknologi i de deltagende institusjoner og tilstøtende humanistiske fag- og forskningsmiljøer. Også institusjoner og bedrifter utenfor humaniora har hatt slik nytte av prosjektet.

Videoplateprosjektet som sådan er nå avsluttet, men aktiviteter i forlengelse av prosjektet vil fortsette i enkelte av de deltagende institusjoner. En forventer at enkelte av de deltagende institusjoner vil starte egen videoplateproduksjon på bakgrunn av de erfaringer som er gjort. Andre vil ta med seg disse erfaringene i arbeid med digital optisk lagringsteknologi.

EDBH for sin del vil på den ene siden sikte seg inn mot en funksjon som nasjonal koordinator og informasjonsorgan for arbeid med optisk plateteknologi og interaktiv video i de humanistiske fag, på den annen side fortsette arbeidet med digital optisk lagringsteknologi i form av prosjekter rettet inn mot bruk av masseproduserbare read-only-plater (bl.a. CD-ROM) og write-once-plater (WORM).

EDBH vil i oktober -87 arrangere en nasjonal konferanse om bruk av optiske lagringsmedier i forskning, formidling og undervisning. Forøvrig vil EDBH stå til disposisjon for de humanistiske fag- og forskningsinstitusjoner for generell konsulentbistand og kontaktformidling i forbindelse med bruk av videoplater i den utstrekning det finnes kapasitet til det.

*Roger Erlandsen er arkivar ved Sekretariatet for fotoregistrering, Oslo.
Claus Huitfeldt er konsulent ved NAVFs edb-senter for humanistisk forskning/daglig leder av Wittgenstein-prosjektet. Øystein Reigem er førstekonsulent ved Senteret og leder av videoplateprosjektet.*

HUMANISTISKE BIBLIOGRAFISKE DATABASER

Annema Hasund Langballe

FLERE AMERIKANSKE DATABASER

Kristin Natvig gav i forrige nummer en nesten komplett oversikt over humanistiske databaser som kan nås med terminal fra Norge. Noen få titler kan tilføyes. Det nevnes (s. 18) at Dialog har åtte humanistiske baser, hvoretter det føres opp seks. Den ene som mangler, må være *RILA* (*Art literature international*) som ble åpnet i 1985 men er lagt inn retrospektivt slik at den starter med 1973. Den andre kan være *RILM*, en base med musikk-literatur som oppdateres uregelmessig og har stort tidsetter-slep, eller kanskje *Religion index*, åpnet 1985, med noe varierende tids-omfang for forskjellige dokumentformer, den eldste delen fra 1949. Som navnet tilsier rommer basen litteratur om religion, men en del litteratur om kunst, film, litteratur og filosofi sett i relasjon til religion er med.

En ny databasevert av interesse for humaniora er ikke nevnt i artikkel-en. Forlaget Wilson i New York åpnet sin egen online-tjeneste Wilson-line i november 1984 i USA og 1986 i Europa. Alle Wilsons trykte bibliografier er med, lagt inn tilbake til år som varierer mellom 1981 og 1984. Her finnes bl.a. *Art index*, *Biography index*, *Book review digest*, *Cumulative book index* (den amerikanske bokfortegnelsen) og *Humanities index*.

Wilsonline's adresse er: *950 University Avenue, Bronx, NY 10452*.

NOEN PRAKTISKE ERFARINGER

Siden norske humanister foreløpig later til å bruke databaser i liten grad, kan noen praktiske erfaringer kanskje være av interesse. Jeg har brukt slike baser både i egen litteratursøking og i undervisningen ved Statens bibliotekhøgskole fra 1983, da databasesøking ble innlemmet i Referansekunnskap siste studieår. Mellom 30 og 50 studenter har fulgt humaniorakurset hvert år. Dvs. at et ganske stort antall bibliotekarer utdannet ved SBH burde være i stand til å hjelpe lånere med slik søking, dersom biblioteket har en terminal. Blant universitetsbibliotekarer finnes også flere med kjennskap til de humanistiske basene.

Prisen for søking varierer fra vert til vert og base til base. Vår erfaring ved SBH med Dialog var at tre kvarter ved terminalen kom på ca. 30 dollar. I tillegg kommer kostnader for telelinje. Men dette var utrenede søkere som tok en god del utskrifter av referanser online – det faller rimeligere ut enn dette med mer erfaring og utskrifter som tas ut offline og blir tilsendt etterpå (gjørne i løpet av 5-10 dager). Og manuelt arbeid koster også! Skal man gjennomgå alle årgangene av *RILA* og *Art bibliographies modern* for å se om det er skrevet noe om Erling Enger i de analyserte utenlandske tidsskriftene, vil det si å slå opp i 54 enkelthefter, mens det tar to minutter å søke i de to databasene og få vite at det er null treff. Jeg har bladd manuelt i dagevis i en rekke indeksorganer for å finne frem til en detaljert presentasjon av litteraturhenvisningssystemet «Harvard-systemet» – det tok tre minutter å søke fritekst gjennom tre baser hos Dialog og finne referanse til en australsk lærebok som ifølge referatet la hovedvekt på dette systemet (og som lot seg låne fra British Lending Library).

OPPLÆRING OG HJELPEMIDLER

De fleste søkere har startet med innføringskurs hos basevertene. Dialog og BRS holder kurs i Norge, mens informasjonsmegleren Axess i Oslo gir kurs i den franske basen FRANCIS. Dette bringer en opp på selvstudiumsnivå. Vertene har grundige og pedagogiske lærebøker i sine søkespråk og oftest spesielle hefter i tillegg som gir detaljer angående de enkelte basene. For selv om søkespråket er felles for alle baser hos samme vert, er søkemulighetene svært varierende. I én base kan det søkes frem spesielle publikasjonstyper som doktoravhandlinger eller kongressrapporter, i en annen ikke, i én base kan en holde utenfor alt som er skrevet på f.eks. slavisk, i en annen ikke. Måten det gis emneord (deskriptorer) på, er også meget varierende. Det er helt uholdbart å søke uten å ha en trykt tesaurus (emneordsfortegnelse) for hånden, eller – om en slik ikke finnes – en årgang av den trykte utgaven av bibliografien å kontrollere emneordene mot. Og et godt resultat er også avhengig av de detaljkunnskapene om basen som oppnås ved hyppig bruk.

DATABASESØKING SOM BIBLIOTEKTJENESTE

Den som ikke regner med å søke mye, vil derfor neppe finne det regningssvarende å ta seg alt dette arbeidet. Derfor bør det være en selv-

følge at universitets-, høyskole- og fagbibliotekene tilbyr databasesøking.

Alle universitetsbibliotekene gjør det i dag, men det er en tjeneste som foreløpig er lite brukt innen humaniora. Derfor er det også vanskelig å fremskaffe sikre tall for omkostningene. Men et par bibliotek oppgir visse rammer. UB Oslo sier at et «normalsøk» på ca. 10 minutter med ca. 50 referanser utskrevet offline (og tilsendt etterpå) koster fra 300 kr. og oppover, hvis basen er spesielt dyr eller søket større. NTH har tilbudt databasesøking i mange år og har stor erfaring innen tekniske fag, men naturlig nok mindre innen humaniora. Prisen pr. oppdrag ligger gjerne mellom 500 og 2000 kr. De øvrige universitetsbibliotekene har foreløpig hatt for lite søking til å kunne gi prisanslag.

Derfor bør bibliotekarene gå mer aktivt ut for å informere om de mulighetene databasesøking byr på. Scott D. Stebelman ved University of Nebraska stilte seg opp med en terminal i hallen utenfor universitetskantinen og vinket til seg dem som kom forbi til en demonstrasjon. Brukertallet steg kraftig etterpå.

Annema Hasund Langballe er avdelingsleder ved Bibliografisk avdeling, A/L Biblioteksentralen.

RAPPORTER

EDB FOR HUMANISTER

ERFARINGER FRA ET NYSTARTET GRUNNFAG VED DET HISTORISK-FILOSOFISKE FAKULTET, UNIVERSITETET I OSLO

Asbjørn Brændeland og Jon Lanestedt

HISTORIKK

Ved HF-data (Edb-tjenesten ved HF, UiO) har det fra høsten 1981 blitt gitt undervisning i et fag kalt «Edb for humanister» (EfH). Undervisningstilbudet var de 5 første årene begrenset til et semesteremne. Høsten 1986 ble tilbudet utvidet med et tilleggsemne som sammen med semesteremnet gir et grunnfag. Mens det opprinnelige semesteremnet hele tiden har vært et halvtidsstudium med eksamen om våren, går grunnfagstillegget på heltid med eksamen om høsten. I tillegg til det opprinnelige semesteremnet er det fra våren 1987 satt i gang et heltids semesteremne med eksamen sammen med halvtidsemnet. Dette for å gjøre det mulig å ta hele grunnfaget i løpet av ett år.

I løpet av siste halvdel av 70-åra ble det ved HF-data tilbudt stadig flere kurs i edb-anvendelser og programmering rettet mot fakultetets forskere og hovedfagsstudenter. Semesteremnet i EfH oppsto bl.a. ut fra et ønske om å samle de fleste av disse kursene i et helhetlig undervisningstilbud som også kunne gå inn som en del av cand.mag.-graden. Faget var dermed i utgangspunktet først og fremst tenkt som et redskapsfag, men man skjelte også til mulighetene for at humanistiske kandidater skulle kunne kvalifisere seg til edb-undervisning i skolen eller til mer eller mindre edb-relaterte yrker i arbeidsmarkedet for øvrig.

Hensynet til undervisningskompetanse og arbeidsmarked var kommet helt i forgrunnen da fakultetet i 1983 oppfordret HF-data om å utrede et fremtidig grunnfagstilbud i EfH. Samtidig hadde studiet av edb-baserte forskningsmetoder begynt å få karakter av et eget fagområde – altså noe mer enn et rent redskapsfag. Og innenfor deler av det humanistiske forskningsmiljøet økte bevisstheten omkring betydningen av humanistisk kompetanse innenfor informasjonsteknologisk forskning.

Den etablerte humanistiske forskningstradisjon utgjør dermed en vesentlig del av grunnfagets basis sammen med de erfaringer man har skaffet seg fra edb-basert humanistisk forskning i løpet av de siste tiår. Dette bestemmer også de innholdsmessige målsetningene for undervisningen, mens de praktiske mål fremdeles er knyttet til fagets nytte som redskap innenfor andre humanistiske fag og som kompetansegivende fag i forhold til skoleundervisning og til arbeidsmarkedet generelt.

Den økende bevisstheten om betydningen av humanistisk kompetanse innenfor informasjonsteknologisk forskning har ført til at en ytterligere utvidelse av undervisningstilbudet i det som vel blir hetende «Humanistisk informatikk», nå er satt på fakultetets program – først og fremst i retning av datalingvistik og anvendt logikk.

GENERELT OM FAGETS INNHOLD

Særlig semesteremnet, men også til en viss grad grunnfaget, er rettet mot programmering. På dette punktet går EfH lenger enn noe tilsvarende humanistisk undervisningstilbud ved andre norske universiteter. Det at programmering har fått en såvidt sentral plass innenfor faget er først og fremst begrunnet med at den typen trening i bruk av et formelt språk som programmeringsundervisningen gir, har langt større generell overføringsverdi enn trening i bruk av ferdige programmer for mer eller mindre spesifikke anvendelser. De fleste studentene vil kanskje i liten grad få direkte bruk for sine programmeringsferdigheter, men de får en grunnleggende kjennskap til hvorledes maskiner og programmer virker og blir i vesentlig mindre grad enn rene brukere prisgitt dårlig formidlede bruksanvisninger og hemmende grensesnitt. Samtidig får de god trening i problemløsning både mht. å lage og forstå formelle løsningsprosedyrer (algoritmer) og å velge mellom formaliserbare og ikke-formaliserbare løsninger stilt overfor ulike problemer. (Er det statistisk grunnlag for å automatisere analysen av bruken av visse stilistiske figurer hos en viss forfatter? Er det mulig og evt. hensiktsmessig å automatisere koblingen av visse typer opplysninger i ulike historiske kilder?)

Erfaringene fra grunnfagsundervisningen viser tydelig hvor nyttig studentenes programmeringskunnskaper er også utenfor programmeringssituasjonen – f.eks. når det gjelder å forstå prinsippene for formell analyse av uttrykk i naturlige språk.

Semesteremnestudentene skal i tillegg til generell programmering selvsagt også lære om bruk av edb i humanistisk forskning spesielt. Dette formidles delvis gjennom egne forelesninger, men er ellers integrert i programmeringsundervisningen gjennom valg av eksempler. Tilsvarende gjelder for grunnfagstillegget, men her har den teoretiske delen av undervisningen fått vesentlig større plass.

Grunnfagstillegget omfatter 4 deler:

1. Systemarbeid, programmeringsmetodikk og edb-anvendelser.
 - 1.a: Filsystemer og databaser: ulike data- og filstrukturer, databaser og applikasjonsgeneratorer.
 - 1.b: Programmeringsmetodikk: datastrukturer og algoritmer, innføring i et 2. programmeringsspråk.
2. Edb-metoder i humanistisk forskning. Fagdidaktikk.
3. Databehandlingens historie og samfunnsmessige aspekter.
4. Ett av følgende valgfrie emner:
 - datalingvistikk
 - kunstig intelligens (AI) og ekspertsystemer
 - datamaskinstøttet læring

Karakteren fastsettes på grunnlag av en 8 timers skriftlig prøve og en semesteroppgave.







KONKRETE ERFARINGER FRA GRUNNFAGETS DEBUT







Den uten videre mest positive erfaringen gjelder studentenes motivasjon og arbeidsinnsats og det faglige og sosiale miljøet som ble utviklet allerede tidlig i semesteret. Mens semesteremneundervisningen i alle år har vært preget av stort frafall og en nokså ujevn arbeidsinnsats blant studentene, var frafallet i praksis lik null blant studentene til grunnfagstillegget. Det ble også oppnådd forholdsvis gode karakterer. Dette henger uten tvil sammen med at grunnfagsstudentene i motsetning til flertallet av semesteremnestudentene har visst hva de gikk til.

Studentene til grunnfagstillegget bidro også i løpet av semesteret og i et avsluttende evalueringsseminar med kritikk og forslag til forbedringer av undervisningsopplegget.

Det oppsto imidlertid en del problemer i forbindelse med fordelingen av arbeidet mellom de ulike delene av studiet – først og fremst ved at

```

void  ( ,  )
int * , *  ;
{
    int  ;

     = *  ;
    *  = *  ;
    *  =  ;

}

```

Illustrasjon: Øystein Reigem.

arbeidet med *programmeringsoppgavene* fra uke til uke tok svært mye tid. Nå er det vel slik at programmering i seg selv periodevis kan få preg av noe i nærheten av en besettelse, og her er nok mange av de som går videre til grunnfagstillegget sterkere disponert enn den gjennomsnittlige semesteremnestudent. Men mye av den skjeve arbeidsfordelingen skyldtes nok også utformingen av programmeringsoppgavene.

For å unngå tidkrevende rutineprogrammering i forbindelse med bl.a. utarbeidelse av menyer, kontroll av inndata og administrasjon av filer,

bør nok oppgavene i større grad gis som «skall» eller delvis ferdige besvarelser enn det som var tilfelle denne gangen.

Et lignende problem oppsto i forbindelse med løsningen av *semesteroppgaven*. Men her dukket det også opp et annet problem: Ifølge studieplanen skal studentene her i løpet av 5 uker i siste halvdel av semesteret «løse en humanistisk-faglig oppgave ved hjelp av edb». Man har med denne formuleringen villet poengtere at den egentlige oppgaven skulle ligge utenfor programmeringen. Ifølge den utredningen som ligger til grunn for studieplanen, var det også meningen å etablere et veiledningsapparat vha. lærere fra de humanistiske fagområder der de ulike oppgavene hørte hjemme, under oppgaveløsningsperioden. Forventningen om at en faglig interessant humanistisk oppgave skulle kunne løses i løpet av 5 uker, viste seg imidlertid å være nokså illusorisk, og det viste seg også i praksis vanskelig å etablere det nødvendige veiledningsapparat. Det ble gitt flere ulike oppgaver, og ambisjonsnivået for de fleste av disse ble redusert til å omfatte system- og programtekniske løsninger, samt en drøftelse av bruksområde for løsningen innenfor humanistisk forskning.

Den delen av grunnfagstillegget som skapte størst problemer under utarbeidelsen av pensumlista og forelesningene, var del 2: *Edb-metoder i humanistisk forskning*. Dette bunner først og fremst i at litteraturen på dette området enten er foreldet, konsentrerer seg om hovedsakelig kvantitative undersøkelser og mangler teoretiske betraktninger om metodeutvikling og begrepsdannelse, eller rett og slett er for overflattisk. Dette burde være et tankekors for de som til daglig arbeider med edb-anvendelser i humaniora, og en spore til innsats på området.

Et annet forhold er at det har vært nødvendig å styrke semesteremnets del 2, Edb-anvendelser i humanistiske fag. Grunnboka i humanistisk databehandling (Susan Hockey) for grunnfagstillegget blir nå også tatt i bruk i semesteremnet, og det er gode grunner til at grunnfagstilleggets del 2 i den form det nå er stryket av studieplanen og overføres til semesteremnet. Dette muligens til fordel for kulturteoretiske eller filosofiske temaer.

I grunnfagstilleggets del 3: *Databehandlingens historie og samfunnsmessige aspekter*, ble en rekke forelesere fra fakultetet og utenfra trukket inn i undervisningen. Dette fordi denne delen tilsikter en bred overflateorientering innen relevante temaer (telematikk, juridiske spørsmål, kommersielle anvendelsesområder for IT osv.). I denne delen skal imidlertid også det humanistiske perspektiv på informasjonsteknologien presenteres – dennes betydninger for menneskelig tenkning og kulturdannelse, lærings- og kunnskapsstrukturer etc. Dette bør være en substansiell del av grunnfaget, og det er ønskelig at denne delen i større grad enn det som

var tilfelle i fjor, dekkes av HF-data selv og av forelesere fra fakultetet. For det andre har erfaringen vist at deler av de valgfrie emnene også er av allmenn interesse, og således burde integreres mer i fellesforelesningene. Særlig gjelder dette høstens forelesningsserie om kunstig intelligens og ekspertsystemer.

I grunnfagstilleggets del 4 – *Valgfrie emner* – ble følgende 3 emner tilbudt: 1. Kunstig intelligens og ekspertsystemer, 2. Datalingvistik og 3. Datastøttet læring. Disse var først og fremst ment som smaksprøver fra de respektive områdene; og for noen av studentene gav de åpenbart også mer smak (bl.a. arbeider nå to av fjorårets grunnfagsstudenter med datalingvistiske emner i sine hovedfagsoppgaver). Studieplanen inneholder flere emner enn de tre som ble tilbudt (f.eks. informasjonsteori og telematikk). I prinsippet skal disse kunne tilbys i den grad HF-data selv rår over eller ser seg i stand til å hente inn den nødvendige kompetanse. Tatt i betraktning at studieplanen sannsynligvis vil bli endret nokså snart (se under), om ikke inneværende år, vil imidlertid de øvrige emnene neppe noen gang bli tilbudt som valgfrie. Men deler av dem vil kunne bli innarbeidet i fellespensum.

Forelesningene om *kunstig intelligens og ekspertsystemer* dreide seg om sentrale spørsmål innen feltet cognitive science, noen viktige standpunkter ble belyst, og linjer fra debatten mellom disse ble trukket opp. Hovedforeleser var professor Andrew Jones fra Institutt for filosofi. Spørsmål om formelle systemer og beregnbarhet, intensjonalitet, teorier om læring og lagring av kunnskap ble belyst som kriterier for forståelse av mennesket kontra maskin. Dessuten ble noen trekk ved ekspertsystemer belyst ved foredrag av Rolf Nossun fra Computas Expert Systems, og professor Stein Bråten fra Institutt for sosiologi.

Forelesningene i *datalingvistik*, som ble holdt av vit.ass. Helge Lødrup fra Institutt for nordisk språk og litteratur, dreide seg bl.a. om problemer vedrørende maskinell analyse av naturlig språk. Man studerte hvilke egenskaper ved naturlig språk som skaper slike problemer, slik som ulike former for flertydighet. Metoder/algoritmer for slik analyse ble belyst. En interessant erfaring fra disse forelesninger var at studentenes kjennskap til programmering og algoritmetenkning klart var befordrende på deres evne til å tilegne seg stoffet. Lødrup, som har erfaring fra tilsvarende undervisning ved Universitetet i Bergen, mente at progresjonstakten godt kunne ha ligget høyere enn det han hadde forberedt.

Kurset om *datamaskinstøttet læring*, som ble ledet av Ivar Solheim, tok utgangspunkt bl.a. i Skinners og Piagets pedagogiske teorier. Man så på et par verktøy/modeller utviklet av Datasekretariatet – Mosaikk og Torgmodellen – for utvikling av læreprogrammer, og på noen programmer

som brukes i skolen. Til tross for at litteraturen i dette emnet tok opp flere læringsteoretiske temaer mente studentene på evalueringsseminaret at kurset burde inneholdt mer pedagogisk teori.

SAMLET ERFARING

Som det fremgår av ovenstående har gjennomføringen av det første semesteret med undervisning i grunnfagstillegget i EfH gitt både positive og negative erfaringer. Sett under ett må erfaringene allikevel sies å være positive, og de bekrefter uten tvil tilbudets berettigelse. Opplegget for undervisningen må også sies å ha vært faglig forsvarlig, selv om erfaringene også gir grunnlag for en del justeringer innenfor studieplanens rammer.

Det arbeides nå med planer om en integrering av de eksisterende informasjonsteknologisk relevante undervisningstilbud ved fakultetet (foruten EfH dreier dette seg om grunnfaget i kommunikasjonsteori og semantikk (KS) som gis ved Institutt for filosofi) og en utvidelse av tilbudet. Det er særlig emner som datalingvistik og logikkprogrammering man ønsker å få dekket. Man ser imidlertid også på mulighetene for å få i gang et eget semesteremne som tar opp kvantitative og edb-baserte humanistiske forskningsmetoder i lys av generell vitenskapsteori og metode-lære. Man vil i så fall gå vesentlig lenger inn både på selve metoden, f. eks. når det gjelder anvendt statistikk, og på filosofisk begrunnede antagelser om det humanistiske gjenstandsområdet og om begrepsdannelsen for dette området, enn det som nå gjøres i grunnfaget.

I tråd med disse planene vil grunnfaget – særlig da tillegget – rimeligvis måtte endres noe. Todelingen mellom praktisk programmering og teori bør imidlertid beholdes, men kanskje slik at de teoretiske emnene samles under overskriften «samfunnmessige og kulturelle aspekter ved informasjonsteknologien».

I skrivende stund truer dystre skyer i horisonten. Regjeringen har forlangt at Universitetet i Oslo skal spare inn bl.a. på bruk av lønnsmidler, noe som i første rekke rammer time- og hjelpelærerbudsjettet; og siden EfH-undervisningen i stor grad er basert på midler fra dette budsjettet, må i alle fall semesteremnet, og i verste fall også grunnfaget, stoppes inn-til videre. Vi er imidlertid allikevel optimister, og ser på den aktuelle krisa som en forbigående misforståelse. Informasjonsteknologi – også humanistisk sådan – er jo tross alt et nasjonalt satsningsområde både når det gjelder undervisning, forskning og utvikling, og innstramningene var neppe ment å skulle ramme dette – ikke minst tatt i betraktning at regje-

ringen har satt av en god del millioner til undervisningsutstyr på området.

På denne bakgrunn og med forventningene om en rask utvikling både i mangfold og kvalitet innenfor fagområdet humanistisk informatikk, ser vi nokså lyst på fremtiden for vårt eget fag Edb for humanister.

Asbjørn Brændeland og Jon Lanestedt er hhv. amanuensis og vit.ass. ved HF-data, Universitetet i Oslo.

COMMUNICATION THEORY AND SEMANTICS – A PROGRESS REPORT

Andrew J.I. Jones

The new *grunnfag* programme in Communication Theory and Semantics (CS), at the Department of Philosophy, University of Oslo, received its first intake of students in August 1985. So far, seven students (three female, four male) have completed the whole *grunnfag* examination, with final grades ranging from 1.6 to 2.7, and some ten others have sat the examination in one or two (5-credits or 10-credits) of the four component courses of which the programme consists. (See *HD 1-85*, or the *Faculty of Humanities Handbook*, University of Oslo, for a description of these courses.) A further seven to ten students are expected to complete the *grunnfag* during 1987, and a number of others will probably complete parts of it.

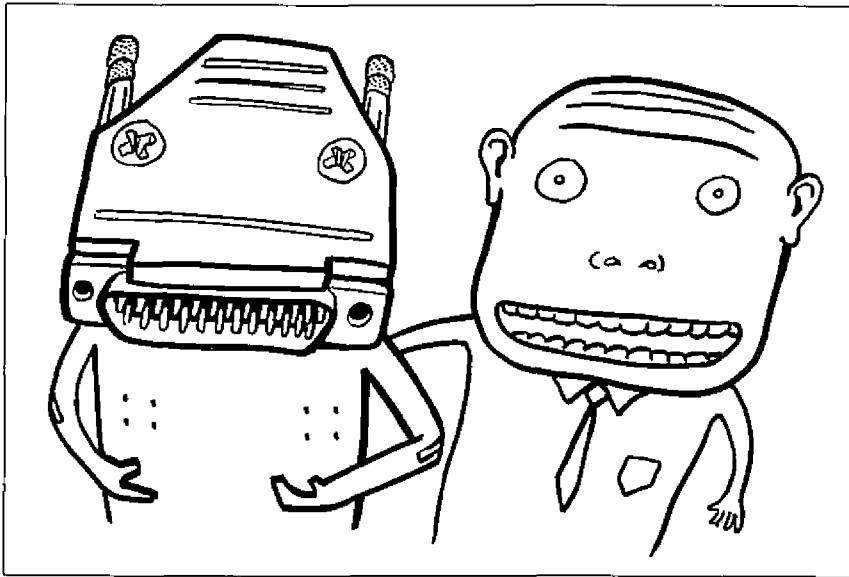
All those entering the CS-programme are required to have an at least *grunnfag*-level background in some subject or other; however, no particular background has so far dominated: students have come to CS from languages, literature, sociology, psychology, philosophy, social anthropology, computer science, mathematics and biology. I take this to be a very positive aspect of the present situation, and a clear indication of the fact that CS cuts across traditional academic boundaries.

As the person responsible for running the programme, and for doing

most of the teaching, I can hardly be said to have a distanced, and in that sense objective, view of the students' attitudes, so far, towards CS. But let me, at any rate, tell the story as I see it. Perhaps the most striking impression is that, compared to a number of other courses, CS appears to be very demanding on the student. (One of the first group, who had previously taken philosophy, described the ordinary philosophy *grunnfag* at this university as «child's play» by comparison.) CS-students are required to take a rather thorough introduction to formal logic, and to become acquainted with aspects of some recent theories on the formal structure of natural language; and, in addition to the formal-logical work, they are expected to penetrate a good deal of literature in the philosophy of language and philosophical psychology – some of it quite sophisticated – and to learn to present in their own terms, and to discuss, some of the leading contemporary theories in these areas. By and large, it is true to say that those CS-students who, for one reason or another, have been unable to study on a full-time, intensive basis, and those students who have come to CS after a long period away from academic work, have found the *grunnfag*-programme difficult to cope with. Whereas those who regularly followed the lectures and seminars, and regularly produced written work, in the form of essays and logic exercises, took the course more or less in their stride and achieved good final results. (It is worth noting that the better exam results also indicate that the formal-logical work can be managed by those who do not have a strong background in mathematical disciplines: indeed all of the best students so far have come from the humanities; so the problem cannot be located *there*. It is of course well known that, for instance, departments of philosophy and departments of linguistics have for a number of years been producing students who have an adequate level of understanding of formal techniques, despite the lack of previous mathematical training beyond the ordinary school level.)

Several of our CS-students have expressed an interest in further studies, and three of them are currently engaged in work designed to lead to the *magistergrad*. (More specifically, two of the topics concerned are: i) the formalization of speech acts and ii) the logic and explanation of action, intention and planning. The third project, still very much in its infancy, is expected to involve cooperation with the Norwegian Research Centre for Computers and Law (*Inst. for retsinformatikk*). The philosophical contribution here will concern legal philosophy and deontic logic.)

Each of these three projects is intimately connected with current research interests in Artificial Intelligence. There is no doubt that CS is providing an introduction to some of the areas of analytical philosophy which are making an important contribution to much current AI re-



Drawing: Øystein Reigem.

search, and that, in virtue of this, it would be natural for students to combine CS-studies with courses in computer science and programming. In fact some of our students have already been doing just this.

Another «connection», which is attracting the attention of an increasing number of psychology students, is with modern AI-inspired Cognitive Science. Unfortunately, the Dept. of Psychology at this university has been extraordinarily slow to take up the challenge of providing for its students an adequate course option in modern Cognitive Science; however, some of their more theoretically inclined students have been attending CS-courses, and/or a seminar which I have been helping to organize together with two research psychologists (Jan Heim and Gunnvald Svendsen) at *Norsk Regnesentral*. (It is planned that this cooperation with NR will lead in due course to an interdisciplinary research project, concerning the design of «user-friendly» interfaces between man and machine ... but that will be another story.)

I have outlined two, not unrelated, ways in which CS connects with other areas of enquiry, and thus two ways in which CS can function within the totality of a student's university study-plan. It is to be noted that both of these pertain to the priority area picked out in *NAVF-RHFs Innstilling om humanistisk informasjonsteknologisk forskning* (NAVF, Oslo,

April 1986): namely, the application of results in research in the humanities to the development of software technology.

But these functions CS performs are of course perfectly compatible with one other role specified at the time CS was first proposed: that CS can provide the interested (and diligent!) student with the opportunity to gain an informed picture of the current debate concerning the relationship between AI and our understanding of the nature of man. It would seem reasonable to assume that the schools will need teachers who can supply this kind of broader perspective on the state of information technology, and thus that CS – combined, ideally, with some computer science and programming – may also fulfil a useful function in the training of school-teachers.

One very natural response to these comments would be to suggest that CS would be at home in a broader, more comprehensive programme of studies where, for instance, options were also available in programming, computational linguistics, applied logic, artificial intelligence and cognitive science. In fact discussions which may lead to the establishing of a wider programme of this kind are already in progress: in December 1986 the Faculty of Humanities set up a committee whose mandate includes the formulation of proposals for teaching and research programmes in computational linguistics, communication theory and semantics, applied logic and programming. Representatives of CS, Linguistics and the Faculty's Computer Science group (*HF-Data*), among others, are now designing courses for a possible new programme. It is naturally too soon to say anything very definite about these plans, which would of course have to be accepted by the powers that be, and funded, before they could be implemented: so there is rather a long way to go yet! But already we seem to have reached agreement, in our preliminary discussions, on some basic points: for instance, the core components of the programme would be 5-credit courses, students perhaps being allowed to take up to as many as eight of these, thereby gaining the equivalent of a *storfag*. Any student taking 20-credits or more would be required to include an introduction to formal logic, a 5-credit course equivalent to the present CS1; one 20-credit «package» would effectively represent a *grunnfag* in computational linguistics; another would correspond to the present *grunnfag* in CS; but, within limits, the programme would allow some flexibility, so that, for example, one could combine some linguistics with some of CS; some of the courses would require, in addition to logic, some background in programming: for example, the proposed courses in LISP and PROLOG, and the proposed introduction to AI. And, finally, some of the courses would effectively represent parts of a *mellomfagstillegg* to the

present *CS-grunnfag* (e.g., courses in modal logic and its applications, courses on the foundations of cognitive science).

It would be nice to think that, within a year or so, *HD* will be able to announce the launching of this new programme; there must surely be some reasonable grounds for optimism, despite the need for funding, and despite the deep-seated prejudices against the very idea of AI which, regrettably, still exist in a number of not uninfluential quarters. After all, not *that* much by way of funding, in addition to present resources, will be needed to get the programme off the ground, and we *are* dealing with a «priority area». We shall see.

Provisionally, we have hit upon the title «Language, Logic and Information Technology» for the new programme. In Norwegian, «Språk, Logikk og Informasjonsteknologi», which bears the rather appropriate acronym «SLIT».

Andrew J.I. Jones is Professor of Philosophy at the University of Oslo.

TRE RESIDENTE PROGRAMMER FOR MS-DOS MASKINER

Espen S. Ore

Residente programmer (strengt tatt en forkortelse for primærlager- eller RAM-residente programmer) virker i praksis som en utvidelse av maskinens operativsystem. IBM PC-familien og kopier (kloner) av denne har en kombinasjon av maskinutrustning (Intel 808x og Intel 80x8x prosessorer) og systemprogramvare som gjør det svært lett å lage slike programmer. I tillegg til denne rapporten, «Kort om residente programmer og MS-DOS», går jeg litt nærmere inn på tekniske og programmeringsmessige aspekter og prøver å gi noen idéer om muligheter som finnes. Forøvrig ber jeg purister og bitfiklere ha meg unnskyldt de forenklinger og unøyaktigheter som forekommer her. (Ja, jeg *vet* at det er forskjell på BIOS- og DOS-utvidelser.)

De tre programmene jeg har prøvd, er *Sidekick* og *Turbo Lightning* fra Borland International og *Smart Notes* fra Personics Corporation.

Sidekick er en samling hjelpeprogrammer som kan minne om Des Accessories på Apple's MacIntosh for de som kjenner denne. Programmene er: *Notepad*, en notisbok med tekstredigeringsmuligheter à la Wordstar, *Calculator*, en lommekalkulator på skjermen, *Calendar*, en avtale/huskebok med datoer fra 1901 (hvorfor vet jeg ikke) til 2099, *Dialer*, et program for automatisk oppringing (dersom maskinen er koblet til et modem) kombinert med en telefonliste med søkemuligheter og *ASCII table*, en ASCII-tabell på skjermen som både viser grafiske tegn for all ASCII-verdier fra 0 til 255 og mnemonisk kode for kontrolltegn.

Turbo Lightning er et automatisk korrekturprogram. Det skal virke sammen med forskjellige tekstbehandlings- og databaseprogrammer f.eks. Wordstar, MS-Word, dBase. Ordlistene er (naturlig nok) basert på amerikansk-engelsk, men det er mulig å lage egne alternative ordlister.

Smart Notes er et elektronisk alternativ til gule klebelapper. Man kan feste en lapp til et hvilket som helst skjermbilde (f.eks. en filkatalog eller en tekstfil). Lappene blir lagret i en egen fil sammen med informasjon om hvilken kontekst (del av skjermbilde) de hører sammen med.

ERFARINGER FRA DE ENKELTE PROGRAMMENE

Jeg vil gjerne understreke at jeg *ikke* har prøvd alle muligheter i de forskjellige programmene. Jeg har heller ikke testet grundig hvilke andre programmer de er kompatible/inkompatible med. Derimot har jeg prøvd å bruke alle tre i mitt daglige arbeid på maskinen (tekstbehandling, programmering og nettverkskommunikasjon). Denne metoden er valgt fordi jeg da håper å se hvilke effekter som virker nyttige (for meg), og hvilke som er ytre staffasje. Andre brukere vil helt sikkert ha andre meninger om dette.

Smart Notes er, såvidt jeg vet, ikke til salgs i Norge, mens flere firmaer selger Borlands produkter. (NAVFs EDB-senter for humanistisk forskning kjøper en god del programmer pr. postordre fra utlandet bl.a. fordi det er billigere, men det kan være mer betryggende å kjøpe fra norsk leverandør med tanke på brukerstøtte dersom det oppstår problemer. Sidekick og Turbo Lightning er ikke tilgjengelige i norsk versjon.

SIDEKICK (SK)

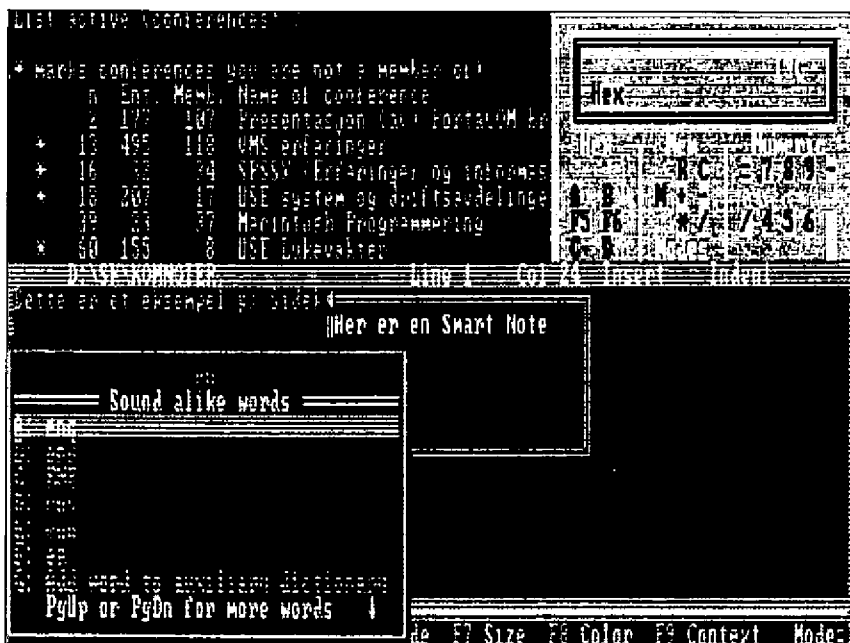
Denne programpakken har vært på markedet i flere år nå, og det fir

nes andre programmer som gjør mye av det samme som SK. I bruksanvisningen står det at SK må være det siste residente programmet som lastes inn i maskinen dersom man bruker flere. Dette henger sammen med måten residente programmer virker sammen med operativsystemet. Rekkefølgen i innlastning er viktig for å unngå at programmene kommer i veien for hverandre, f.eks. ved å fange opp hverandres kommandoer (se tillegget for mer informasjon).

Bruksanvisningen gir en grei innføring i hvorledes man installerer SK. Den er også god når det gjelder tilpasning til spesielle maskiner, tekstbehandlingsprogrammer etc. Derimot kan det være litt vrient å finne informasjon om bruken av programmene. (Dette er et eksempel på subjektiv erfaring – dersom man leser *hele* bruksanvisningen grundig og har god hukommelse, kan man alltid finne frem. Bruksanvisningen er på 111 sider.)

Jeg har ikke brukt kalenderprogrammet. Det er fordi det har virket enklere og raskere for meg å holde meg til min papir-bordkalender.

Kalkulatorprogrammet har noen spesialeffekter for programmerere.



Her er alle tre programmene aktivisert på én gang.

Man kan velge om man vil regne desimalt, hexadesimalt eller binært. Dette er det mulig å endre midt i en utregning. Man kan også overføre resultater til en tekst. Dette programmet har ofte vært nyttig å ha.

Programmet for oppringing har jeg ikke prøvd fullt ut (fordi jeg vanligvis ikke har modem koblet til maskinen). Men den elektroniske telefonlisten og søkemulighetene den gir, virker så bra i seg selv at jeg holde på å overføre min papirbaserte hit.

ASCII-tabellen er mest et verktøy for programmerere. Den kan syne uønsket siden annenhver systemhåndbok inneholder den som et vedlegg, men én vesentlig side ved den er at man kan se tegnene med tegnverdi over 127 slik de er på den maskinen man bruker. Siden disse tegnene ikke er standardiserte (ASCII-standarden dekker tegnene fra 0 til 127), er det ikke gitt at alle maskiner har samme tegnsett som IBM PC.

Notisboken tilsvarer en kombinasjon av Clipboard, Note Pad og Scrapbook på MacIntosh. (NB! Man kan ikke legge grafikk inn i Notepad SK.) Den virker omtrent som Wordstar m.h.p. tekstbehandlingsmuligheter. Man kan starte Notepad når som helst, og det er mulig å hente tekst fra skjermen eller å skrive tekst fra Notepad til det programmet (f.eks. tekstbehandler) som måtte være i gang når Notepad startes. Jeg har funnet programmet spesielt nyttig i forbindelse med telekonferanser (se HI 3-86). Det er sjelden at all tekst som kommer på skjermen under en konferansesesjon, er av interesse, og i stedet for å lage en komplett log (f.eks. gjennom LOG-kommandoen i KERMIT) kan man bruke Notepad til bare å hente inn det man ønsker å ha skriftlig tilgjengelig.

TURBO LIGHTNING (TL)

Korrekturprogrammer er etterhvert blitt temmelig vanlige. Mange tekstbehandlingsprogrammer har også innebygget korrekturfunksjon. TL er imidlertid ikke knyttet til noe spesielt tekstbehandlingsprogram, istedet kan det settes opp så det enten sjekker hvert ord man skriver (et ord oppfattes som ferdigskrevet når neste tegn er et ordskilletegn eller linjeskift), eller på kommando leser og sjekker alle ord på skjermen.

Installasjon gikk like greit som med SK, og det kreves ikke at TL skal være det siste residente programmet som lastes inn. Derimot står det både i SK- og TL-håndboken at man først må laste inn alle residente programmer som *ikke* er produsert av Borland, så Borlands programmer og til slutt SK.

Når TL finner et ord som er feilstavet, kan man gi en kommando, og det gale ordet byttes automatisk ut med det riktige. TL kan læres opp t

å tolke meldinger fra tekstbehandlingsprogrammer, slik at jeg, når jeg bruker en norsk versjon av Wordstar, kan gi TL beskjed om å tolke «INNSETT PÅ» i øverste høyre skjermhjørne riktig. Dette er vesentlig når man skal benytte seg av muligheten for automatisk å bytte ut gale ord med riktige fra ordboken. TL foreslår riktige ord etter følgende system: Først blir det lett i en liste over ord (på engelsk) som ofte blir feilskrevet (f.eks. «hte» for «the»). Deretter blir det foreslått ord som lyder likt det gale, og til slutt ord som er like lange som det.

Ved siden av ordboksfunksjonen har TL også et synonymtesaurus. Dette har jeg personlig funnet nyttig når jeg skal skrive på engelsk. Dessverre kan man ikke bygge opp et norsk synonymtesaurus slik man kan lage en norsk ordliste.

Borland leverer forøvrig et prosedyrebibliotek for kobling med Turbo Pascal-programmer i pakken *Lightning Word Wizard* (som også inneholder mye annet). Ved hjelp av denne kan man lage egne korrekturprogrammer som da selvfølgelig kan skreddersys for norsk i større utstrekning enn det ferdige produktet TL. (Jeg håper å komme tilbake i et senere nr. av HD med erfaringer fra LWW.)

Dette var det eneste programmet som skapte «krøll» med kommandotastene. Standardkommandoen for å aktivisere TL er SHIFT-F8. På Olivetti-tastaturet blir dette det samme som F18, og KERMIT-initieringsfilen jeg bruker, setter F18 opp som en kommandotast for IBM-stormaskinen jeg bruker til EARN (se HD 3-86). Det er heldigvis ikke noe problem å endre aktiviseringstast for TL.

SMART NOTES (SN)

SN gjør det mulig å feste notiser til data omtrent som når vi bruker klebelapper til å kommentere tekster på papir. Notisene samles i egne filer, og programmet forandrer ikke de filene vi kommenterer. De elektroniske lappene er strengt tatt ikke knyttet til et bestemt datasett (f.eks. en tekstfil), men derimot bruker de opplysninger om kontekst i skjerm-bildet. Det gjør det mulig å knytte kommentarer til en filkatalog e.l.

Installasjon er, som for de to andre programmene, smertefri. Programmet er enkelt i bruk, og en oversiktlig bruksanvisning på 51 sider gjør det vanskelig å få problemer når man bruker det. Registeret (indekset) er etter min subjektive oppfatning bedre enn det vi finner for TL. Med residente programmer i maskinen blir det av og til for lite plass igjen til de (vanlige) programmene vi måtte ønske å bruke. Det kan derfor være aktuelt å fjerne ett eller flere av de residente. I registeret til SN-bruksanvis-

ningen finnes følgende to oppslagsord:

Memory

removing SmartNotes from

og

Removing from memory

I TL-bruksanvisningen finner vi kun:

Killing Turbo Lightning

I SK-bruksanvisningen er oppslagstermen «Stop Sidekick», og selv om dette er greit å finne, ville det etter min mening vært en fordel om Borland var konsekvent i termbruken i sine bruksanvisninger.

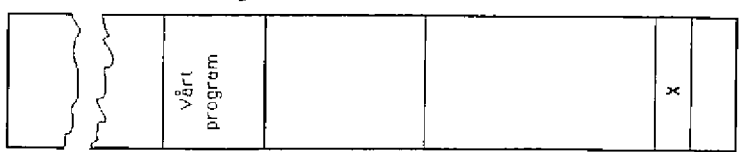
I tillegg til å knytte lapper til tekstlig kontekst kan man i SN også knytte dem til ruter i rutearkprogrammer. Det vil si at lappen ikke er avhengig av at innholdet i ruten forblir konstant, noe som er en fordel her.

Det er vanskelig å gi uttrykk for et klart standpunkt m.h.p. nytten av programmet. Når jeg arbeider med egne tekster eller programmer, viser det seg at jeg i praksis alltid vil ta en papirutskrift og skrive mine arbeidsnotater direkte på den. Det synes mer interessant å kunne koble merknader til filkataloger, ihvertfall når antallet filer blir et par hundre eller flere. Det kan også virke interessant å bruke programmet i forbindelse med et lokalnett der flere personer skal kommentere artikler, rapporter e.l. som skrives. Men også i dette tilfelle er det mulig at papir og gule klebelapper er et vel så godt verktøy.

TILLEGG: KORT OM RESIDENTE PROGRAMMER OG MS-DOS

Et resident program for MS-DOS maskiner består i utgangspunktet av to deler: de prosedyrene som skal ligge i primærlageret, og en liten installasjonsdel som skal legge dem inn. For at de residente prosedyrene skal kunne aktiviseres, knyttes de til systemavbrudd (interrupts). Et avbruddssignal sendes f.eks. når vi trykker på en tast. Når et avbruddssignal kommer, vil sentralenheten stoppe utførelsen av det «vanlige» programmet som måtte være i gang. I stedet for å hente neste instruksjon fra dette programmet (det er det som normalt skjer når et program går), vil sentralenheten hente en verdi fra en gitt lagerplass (lagerplassen er bestemt ut fra hvilket avbrudd som kommer), og bruke denne verdien som adresse til neste instruksjon. På denne adressen vil det ligge en rutine for å behandle avbruddet. Det siste denne rutinen bør gjøre, er å starte programmet som ble avbrutt der det stoppet.

640 K

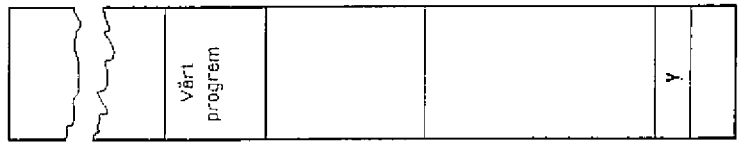


Her slutter MS-DOS, og vårt program begynner

Her begynner den opprinnelige rutinen

Adresse til opprinnelig
nybruksrutine

640 K



Før vi legger inn adressen til vårt program

Etter at vi har lagt inn adressen til vårt program

Hvis vi legger adressen til en resident rutine inn på en av de avbruddsbestemte lagerplassene, kan vi sørge for at den blir aktivisert når vi ønsker det (se figuren). Det finnes mange programmer som gir norsk tegnsatt, og en enkel måte å lage dem på, er å legge adressen til dem inn på plassen for tastaturavbrudd. Hver gang det blir trykket på en tast, kan programmet sjekke om den er «kritisk» og sørge for å bytte ut '[' med 'Æ'.

Vi kan få problemer dersom flere residente programmer er i bruk, og alle har byttet ut den samme adressen. Dette er en av grunnene både til at det angis en spesiell rekkefølge for innlastning, og for at enkelte residente programmer ikke kan brukes sammen.

Et viktig punkt som gjelder alle residente programmer, er at hvis vi ønsker å fjerne ett fra lageret, må vi først fjerne alle de som måtte være lastet inn etterpå. Gjør vi ikke det, får vi «hull» i primærlageret, og maskinen går i frø.

Det er fullt mulig å skrive residente programmer i Turbo Pascal. De som har lyst til å prøve det, bør imidlertid skaffe seg kunnskaper utover det som står i Turbo Pascal-håndboken. Turbo Tutor (fra Borland) og Peter Norton: *Programmer's guide to the IBM PC*, Penguin 1985 gir nyttig tilleggsinformasjon.

NORDISK KONFERANSE OM TEKSTFORSTÅELSE OG INFORMASJONSSØKING

SÄBY SÄTERI, STOCKHOLM 8.-9.12. 1986

Jostein H. Hauge

Institutionen för lingvistik, Stockholms universitet arrangerte den 8.-9.12. en konferanse om tekstforståelse og informasjonssøking. Til arrangementet, som var økonomisk støttet av Nordisk forskningspolitisk råd, var invitert alle de nordiske lingvistiske institusjonene som var inte-

ressert i eller hadde i gang forskning med tilknytning til emnet. Også representanter fra næringslivet var innbudt.

På konferansen deltok 36 personer fra Danmark, Finland, Norge og Sverige fra i alt 26 ulike institusjoner. Målet for konferansen var dels å få en oversikt over aktuell nordisk forskning på dette spesialområdet, dels å utarbeide en konkret samnordisk forskningsplan med det mål å styrke og utdype de faglige kontaktene.

På konferansen ble det gitt et 20-talls presentasjoner av pågående arbeider. Presentasjonene dekket omfattende deler av datalingvistikken: metoder for automatisk analyse av morfologi, syntaks og semantikk; tekstrepresentasjon; korpusarbeid; terminologi; lingvistiske databaser; indeksering; informasjonssøking; kvantitativ datalingvistikk og sosio- og psykolingvistisk orientert forskningsarbeid.

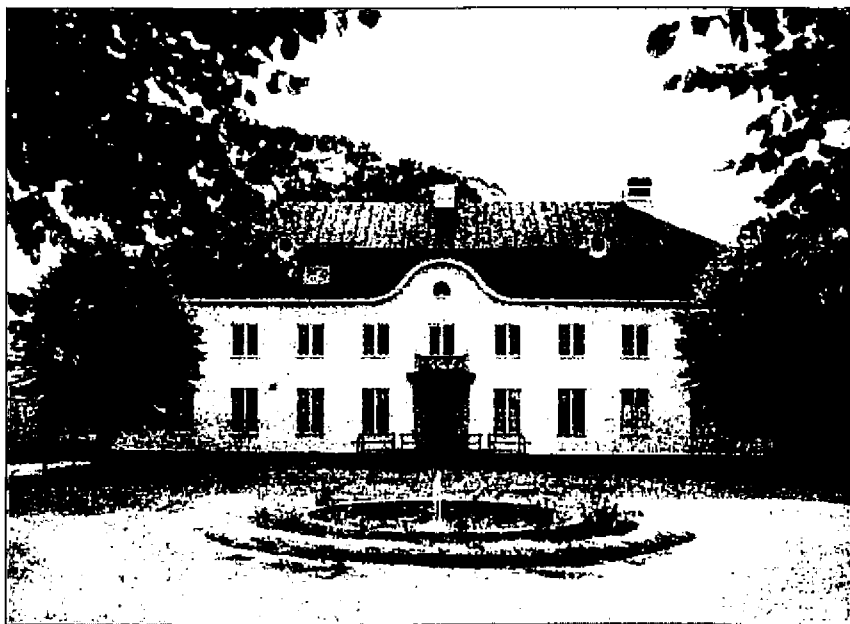
På slutten av konferansen var det reservert tid for prosjektplanlegging. Det var enighet om å starte et samnordisk forskningsprogram som bygger på eksisterende personellmessige og materielle ressurser og som er grunnlagt på de aktivitetene som foregår innenfor de land og institusjoner som ønsker å delta. Det var videre enighet om at alle institusjoner kan søke om tilslutning til forskningsprogrammet, men at en styringsgruppe skal vurdere og godkjenne de enkelte søknadene.

For forskningsprogrammet valgte man *tekstforståelse hos menneske og datamaskin* som samlende tema. Programmet skal vare i tre år fra sommeren 1987. Målet er med samnordiske krefter å produsere en artikkelsamling med originalartikler av internasjonal kvalitet. Det er målet å utgi publikasjonen på et så velansett internasjonalt forlag som mulig.

Professor Östen Dahl, Institutionen för lingvistik, Stockholms universitet vil fungere som koordinator for programmet. Ved hans side vil det være en styringsgruppe som består av Helge Dyvik, Institutt for fonetikk og lingvistik, Universitetet i Bergen, Fred Karlsson, Institutionen för allmän språkvetenskap, Helsingfors universitet, og Hanne Ruus, Institut for nordisk filologi, Københavns universitet.

Temaet for forskningsprogrammet ble ikke fastlagt i detalj på konferansen. Meningen er likevel at programbidragene skal konsentrere seg om tekstsemantikk og tekstrepresentasjon snarere enn på innsats innenfor morfologi og syntaks. De teoretiske datalingvistiske aspektene bør være fremtredende selv om målet ikke i alle tilfeller behøver å være konkret fungerende datamaskinprogrammer.

Forskningsprogrammets styringsgruppe fikk fullmakt til å avgjøre hvilke institusjoner av dem som melder sin interesse for å delta som skal tas med i programmet. Det blir styringsgruppens oppgave å spre informasjon om programmet i sine respektive land og blant aktive potensielle interes-



Süby Säteri.

senter. Det ble vedtatt at prosjektforslag fra de enkelte institusjonene skulle innsendes til styringsgruppens formann før 15. februar i år. I løpet av våren vil styringsgruppen ta stilling til de innkomne forslagene og utarbeide en koordinert plan.

Tanken er at lønn, utstyr o.s.v. skal ordnes av hver deltaker for seg. Styringsgruppen på sin side tar initiativ til å søke om midler til prosjektaktiviteter i forskningsprogrammets navn, til å administrere programmet samt til årlige koordinerende og evaluerende møter, og midler slik at forskere fra ulike institusjoner kan besøke hverandre.

Blant de instanser som man ønsker å søke om midler, er Nordisk forskningspolitisk råd, NOS-H, NORDINFO, NORDFORSK og de nasjonale forskningsråd.

På konferansen ble det vedtatt å oppta temaet *tillempingen av informasjonssøking* som et spesialfelt. Dette delprogrammet styres av Benny Brodda, Institutionen för lingvistik, Stockholms universitet, Anna-Lena Sågvall-Hein, Språkdata, Göteborg, og Lennart Lönnngren, Centrum för datorlingvistik, Uppsala universitet.

Nærmere opplysninger om rammeprogrammet og om planene frem-

over kan gis av de nasjonale medlemmene i koordineringsgruppen.

Ovenstående referat bygger på en rapport utarbeidet av professor Fred Karlsson.

INFORMASJONSMØTE OM SENTERETS VIDEOPLATEPROSJEKT

BERGEN, 24. FEBRUAR 1987

Einar-Arne Drivenes

Møtet ble innledet med en kort generell orientering om NAVFs edb-senters videoplateprosjekt. Denne orienteringen besto bl.a. i at deltagerne ble vist et TV-program om prosjektet, produsert av Audiovisuell avdeling ved Universitetet i Bergen. Deretter ga prosjektarbeiderne ved Senteret en orientering om og demonstrasjon av programvare utviklet ved prosjektet.

*

I møtets andre del ga deltakerne orienteringer om egne prosjekter eller planer i forbindelse med interaktiv video.

Olle Eriksen, Trondheim offentlige skole for døve: En arbeider med å finne fram til ulike datastøttede undervisningssystemer for døve når det gjelder norskopplæring. I den forbindelse har en i samarbeid med Datasekretariatet også starta et prosjekt hvor en skal vurdere interaktiv video. En ønsker samarbeid med andre om utprøving av deler av materialet på videoplate, ikke minst av kostnadshensyn. Presentasjonsformen vil være viktig i dette prosjektet, bl.a. bør både tekst og bilde presenteres på samme skjerm. Målsettinga er å presse ei plate i løpet av året. Det eksperimenteres i dag med en versjon på interaktivt videobånd.

Thor Berntsen, Norsk Arbeidsgiverforening: NAF ønsker å holde seg ajour med den tekniske utviklinga m.h.t. interaktiv video av to grunner:

(1) Som høringsinstans f.eks. i utdanningsspørsmål må en kjenne dette feltet. (2) Dette kan bli et viktig redskap i å spre informasjon internt i organisasjonen sentralt og til medlemsbedriftene.

Alvar Löfskog, DataProgramGruppen, Utbildningsdepartementet: En arbeider med et prosjekt kalt «Gamla Stan» for å se hvordan interaktiv video kan utnyttes i undervisning (for en nærmere orientering se nedenfor). Det arbeides spesielt med å finne frem til systemer for bildebeskrivelse som er anvendelige i en slik forbindelse. En samarbeider med Stockholms Stadsmuseum, som er kilde for alt bildematerialet, og Stockholms historiska databas.

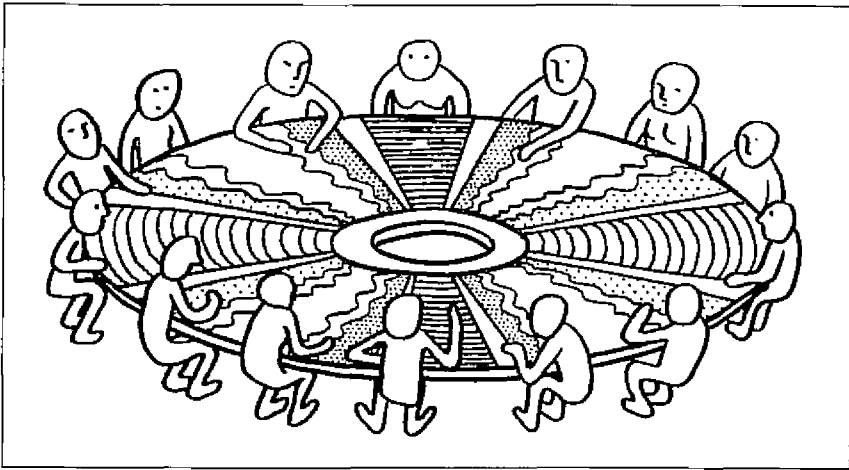
Erik Meistrup, Århus Amts Informatikcenter: Domesday-prosjektet skal demonstreres ved senteret i nærmeste framtid. Danmarks Radio har planer om et videoplateprosjekt finansiert sammen med en del undervisningsinstitusjoner, bl.a. Sygehusenes Uddannelsescenter. Meistrup mente at en i langt sterkere grad burde satse på nordiske samarbeidsprosjekter innafør dette feltet.

Gunn Nordal, Sekretariatet for registrering av faste kulturminner (SE-FRAK), Miljøverndepartementet: Innafor SEFRAK registreres alle faste kulturminner fra tiden mellom år 1537 og 1900, registeret består av tekst, kart og bilder. En har utarbeidet et system for tekstdelen av prosjektet og er i gang med overføring av tekstregisteret til edb. En eksperimenterer med digitalisering av kart. En ønsker også å få det omfattende bildematerialet over på et optisk medium.

Eivind Osnes og Fred Sollie, ASEC (Sony): ASEC Bergen samarbeider med bl.a. Haukeland sykehus, politi og brannvesen om et alarmsystem. En jobber ut fra en standard for slike systemer for EF-landene som er anbefalt av Europarådet. Lagring av bl.a. bilder og kart på videoplate utgjør en del av dette prosjektet, og skal gjøre det mulig hurtig å gjøre seg kjent med lokaliteter og terreng på ulykkessteder. En regner med å ha platen ferdig om ca. 1/2 år.

Hasse Söderström, PPS: PPS, et datterselskap av Mercuri Vekst A/S, er et produsenthus for film, video og interaktiv video. De er den tekniske produsent for NAF's TV-10-program. De bygger nå opp et komplett team for interaktiv video.

Magne Velure, De Sandvigske Samlinger, Maihaugen: Museumsetaten har ansvar for store mengder kulturhistorisk materiale og informasjon, men mangler muligheter for å ha en god nok oversikt over dette materialet. Dermed blir det vanskelig å aktivisere og formidle det verdifulle innholdet i magasinene. Maihaugen har ikke konkrete planer om å produsere videoplate, men håper å komme med i et prosjekt basert på bruk av digitale optiske plater. Velure var også opptatt av hvordan interaktiv vi-



Tegning: Øystein Reigem.

deo kan bli brukt i publikumstjenesten.

Manne Lidén, Rikscentralen för pedagogiska hjälpmedel i Särskolan: Sentralen arbeider med læremidler for psykisk utviklingshemmede. Videomediet er anvendelig i denne sammenhengen fordi mange av de psykisk utviklingshemmede ikke kan lese. Gjennom video kan en simulere virkelige situasjoner og dermed på en helt spesiell måte konkretisere undervisninga. En satser på bruk av berøringsfølsomme skjermer og legger vekt på en pedagogisk tilnærming på grunnlag av studier av «videospråkets grammatikk».

Olle Cederbrand, Den Norske Creditbank: En har funnet at vanlig datamaskinstøttet læring virker lite motiverende og bare har begrenset nytte i intern opplæring. Det har derfor vokst frem en interesse for interaktiv video som alternativ til dette. Det er gjennomført et forstudie som konkluderer med at interaktiv video egner seg i undervisning i holdningstrening for skrankepersonale. Et pilotprosjekt starter 1. mars, og en regner med å ha ferdig en videoplate mot slutten av året.

Reidun Oanes Andersen, Det pedagogiske seminar, Universitetet i Bergen: Andersen har fått i oppdrag fra Datasekretariatet å skaffe oversikt over ulike program når det gjelder bruk av interaktiv video i språk-opplæringa (fransk). Hun arbeider nå med å få fatt i videoplater som er produsert for dette formålet i Frankrike.

Lars Skjold Wilhelmsen, Det pedagogiske seminar, Universitetet i Bergen (tidligere ved Audiovisuell avdeling, Universitetet i Bergen): Ved

Det pedagogiske seminar arbeider man med pedagogiske sider ved ulike typer datastøttet undervisning og interaktiv video. Spesielt arbeider man med studier av videomediets dramaturgi, et emne nært knyttet til hva *Manne Lidén* omtalte som videospråkets grammatikk.

Etter denne presentasjonsrunden samlet deltagerne seg til diskusjon rundt følgende tema: (1) erfaringer fra igangværende eller avsluttede prosjekter (2) prosjektplanlegging (3) interaktiv video i undervisning og (4) tekniske spørsmål i tilknytning til utstyr og programvare.

(1) ERFARINGER FRA IGANGVÆRENDE ELLER AVSLUTTEDE PROSJEKTER

Alvar Löfskog demonstrerte programvare for «Gamla Stan»-prosjektet. Programmet inneholder en innledningssekvens som består av (a) demonstrasjon av programmet, (b) beskrivelse av programmet i tekst, (c) introduksjon til hvordan en skal tolke bilder og befolkningsdata. Selve programmet er delt inn i fire avdelinger: (a) en sekvensiell serie om hvordan Gamla Stan er vokst fram, (b) turistvandring rundt i Gamla Stan, fra bilde til bilde, men først og fremst via kart, (c) menneskeskjebner, er velger ut to mennesker og følger deres historie (demografi, boforhold, yrkesvalg etc.) (d) en modul som setter brukeren i stand til å gjøre egne søk og å definere egne ruter gjennom materialet.

Hasse Söderström orienterte om en Sony-workshop i USA hvor Circus Associates demonstrerte et uhyre brukervennlig system for søking i materiale fra en bok om gresk mytologi. En tok utgangspunkt i hvordan er leser orienterer seg i sin egen bokhylle og en hadde satsa på en løsning der leseren kunne bla seg gjennom boken og søke i materialet ved å peke på en berøringsfølsom skjerm. Materialet som var lagt inn var fra Getty's Art Museums. Systemet finnes i to versjoner, – ett med en og ett med to plater. Utviklingskostnadene har vært ca \$250.000. (Circus Associates Interactive Media and Space, Design/Production, 10766 Chushdon Av. Los Angeles, CA 90064.)

(2) PROSJEKTPLANLEGGING

Diskusjonen dreide seg stort sett om muligheter for samarbeid på nordisk plan. Thor Berntsen påpekte at slike samarbeidsprosjekt måtte være

konkrete og anvendelsesorienterte for å ha noen mulighet til å lykkes.

Magne Velure opplyste at videoplateteknologi er aktuell i igangværende nordiske samarbeidsprosjekt innafor museumssektoren, bl.a. en kunnskapsbank om håndverksteknikker.

Erik Meistrup påpekte at de beskjedne midlene som i dag ble avsatt i de nordiske landene nå ble spredt på en rekke prosjekter og ikke konsentrert til noen få tunge prosjekter. Prosjekter som består i tilrettelegging av data eller programvare for fellesnordisk bruk har oftest svært begrenset nytteverdi p.g.a. manglende standardisering både på software- og hardwarensiden såvel mellom de nordiske land som internt i hvert enkelt land. I stedet for slike prosjekter kan det derfor være en god ide å satse på prosjekter der en tilrettelegger bildemateriale til felles bruk på videoplate, ettersom en i så fall ikke vil støte på slike standardiseringsproblemer i samme grad.

Alvar Löfskog meddelte at Nordisk ministerråd vurderer et nordisk Domesday-prosjekt. Datasekretariatet er kontaktpunktet i Norge. Han fremsatte for øvrig et forslag om å arbeide for et fellesnordisk universitetslaboratorium for produksjon av mastertapes.

(3) INTERAKTIV VIDEO I UNDERVISNING

Reidun Oanes Andersen kunne fortelle at Europarådet legger vekt på at språkundervisningen skal være «kommunikativ». Mens man i språkundervisningen tidligere har tatt utgangspunkt i strukturelle aspekter ved språk, skal man nå ta utgangspunkt i at språkbruk er sosiale handlinger som formidler budskap. Undervisningen bør derfor i større grad være kontekstbasert. Interaktiv video gir nettopp mulighet til å simulere språkbruk i tilknytning til de sosiale sammenhenger og konkrete situasjoner hvor den forekommer i det praktiske liv, og mediet er derfor interessant i en slik undervisningssammenheng.

Erik Meistrup understreket at den feil som ofte begås i utvikling av undervisningsopplegg med interaktiv video, er at det legges for lite vekt på en realistisk utprøving av systemene. Utprøvingen må foregå over lang tid og basere seg på bruk overfor personer som har et reelt etterutdanningsbehov.

(4) TEKNISKE SPØRSMÅL I TILKNYTNING TIL UTSTYR OG PROGRAMVARE

UTSTYR

Alvar Löfskog fortalte at man ved «Gamla Stan»-prosjektet har ønske å knytte lyd til stillbilder, men at en ikke har funnet noe tilfredsstillende system for å ivareta denne funksjonen. Fred Sollie opplyste at Sony har utarbeidet et system for lyd-bilde-kombinasjon og at man har kommet så langt at det er grunn til forhåpninger om at dette systemet skal sette en standard. I dette systemet legges digitalisert lyd ut i reserverte frames bildesporet på videoplaten. Man kan velge mellom forskjellige kvaliteter i lydgjengivelsen. 20 sekunder lyd av «telefonkvalitet» pr. bilde gir 40.000 bilder pr. plateside, 40 sekunder gir ca. 20.000 bilder pr. plateside. For avspilling kreves at man har et spesielt kort for dekoding av den digitaliserte lyden. For tilrettelegging av materialet før mastering kan digitale lydbånd benyttes.

Olle Cederbrand opplyste at produksjonsutstyr for PAL-format videoplater i «lollypop»-versjon snart blir tilgjengelig på det kommersielle marked for ca. kr 200.000-300.000. («Lollypop» – en rimeligere, men mindre holdbar plate produsert i ett eksemplar) Fred Sollie bemerket i dette at også utstyr for skrivbare (write-once) videoplater snart vil komme i vanlig salg. Overskrivbare (erasable) videoplater finnes, men medier har vist seg ustabil og så kostbart at det neppe vil være tilgjengelig på det kommersielle marked innen overskuelig fremtid.

Fred Sollie opplyste for øvrig at avspillingsutstyr til bruk i forbindelse med videoplater i interaktiv video-sammenheng ser ut til å bevege seg i retning av en standard bestående av en enskjerms-løsning med IBM-kompatible PC'er og CGA-grafikk.

PROGRAMVARE

Ved «Gamla Stan»-prosjektet er all programvare utviklet i Turbo Pascal ved Stockholms historiska databas. Den baserer seg på EGA-grafikk.

Den Norske Creditbank har prøvd ut en rekke forfatterspråk som ledet i arbeidet med å finne et språk som egner seg for bruk i prosjektet. De vurderer forfatterspråk ut fra følgende hensyn: Systemet skal være tilgjengelig, ha leverandørstøtte, ikke være for dyrt, og det må lett kunne aksepteres av dem som skal bruke det. Til nå har en stort sett benyttet språket Topclass, men løpende vurderinger ser ut til å gå i favør av Pile 2.

Lars Skjold Wilhelmsen har prøvd ut et forfatterspråk utviklet ved Heidelberg Universitet i samarbeid med Sony. Hans erfaringer med dette var overveiende positive, men han reiste likevel spørsmål om forholdet mellom de rammer bruk av slike systemer setter og den noe friere stilling egenutvikling av programvare i vanlige generelle programmeringsspråk setter en i.

Det lot til å være utbredt enighet om at forfatterspråk ofte gir sterke føringer under utvikling av programvare for interaktiv video, den sterke vekten som legges på menybaserte løsninger kan f.eks. gi uheldige resultater. På den annen side er dette noe som må vurderes i hvert enkelt tilfelle, i mange sammenhenger kan bruk av forfatterspråk være meget funksjonelt og arbeidsbesparende.

MELDINGER

TRENGER DU EDB-ASSISTANSE?

NAVFs edb-senter for humanistisk forskning påtar seg betalte oppdrag innenfor sitt arbeidsområde. Oppdrag kan innbefatte bl.a. utviklingsarbeid når det gjelder bruk av edb i humaniora, registrering av forskningsmateriale, installasjon av programvare og produksjon av laserutskrift og fotosats. Nærmere opplysninger fås ved henvendelse til: *Direktør Jo Stein H. Hauge, NAVFs edb-senter for humanistisk forskning, Boks 53 Universitetet, 5027 Bergen. Tlf. 05-212954/55/56.*

ELEKTRONISKE NYHETER FRA SENTERET

NAVFs edb-senter for humanistisk forskning skal sette i gang en elektronisk nyhetsformidlingstjeneste. Ca. én gang i måneden kommer vi til å distribuere meldinger via elektronisk post til alle interesserte som har nettsadresse. De fleste meldinger som trykkes i Humanistiske Data vil inngå i systemet, i tillegg til nyheter som krever en rask kunngjøring. Annen informasjon om f.eks. programvare og utstyr blir også inkludert. Bidrag fra personer utenfor Senteret ønskes velkommen.

Meldingene vil for øvrig også bli tilgjengelige på HUMKOM (se HI 2-86, s. 127).

Er du interessert i å bli abonnent, send nettsadressen din enten skriftlig til Kristin Natvig, eller via elektronisk post til Knut Hofland FAFKH@NOBERGEN.EARN.

PROGRAMMEVALUERING

I hvert nummer av Humanistiske Data vil vi nå prøve å gi en vurdering av forskjellige programmer. (I HD 3-86 vurderte Torrey Seland *NotaBene*). Ved NAVFs edb-senter er det anskaffet programmer som Senterets personale vil vurdere, men det kjøpes også programmer andre steder (!), og dette er herved en oppfordring til miljøene om å komme frem med sine erfaringer.

Vi har foreløpig satt opp en liste over punkter vi mener en slik evaluering bør inneholde:

- En generell beskrivelse av programmet og hensikten med det slik dette er oppgitt i reklame og/eller dokumentasjon.
- En beskrivelse av utprøvingen og de erfaringene man sitter igjen med.

Under dette siste kan følgende være interessant:

Hva slags maskinutrustning krever programmet?

Pris, produsent, leverandør

Hvor vanskelig var det å komme i gang (installere)?

Hvor vanskelig var det å bruke programmet etter hvert?

Hvorledes er dokumentasjonen?

Driftssikkerhet. (Gikk maskinen i stå?)

Oppfyller programmet løftene fra reklame/dokumentasjon?

Hvorledes er programmet sammenlignet med lignende programmer?

En subjektiv oppsummering av erfaringene

Denne listen er ikke ment å føles som en tvangstrøye, det er f.eks. lite interessant å prøve å sammenligne et program med tilsvarende programmer dersom man ikke har brukt slike tidligere. I alle fall er det bedre med en kort notis som informerer om at programmet eksisterer enn ingenting i det hele tatt. Det finnes utrolig mange programmer tilgjengelig, og enhver veiviser i jungelen vil være av det gode.

ØNSKES: BIDRAG OM EDB I MUSIKKFAGET

Edb i musikk – både forskning og undervisning – har vært et forsømt felt i Humanistiske Data. Redaksjonen ønsker å bøte på dette i framtidige nummer – men vi trenger lesernes hjelp til å skaffe stoff. Vi ta gjerne imot bidrag av alle typer og omfang – ideer og tips, meldinger, rapporter og artikler. Bidrag kan f.eks. gjelde litteratur om feltet, programvare, utstyr, prosjekter, praktiske forsøk, faglige og pedagogiske overveieringer, m.m. Både ferdige bidrag og stoff som vi selv kan redigere ønskes velkommen. Ring eller skriv til redaktøren: *Kristin Natvig NAVFs edb-senter for humanistisk forskning, Boks 53 – Universitetet 5027 Bergen. Tlf. 05-212954/55/56.*

THE TAGGED LOB CORPUS – USERS' MANUAL

A users' manual for the tagged version of the Lancaster-Oslo/Berger (LOB) Corpus is now available. The manual has been prepared by Stig Johansson in collaboration with Eric Atwell, Roger Garside, and Geoffrey Leach. Price: 100 N.kr. Orders may be placed with: *The Norwegian Computing Centre for the Humanities, P.O. Box 53 – Universitetet N-5027 Bergen, Norway.*

NY PROSJEKTOVERSIKT OVER HUMANISTISK FORSKNING

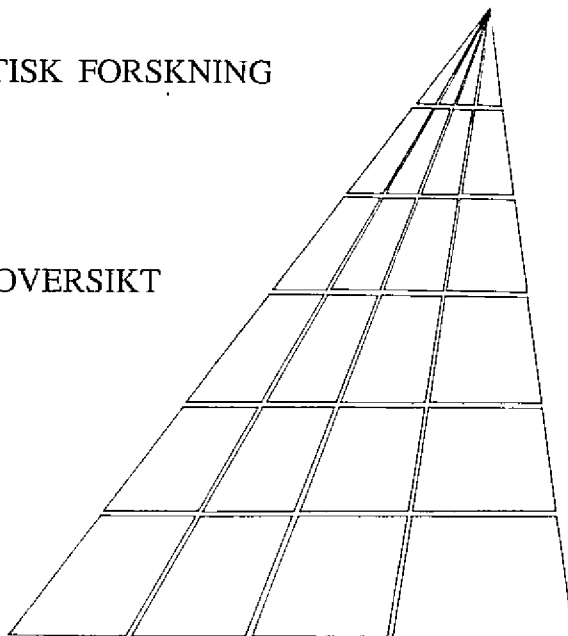
NAVFs informasjonstjeneste for forskningsprosjekter har gitt ut en katalog over nylig avsluttede og igangværende forsknings- og utviklings-

HUMANISTISK FORSKNING

I NORGE

PROSJEKTOVERSIKT

1985-86



NAVF's INFORMASJONSTJENESTE FOR FORSKNINGSPROSJEKTER

NAVF

prosjekter innenfor de humanistiske fag i Norge. Oversikten gir kortfattede opplysninger om 1006 prosjekter, som også fins i en mer utførlig versjon i Informasjonstjenestens prosjektdatabase. Katalogen dekker prosjekter finansiert av universiteter, høyskoler og andre forskningsinstitusjoner i perioden 1985-86.

Katalogen er inndelt etter de humanistiske fagområder og har en egen avdeling for tverrfaglige prosjekter. Det gis opplysninger om bl.a. prosjektittel, institusjon, prosjektansvarlig og medarbeidere, fagområde(r), faglig innhold, og hvorvidt det fins publikasjoner i tilknytning til prosjektet. Det fins også person-, emneord- og institusjonsregistre.

Prosjektkatalogen koster kr 100,- og kan bestilles fra nedenstående adresse.

Prosjektdatabasen inneholder mer detaljerte opplysninger om prosjektene og tilknyttede publikasjoner. I en prøveperiode er bruk av databa-

sen gratis. Brukere som ikke selv ønsker å søke i databasen, kan få hjelp av Informasjonstjenesten til å ta ut spesialutskrifter av prosjektinformasjon.

Interesserte kan henvende seg til: *NAVFs informasjonstjeneste for forskningsprosjekter, Boks 53 – Universitetet, 5027 Bergen.*

CORRECTION

Unfortunately, illustration no. 4 on page 39 in *Humanistiske Data* no 3-86 was a misprint. Correct illustration below:

51

\ 'Die Beschreibung der *Phänomene* der Farbenblindheit 24 gehört zur Psychologie.³ Also auch die der Phänomene des normalen Farbsehens? Gewiß, - aber was nennt man "Phänomene des normalen Farbsehens", was ihre Beschreibung, und für wen ist sie bestimmt? Oder⁴ besser: welches Hilfsmittel bedient sie sich? Wenn ich sage "Was setzt sie voraus? ", so heißt das: "Wie muß einer auf *sie* schon reagieren, um sie zu *verstehen*." Jene erste Beschreibung, an die ich dachte, war eine mittels der Begriffe der Sehenden.⁵

52

Dieses Papier ist an verschiedenen Stellen verschieden hell; aber kann ich sagen, es sei nur an gewissen Stellen weiß, an den \ andern aber grau?? - Ja, wenn ich es malte,⁶ 25

1 Hinzufügung im oberen Rand, links: Ich kann sagen: Dieser Mann unterscheidet nicht Rot und Grün; Kann ich aber sagen: Wir Normalen unterscheiden Rot und Grün? Wir könnten aber sagen: "Wir sehen *hier* 2 Farben, jener nur *eine*."

2 Hinzufügung im oberen Rand, rechts: 13

3 Alternative: [ist eine psychologische / gehört [in die / zur] Psychologie]

4 Alternative: [aber was setzt so eine Beschreibung voraus, für wen ist es eine Beschreibung, oder / aber was nennt man "Phänomene des normalen Farbsehens", was ihre Beschreibung, und für wen ist sie bestimmt? Oder]

5 Alternative: [Wer in einem Buch die Phänomene der Farbenblindheit beschreibt, beschreibt sie mit den Begriffen der Sehenden / Jene erste Beschreibung, an die ich dachte, war eine mittels der Begriffe der Sehenden]

6 Überstrichen: so

DOVRE DATA ARKIV

For vel ett år siden ble Dovre Data Arkiv (DDA) opprettet for en prøveperiode. Arkivet har flere formål: å registrere og databehandle historiske kilder som folketellinger og kirkebøker fra 1800-tallet for kommuner i Hedmark og Oppland, å produsere lett leselige og oversiktlige utgaver av originalkildene samt alfabetiske registre til disse, å utarbeide statistikk for ulike områder innen kommunene, og å drive forsøk med lenking av kildene. Arbeidet vil være en del av oppbyggingen av en fullstendig persondatabase fra Norge på 1800-tallet.

For tida foregår registrering av folketellingene for 14 kommuner, mens 35 gjenstår. Registrering av kirkebøkene er påbegynt for 2-3 kommuner.

Bakgrunnen for opprettelsen av arkivet er dels at Dovre kommune var interessert i å skape flere arbeidsplasser, dels et forsøk ved Hjerleid videregående skole på å bygge opp en lokalhistorisk database til bruk i undervisning, og endelig at Opplandsarkivet var interessert i et pilotprosjekt om innsamling og ordning av lokalhistoriske privatarkiver i én kommune.

DDA sysselsetter 1 leder og 5 registratorer og ei styringsgruppe har det overordnede ansvaret. Arkivet samarbeider med bl.a. Registreringssentralen for historiske data i Troms, og registreringsopplegget er kompatibelt med RHDs.

I tillegg til registreringsoppdrag for kommunene i Hedmark og Oppland, har DDA følgende tilbud: disketter med tekstfiler og ferdig innlagte databaser i RUBREG, og kopier av originalkildene.

Forespørsler rettes til: *Dovre Data Arkiv, Boks 2, 2660 Dombås. Tlf. 062-41364.*

NORSK ONLINE BRUKERGRUPPE – NOLUG

Norsk Online Brukergruppe (NOLUG – Norwegian Online User Group) er en interessegruppe for brukere av online informasjonsbaser.

NOLUG er en interessegruppe under Norsk Dokumentasjonsgruppe, et uavhengig forum for personer med tilknytning til norsk informasjons- og litteraturformidling.

Brukergruppen har til formål å fremme medlemmenes felles interesser som brukere av online systemer, herunder:

- a. arbeide for erfaringsutveksling mellom medlemmene
- b. være initiativtaker til kurs og møter
- c. representere medlemmenes interesser på området utad, nasjonalt og internasjonalt, så som deltakelse i EUROLUG (den europeiske brukergruppen), kanalisere henvendelser til databaseprodusenter, verter og nettverk
- d. være referansegruppe for Norsk DIANE-senter (servicesenter for databasesøking).

Brukergruppen vil holde 2-3 møter i halvåret.

Foreløpig dekkes medlemskapet i NOLUG av medlemsavgiften i Norsk Dokumentasjonsgruppe, for tiden kr. 100 pr. år for personlige medlemmer. Det kan bli aktuelt med en mindre tilleggskontingent for Norsk Online Brukergruppe.

Arbeidsutvalget som leder brukergruppen, er for tiden:

Per B. Kristensen

Norsk Hydros Forskningscenter, Porsgrunn

Ane Cecilie Røed

Axess Information Services, Oslo

Jon Anjer

Universitetsbiblioteket i Oslo

Er du interessert i medlemskap? Kontakt *Per B. Kristensen, Norsk Hydros Forskningscenter, P.B. 110, 3901 Porsgrunn. Telefon 035-51120/ linje 2765.*

NYE STUDIETILBUD I DATA- FAG/INFORMASJONSTEKNOLOGI

Fra høsten 1987 vil Nord-Trøndelag distriktshøgskole ha to studietilbud innenfor datafag/informasjontechnologi – et toårig studium med opp-tak fra videregående skole og et påbyggingsstudium for høgskolekandida-ter.

Formålet med det toårige studiet er å utdanne høgskolekandidater som vil være godt kvalifisert til oppgaver innenfor informasjonsbehandling i den offentlige og private sektor. Det toårige studiet tar utgangspunkt i

en innføring i systemering, programmering og databaser, med fordypning innenfor emnene beslutningsstøtte/ekspertsystemer, simulering og samfunnsinformatikk. Innenfor samfunnsinformatikk blir samfunnsvitenskapelige vurderinger av informasjonsteknologi drøftet, og konsekvensene av informasjonsteknologi blir belyst gjennom en analyse av endringsprosesser i samfunnet.

Det ettårige studiet er et spesialstudium, med formålet å gi studentene en grundig innsikt i viktige områder innenfor administrativ databehandling.

Flere opplysninger kan fås fra: *Nord-Trøndelag distriktshøgskole, Postboks 145, 7701 Steinkjer.*

SCANNET

SCANNET ble etablert for å fremme bruken av nordiske online databaser. Prosjektet blir finansiert av NORDINFO, Nordiska samarbetsorganet för vetenskaplig information. Styret består av ett medlem for hvert av de nordiske landene.

SCANNETs formål er å gjøre lettere tilgjengelig og oppmuntre til effektiv bruk av nordiske edb-baserte informasjonstjenester og resursser. Hovedoppgavene består i å: formidle informasjon om nordiske databaser, gi råd om bruk av disse, arbeide for koordinering av online systemer og kommandospråk, gi råd om planlegging og oppbygging av databaser, og utvikle og støtte prosjekter om nordiske databaser.

Det var SCANNET som tok initiativet til *Nordisk databasguide 1985*, som ble omtalt i *Humanistiske Data* 3-86. Meldingsbladet *SCANNET TODAY* utgis to ganger i året og distribueres gratis.

SCANNETs øvrige tjenester er også gratis. Problemer, spørsmål, ideer og ønsker vedr. nordiske databaser kan rettes til koordinatoren: *Elisabet Mickos, SCANNET, c/o Helsinki University of Technology Library, Ot-näsvägen 9, SF-02150 Esbo, Finland.*

ÅRHUS AMTS INFORMATIKCENTER (AAIC)

AAIC er et forsøksprosjekt som er opprettet på initiativ av Århus Amtsråd. Senterets formål er å sikre at lærere, ledere og lokale skolemyndigheter får mulighet til å følge utviklingen innenfor edb og undervisning og til selv å utvikle undervisningsprogrammer av høy kvalitet. Ved senteret foregår bl.a. programutvikling og bearbeiding og vurdering av eksisterende programvare, også gjennom utprøving i konkrete undervisningssituasjoner.

Blant senterets aktiviteter er en pågående undersøkelse av forskjellige muligheter vedr. anvendelsen av eksterne informasjonsbaser, med tanke på opprettelse av en egen base for skolene. Senteret er koblet til informasjonsbaser og konferansesystemer som alle lærere har anledning til å utprøve.

Fleire programutviklingsprosjekter som er initiert av lærere er blitt støttet og koordinert. Disse omfatter: et diktatprogram til anvendelse i stave- og lydtraining med delvis bruk av båndopptaker, et program til blindeskrifttraining med tanke på undervisning i tekstbehandling og maskinskriving, et kommunikasjonsprogram for funksjonshemmede, og et simuleringssystem som kan framstille og analysere modeller til bruk i undervisning på flere forskjellige fagområder.

For øvrig låner AAIC ut maskiner og programvare, foretar rådgiving og konsulentvirksomhet, og holder foredrag og kurs.

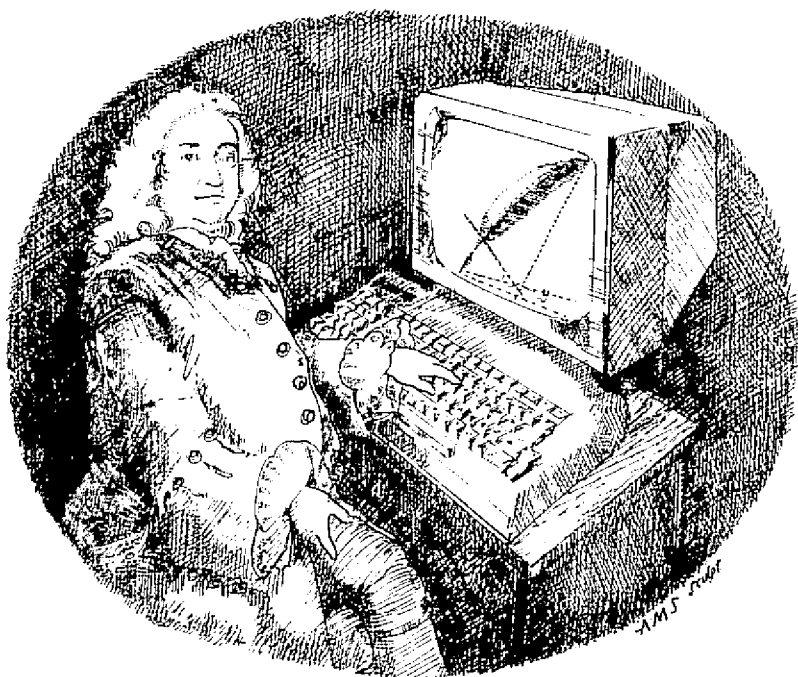
Adresse: *Århus Amts Informatikcenter, Halmstadgade 6, DK-8200 Århus N, Danmark.*

THE ASSOCIATION FOR HISTORY AND COMPUTING

The Association for History and Computing, en internasjonal organisasjon, er blitt opprettet i Storbritannia for å fremme bruken av edb i historieforskning og -undervisning. Organisasjonen holder årlige konferanser,

COMPUTING *and* HISTORY TODAY

AHC NEWSLETTER



NO 1

February 1987

utgir det fyldige meldingsbladet *Computing and History Today*, og har initiert både flere store samarbeidsprosjekter og undervisningsopplegg.

Medlemskap koster £8 pr. år (£5 for studenter) og gir bl.a. adgang til rabatt på foredragssamlingen *History and Computing* (Manchester University Press, 1987). Interesserte kan henvende seg til: *The Association*

for History and Computing, Peter Denley, Dept. of History, Westfield College, University of London, London NW3 7ST, UK.

COMPUTERS AND MUSIC RESEARCH

Et system er etablert i England for distribusjon av meldinger om musikkforskning via elektronisk post. Det er meningen at systemet skal fungere som en møteplass for alle som arbeider med edb-anvendelser innenfor musikkforskning.

Alle som har tilgang til internasjonale datanettverk kan bidra med materiale for videre distribusjon. Bidrag kan bestå av ideer/ønsker/problemer, prosjektbeskrivelser, utdrag/sammendrag av publisert eller upublisert materiale, osv. Det er selvsagt også mulig å bare være «passiv» deltaker.

Nærmere opplysninger og innmeldingsskjema fås ved henvendelse til: *Stephen Page, Eclipse Data Systems, New College, Oxford OX1 3BN, UK.*

INSTITUT NATIONAL DE LA LANGUE FRANÇAISE (INALF)

INALF består av 10 forskningsenheter i Frankrike som utfører hver sine forskningsprogrammer i tilknytning til fransk språk og lingvistikk. Instituttet har 250 forskningsmedarbeidere og er dermed antakelig Europas største forskningssenter innenfor humaniora. Det var også ett av de første sentrene til å anvende edb i humaniora i Frankrike.

Forskningsenhetene arbeider med bl.a. en ordbok over fransk språk i det 19. og 20. århundre, leksikologi og politiske tekster, undervisning av moderne fransk, og gammelfransk. De deler edb-resursser og benytter seg av et sentralt dokumentasjonssenter, Trésor Général des Langues et Parlers Française (URL 1).

URL 1 har tre seksjoner som arbeider på ulike forskningsfelt. Den største seksjonen, i Nancy, skaffer eksempler og dokumentasjon for ordboka nevnt over, den andre vedlikeholder en edb-lagret historisk fortegning over franske ord, mens den tredje konsentrerer seg om moderne fransk og registrerer neologismer både i Frankrike og utenlands. Seksjonene samarbeider også om fellesprogrammer vedr. generell dokumentasjon og bibliografier.

Underavdelingen av URL 1 i Nancy, Centre de Recherche Documentaire (CRD), utfører også dokumentasjonstjenester for eksterne brukere i Frankrike og utenlands. Senterets tre hovedaktiviteter består for øvrig av: 1. Etablering av en tekstdatabase som består hovedsakelig av skjønnlitterære tekster fra 16.-20. århundre samt tilhørende konkordanser og indekser, 2. Utvikling av en leksikografisk database som dekker historisk og moderne fransk, 3. Utvikling av en bibliografisk database over fransk språk og lingvistikk.

Redigering av ordboka *Trésor de la Langue Française* (TLF) foregår også i Nancy, men ved et av INALFS forskningsenheter. TLF er ei historisk ordbok som dekker perioden 1789-1964 i 16 bind. Ti av bindene er utgitt hittil og hele prosjektet skal ferdigstilles i 1991. Ordboka er basert på databasene utviklet ved CRD.

ARCHIVES AND MUSEUMS INFORMATICS

Amerikanske museer og arkiv uten økonomiske bekymringer har fått et nytt og sikkert utmerket tilbud i form av *Archives and Museum Informatics*, et firma som ble opprettet i 1986. Firmaet yter konsulenttjenester vedr. informasjonsbehandling ved arkiv og museer (inkl. valg og implementering av edb-systemer), utfører forskning på oppdragsbasis om anvendelser av informasjonsteknologi innenfor sektoren, utgir *Quarterly Update on Archival Informatics* (se *Aktuelle tidsskrifter*) og tilbyr (rådyre!) kurs og seminar om emner innenfor edb i arkiv og museer.

Flere opplysninger kan fås fra: *David Bearman, Archives and Museums Informatics, 5600 Northumberland Street, Pittsburgh, Pennsylvania 15217, USA.*

GRESKE VASER PÅ VIDEOPLATE

Et interaktivt program om greske vaser er blitt produsert av firmaet Interac for The J. Paul Getty Museum i USA. Programmet har vært tilgjengelig for besøkende siden august 1986. Utstyret består av en 13" berøringsfølsom skjerm, en NCR datamaskin med harddisk og to videoplatespillere. Stillbilder og levende film fyller hele platen sammen med én time lyd og tekstinformasjon. Kombinasjonen av to videoplatespillere tillater utstrakt bruk av lyd sammen med stillbildene og reduserer aksestida til hvert bilde. Over 100 av museets greske vaser er avbildet.

Programmet består av seks deler og starter med en introduksjon til greske vaser. Deretter følger fire hoved«forgreininger» i programmet: emner for dekoren, utførelse, form og funksjon, og kunstnerisk kvalitet. Hver forgreining gir brukeren anledning til å velge ulike typer fordypninger.

Gjennom hele programmet har brukeren adgang til en hovedmeny på den ene av spillerne som via en vindusfunksjon kan styre bildene på den andre spilleren. Denne funksjonen virker sammen med farge- og symbolkodete felt på skjermen på en slik måte at nye valg kan foretas uten at bildet forsvinner fra skjermen.

En av programmets finesser gjør det mulig å betrakte fire av vasene i detalj. På skjermen vises en roterende vase i helbilde samtidig med et synkronisert nærbilde. Brukeren kan når som helst stoppe og starte roteringen og få tilgang til forstørrede detaljer ved å berøre skjermen på det ønskede sted. Det er meningen at brukerens kontroll over bildene skal simulere en situasjon hvor en tar vasen i hendene og undersøker den selv.

AKTUELLE TIDSSKRIFTER

WINDOW ON THE HUMANITIES

Dette meldingsbladet utgis av Center for Computer Applications in the Humanities, University of Nevada. Bladet inneholder nasjonale og internasjonale nyheter, anmeldelser av bøker og programvare, meldinger om

kommende konferanser og tips for Macintosh-brukere.

Window on the Humanities utgis nokså uregelmessig, men til gjengjeld distribueres det gratis. Adresse: *Center for Computer Applications in the Humanities, University of Nevada, Las Vegas, Nevada 89154, USA.*

MDA INFORMATION

MDA Information utgis av the Museum Documentation Association i Cambridge. Bladet inneholder artikler, rapporter og meldinger av relevans for museumsfaglig dokumentasjonsarbeid. Det utgis kvartalsvis og koster £5 + porto pr. år for ikke-medlemmer. Adresse: *The Museum Documentation Association, Building O, 347 Cherry Hinton Road, Cambridge CB1 4DH, England.*

QUARTERLY UPDATE IN ARCHIVAL INFORMATICS

Denne nye publikasjonen består av to deler, *Archival Information Newsletter* og *Archives and Museums Technical Report*, som inneholder omfattende tekniske rapporter i form av monografier. Det er mulig å abonnere bare på meldingsbladet (\$30 pr. år), eller begge delene sammen (\$180 pr. år).

Archival Information Newsletter tar opp bruk av datateknologi i arkiv og museer i form av artikler, prosjektbeskrivelser og kritiske rapporter. Det gis også omtaler av nye produkter, litteratur og kommende konferanser osv. av interesse for arkiv- og museumsfaglig personale. Bladet inneholder i tillegg sammendrag av rapportene som publiseres i *Archives and Museums Technical Report*.

Quarterly Update in Archival Informatics utgis av: *Archives and Museums Informatics, 5600 Northumberland Street, Pittsburgh, Pennsylvania 15217, USA.*

Se for øvrig egen melding om *Archives and Museums Informatics*.

COMPUTING FOR THE HUMANITIES

Dette meldingsbladet (må ikke forveksles med tidsskriftet *Computers in the Humanities*) fyller kun fire sider, men det utgis annenhver måned og distribueres gratis. Adresse: *College of Arts and Sciences, Cornell University, Ithaca, New York 14853, USA.*

LEONARDO

Leonardo er et internasjonalt tidsskrift som særlig omhandler syntese av kunstartene, vitenskap og teknologi. Fokus er på bildende kunst, men også media, musikk, utøvende kunst og språk får bred omtale. Tidsskriftet inneholder artikler, anmeldelser av bøker og programvare og meldinger fra utgiveren, The International Society for the Arts, Sciences and Technology.

Leonardo utgis kvartalsvis og koster DM 112 pr. år for enkeltpersoner. Nye abonnenter får 25% rabatt på tidligere utgivelser. Gratisnummer og flere opplysninger fås fra: *Pergamon Journals Limited, Headington Hill Hall, Oxford OX3 0BW, England.*

JOURNAL OF THE INTERNATIONAL SOCIETY

LEONARDO

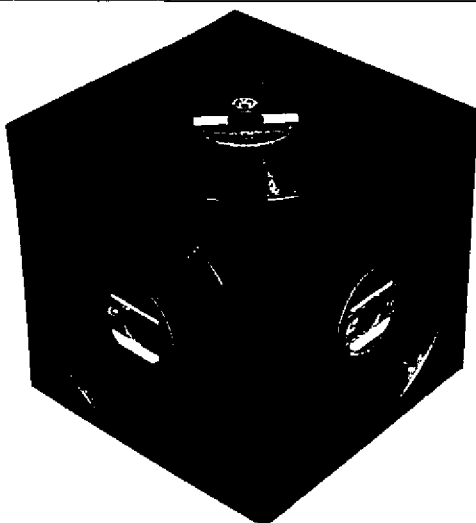
FOR THE ARTS SCIENCES AND TECHNOLOGY

KONFERANSER

TEKST – LYD – BILDE

NAVFs edb-senter for humanistisk forskning skal holde en konferans om bruk av optiske lagringsmedier i forskning, undervisning og formidling i Bergen 14.-16. oktober 1987.

På konferansen, som arrangeres i et faglig samarbeid med Pedagogis seminar, Universitetet i Bergen, vil vi både ta opp det tekniske grunnlaget for analoge og digitale laserplater m.v. og i foredrag og demonstrasjoner vise bruk av den nye teknologien.



*Tekst
Lyd
Bilde*

Konferanse om bruk av optiske lagringsmedier i forskning, undervisning og formidling

Bergen 14.-16. oktober 1987

- Anvendelser i
 - undervisning
 - dokumentasjon
 - fag- og forskningsarbeid
(med hovedvekt på humaniora)
- Prosjektplanlegging
- Pedagogiske metoder/programvarer
- Tekniske spesialsesjoner



Nærmere opplysninger hos
NAVF's edb-senter for humanistisk forskning
Harald Hårfagresgt. 31, Boks 53, Universitetet
5027 Bergen
Tlf. (05) 21 29 54/55/56

Konferansen vil ha sitt fokus på anvendelser innenfor humanistisk fag, men en rekke av de emner som tas opp, vil være av stor interesse også for andre fagfelt, f.eks. undervisningssektoren. Vi ønsker derfor deltakere også fra andre fag- og forskningsfelt velkommen.

For at deltakerne skal få en internasjonal oversikt over arbeid på feltet, er det invitert en rekke gjesteforelesere fra USA og europeiske land. I tillegg vil eksperter fra vårt eget land delta.

Flere datamaskinleverandører og spesialfirma for nye medier vil være representert med utstyr og gi en presentasjon av sine tjenestetilbud.

Det er alt en rekke fag- og forskningsmiljøer i Norge som har tatt bruk eller som planlegger å nytte de nye optiske medier i prosjektarbeid. For dem vil det bli arrangert egne sesjoner hvor planlagte og pågående prosjekt kan presenteres og drøftes.

Deltakeravgift er kr. 700,- for deltakere fra offentlige fag- og forskningsmiljøer og kr. 1200,- for andre. Innkvartering på konferansehotell (Scandic Hotel) koster kr. 1440,- for to døgn.

Frist for påmelding er 1. september. Rammeprogrammet og flere opplysninger kan fås ved henvendelse til: NAVFs edb-senter for humanistisk forskning, Postboks 53 - Universitetet, 5027 Bergen. Tlf. (05) 212954/5:56.

CIDOC ANNUAL MEETING – MDA STUDY TOUR – MDA ANNUAL CONFERENCE

The Documentation Committee of the International Council of Museums (CIDOC) skal avholde årsmøte i Cambridge 20.-23. september 1987. Møtet er åpent også for interesserte som ikke er CIDOC-medlemmer. I tillegg til to plenumsesjoner, skal det holdes «working sessions» knyttet til CIDOCs ulike arbeidsgrupper.

I tida 23.-26. september arrangerer the Museum Documentation Association (MDA) en studietur til seks museer i London og Cambridge. På museene vil deltakerne få demonstrert både manuelle og datamaskinell teknikker for håndtering av samlinger.

Like etterpå, 26.-29. september, skal MDA holde i Cambridge sin første årlige konferanse, om «Collections Management for Museums». Ekspertene fra museer og støtteinstitusjoner i Europa, USA og Australia skal holde foredrag innenfor dette temaet. Det vil også finne sted en utstilling over integrerte edb-systemer.

Nærmere opplysninger om alle disse arrangementene fås ved henvendelse til: *D. Andrew Roberts, Secretary, The Museum Documentation*

Association, Building D, 347 Cherry Hinton Road, Cambridge CB1 4DH, UK.

SUMMER SCHOOL FOR INFORMATION TECHNOLOGY IN THE HUMANITIES

I uka 21.-25. september 1987 skal Faculty of Humanities ved Polytechnic of North London holde en serie dagskurs innenfor feltet edb i humaniora. Emnene som dekkes, er: Oxford Concordance Program, tekstbehandling, kvantitative metoder i analysen av litterære tekster, historiske databaser, maskinassistert språkundervisning, Prolog og Logic programmering, on-line informasjonsgjenfinning, edb-systemer, kommunikasjon i informasjonssystemer, og edb og lingvistik.

Kursene krever ingen forhåndskunnskaper i edb, og koster £25 pr. dag. Flere opplysninger kan fås fra: *Dr. Noel Heather, Humanities Computing Centre, Polytechnic of North London, Faculty of Humanities, Prince of Wales Road, Kentish Town, London NW5, UK.*

INTERNATIONAL CONFERENCE ON COMPUTERS IN THE HISTORY CLASSROOM

Denne konferansen skal holdes ved University of Leeds 6.-8. juli 1988. Formålet er tredelt: 1. å gi en introduksjon til anvendelsen av edb i historieundervisning, 2. å gjøre opp status for utviklingen på feltet samt utveksle og formidle ideer og informasjon, 3. diskutere tilgjengelig programvare. Hovedtemaene vil være: utvikling av pensum, undervisningsmiljøet, og design og evaluering.

Arrangørene ønsker foredrag innenfor disse temaene som er knyttet til undervisning både i grunn- og videregående skole og i høyere utdanning. I tillegg til plenumssesjoner vil programmet omfatte bl.a. seminarer, demonstrasjoner og «workshops».

Sammendrag av foredrag sendes til: *Allan Martin, CHC 88 proposals, School of Education, University of Leeds, Leeds LS2 9JT, UK* innen 31. desember 1987.

Den 12. internasjonale konferansen om datalingvistikk skal holde 22.-27. august 1988 i Budapest. Arrangørene ønsker foredrag om all aspekter ved datalingvistikk. Det skal også arrangeres «workshops umiddelbart før selve konferansen.

Sammendrag av foredrag sendes til: *Dr. Eva Hajicova, Charles University, Faculty of Mathematics, Linguistics, Malostranské n. 25, 118 00 Praha 1, Czechoslovakia* innen 10. desember 1987.

NYTT I BIBLIOTEKET

EDB-EMNER

Hendley, Tony: Video discs, compact discs and digital optical disk systems. Cimtech, Hatfield, 1985. 208 s.

Huxham, Fred A./Burnard, David/Takatsuka, Jim: Using the Macintosh toolbox with C. SYBEX, Berkeley, 1986. 559 s.

Inside Macintosh. Volumes I, II, III. Addison-Wesley, Reading, Mass 1985.

Inside Macintosh. Volume IV. Apple Computer Inc., Addison-Wesley Reading, Mass. 1986. 326 s.

Mathews, Keith: Assembly language primer for the Macintosh. Plume Waite, New York, 1986. 412 s.

Parsloe, Eric: Interactive video. Sigma Technical Press, Wilmslow, 1984 290 s.

Schwerin, Julie B. et. al.: CD-ROM standards: The book. Learned Information, Oxford, 1986.

Smith, Hugh T.: IFIP WG 6.5 Working conference on computer-based message services. Nottingham, 1984. North-Holland, Amsterdam 1984. 340 s.

Uhlig, Ronald P.: IFIP TC-6 International symposium on computer message systems 6. Washington, D.C., 1985. North-Holland, Amsterdam 1984. 425 s.

EDB I SKOLEN

- Adams, Anthony/Jones, Esmor: Teaching humanities in the microelectronic age. Open University Press, Milton Keynes, 1983. 150 s.
- Handlingsprogrammet «Programvara i skolan i Norden». Rapport från den nordiska kursen i programdesign Uddevalla 1986. Nordisk ministerråd, København, 1986.

SPRÅK OG LITTERATUR

- Maegaard, Bente/Ruus, Hanne: Hyppige ord i danske aviser, ugeblade og fagblade. Gyldendal, København, 1986. 425 s.
- Picken, Catriona (red.): Translation and communication: translation and the computer 6: proceedings of a conference. London 1984. Aslib, London, 1985. 190 s.

DIVERSE

- Berkowitz, Luci/Squitier, K.A.: Thesaurus linguae graecae. Canon of Greek authors and works. Oxford UPR, New York, 1986. 341 s.
- Bore, Ove Magnus: De antikvariske myndigheters bruk av SEFRAK-registeret. Drift av kulturminneregisteret. On-line ajourføring. Innstilling fra en arbeidsgruppe. SEFRAK, 1986. 45 s.
- Brachman, Ronald J./Levesque, Hector J.: Readings in knowledge representation. Morgan Kaufmann, Los Altos, Cal., 1985. 571 s.
- Hart, Anna: Knowledge acquisition for expert systems. Kogan Page, London, 1986. 180 s.
- Lönnqvist, Harriet: Nord IoD 6. Helsingfors 1985. Information resources management. 6. nordiska IoD-konferansen. Samfundet för informationstjänst i Finland, Helsingfors, 1985. 457 s.
- Magnusson, Cecilia/Torvund, Olav: EDB, kontrakter og opphavsrett. TANO, Oslo, 1986. 197 s.
- Rudall, B.H./Corns, T.N.: Computers and literature. A practical guide. Abacus Press, USA, 1987. 129 s.

ARTIKLER/RAPPORTER I 1986

ARKIVVERKET

1-86

Randsborg, Hege Britt: Forskerstipend i «Automatisert arkivinformatjon»

Thime, Torkel: Arkivregistrering og edb i de nordiske lokalarkiver

2-86

Fonnes, Ivar: Edb-oppbygging og edb-virksomhet i det offentlige arkivverk

EDB I UNDERVISNING

1-86

Brekke, Magnar: Engelsk grammatikk med nogo attåt

Eritslund, Alf Gunnar: Edb og norskundervisninga i grunnskolen

2-86

Toubro, Charlotte: Educational Uses of Existing Linguistic Software

3-86

Johansen, Harald: Edb i samfunnsfag

Natvig, Kristin: Datasekretariatet og humanistiske fag

Natvig, Kristin: Kurs i design av pedagogisk programvare

Ore, Espen S.: Datasekretariatets seminar i Halden

EDB-TEKNOLOGI

2-86

Reigem, Øystein: CD-ROM – en halv gigabyte i lommen

3-86

Ore, Espen S.: Datakommunikasjon

MUSEUMSFAG

1-86

Pedersen, Margrethe: Kunsten at registrere kunst
Vestbøstad, Per: Årskonferansen til Museum Computer Network 1985

2-86

Østby, Jon Birger: Status for edb-arbeidet i de norske kunst- og kulturhistoriske museer
Div. forfattere: Nordisk edb-konferanse for kunst- og kulturhistoriske museer

SPRÅK

1-86

Hofland, Knut/Reigem, Øystein: De femte nordiske datalingvistikkdage-
ne
Slethei, Kolbjørn: Tekst-til-tale-syntetisering

2-86

Hofland, Knut: ICAME 7th

3-86

Hauge, Jostein H.: Nordisk seminar om maskinoversettelse
Helleland, Botolv: «Norsk stadnamnbase»
Johansson, Stig: Machine-Readable Texts in English Language Research
Slethei, Kolbjørn: IBM-seminar om taleprosessering

DIVERSE

1-86

Hauge, Jostein H.: Edb i humaniora – i fortid og fremtid
Hauge, Jostein H.: Oppvekst i den elektroniske framtid
Huitfeldt, Claus: Det norske Wittgensteinprosjektet
Myhre, Bjørn T.: Edb og veving
Natvig, Kristin: Edb-tilbud for humanister i Storbritannia

2-86

Brændeland, Asbjørn: Utviklingsseminaret på Ustaoset april 1986
Hauge, Jostein H.: Training the Information Researcher for the Future
Natvig, Kristin: «Project Emperor-I» – kinesiske skatter på videoplate
Ore, Espen S.: XVIII International Congress of Papyrology

3-86

Fagertun, Fredrik: «Kommunikasjons- og datateknologi som infrastruktur i nord-norsk næringsliv og forvaltning»
Fjornes, Erik: Humanistiske fag og kunnskapsbaserte systemer
Huitfeldt, Claus: Computers and Philosophical Manuscripts
Natvig, Kristin: Humanistiske databaser
Natvig, Kristin: Senterets videoplateprosjekt – en statusrapport
Seland, Torry: nota bene

SUMMARY

PÅ KRYSS OG TVERS. EDB-REGISTRERING FOR TVERRFAGLIGE ANALYSER AV ARKEOLOGISKE FUNN AV MENNESKESKEJLETTER

COMPUTER ENTRY FOR INTERDISCIPLINARY ANALYSIS OF ARCHAEOLOGICAL FINDS OF HUMAN SKELETONS

The research project «The People of the Iron Age» is an interdisciplinary project funded by the Research Council for the Humanities. It started in 1984 as a half-year pilot project. In this phase it was established that archaeological finds of human skeletons can be used as basic materials in the analysis of people and living conditions in the Iron Age. However, unambiguous documentation of the skeletons' find context was found to be necessary. The large number of finds made computer entry of the material desirable, also in connection with documentation, sorting, and analysis.

A three-year project started in 1985 is currently being carried out by Research Fellow Berit J. Sellevold and Senior Curator Jenny-Rita Næss. The project is divided in two: Computer entry of data on grave finds in Norwegian collections, and an analysis of the parts of the material that are most suitable as a basis for drawing conclusions of a cultural historical nature.

Apart from the skeletons and parts of skeletons, the project's basic material consists of all of the documentation connected to the finds. The cultural historical data must be registered for each find.

The anthropological data are mainly quantitative, whereas the archaeological ones are qualitative. These data must be freely combined for purposes of analysis. For this reason it was necessary to develop both a database structure and forms for computer entry in fields for the data collection. The authors give details of these two systems.

BRUKEN AV EDB VED DE ARKEOLOGISKE MUSEENE

THE USE OF COMPUTING AT THE ARCHAEOLOGICAL MUSEUMS

In this article Senior Curator Egil Mikkelsen describes cooperative ventures between Norway's five archaeological museums concerning the production of general indexes of antiquities and archaeological finds.

Since 1980 a national committee has been working on the development of a computer system for entry of data on such material. The committee's chief aim is to prepare data entry forms that contain certain types of clearly-defined information that can be organized in local databases. As a minimum, museum staff should be able to retrieve information in each other's databases. However, existing resources have not allowed for coordinated purchase of hardware and software at the museums, and sufficient computing expertise has been lacking. The large amount of data involved requires powerful programs and equipment.

So far work on computer entry of data on antiquities has progressed the farthest. Forms with specifications of certain fields have been developed at three of the museums.

The national committee's work will be finished this year. Remaining tasks concern retrieval of data and storage capacity. As for the museums, much information has yet to be entered.

UTGRAVNINGER I MIDDELALDERBYENE

EXCAVATIONS IN TOWNS FROM THE MIDDLE AGES

Attempts are now being made to coordinate computer entry of archaeological material at Norway's five archaeological museums and the Directorate of Ancient Monuments and Historic Buildings. The aim is to be able to employ a joint catalog system and subject headings. In this article Petter B. Molaug describes some previous attempts at computer entry of material from the Middle Ages and points to some future tendencies.

Various systems for data entry have been developed since the 1960's for excavations in five Norwegian towns and at several church sites. A joint form is desirable due to the need for searching in databases at each archaeological institution. An important task in the near future is to combine accessibility with flexibility and solutions to special problems. This task is particularly challenging because it is likely that automatic entry systems will be introduced both for field documentation and object de-

scription based on digitalization of shapes and integrated program packages.

NATURLIG SPRÅK I KUNNSKAPSSYSTEMER

NATURAL LANGUAGE IN KNOWLEDGE-BASED SYSTEMS

University Lecturer Ivar Utne starts this article with an outline of knowledge-based systems and topics that relate them to natural language. In this connection Utne briefly touches on applications of the technology with regard to the entry and retrieval of information, documentation, and machine translation.

The main part of the article consists of a presentation of how natural language is processed by using sets of rules and in representation. Here Utne concentrates on two of the main components of a knowledge-based system that communicates with users via natural language: the analysis of language, representation of meaning and reasoning processes; and the production of language. The main emphasis is on the contents of language.

In Utne's opinion, the subject of his article is an important field in which philologists can contribute to the development of the wider field of artificial intelligence.

GRUNNLAGSMATERIALE FOR INFORMASJONSSØKING VED NORSK TERMBANK

BASIC MATERIAL FOR INFORMATION RETRIEVAL AT THE NORWEGIAN TERM BANK

In this article University Lecturer Ivar Utne describes how terminological material stored at the Norwegian Term Bank can be used as basic data in information retrieval systems. This material consists of terms and expressions with varying spelling, abbreviations, synonyms within the same language, and thesauruses.

The Term Bank's emphasis is on the development of technical terminology, mainly related to the oil industry. To a large extent, computer methods are applied to the storage and retrieval of lexical material, language control, and computer-assisted translation. These methods primarily provide support in terminological work, but are also part of plans for development work within the general field of information technology, including knowledge-based systems and automatic translation.

Word lists can be built into retrieval systems, so that a retrieval program can have access to the information in the form of a table or file. In his article Utne describes how this data can be utilized for various purposes, with emphasis on the linguistic aspects. His aim is to shed light on the linguistic basis of information retrieval.

*FORFATTARENS HØVELBENK –
INTELLIGENT TEKSTBEHANDLING?*

*A NORWEGIAN WRITER'S WORKBENCH –
INTELLIGENT WORD PROCESSING?*

Senior Computing Officer Per Vestbøstad of the Centre has taken part in the development of a Norwegian test version of the American program package «Writer's Workbench». The package consists of a series of programs that can assist in the control of language usage in the process of text production. The project is a cooperative venture between the Centre and the Section for Norwegian Lexicology at the University of Bergen.

The programs can perform proofreading, analyze some stylistic elements, give a compressed printout of some aspects of the composition of a text, give some statistical calculations based on both words and sentences, evaluate which educational level a text is suitable for, and produce concordances and vocabularies.

This type of software can be of use to authors of technical texts, school pupils, students, public administrators, researchers, and journalists. Therefore the system can be easily adapted to the special needs of the various user-groups.

The programs do not aim at standardizing language – they merely point out linguistic characteristics that hamper efficient communication. They are not meant to perform as «substitute teachers», but as a supplement that allows for greater individualization and linguistic awareness.

The planned standard version of the Norwegian system will feature additional functions to the ones mentioned above. Projects concerning similar software are currently taking place at several other Norwegian institutions. In order to coordinate this work, the Centre plans to arrange a seminar for all of the involved parties in September.

COMPUTERS IN LANGUAGE TEACHING

Per M. Mathisen teaches languages and psychology at a Norwegian high school that has received government funding for testing computing methods in instruction. In this article he discusses the pedagogical and practical implications of introducing computers into the language classroom.

In Norway there is a tendency towards using computers primarily as a means of either improving teaching methods or making them more efficient, rather than as «substitute teachers». This goal puts great demands on both hardware and software, but there are only a few (good) programs on the market for language teaching. These often represent merely a means of variation.

Word processing programs can be used as *tools* in language teaching, but so far few teachers have experience in applying such programs. Mathisen gives possible reasons for this situation, and outlines sensible applications of word processing equipment, depending on whether pupils can use a keyboard functionally or not. In Mathisen's opinion, efficient utilization of computers requires a minimum of keyboard skills.

Word processing has been made a main project at Mathisen's school, as experiments have been successful. In this project the teaching of written Norwegian is process-orientated.

A drill program has also been developed and tried out with great success. This program can be used in various languages and in many different ways.

ERFARINGER FRA SKRIVING AV INTERAKTIVE LINJE-ORIENTERTE UNDERVISNINGSPROGRAMMER I SPRÅK

EXPERIENCES WITH WRITING INTERACTIVE LINE-ORIENTATED LANGUAGE TEACHING PROGRAMS

University Lecturer Signe Marie Sanne has written FORTRAN programs for exercises aimed at students of Italian. The programs are run interactively on the University of Bergen's Sperry UNIVAC mainframe. Sanne describes in detail how she has solved both editorial problems posed by line orientation, and technical problems posed by the hardware used.

Sanne has written two types of exercise programs: for grammatical

analysis and correct use of grammatical forms, plus combinations of these two. Both types have been designed for students at various levels, and some programs allow for several alternative answers. Details are given of how students use these programs.

SENTERETS VIDEOPLATEPROSJEKT

THE CENTRE'S VIDEO DISC PROJECT

Archivist Roger Erlandsen of the National Institute for Historical Photography and Computing Officers Claus Huitfeldt and Øystein Reigem of the Centre have written a lengthy final report on the Centre's videodisc project. The present version consists of excerpts of the report, which is due to be published in the Centre's report series.

The goals of the project are outlined (see HD 3-86), as well as the reasons for producing a videodisc as a means of testing optical storage technology. The production process of the disc and the adaptation of the material's reference data is described, along with the software developed for the project. Three previously developed programs for the entry and retrieval of museum object data have been adapted for linking pictures and their corresponding databases.

Other programs that have been designed are: a general program for linking the videodisc with external data sources via a modem, a residential program that «snaps up» picture references from the stream of data to the screen, and a menu-based program for «wandering» in the collection of stills on the disc.

The authors proceed to evaluate the different aspects of the project and conclude that videodisc technology is suitable for picture databases that do not put strict demands on image quality and ability for manipulation. Otherwise, the digital optical media currently being developed probably are more suitable.

Transferral to both types of media means increased accessibility, ease of retrieval and less wear and tear on the original material. The project group therefore recommends that humanities institutions as soon as possible start to prepare material in their collections and connected reference data for optical storage.

HUMANISTISKE BIBLIOGRAFISKE DATABASER

BIBLIOGRAPHIC DATABASES IN THE HUMANITIES

Annema Hasund Langballe, Head of the Bibliographical Department at the Norwegian Library Bureau, adds more titles to the list of bibliographical databases in the humanities given in HD 3-86. *Dialog* also runs *RI-LA* (Art literature international), *RILM* (music literature), and *Religion index*. A new database is *Wilsonline*, which includes all of the publishing firm Wilson's printed bibliographies 1981-84: *Art index*, *Biography index*, *Book Review digest*, *Cumulative book index*, and *Humanities index*.

Langballe's own experience with searching in humanities databases is outlined. She also gives advice on how to learn to use these databases. In her opinion, however, occasional users should take advantage of searching services offered by university libraries.

EDB FOR HUMANISTER. ERFARINGER FRA ET NYSTARTET GRUNNFAG VED DET HISTORISK-FILOSOFISKE FAKULTET, UNIVERSITETET I OSLO

COMPUTING FOR HUMANITIES STUDENTS. EXPERIENCES WITH A NEW «GRUNNFAG» AT THE UNIVERSITY OF OSLO

A course on computing for humanities students has been given by the computing service for the humanities at the University of Oslo since 1981. The course has now been expanded with an additional term, so it can be taken also as a «grunnfag».

The established humanistic research tradition constitutes a major part of the «grunnfag's» basis along with experiences made within computer-based humanities research in the past 10 years. In practical terms the course is to serve as a tool within other humanities disciplines and give students competence as teachers and in other professions.

Lecturer Asbjørn Brøndeland and Research Fellow Jon Lanestedt of the computing service outline the contents of the course. Strong emphasis is put on general programming skills, and students learn about computing in the humanities in particular.

The second term consists of four parts: 1. computer systems, programming methodology, and computer applications, 2. computing methods in humanities research, 3. the history of computing and societal aspects, 4. a choice of computational linguistics, AI and expert systems, or computer-assisted instruction. The authors sum up both negative and positive

experiences made after teaching the second term of the course for the first time.

A further expansion of this course, geared mainly towards computational linguistics and applied logic, is currently being considered by the Faculty of Arts. Plans are also being made to integrate all existing courses related to information technology at the faculty.

TRE RESIDENTE PROGRAMMER FOR MS-DOS MASKINER

THREE RESIDENT PROGRAMS FOR MS-DOS COMPUTERS

Resident programs work as expansions of a computer's operating system. Senior Computing Officer Espen S. Ore of the Centre reviews three such programs: Sidekick, Turbo Lightning, and Smart Notes.

Sidekick is a collection of aids similar to Apple's Desk Accessories, Turbo Lightning is a program for automatic proofreading, and Smart Notes allows the user to «fasten» notes onto a file.

Ore has tested all three in his daily work on a PC (word processing, programming, and network communication). He finds the first two programs useful, but feels that an «old-fashioned» notepad and pen is as good a tool as Smart Notes.

Ore also discusses technical issues and do-it-yourself design of resident programs.

NORDISK KONFERANSE OM TEKSTFORSTÅELSE OG INFORMASJONSSØKING

NORDIC CONFERENCE ON TEXT COMPREHENSION AND INFORMATION RETRIEVAL

The Department of Linguistics at the University of Stockholm arranged a conference in December on text comprehension and information retrieval. Director Jostein H. Hauge of the Centre reports that 36 people from 26 different institutions in the Nordic countries took part. The aim of the conference was twofold: to get an overview of Nordic research in this field, and to work out a joint Nordic research plan in order to strengthen and extend existing contacts.

At the conference about 20 presentations were given of on-going research. These dealt with computational linguistics (methods for automatic analysis of morphology, syntax, and semantics), text representation,

corpus work, terminology, linguistic databases, indexing, information retrieval, quantitative computational linguistics, and socio- and psycholinguistical-oriented research work.

The participants agreed on a joint Nordic research program based on existing resources and activities. The theme of the program, which is intended to last for three years from the summer of 1987, will be text comprehension by man and machine. The aim is to produce a volume of articles intended for an international readership. A subfield within the program will be information retrieval systems involving linguistic competence.

INFORMASJONSMØTE OM SENTERETS VIDEOPLATEPROSJEKT

INFORMATION MEETING ON THE CENTRE'S VIDEODISC PROJECT

Office Manager Einar-Arne Drivenes of the Centre reports on a Nordic information meeting held in February on the Centre's videodisc project. Representatives of the Centre demonstrated the disc, and the other participants presented their own projects and plans in connection with interactive video.

Afterwards a discussion took place of experiences with on-going and completed projects, project planning, interactive video in education, and technical questions in connection with hardware and software. In this session a demonstration was given of software developed for a planned videodisc on the Old Town in Stockholm, and possibilities for Nordic co-operation were discussed.

MELDINGER

NEWS

The Centre is to start an electronic news service. About once a month news will be distributed via electronic mail to all interested parties in Norway who have network addresses.

Software reviews will be a set feature in future issues of Humanistiske Data. We also hope to print more material on computers in music. Contributions of all kinds concerning this field are welcome.

The Norwegian documentation service for research in progress has published a catalog of recently completed and on-going humanities pro-

jects in Norway. Brief information is given on 1006 projects, which are also described in more detail in the service's project database, along with connected publications.

Dovre Data Archive was established on a trial basis a year ago. The archive has several aims: 1. to enter and data process historical sources from the 19th century for the Norwegian counties of Hedmark and Oppland, 2. to produce easily accessible versions of the original sources and alphabetical indexes, 3. to produce various types of statistics, and 4. to experiment with record linkage.

The Norwegian Online User Group is a special interest group for users of online databases. Among the group's aims are: 1. to work for the exchange of information between members, 2. to arrange courses and meetings, and 3. to represent the interests of its members both nationally and internationally. The group meets 4-6 times a year.

From this autumn the regional college in North Trøndelag, Norway, will offer both one and two-year degrees in information technology. In the two-year course much emphasis will be put on societal informatics. Societal evaluations of information technology will be discussed, and light will be shed on consequences of information technology via an analysis of the processes of change in society.

SCANNET was established in order to promote the use of Nordic online databases. SCANNET's aims are to facilitate the accessibility and encourage efficient use of Nordic computer-based information services and resources. The organisation's main tasks are to: disseminate information on Nordic databases, give advice on their use, work for the coordination of online systems and command languages, advise on the planning and development of databases, and develop and support projects on Nordic databases. The newsletter SCANNET TODAY is published twice a year and distributed free of charge.

A centre for informatics has been established in the Danish city of Århus on a trial basis. The centre's aim is to enable teachers, administrators, and local school authorities to keep up with developments within computing and teaching and to develop their own high-quality teaching programs. At the centre programs are developed, adapted, and evaluated. Several software development projects initiated by teachers have been supported and coordinated. The centre also lends hardware and software, gives advice and consultancy services, and arranges lectures and courses.

The Association for History and Computing has been established in England in order to promote the use of computing in historical research and teaching. The association arranges annual conferences, publishes the

newsletter *Computing and History Today*, and has initiated several comprehensive projects and teaching programs.

A system has been established in England for the distribution of news on music research via electronic mail. Users of «Computers and Music Research» are welcome to submit their own material for distribution.

Institut National de la Langue Française (INALF) consists of 10 units that carry out research programs in connection with French language and linguistics. The institute employs a research staff of 250 and was one of the first in France to apply computing methods in the humanities. The research units are currently at work on a dictionary of French language in the 19th and 20th centuries, lexicology and political texts, the teaching of modern French, and Old French. The units share computing resources and a central documentation centre.

Archives and Museum Informatics is an American firm established last year. The firm supplies consultancy services concerning information handling at archives and museums, carries out commissioned research on applications of information technology within the sector, publishes *Quarterly Update on Archival Informatics*, and offers courses and seminars on topics within the fields of computing in archives and museums.

An interactive program on Greek vases has been produced by the firm Interac for the J. Paul Getty Museum in California. Stills and live film fill the whole videodisc together with a 1-hour soundtrack and textual information. A combination of two videodisc players allows extensive use of sound together with the stills and reduces access time to the images. More than 100 of the museum's Greek vases are portrayed on the disc.

The following journals are presented: 1. *Window on the Humanities* – national and international news, reviews of books and software, tips for Macintosh users. Irregular publication, free of charge. 2. *MDA Information* – articles, reports, and news concerning museum documentation, quarterly. £5 per year + postage. 3. *Quarterly Update in Archival Informatics* – consists of two parts: *Archival Information Newsletter* (\$30) and *Archives and Museums Technical Report* (\$180). 4. *Computing for the Humanities* – newsletter published every other month, free of charge. 5. *Leonardo* – international journal on the synthesis of the arts, sciences, and technology. Quarterly, DM 112 per year, trial copy available.

Forthcoming conferences:

- *CIDOC Annual Meeting – MDA Study Tour – MDA Annual Conference* – Cambridge, England 20-29 September
- *Summer School for Information Technology in the Humanities* – London, 21-25 September

- *International Conference on Computers in the History Classroom* – Leeds, England, 6-8 July 1988
- *COLING 88* – Budapest, 22-27 August 1988

The Centre will be arranging a national conference in Bergen 14-16 October on the use of optical storage media in research, education, and mediation. The focus of the conference will be on the humanities. A number of guest lecturers from the US and Europe have been invited along with Norwegian experts.

Forts. fra 2. omslagsside.

RAPPORT nr. 29, 30, 31, 32: *Stig Welinder et al.: STAR I-IV. A program package for archaeological use.* Bergen 1983.

Nr. 29 *STAR I. Introduction and STAR manual.* ISBN 82-7283-033-7
Pris kr. 50.

Nr. 30 *STAR II. Student textbook and STAR examples.* ISBN 82-7283-034-5
Pris kr. 60.

Nr. 31 *STAR III. Archaeology for statisticians.* ISBN 82-7283-035-3
Pris kr. 60.

Nr. 32 *STAR IV. STAR algorithms.* ISBN 82-7283-036-1
Pris kr. 30.

RAPPORT nr. 33. *Årsmelding 1983.* NAVFs edb-senter for humanistisk forskning. ISBN 82-7283-038-8
Gratis.

RAPPORT nr. 34. *Jostein H. Hauge: Tutorial on Machine Translation.* Rapport fra en konferanse i Lugano 2.-6. april 1984. Juni 1984. ISBN 82-7283-039-6
Pris kr. 60.

RAPPORT nr. 35. *Ole Lauvskar: Bruk av statistiske metodar i språk- og litteraturforskninga. Rapport frå ei spørjeundersøking.* September 1984. ISBN 82-7283-041-8
Pris kr. 50.

RAPPORT nr. 36. *Årsmelding 1984.* NAVFs edb-senter for humanistisk forskning. ISBN 82-7283-042-6
Gratis.

RAPPORT nr. 37. *Årsmelding 1985.* NAVFs edb-senter for humanistisk forskning. ISBN 82-7283-043-4
Gratis.

RAPPORT nr. 38. *Jon Birger Østby: Edb-metoder for kunst- og kulturhistoriske museer.* ISBN 82-7283-045-0
Pris kr. 60.

RAPPORT nr. 39. *Årsmelding 1986.* NAVFs edb-senter for humanistisk forskning. ISBN 82-7283-04-7
Gratis.

Øvrige rapporter er utsolgt.

C

HOVEDTEMA I DETTE NUMMER:

**ARKEOLOGI
SPRÅK
SPRÅKUNDERVISNING
VIDEOPLATETEKNOLOGI**

RETURADRESSE:
*NAVF's edb-senter for humanistisk forskning
Boks 53 – Universitetet
5027 Bergen*

