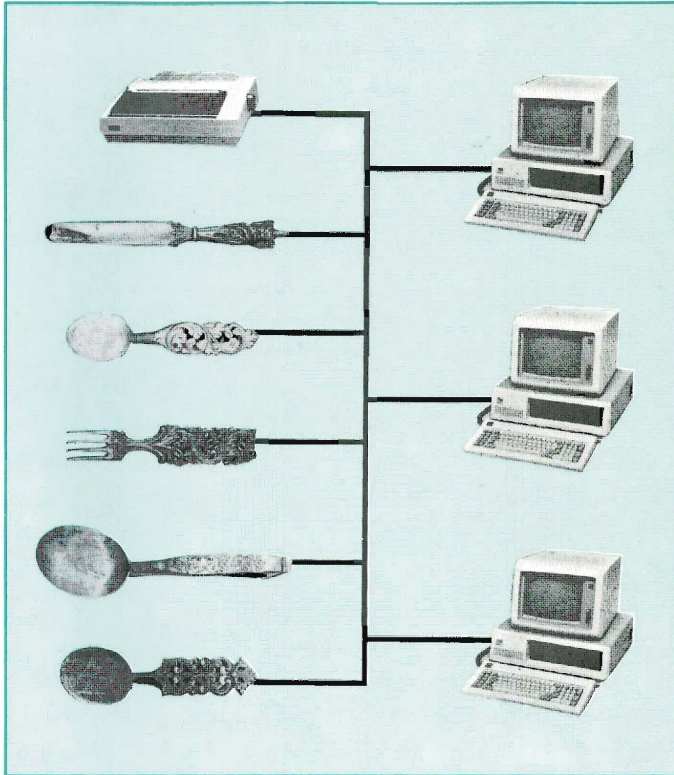


# humanistiske data



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

**2-86**

# SENTERETS RAPPORTSERIE

## Rapporter utgitt f.o.m. 1980

RAPPORT nr. 13. *Datatjenester for og datasamarbeid mellom kunst- og kulturhistoriske museer.* Februar 1980. 2. opptrykk november 1981. ISBN 82-7283-010-8 Pris kr. 50.

RAPPORT nr. 14. *NOVA\*STATUS HÅNDBOK*

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

Del 2: Fil-beskrivelser. Systemdokumentasjon. Utsolgt.

Del 3: Generering og oppdatering av databaser. Utsolgt.

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

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

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

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

RAPPORT nr. 19. *Statistiske metoder på arkeologisk materiale.* Rapport fra et seminar på Bryggens museum, Bergen 24.-26. november 1980. Mars 1981. ISBN 82-7283-017-5 Pris kr. 35.

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

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

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

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

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

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

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

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

RAPPORT nr. 28 *Årsmelding 1982.* NAVFs EDB-senter for humanistisk forskning. ISBN 82-7284-31-0. Utgått.

*Forts. 3. omslagsside.*

# humanistiske data 2-86

---

NAVFs EDB-senter for  
humanistisk forskning

---

The Norwegian Computing  
Centre for the Humanities

---

**NAVF** NORGES  
ALLMENNVIITENSKAPELIGE  
FORSKNINGSRÅD

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

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

Senteret utgir tidsskriftet *Humanistiske Data* (3 nr. pr. år) og en rapportserie (37 er utkommet pr. 1.9.86).

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

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

Interesserte kan kostnadsfritt bestille årsmelding og *Humanistiske Data* (kr. 60,- for institusjoner).

---

*Humanistiske Data* blir utgitt av NAVFs EDB-senter for humanistisk forskning. Redaksjon: Jostein H. Hauge, Kristin Natvig (red.).

Senterets adresse: Harald Hårfagresgt. 31, Boks 53, 5014 Bergen-Universitetet. Tlf. (05) 212954/55/56

Artikler, rapporter, meldinger mottas. Redaksjonen avsluttet 1. september.

---

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

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

---

Medarbeidere fra Senteret i dette nummer:

*Jostein H. Hauge, Knut Hofland, Kristin Natvig, Espen Ore, Øystein Reigem.*

---

*Fotosats i kommunikasjon med Univac 1100/82.*

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

*Grafisk design og montasje: Kristin Natvig.*

*Trykk: John Grieg A/S*

*Forsidebildet:*

*Collage av Øystein Reigem, med utgangspunkt i foto fra Norsk Folkemuseum.*

# Innhold

## Artikler

Nordisk edb-konferanse for kunst- og kulturhistoriske museer. <i>Div. forfattere</i>	s. 4
Status for edb-arbeidet i de norske kunst- og kulturhistoriske museer. <i>Jon Birger Østby</i>	s. 41
CD-ROM – en halv gigabyte i lommen. <i>Øystein Reigem</i>	s. 48
Educational Uses of Existing Linguistic Software. <i>Charlotte</i> <i>Toubro</i>	s. 74
Edb-oppbygging og edb-virksomhet i det offentlige arkivverk. <i>Ivar Fønnes</i>	s. 85
«Project Emperor-I» – kinesiske skatter på videoplate. <i>Kristin</i> <i>Natvig</i>	s. 94

## Rapporter

Training the Information Researcher for the Future. <i>Jostein H.</i> <i>Hauge</i>	s. 98
Utviklingsseminaret på Ustaoset april 1986. <i>Asbjørn Brændeland</i>	s. 104
XVIII International Congress of Papyrology. <i>Espen S. Ore</i>	s. 110
ICAME 7th. <i>Knut Hofland</i>	s. 113
Nytt fra RHF/NAVF	s. 117
Fra bokhylla	s. 125

<b>Meldinger</b>	s. 127
------------------	--------

<b>Summary</b>	s. 131
----------------	--------

## Nordisk edb-konferanse for kunst- og kulturhistoriske museer

*Karl Ragnar Gjertsen*

NAVFs EDB-senter for humanistisk forskning arrangerte 4.-6. juni en nordisk konferanse på Os ved Bergen om bruken av edb ved de kunst- og kulturhistoriske museene. At det var «Senteret» i Bergen som arrangerte konferanse om dette tema er ganske naturlig. Senteret har vært den viktigste samarbeidspartneren for norske museer når det gjelder innføring av edb helt fra dette temaet ble brakt på bane her hjemme for ca. 10 år siden. Debatten om også museer kan ha nytte av edb er nå definitivt slutt.

Konferansen viste at samarbeid mellom museer, helst på det nasjonale plan, har vært viktig for å bygge opp datasystemer som kan fungere i museumssammenheng. Det er også interessant å konstatere at de tekniske mulighetene til å kombinere data og bilde nå nærmer seg vår daglige museumsvirkelighet.

Konferansen gikk over tre dager. Første dags program dreide seg om status for bruk av edb ved museene i dag, andre dag var hovedsakelig avsatt til tekniske forhold og demonstrasjoner, mens programmet tredje dag tok opp framtidsprosjekt og arbeidsstrategi.

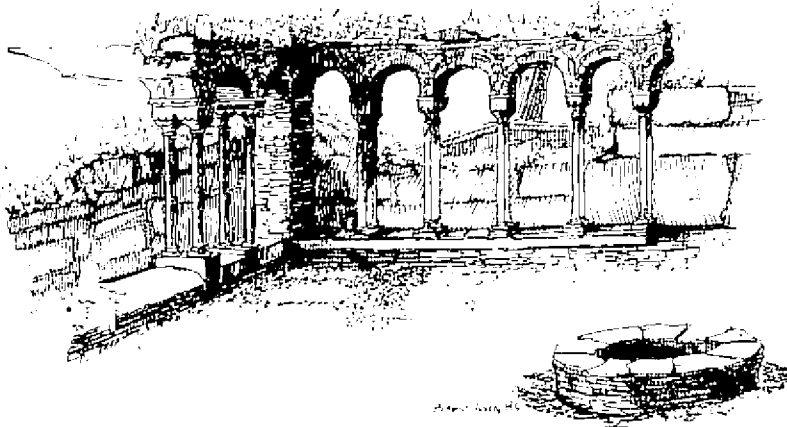
Siden sammendrag av alle innleggene på konferansen trykkes i denne utgave av Humanistiske Data, skal det her bare gis en kort gjennomgang av emnene som ble tatt opp.

Status for bruk av edb ved museene i Norden ble presentert av en innleder fra hvert land. Nordiska museet tok i bruk edb allerede for 20 år siden da man sto overfor oppgaven å registrere Skoklosters slott 20.000 gjenstander. Siden den gang har Nordiska museet ligget i teten når det gjelder bruk av edb i vår nordiske museumsverden. Behovet for effektiv katalogproduksjon på grunnlag av museets 350.000 gjenstander forklarer hvorfor dette museet satset på edb. Museets egen gjenstands-samling samt samlingene til 13 läns museer er gitt maskinlesbar form på grunnlag av de gamle protokollene. Tilsammen 630.000 gjenstander er registrert i regi av Nordiska museet. Andre svenske museer har selv 109.000 gjenstander i database. Til sammenlikning har norske museer ca. 37.000 gjenstander på data – de fleste tilhører Norsk Folkemuseum. De arkeologiske museene er her holdt utenfor.

Behovet for felles nasjonale registreringsnormer har meldt seg i flere land med innføring av edb. Flere av foredragsholderne var inne på dette spørsmålet. I Norge oppnevnte NKKM i 1977 en egen edb-komite som utarbeidet et gjenstandskort med veiledning. Nylig har NAVF avsluttet et toårig prosjekt som har utarbeidet en programpakke for registrering av

**NORDISK EDB-KONFERANSE  
FOR  
KUNST- OG KULTURHISTORISKE MUSEER**

*4. - 6. juni 1986 på Solstrand Fjord Hotel  
Os ved Bergen*



**NAVFs EDB-senter for humanistisk forskning  
Bergen, Norge**

gjenstander og fotografier (arkivmateriale kommer), i applikasjonsprogrammene Datastar og Dataflex. (Kunstmuseene har laget sitt eget registreringsprogram i dBase II.) Dette programmet ble demonstrert på konferansen. I Sverige er man i avslutningsfasen for prosjektet SAMOREG, som har utviklet registreringsskjema for ulike typer materiale. Skjemaene skal sikre et felles minimumsnivå av opplysninger, standardisere opplysningenes utforming og ved oppbygging av skjema i blokker gi en enhetlig disposisjon. I Danmark arbeider et utvalg under Statens Museumsnævn (SMN) med et felles dokumentasjonssystem for museene, og en skisse av dette skal være ferdig i 1987. Man legger vekt på å få et integrert system for både de lokale og nasjonale museer. To sentralregistre for hele Danmark skal motta et utvalg av feltene til sine registre.

Nomenklaturarbeid, som *Jon Birger Østby* redegjorde for, er en del av

tilretteleggingen for bruk av edb. NKKM's edb-komite ble tidlig klar over at alle synonymer og varianter av gjenstandsbetegnelser ville skape problemer ved maskinell sortering. Dette er ikke bare et norsk problem – selv om vi har ekstra problemer med våre to svært like skriftspråk og mange dialekter. I Norge er det gjennomført et treårig prøveprosjekt der man har tatt for seg to utvalgte gjenstandsgrupper.

Hvordan få med våre gamle data inn i edb-arkiv? Dette spørsmålet var flere foredragsholdere inne på. Problemene er flere. Ett av de nærliggende er selve arbeidsmengden. Spesiell interesse var det for måten Nordiska museet løste dette på. *Göran Bergengren* fortalte om dette i sitt innlegg. Ved bruk av sysselsettingsmidler er det i Norrland bygd opp en registreringsentral med 20 arbeidsplasser. På grunnlag av kopier av hovedprotokollene skrives data inn i Norrland, og sendes på diskett pr. post til Stockholm.

*Magne Velure* tok opp problemet fra en annen synsvinkel, nemlig avstanden mellom våre ideelle forestillinger om hvordan gjenstander burde registreres og hvilke ressurser museene kan avsette til gjenstandsregistrering. Eksempelvis har De Sandvigske Samlinger 35.000 gjenstander. Det vil ta 20 årsverk å gjennomføre en fullgod registrering etter NKKM's gjenstandskort. Det vil da gå lang tid før en får oversikt over gjenstandsmassen. Velure foreslo at man burde bygge opp informasjon stegvis og i første omgang nøye seg med et utvalg av opplysninger fra de gamle hovedprotokollene, opplysninger som vil tilfredstille 80% av alle spørsmål. En slik strategi kan bl.a. på kort tid hjelpe museet å finne hull i samlingene og gi en offensiv innsamlingspolitikk.

Hva med de store nasjonale museumsdatabaser? Konferansedeltakerne fikk stifte nærmere bekjentskap med The Canadian Heritage Information Network (CHIN) representert ved *Stephen H. Delroy*. CHIN har som oppgaver: å drive en nasjonal museumsdatabase, gi veiledning til museer, rådgivning innen museologi, data og ny teknologi. Medlemsmuseer får de fleste tjenester gratis, men må selv legge inn data og holde egne modem og terminaler. CHIN disponerer en Cyber 835 med over 8 gigabytes (8000 Mb) lagerkapasitet. 120 terminaler i 17 byer er tilkoblet og de kan legge inn, redigere, endre, analysere og ta ut rapporter – alt on-line. Medlemsmuseene strekker seg fra kunstgallerier med 1000 gjenstander til Royal Ontario Museum med 750.000. Ca. 25% av alle museums-gjenstander i Canada er i basen. Nesten alle felt er indekserte, det betyr at bare 1/3 av lagerkapasiteten utnyttes, mens 2/3 er indeksfiler. Egentlig dreier det seg om flere ulike databaser for ulike typer objekter: kulturhistorisk database, naturhistorisk database, tyveri db., kunstnerbiografi db., bibliografi db., m.m.

Delroy så en klar trend fra kortkopier på papir og kartong til fullstendig on-line søking. En av retningslinjene som CHIN driver sin database etter er at opplysningene bør lagres på det nivå de brukes. F.eks. bør



opplysninger om hvor gjenstanden befinner seg lagres på institusjonsnivå.

Et spørsmål man kan stille er om telefon er billig i Canada. Forsøk med on-line database i Danmark er mislykket p.g.a. kostnadene. Som ledd i demonstrasjonene ble databasen i Ottawa oppringt fra konferansehotellet. Vellykket, men kostbart!

Kunst Index Danmark bruker CHIN som ett av sine forbilder. *Margrethe Pedersen* fra Statens Museum for Kunst i København presenterte i sitt innlegg det registreringsprogram Kunst Index Danmark vil bruke. Man legger vekt på å integrere databehovet på det lokale versus det nasjonale plan, både for de faglige og administrative interesser. Man legger videre vekt på at kunsthistorikere utenfor museene skal ha nytte av sentralarkivet. Kunstverkene skal kodes etter internasjonale systemer, f.eks. Iconclass.

Ved Nordiska museet finnes kimen til en svensk nasjonal database. Museets egen samling, Livrustkammaren, Skokloster slott og 2 läns museer ligger i dag på Nordiska museets maskin – hele kapasiteten på 524 Mb er utnyttet. Västerbotten museum i Umeå har PC'er som brukes til søking i deres database ved Nordiska museet. De kan hente ut deler av basen og bearbeide disse. Maskinens kapasitet gjør det umulig å gjøre flere av de 13 registrerte läns museene tilgjengelige on-line.

Under konferansens demonstrasjons-sesjon ble det vist eksempler fra den felles prøve-databasen for noen norske kulturhistoriske museer.

Data og bilde har framtida for seg i museene. Våre «manuelle» registreringskort inneholder gjerne også et foto av gjenstanden. Dette er det store pre for det tradisjonelle arkivkort, mens dataskjermen mangler denne mulighet. Kulturarvet i Falun har allerede et operativt system (beskrevet i *Humanistiske Data 2-83*), der foto av gjenstanden overføres til videoplate via film og videotape. Data og bilde kan betraktes på to atskilte skjermer.

Som ledd i demonstrasjonene ble NAVFs EDB-senters nye prøvevideoplate demonstrert koblet mot gjenstands- og fotoregister. Søkertiden er begrenset til sekunder blant de bildene som er lagret på en plateside. På grunn av prisene og manglende oppdateringsmuligheter er denne type plate lite aktuell i museumssammenheng. Produksjonsprisen er kr. 25.000-30.000 for master og 10 kopier. I museumssammenheng er denne type plate mest aktuell der oppdateringer er sjeldne.

Videoplater har et annet aktuelt bruksområde i museenes publikumsrettede virksomhet. Firmaet Scan Vision, som har produsert Senterets prøveplate, har også laget plate for Norsk Teknisk Museum. Denne platen ble demonstrert tilkoblet pekeskjerm. Publikum velger i menyen på skjermen og styrer spilleren til den interessante sekvens av platen.

Det museene fremfor alt ønsker seg er et datastyrt bildearkiv der det stadig kan legges nye bilder til. Optiske digitale plater (write once) er

## PROGRAM

## ONSDAG 4. JUNI

kl. 9.00	Innledning Direktør <i>Jostein H. Hauge</i> , NAVF's EDB-senter for humanistisk forskning, Bergen
" 9.15	Status for bruk av EDB i museumsarbeidet i de nordiske land: - Danmark v/udvalgssekretær <i>Ole Thundrup</i> , Statens Museumsnavn og museumsinspektør <i>Carsten U. Larsen</i> , Nationalmuseet - Finland v/museisjef <i>Jouko Heinonen</i>
" 10.45	Pause - Island v/nasjonal representant - Sverige v/intendent <i>Göran Bergengren</i> , Nordiska Museet - Norge v/konservator <i>Jon Birger Ostby</i> , Norsk Folkemuseum
" 12.15	EDB-arbeidet ved Nordiska Museet Intendent <i>Göran Bergengren</i> , Nordiska Museet, Stockholm
" 13.00	Lunsj
" 15.00	International survey of use of data processing in museums Dr. <i>Andrew Roberts</i> Museum Documentation Association, United Kingdom
" 15.45	Pause
" 16.00	The museum documentation system PARIS in the Canadian Heritage Information Network Ass. Director <i>Stephen H. Delroy</i> , The Canadian Heritage Information Network, Ottawa
" 17.00	Erfaringer fra nomenklaturarbeid Konservator <i>Jon Birger Ostby</i> , Norsk Folkemuseum
" 17.45	Evt. demonstrasjon
- 19.30	Middag
- 21.00	Ole Bull, Lysøen og folkelege prydtadisjoner i Midthordlands-bygdene. Kulturhistorisk inventering ved hjelp av dataregistrering Fylkeskonservator <i>Nils Georg Brekke</i> , Hordaland Fylke

## TORS DAG 5. JUNI

kl. 9.00	EDB i kunstmuseer Museumsinspektør <i>Margrethe Pedersen</i> , Statens Museum for Kunst, København
" 9.45	EDB-virksomheten ved Kulturarvet Direktør <i>Rune Hjernmann</i> Kulturarvet, Falun
" 10.30	Produksjon av videoplate. Trinn i arbeidsprosessen Daglig leder <i>Anne Leer</i> , Scan Vision, Oslo
" 11.00	Pause
" 11.15	Optiske digitale plater for tekst- og bildelagring (CD-ROM og WORM) Direktør <i>Tor Thomassen</i> , Data Visual
" 11.45	Videoplate-teknologi i museumsarbeid Førstekonsulent <i>Oystein Relgen</i> , NAVF's EDB-senter, Bergen
" 12.15-15.30	Demonstrasjoner og lunsjpause
" 16.00	Ekskursjon med innlagt middag
FREDAG 6. JUNI	
kl. 9.00	Nationalmuseets Bevaringsseksjon og bruken af EDB Ledende konservator <i>Knud Holm</i> , Nationalmuseet, København
" 9.45	Nationalmuseets fremtidsprosjekt Museumsinspektør <i>Carsten U. Larsen</i> , Nationalmuseet, København
" 10.30	Pause (tydding av hotellrom)
" 11.00	Utvalgsstrategier ved dataregistrering av eldre gjenstandsmateriale Museumsdirektør <i>Magne Velure</i> , De Sandvigske Samlinger, Mailhaugen, Lillehammer
" 11.45	Avslutningsseksjon: Virkninger av datateknikk på museumsarbeid og for samarbeid mellom museer regionalt, nasjonalt og internasjonalt. Samtale
Avslutning til lunsj ca. kl. 13.00	

relativt nytt på markedet, men svært interessant i museumssammenheng. Firmaet Data Visual orienterte om disse mulighetene.

Fremtidsprosjektet ved Nationalmuseet i København står foran en omfattende utbygging – nye utstillinger, nytt magasinhus – og et omfattende dokumentasjonsprosjekt. Bare til dokumentasjonsprosjektet er det bevilget 50 millioner Dkr. Carsten U. Larsen utdypet planene for dokumentasjonssystemet. Man baserer seg på en nettverksløsning der hver enkelt avdeling får ansvar for sine egne data. Kapasiteten må være så stor at man kan ta både tekstdata og billedata.

Ikke minst for Nationalmuseet blir det en stor og vanskelig oppgave å legge inn sine gamle arkiver, som er ført på forskjellig måte. Man vil la datamaskinen selv gi forslag til nomenklatur og klassifikasjonssystem. Billedata skal legges til rette slik at materialet kan utnyttes i det profesjonelle museumsmiljø og for publikum. Det Kulturhistoriske Centralregister får kontorfellesskap med deler av hardware, software og programmerere i fellesskap med Dokumentasjonsprosjektet. DKC har ansvar overfor alle museer i Danmark.

Avslutningssesjonen var avsatt til diskusjon og samtale. Etter forespørsel fra Magne Velure lovt Carsten U. Larsen å invitere observatører fra de andre nordiske landene til å følge utviklingen av Nationalmuseets dokumentasjonsprosjekt.

Tuula Leimu fra Finlands Museiförbund takket for at det var tatt initiativ til og gjennomført en slik edb-konferanse. Han inviterte til en ny nordisk edb-konferanse i Finland i 1988.

Om konferansen følges opp med et nærmere nordisk samarbeid om museums-edb og en ny nordisk edb-konferanse i 1988, vil jo det i seg selv være bevis for at arrangementet på Solstrand Fjord Hotel var verdifullt.

Det var også avsatt tid til en kulturhistorisk ekskursjon i skjærgården i Sunnhordland.

## **ADB-arbetet vid Nordiska museet**

### *Göran Bergengren*

Nordiska museet har i dag sin andra minidatamaskin, en Hewlett Packard 3000 modell 40. Uppgifter om museets föremålssamlingar, inklusive samlingarna i museets hus och gårdar, finns tillgängliga för sökning on-line i en databas i denna maskin. Arbetet med inmatning av uppgifterna har gjorts vid museets registreringscentral i Älvsbyn i Norrbotten. Med en personal bestående av 2 arbetsledare och 19 kontorsutbildade tog

detta arbete sammanlagt 18 månader under två budgetår (1982/83 och 1983/84).

Grundmaterialet för inmatningen var museets huvudläggare, från vilken uppgifterna överfördes till ett «stansunderlag» (Bild 1). De flesta data måste då plockas ut ur en löpande text och skrivas in i standardiserad form på rätt plats på underlaget. Utöver detta måste varje föremål placeras in i två olika klassifikationssystem, nämligen dels det decimala Outline of Cultural Materials, (Realnr på underlaget) (Bild 2), dels museets eget system, som är ett hierarkiskt system med grupper och undergrupper och i klartext (Klassifikation A1, A2 etc) (Bild 3).

Den databashanterare, som används är Hewlett Packards egen IMAGE. För sökning i databasen används QUERY, INFORM och museets egna applikationsprogram, skrivna i programspråken BASIC och TRANSACT.

När inmatningen av Nordiska museets eget material avslutats övergick centralen i Älvsbyn till att registrera läns museernas samlingar efter samma principer, som gällde för Nordiska museets. Efter två registreringsperioder om 9 månader vardera (budgetåren 1984/85 och 1985/86) finns nu 12 läns museers samlingar registrerade. Vilka läns museer och hur stora samlingar det gäller framgår av tablån till redovisningen av pågående ADB-verksamhet i Sverige. Verksamheten i Älvsbyn har hela tiden beaktats av länsarbetsnämnden i Norrbottens län. Det enda Nordiska museet behövt bidra med är teknisk utrustning, dvs datamaskin, terminaler, skrivare etc samt arbetstid och resor för den systemansvarige intendenten.

Tyvärr har Nordiska museets datamaskin för närvarande inte minneskapacitet för mera än våra egna samlingar, Livrustkammarens och Skoklosters slotts samt två läns museers åt gången. Övriga 10 läns museers material måste alltså tillsvidare förvaras på magnetband.

Syftet med denna registrering av läns museernas samlingar är i första hand att ge bättre möjligheter att söka i informationen om föremålen genom att åstadkomma fler och/eller andra sökingångar än man har i befintliga kortregister. Ser man dataregistreringen i ett större sammanhang, kan man säga att den standardiseringen, som uppgifterna i viss mån utsätts för, blir till stor nytta för den, som söker t.ex en viss kategori föremål i flera museer. Det omedelbara resultatet av dataregistreringen i Älvsbyn är emellertid att varje läns museum gratis får tillgång till alla väsentliga data om den egna samlingen, registrerade på ett maskinläsbart medium. Om sedan fortsättningen blir register i listform, utskrivna i Nordiska museet (Bild 4), en databas i Nordiska museet, tillgänglig för on-line sökning via telefon och modem eller en lokal databas i det egna museets datamaskin, det bestämmer varje läns museum för sig.



- 27 DRYCKER, NJUTNINGSMEDEL OCH LÄKEMEDEL
- 27 DRYCKER, NJUTNINGSMEDEL OCH LÄKEMEDEL  
allmänt ur olika aspekter rörande att släcka törst och stilla särskilda aptitbegär samt rörande tillverkning av produkter för att tillfredsställa dessa begär och beredning av läkemedel.
- 271 VATTEN OCH TÖRST  
vatten som dryck; yttringar och kontroll av törst; föreställningar och sedvänjor kring vatten som dryck; etc
- 272 OJÄSTA DRYCKER  
förtäring av icke alkoholhaltiga drycker (ex. mjölk, kaffe, te, saft, kakao, blod); särskilda tillredningsmetoder (ex. kokning, bryggnng); konsumtion (ex. mängder, tillfällen); näringsvärde och fysiologisk verkan; dryckesseder; etc
- 273 JÄSTA OCH/ELLER DESTILLERADE DRYCKER  
förekommande alkoholhaltiga drycker (ex. öl, vin, sprit); tillredningsmetoder (ex. bryggnng, bränning); apparatur (ex. brännvinsapparat); konsumtion (ex. konsumenter, mängder, tillfällen); dryckesseder; alkoholbegär; omfattningen av onykterhet och onyktert beteende; hithörande föreställningar och sedvänjor; social kontroll och överhetlig reglering; etc
- 274 DRYCKESVARUINDUSTRI  
industriell produktion av drycker (ex. bryggnng, vinberedning, destillering, läskedrycksfabrikation); särskild teknik och utrustning; organisation av arbete och produktion; överhetlig reglering; etc
- 275 UTSKÄNKNING AV JÄSTA OCH OJÄSTA DRYCKER  
särskilda utskänkingsställen; olika typer (ex.

## DRYCKENS SERVERING

### Kaffe

Gräddkannor  
Kaffekannor  
Kaffekoppar  
Kaffeserviser  
Kaffeskedar

### Konjak

### Krus

### Kåsar

### Likör

### Muggar

### Pokaler

### Punsch

### Saft

Saftserviser

### Sifoner

### Skopor

### Snibbskålar

### Stop

Gyckelstop

### Stänkor

### Tappskålar

### Te

Tekannor  
Tekoppar  
Teserviser  
Teskedar

### Tillbringare

### Tröjser

### Vin

Istänger  
Karaffskyltar  
Korgar  
Pokaler  
Remmare  
Vinbålar  
Vinglas  
Vinkannor  
Vinkaraffer  
Vinkaraffiner  
Vinkorkar  
Vinkylare

### Öl

Ölbollar  
Ölkannor  
Ölsejdlar

TILLVERKNINGSTID	SAKORD	TAVNR	LITTERA
1920	REMSKIVA	23415	
1920	TVING	23382	
1920 CA	KAFFE, BURK FÖR	23455	A-B
1920-1930	GRAMMOPONSKIVA	23422	
1920-1940	BARNFILT	23648	
1920-1940	DRAPERIPROV	23652	A-B
1920-1940	DRAPERIPROV	23657	A-E
1920-1940	DRAPERIPROV	23669	A-H
1920-1940	DRAPERIPROV	23670	A-B
1920-1940	DRAPERITYG, PROV TILL	23650	A-B
1920-1940	GARDINPROV	23671	
1920-1940	HALSDUK	23644	A
1920-1940	MATTPROV	23647	A-C
1920-1940	MÖBELTYGPROV	23651	
1920-1940	MÖBELTYGSPROV	23645	A-E
1920-1940	MÖBELTYGSPROV	23646	A-E
1920-TAL	ELÖTKÄNGOR, PAR	23365	
1920-TAL	BRÄTTSKARARE	23765	
1920-TAL	BUTELJ, GLASFLASKA	23400	
1920-TAL	BÄRENHINGSSTALL	23442	
1920-TAL	KARAMELLER, BURK MED	23400	
1920-TAL	LINNELOPARE	21025	
1920-TAL	PLATFORH	23734	B
1920-TAL	PLATFORH	23735	
1920-TAL	PRESSHÄST	23378	
1920-TAL	REMSKIVA	23414	
1920-TAL	REMSKIVA	23417	
1920-TAL	TARA	23577	
1920-TAL	TÖBÅK, BURK FÖR	23407	
1920-TAL	UTTERLÅDA	23732	
1920-TAL	VINDA FÖR UTTERFISKE	23753	
1921	PÅLSKRAGE	20992	C
1922	SKED	23771	A-B
1925	RADIOAPPARAT	15093	
1925	RADIOAPPARAT	23379	A-B
1927	DUK	23664	
1927	DUK	23665	
1927	DUK	23666	
1927	NÄSDUK	21022	A
1929	NÄTBRINDNINGSLÅDA, MED TILLBEHÖR	23421	I-3B
1929 CA	SIDENKAPPA	21024	C
1930	MASKDOSA	23369	
1930	POTATISKVARN	23395	
1930-1940T	KAFFEKOPP	23749	
1930-TAL	FISKNÄT	23439	A-C
1930-TAL	FISKNÄT	23439	C
1930-TAL	FÖRKLÅDE, LOTTERIVINST	21020	B
1930-TAL	FÖRVARINGSBOX	23761	A-B
1930-TAL	FÖRVARINGSBOX	23762	
1930-TAL	PRESSBRÅDA	23747	
1930-TAL	REMSKIVA	23368	

Bilde 4.



## Survey of the use of data processing in museums

*Andrew Roberts*

Five areas of use of automated (computer) systems in museums were discussed:

information management uses (including the storage and processing of collection information as a basis for catalogues, information retrieval and collection control);

management control uses (including controlling financial, staff, visitor and shop stock information);

word processing uses (including the production of reports and manuals);

analytical research uses (including the analysis of statistical results of an excavation or technical examination);

education uses (including interactive displays in public galleries).

The main emphasis was on information management uses. Influences behind the growing adoption of automated systems for information management include the declining cost and increasing availability of suitable systems, the gradual development of staff expertise and the powerful advantages which they have over a manual approach (such as their support of a higher standard of collections management and their ability to maintain, update and verify records, to produce catalogues and to provide effective retrieval facilities). Against these positive influences must be balanced the need for expertise (rarely available in many museums), the short-term and long-term cost, the shortage of appropriate hardware, the scale of museum collections and the nature of museum records. Some of the hardware and software options available to museums were considered, including the different methods for gaining access to computer facilities and the various approaches to software development. Reference was made to the need for staff for system management, system development and operations and to problems associated with data.

With this background, work on an international scale and within a number of different countries was reviewed. The main coordinating museum committee is CIDOC, the Documentation Committee of the International Council of Museums. This has been developing data standards and exchange standards for museums, and acting as a forum for the various national documentation committees. Work by national organisations and individual museums in different parts of the world were contrasted. The approach in a few countries (Canada, France and Italy) has concentrated on the provision of a national computing centre, funded to assist individual museums and to build up a national inventory. (In

Canada, this initial emphasis has now been superceded by the national centre supporting the collections management requirements of its individual users, together with the semi-independent local use of systems within individual museums.) The emphasis elsewhere has been on the requirements of individual museums, with plans for cooperative inventories being considered to be of secondary importance. (The varying roles of the Museum Documentation Association, including advice and training in both manual and computer systems, system development, publications and computing services, was noted.)

Future trends include the likely availability of a growing range of standard museum applications of computer packages, developed by both commercial and museum organisations; the expanded use of microcomputers for a wide range of functions, including administration, word processing and information management; and the need for more staff with relevant expertise. The effective use of computer systems for information management purposes will depend on the quality of system developments, standards and terminology control. There will be an urgent need for staff training and the availability of advisory support for individual museums. National or regional organisations will play a crucial role. In an area like Scandinavia, a multi-national coordinating approach may prove very effective.

Three sources for further information are available from the MDA:

Microcomputers in museums (1984)

Planning the documentation of museum collections (1985)

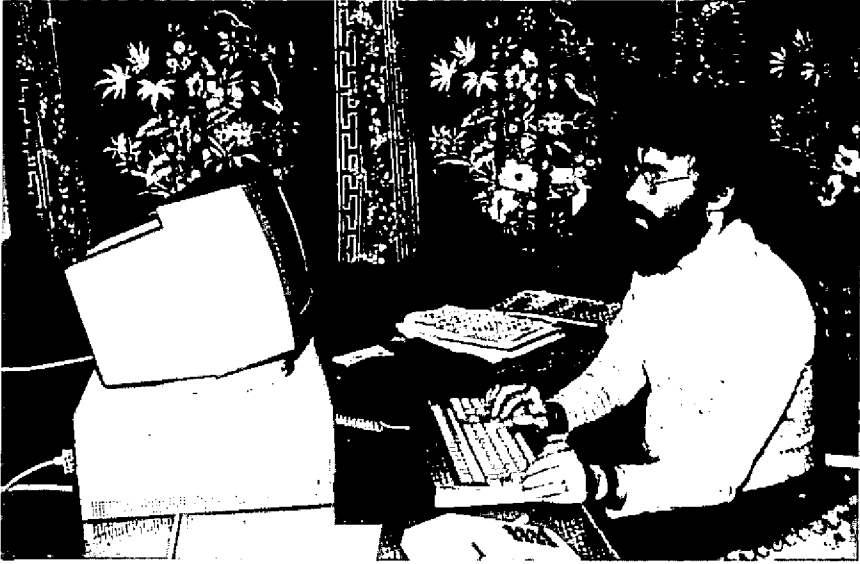
Museum documentation systems: developments and applications (1986).

## **The Canadian Heritage Information Network**

*Stephen H. Delroy*

## **The Canadian Heritage Information Network**

The Canadian Heritage Information Network (CHIN) was created in 1972 although data entry did not begin until 1975 and our final name was not assigned until 1982. Between 1972 and 1975, various task forces and committees set the goals of the programme which is one of the National Programmes administered by the National Museums of Canada, a federal crown corporation reporting to Parliament via a board of trustees and the minister of the Department of Communications. The



*Stephen H. Delroy demonstrerer databasen til The Canadian Heritage Information Network i Ottawa.*

original goal was to set up a national inventory of public museum collections. Data bases were set up for each discipline such as History or Ornithology. On the basis of museum declarations, it was assumed that the collections were completely documented and the information seldom changed.

Between 1975 and 1978, we encountered the reality of undocumented collections with rapidly changing information. A moratorium on the acceptance of new client museums was put into place and attempts were made to get our software to handle changing information. In 1982, a joint venture with the Control Data Corporation was begun. The new hardware could handle vast amounts of data with on-line access from across the country. The software, BASIS from the Batelle Memorial Institute, could process the numerous, large and many-fielded museum records we then had.

Since 1982, our mandate has been to provide:

1. a national inventory of public collections;
2. support for collections management; and
3. advice on museology, computerization and new technology.

Museums that are members of the network receive these services for

free, including initial training, data storage and retrieval, and communications costs. Museums supply their own entry and retrieval staff, and buy their own modems and terminals. Museums not covered by the original mandate, such as other government departments, private and foreign museums, are paying clients. Paying clients pay for all their costs including communications, training, computer usage and on-going support.

CHIN has recently replaced its old CYBER with another larger CYBER 835 with over 8 gigabytes of storage. In addition, we have more than 10 microcomputers of various types, mostly IBM XT's and AT's. Over 120 terminals and more than 20 user microcomputers are directly connected to the CYBER using DATAPAC and local nodes in 17 Canadian cities. This network is being improved by the addition of multiplexors (5 installed, 9 more planned by 1989).

The BASIS software, which together with the network and the CYBER form the PARIS system, allows for 3000 data fields. We are using almost 720 in the Humanities and 620 in the Natural Sciences. Records can be 10,000 characters long and record types are supported. All fields can be indexed, have multiple occurrences, be multi-indexed and contain alphabetic and/or numeric data. Entry, editing, retrieval, analysis, and reporting are all on-line functions using a naturalistic data base language in either English or French. BASIS allows thousands of records to be changed at once and on-line by the user; this greatly aids the standardization process. CHIN has developed a micro to mainframe link that allows up and down-loading of data. CHIN has also developed three modes of rapid data entry with which some users are entering over 230 records of 15 fields per day, per operator.

Participants in the network are of all sizes from art galleries with 1,000 records to the Royal Ontario Museum with 3/4 million records. Over 20 humanities and natural science disciplines, from archaeology to parasitology, have data on the network. There are more than 65 user institutional data bases comprising 2.2 million records and representing 4.4 million specimens and objects. This represents approximately 25% of the potential 9 million museum records in Canada. There are also a number of National Reference Data Bases, including the national inventories embodied in the National Humanities and National Natural Science Data Bases. Two networks of regional museums are about to become part of these national data bases. Three (soon to be four) museums have their Local Area Networks linked on-line to CHIN and the Conservation Information Network, involving the J.P. Getty Trust, the Art and Archaeology Technical Abstracts, etc. has also been formed as part of the CHIN system.

To a member institution, all this means that they can register their collections on-line, track object movement, plan exhibits, produce admi-

nistrative reports, monitor conservation and loans, and conduct object and object-associated research from the same terminal or microcomputer. Participants are buying additional terminals for research and managerial staff, and several are placing terminals in exhibit areas (with the menu driven options of BASIS). We see a trend away from any hard copy, such as cards, and toward complete on-line access. There is also a trend, substantiated by surveys of all our members, toward the creation of reference data bases and information sharing. We are encouraging institutions to store information at the level it is being used. In other words, local information, for example where the object is, is stored at the museum on a microcomputer while national information, such as what the object is, is stored at CHIN. With micro to mainframe links, the two kinds of data can always be relinked. This has given a tremendous focus to our efforts to standardize museum data. It is no longer a question of simply defining new fields in our two on-line data dictionaries, now the emphasis is on standardizing the content of those fields used as retrieval keys for national information. To this end, user working groups have been formed along disciplinary lines to set and authorize standards with support from CHIN's documentation research. Nonetheless, CHIN at the same time is promoting the sharing of collections management expertise and the rational introduction of new technology. This can be seen in the twin themes of this year's PARIS User Representatives Seminar which are Loans and the Microcomputer.

## **Erfaringer fra nomenklaturarbeid**

*Jon Birger Østby*

NKKMs edb-komite ble tidlig klar over at alle synonymer og varianter av gjenstandsbetegnelser ville skape spesielle problemer ved innføring av edb til gjenstandsregistrering. Komiteen tok derfor initiativ til å få i stand et prøveprosjekt for å arbeide med «nomenklatur for etnologisk gjenstandsmateriale fra etterreformatork tid». NAVF (Norges Allmennvitenskapelige Forskningsråd) har stilt midler til rådighet, og etter tre år ble prøveprosjektet fullført i 1984. NKKMs edb-komite supplert med dosent *Finn Hodnebo* fra Leksikografisk institutt ved Universitetet i Oslo har fungert som styringsgruppe, og undertegnede var ansatt som forsker/prosjektleder.

I prøveprosjektet er det utarbeidet eksempler på gjenstandsomenklatur, og det er trukket opp retningslinjer for organisering av fortsatt nomenklaturarbeid. Det er utarbeidet nomenklatur for «Spise- og serveringsbestikk», «Drikkestell», «Melkestell» og «Tiner, esker og skrin».

Disse gjenstandsgruppene ble valgt fordi de står sentralt i samlingene ved de fleste kulturhistoriske museene. Samtidig skulle de gi erfaringer med nomenklaturarbeid for gjenstandsgrupper med tilknytning til forskjellige miljøer og tidsepoker.

Nomenklaturlistene inneholder:

Normalbetegnelser på nynorsk og bokmål

Synonymordlister med dialektbetegnelser

Illustrasjoner

Definisjoner

Henvising til klassifikasjonskode i «Outline of Cultural Materials»

Oversikt over sentral litteratur for gjenstandsgruppene

Stikkordregister for termer og synonymer

De tre første listene har også en oversikt over delbetegnelser.

I nomenklaturarbeidet har det vært viktig å få reaksjoner fra museene. Før publisering har derfor listene vært utsendt til høring ved et utvalg av museer der alle landets fylker er representert. Alle innkomne merknader har vært forelagt for Styringsgruppa, og listene er gjennomgått av Norsk språkråd. Nomenklaturen er utarbeidet på grunnlag av et omfattende registreringsarbeid.

Det har vist seg at nomenklaturlistene er nyttige i bruk ved registreringsarbeidet også for museer som ikke har planer om å ta i bruk edb. Det har vært særlig hyggelig å få mange positive kommentarer fra mindre museer, og det virker som det er en utbredt oppfatning at et fortsatt nomenklaturarbeid er av stor betydning både for små og store institusjoner.

I prøveprosjektet har det vært viktig å velge ut gjenstandsgrupper som har gitt erfaringer med å arbeide med gjenstander fra forskjellige miljøer og tidsepoker. Men det har også vært et mål å velge gjenstandsgrupper som er godt representert i museenes samlinger slik at nomenklaturlistene skulle komme til størst mulig nytte. Det er årsaken til at alle emnene er valgt innen husholdning.

Erfaringene fra arbeidet med listen med nomenklatur for esker, tiner og skrin viser at det er svært vanskelig å utarbeide nomenklatur med utgangspunkt i form, konstruksjon, eller materiale. Det vil alltid være gjenstander der normalbetegnelsen vil måtte avvike fra de generelle retningslinjene, og i det fortsatte nomenklaturarbeidet bør man derfor ta utgangspunkt i gjenstandenes bruk slik som det er gjort i de øvrige listene.

Det var en forutsetning at det skulle benyttes edb i Nomenklaturprosjektet. For de to første listene ble alle registreringer ført på kladdekort og sendt til NAVFs EDB-senter i Bergen for innlesing, og Senteret leverte sorterte lister. Senere fikk prosjektet disponere egen dataterminal på Norsk Folkemuseum, og dette betød en vesentlig arbeidsbesparelse.

All registrering kunne da utføres direkte ved terminal, og ajourførte utskrifter og sorterte lister kunne produseres etter som det var behov for det. Samtidig ble alle manuskripter skrevet på tekstbehandling. Dette gjorde det enkelt å rette opp forslag til endringer, og utskrift av den endelige versjonen ble brukt som grunnlag for trykking av listene i offset.

NKKM har foreslått at et framtidig nomenklaturarbeid bør knyttes til Norsk Folkemuseum, og at det etableres skiftende arbeidsgrupper for de forskjellige gjenstandsområdene som skal behandles. Etter anmodning fra NKKM har Norsk Folkemuseum derfor søkt om en fast stilling over statsbudsjettet til denne virksomheten. I søknaden er det understreket at dette ikke er en stilling primært for Norsk Folkemuseum, men en stilling som skal tjene alle de kulturhistoriske museene. Plasseringen er foreslått utfra praktiske og tekniske hensyn. Men søknaden er foreløpig avslått.

I prøveperioden har prosjektet hatt arbeidsplass ved Norsk Folkemuseum. NKKM fant det naturlig å knytte prosjektet til landets største kulturhistoriske museum for å få lett adgang til museets kataloger, bibliotek og faglige ekspertise. Dette har vist seg å være en svært verdifull tilknytning. Men det finnes også mange museumssamlinger utover i landet som er viktige kilder i et nomenklaturarbeid, og det vil derfor alltid være viktig både med arbeidsgrupper, godt kontaktnett og med studiereiser. Det er uvisst hvordan det nå vil gå med nomenklaturarbeidet. NKKM og Norsk Folkemuseum vil i fellesskap fortsette å arbeide for å finne en løsning, og vi håper at de bevilgende myndigheter vil forstå hvor viktig det er at nomenklaturarbeidet får fortsette.

## **Data Processing In Art Museums Computers For Art History's Sake**

*Margrethe Pedersen*

In this paper, I should like to introduce the Danish art museums' activities and plans concerning data processing, while emphasizing the impact of the community's scholarly and administrative interests. I also want to stress the importance of an intensive dialogue between curatorial art historians, on one hand, and faculty and library art historians, on the other hand. Finally, I shall touch on some methodological questions that may seem more or less intertwined with art historians' involvement with database technology.

The development of a database of the Danish museums' collections of

visual art is under the auspices of the Ministry of Cultural Affairs. As such, the project is stipulated in the Museum Act at present in force, dating from June 6, 1984, and was formally initiated in 1985.

The official name of the project is: Kunst Indeks Danmark, and in accordance with the Museum Act, Kunst Indeks Danmark's office and facilities are at the national gallery: Statens Museum for Kunst. It is a project involving the entire Danish museum community as well as the universities and The Royal Academy of Arts. Upon arrangement, all interested parties, national and international, will be able to get access to the future database.

Kunst Indeks Danmark will be responsible for the scientific standard and homogeneity of the database, performing input routines and maintaining and expanding the lists of preferred terms and the thesauri. The museums, on their part, are responsible for continuously presenting Kunst Indeks Danmark with a minimum cataloguing level for each work of art in their collections, be it in manual or electronic form.

At the local level, the museums are getting increasingly involved with computerizing various routines: administration, climate control and specific museum tasks. So far, however, only 3 art museums have their own computer equipment. By 1990, ten out of a total of about forty art museums are expected to have micro- or minicomputers at their disposal.

Kunst Indeks Danmark is currently running a database (PC/Focus) and a word-processing program (Wordperfect) on a microcomputer (Olivetti M24), supplied with a 20MB hard disk. By September 1986, our database (VAX/FOCUS) will be running on a minicomputer (MicroVAX II, 2 x 71MB Winchester disk).

In the light of experiences made within the compass of a series of international art history computer projects, Kunst Indeks Danmark is prototyping its database. Work in progress includes analyzing and refining a representative section of the museums' standard art cataloguing output, developing rules for format and content control (incl. thesauri), and analyzing and deciding upon the application of existing standards and thesauri. The first task to be completed is the systematic recording of the Danish art museums' 28,000-30,000 paintings and sculptures. By the end of 1987, the database is expected to render some practical results. Kunst Indeks Danmark is paying particular attention to the following data categories: iconographic classification, topographical information (artist's record, iconographic record, museum record), art medium and materials, authenticity, function, status, and chronological information. Concerning iconography, it has been decided to use ICON-CLASS, an iconographic classification system developed at the University of Leyden. Abstract works of art, on their part, may at some later point in time be subject to a special kind of form and content classification, possibly based on semiotics. Some of the object information that is



recorded by the art museums will not be input by Kunst Indeks Danmark in the first relay. This applies to bibliographic and exhibition information, provenance data and a narrative characterization of the single work of art.

We are deeply indebted to a number of colleagues abroad for their having inspired us and provided us with abundant documentation on their computerized catalogues and other applications. We collaborate with Bildarchiv Foto Marburg, to whom we are extremely grateful for their having initiated us into the archive's principles of art documentation. Other sources to be mentioned here are La Joconde (Paris), The Canadian Heritage Information Network (Ottawa), The Getty Art History Information Program (Santa Monica). We hope that this list will never be closed. Indeed, we are convinced of the necessity and the many advantages of a global dialogue on art history and data processing. Let us therefore consider this conference a creative opportunity for the Nordic art museum community to intensify the exchange of facts, ideas, and hopefully also data.

## **EDB-virksomheten ved Kulturarvet**

*Rune Hermansson*

Utveckling av informationssystem för föremål och bilder har bedrivits under ett antal år vid Kulturarvet. Mindre satsningar har gjorts för att skapa system för bibliotek och administrativa rutiner. Föremålsdokumentation hanteras av system REFORM. Föremålsregistrering görs på en blankett eller matas direkt in på mikrodator. Enkla tester görs i samband med inmatning för att undvika fel i data. Från mikrodator överförs data via tråd eller post till Medicindatas DEC 10 i Göteborg. En rad testrutiner finns inlagda i systemet som garanterar att många typer av fel inte förekommer i slutresultatet. Normering av sakord och realnummer görs i en speciell rutin. På det sättet får man en enhetlig nomenklatur och undviker synonymer. Kontroll sker hela tiden mot tidigare inmatade sakord och i kombination med realnummer. Koordinater för geografiska orter, latituder och longituder lagras i en speciell tabell, varifrån data hämtas för automatisk koordinatsättning.

Registreringsblanketter skrivs efter rättning ut vid datacentralen och kommer att fungera som originalhandlingar. Listor med olika ingångar och sorteringsordning tas fram, t ex med ingång från sakord, material och namn. Alla listor kan också framställas på mikrofisch.

Alla data samlas i en databas med fullständig fritextsökning. Som databashanterare används 3RIP. Sökning kan göras på ord eller delar av ord i hela textmassan eller bundet till specifika fält och kombineras med OCH ELLER INTE (AND OR NOT). Dialog med systemet kan ske på svenska eller engelska. I databasen kan man välja att söka i sitt eget material eller i den totala databasen.

Arbete pågår med överföring till dBASEIII på mikrodator. Det kommer då bli möjligt att lokalt bygga upp en egen databas.

Foton av föremål överförs till bildskiva. Första steget är överföring till 35 mm spelfilm i en speciell kamera. Från filmen sker skanning till videotape, varifrån bildskivor sedan framställs. Bilderna visas på teve med hjälp av en laserbildspelare. Den kan användas fristående eller kopplad till datasystemet.

Olika typer av bilder hanteras i system BILDREFORM. Det har tagits fram i samarbete med Armemuseum. I princip är det uppbyggt som REFORM. Data finns i två skilda databaser, den ena med uppgifter om den fysiska bilden, fotograf m m, den andra med uppgifter om bildinnehåll, t ex sakord, realnummer och personnamn. För samiskt material finns en speciell variant framtagen för Nordiska museet. Även länsmuseum i Skara har en variant med något färre uppgifter vid inmatningen. Den slutliga databasen är dock lika.

Utvecklingen av systemen har finansierats genom ekonomiskt stöd från DFI, Delegationen för vetenskaplig och teknisk informationsförsörjning. Projektmedel har beviljats till Armemuseum, Kulturarvet och Medicindata.

Systemet finns publicerat i Svenska museer 4/80 och Humanistiske Data 2-83.

## **Nationalmuseets bevaringssektion og brugen af edb**

*Knud Holm*

Nationalmuseets bevaringssektion er dannet ved sammenlægning af konserveringsafdelingerne ved Nationalmuseet, og planen er at samle aktiviteterne i Brede nord for København, hvor konserveringsanstalten allerede i en årrække har omfattet konservering af alt recent materiale, mens det arkæologiske materiale endnu fortrinsvis er blevet behandlet i museets gamle konserveringsanstalt i Ny Vestergade d.v.s. lige ved siden af museumskomplekset i Prinsens Palæ.

Bevaringssektionen varetager således konserveringen af Nationalmuseets samlinger, men derudover udøves konsultativ virksomhed ligesom

konservering og undersøgelser af teknisk og analytisk art tilbydes mod betaling. Kunderne er andre museer og institutioner, danske kirker og ministerier som boligministeriet og miljøministeriet m.fl.

Personalet er for tiden ca 110 personer. Budgettet omfatter faste lønninger på godt 11 mill. kr, driftsbevillinger på omkring 2 mill. og omsætningen indenfor den rekvirerede virksomhed er på ca 5 mill., hvoraf hovedparten naturligtvis er lønninger til det personale, som er ansat i sammenhæng hermed.

Sekretariatet er bemandet med 10 personer, og det er indtil videre her, at bevaringssektionen har investeret i edb-udstyr.

Vi har anskaffet 4 Olivetti PC'ere M24, den ene af dem nyeste model, den såkaldte SP (special performance) model. To af de almindelige M24'ere er udbygget med en extern 10 MB hard disk enhed. Tre af maskinerne forsøges nu koblet sammen i en netværksløsning af nrk. 10-net. De deler en skønskriver og en matrixskriver. Herudover er en matrixskriver knyttet til den sidste enlige maskine.

Maskinerne er indtil nu udelukkende blevet brugt til tekstbehandling, - som omgående er blevet meget populært hos brugerne - og til økonomistyring (time/sags-styring). Vi benytter Dantekst-programmet til tekstbehandling og et time/sags-program fra PolySoft på regnskabsiden.

Tanken er imidlertid snarest at gå over til at føre den detaljerede registrering af de genstande, der passerer konserveringsanstalten, over i et edb-baseret medium.

Det kommende system kommer til at bygge på en registrering af «sager», hvor vi ved sagsoprettelsen definerer (afgrænser), navngiver og nummererer (d.v.s. daterer og eventuelt undernummererer) sagen. Hver sag kan vedrøre flere eller færre genstande, som naturligvis hver for sig registreres og beskrives på behørig vis. Konserveringssager udgør i reglen en afsluttet og afrundet enhed, og genstande, der dukker op igen senere i en ny sammenhæng, kommer altså til at indgå i en ny konserveringssag. Genstandsregistret vil imidlertid til hver en tid kunne fortælle, i hvilke sagsammenhænge en genstand har optrådt. Til en konserveringssag knyttes iøvrigt alle relevante oplysninger - herunder henvisninger til fotoarkiver, prøvesamlinger m.v.

Som det har været hidtil, vil det også være i en startfase i det nye system, at indgangen vil være sagsbetegnelse og sagsnummer samt genstandsnummer, hvor det drejer sig om museumsgenstande, som fortrinsvis bør være inventariseret forud, og ellers bør blive det, som led i indføringen på konserveringsanstalten. Her tænkes særlig på arkæologiske fund, som ofte vil arrive direkte fra udgravning med foreløbige numre men ikke med endeligt inventarnummer.

Men det er klart, at et endemål vil være, at vi kan søge ud fra et antal saglige kriterier, som kunne være genstandsbetegnelse, materiale og fremstillingsteknik måske også tilstand og konserveringsbehandling - men det

vil kræve et system for beskrivelse af disse ting, som vi ikke råder over endnu – endvidere formentlig fundsted, indsamlingssted eller hjemsted, produktionssted og datering. Såfremt mester eller fabrikant er kendt også denne. For administrative formål kan det være praktisk at kunne søge på rekvirerende institution, afdeling og kontaktperson. Dette også selvom sagsbetegnelsen i reglen vil afsløre hvorfra genstanden kommer, idet sagsbetegnelsen typisk vil bestå af en steds- eller ejerbetegnelse + en genstandsbetegnelse eller en rammebetegnelse for det sagen drejer sig om.

Genstandsnumre – nogen gange kaldt inventarnumre andre gange kaldt museumsnumre – må være fuldstændig entydige, hvad genstandsbetegnelsen i følge sagens natur ikke kan være. På samme måde må sagsnumre – der som før sagt indeholder en datokode + eventuel undernummerering – være fuldstændig entydige, hvad sagsbetegnelsen ikke nødvendigvis er. En sagsbetegnelse hos os kunne være «Roskilde domkirke, kiste», betydende, at denne sag omhandler restaurering af én kiste fra Roskilde domkirke. Den kunne være «Roskilde domkirke, kister», betydende, at sagen omfatter et antal kister – måske 10 – optaget til behandling på en gang. År senere opstår måske en ny sag, der også hedder «Roskilde domkirke, kister» f.eks. fire andre kister. Men de to sager er adskilt i tid, bærer hvert sit sagsnummer og kan derfor ikke sammenblandes.

Som en del af Nationalmuseet vil vi med tiden skulle kobles op mod museets centrale registre og DKC, sådan at vi både bidrager med data og kan trække på data der. Men vi vil formentlig stadig skulle have vore egne registre, hvor vi kan holde styr på, hvad der passerer gennem vore hænder fra alle mulige steder. Og vi vil have brug for et system, der kan håndtere alle mulige typer af genstande, hvadenten der er tale om kunst, arkæologi, kulturhistorie, etnografi eller naturvidenskabelige samlinger.

Vi behøver et system, der tillader os at registrere besigtigelser af genstande i museets samlinger, således at et opkald på et inventarnummer kan fortælle os, hvornår genstanden sidst blev besigtiget af en konservator, hvem han var, og om han fandt den OK, eller om han udtog den til behandling.

## Nationalmuseets fremtidsprojekt

*Carsten U. Larsen*

Et dansk fond – Egmont Fonden – tilbød i 1982 at bekoste utarbejdelsen

af et forslag til en helhedsplan for Nationalmuseets udvikling frem til år 2000, baseret på en analyse af det danske samfunds forventninger og krav til sit centrale kulturhistoriske museum.

I første omgang stillede fonden 2.3 mill. kroner til rådighed til et studieprojekt, som fik navnet «Nationalmuseet & Fremtiden».

Resultatet af dette projekt var et forslag om tilskud til Nationalmuseet på kr. 250 mill., som skal fordele sig på byggearbejder, nyopstilling af samlingerne, magasinhus og indførelse af ny teknologi.

Egmont Fonden yder kr. 100 mill. og d. 22/6 vedtog Folketinget at yde de resterende kr. 150 mill.

Hvad der har særlig interesse i denne forbindelse er anvendelsen af ny teknologi. Der er afsat kr. 50 mill. til formålet frem til 1994 og årligt ca. 5 mill. til en fortsat drift efter projektets udløb.

Organisationsmæssigt vil indførelsen af ny teknologi blive henlagt til en projektgruppe (Dokumentations-gruppen), som refererer til direktio-  
nen.

Driftsmæssigt vil den ny teknologi blive indført som et distribueret system, hvor Det kulturhistoriske Centralregister (DKC) vil udgøre den centrale del og de enkelte afdelinger de decentrale dele. Kommunikationsmæssigt vil huset blive bundet sammen af et lokalnet (bredbånd), som sikrer, at alle kan få tilgang til dyre, fælles maskiner (f.eks. plottere, tape-stationer).

Groft set kan der skelnes mellem to former for anvendelse af ny teknologi på museet, nemlig 1) administration og 2) museumsdokumentation (herunder antikvarisk administration).

Rent administrativt vil der blive tale om indførelse af decentrale mikrodatamater med tekstbehandlingsprogrammell og økonomiske programmer, som kan føre de enkelte afdelingers regnskaber og over lokalnettet indberette til museets økonomiske forvaltning, som igen kan indberette til staten. Kun en ringe del af midlerne vil blive anvendt hertil, idet der er tale om gængse og afprøvede systemer.

Langt de fleste midler vil blive anvendt til registrering af museumsdokumentation. Der vil først og fremmest blive taget fat på at EDB-registrere museets store bestand af genstande. En arbejdsgruppe har igennem ca. 1 år udarbejdet et tværmusealt registreringssystem for genstande. Dette system vil være færdigudviklet og klart til implementation i de enkelte afdelinger i foråret 1987, og vil senere blive anvendt af landets øvrige kulturhistoriske museer.

Ud over den alfanummeriske EDB-registrering vil der også blive foretaget en billedregistrering af dele af samlingerne. Disse billeder vil blive lagret på optiske plader og vil blive anvendt til bl.a. fremstilling af interaktive videoplader.

Under EDB-registreringen af afdelingernes samlinger vil de enkelte afdelinger få gratis maskinel, programmell og konsulentbistand, ligesom

der vil blive givet undervisning i EDB til museets personale.

Dokumentations-gruppen har frem til midten af 90'erne fået stillet 14 nye stillinger til rådighed. En del af disse (ca. 1/3) vil blive brugt centralt (programmører, systemfolk etc.), mens ca. 2/3 vil være til rådighed for de enkelte afdelinger til datafangst etc.

Det kulturhistoriske Centralregister, som er et fællesmusealt organ med placering ved Nationalmuseet, vil stille de registrerede data til rådighed for den samlede museumsverden. Ligeledes vil data blive bearbejdet således, at publikum, via terminaler i indrettede rum, kan få adgang til selvstændigt at gå på opdagelse i museets elektroniske arkiver.

Det er Nationalmuseets håb, at man både i Norden og bredere internationalt vil følge dokumentations-projektet nøje. Vi håber at modtage hjælp, kritik, misundelige blikke, hep-råb og især ideer i det kommende decennium. Til at sikre dette vil museet jævnligt udsende skriftligt materiale og forsøge at etablere informationsdage på museet.

## **Strategiar for edb-katalogisering av gjenstandsmateriale ved kulturhistoriske museum**

*Magne Velure*

Sidan 1981 har fleire norske museumsmiljø teke i bruk edb i katalogiseringsrutinar for gjenstandar og foto. I april 1986 vart det avslutta eit utviklingsprosjekt, finansiert av NAVF, der ein har arbeidd fram felles metodar og rutinar for edb-katalogisering av gjenstandar og foto. Det har vore lagt ned mykje arbeid i å finna fram til ein felles standard, og så langt har ein lukkast i rimeleg grad.

I denne innleiande fasen har målet vore å nå fram til ein katalogiseringsstandard som skal gje mest mogeleg fullstendig «informasjonskart» for kvart katalogisert objekt. Intensjonen med all katalogisering er å sikra best mogeleg kjeldeverdi for gjenstandane og fotografia. Det langsiktige og ideelle målet vil difor vera at alt aktuelt tilfang skal katalogiserast etter ein fullstendig standard.

I dei musea som etter kvart skaffar seg det første edb-utstyr vil ein snart oppdaga at det er fleire faktorar som gjer det vanskeleg å ta i bruk den fullstendige katalogiseringsstandarden slik utan vidare. I høve til den aktuelle informasjonsmengda som ein gjerne vil ha med, vil platelageret vera for lite på det utstyret som musea har råd til å kjøpa. Men det største hinderet er mangel på personale til å gjera sjølve katalogiseringa, det er nemleg den tidkrevjande faktoren i heile prosessen. Med utgangspunkt i røynsler frå andre museum har vi funne ut at for å få ei

fullstendig edb-katalogisering av Maihaugens ca. 35.000 gjenstandar, vil vi måtte rekna med 20 årsverk. Dersom vi sette to av våre konservatorar til å arbeida med dette på heiltid, ville det altså ta 10 år før vi fekk skikkeleg oversyn over *eksisterande* gjenstandsmasse.

Med små utsikter til å auka personalressursar må vi finna alternative strategiar, som på den eine sida gjev ei rask edb-aktivisering av samlingane, samstundes som ein tek omsyn til tilgjengelege personalressursar. Slik situasjonen er i dag ved musea, vil ei katalogisering etter fullstendig standard berre vera aktuell for mindre omfattande samlinger, opp til 5-6000 gjenstandar. Dei andre vil måtte velja mellom to alternative framgangsmåtar.

- a) Ein brukar fullstendig standard, men konsentrerer seg om tilveksten. Dei eldre samlingane katalogiserer ein berre i den grad ein har tid og pengar.

Ulempen med denne framgangsmåten er at storparten av samlingane ved musea vert liggjande ukatalogiserte og dermed vanskeleg tilgjengelege både for forskning og formidlingsarbeid. Føremønnen er at ein følgjer fullstendig standard, og har dermed sjansen til å få høg kjeldeverdi på det som vert katalogisert.

- b) Den andre måten er å gjennomføra katalogiseringa stegvis, med ulike nivå i oppbygginga av informasjon. Ved De Sandvigske Samlinger, Maihaugen har vi valt ein slik modell og med økonomisk stønad frå NAVF til utstyrskjøp har vi sett i gang eit prøveprosjekt.

Utgangspunktet er det fullstendige katalogiserings skjemaet (Fig. 1). Men vi tek berre ut ein del av dei viktigaste rubrikkane og lagar eit skjema for *oversynsregistrering av alt gjenstandsmateriale* på grunnlag av hovudprotokollen. Det er viktig å passa på at ein definerer og legg til rette skjema for oversynsregistreringa slik at opplysningane i kvart felt utan vidare kan gå inn i hovudskjemaet (Fig. 2). Alt som vert skrive inn skal berre skrivast *ein* gong. Meininga med dette er at ein med utgangspunkt i oversynsregistreringa kan gå inn med fullstendig katalogisering av dei gjenstandsgruppene som det til ei kvar tid er ønskjeleg å arbeide vidare med, t.d. i samband med utstillingar, forskingsprosjekt eller også p.g.a. stor etterspurnad frå publikum.

Utstyret som vi brukar, er ein DISCOVERY fleirbrukarmikro med 70 Mb disk og 55 Mb bandstasjon. Førebels er det kopla til tre Tandberg-terminalar og to Ericsson PC. Kvar terminalbrukar disponerer ein 10 Mz 8086 brukarprosessor. Operativsystemet er Concurrent DOS og Personal C/PM-86. Sjølve registreringsystemet er utvikla i Dataflex.

Så langt har oversynsregistreringa gjeve dei resultat som vi har vona på:

- ein person registrerer 3500-4000 gjenstandar pr. månad. Det kan i stor grad brukast personale på sysselsettingsmidlar.





```

/DSSREG
01 GJENSTAND : ( _____ ) 02 MUSNR: ( _____ )
03 INNK. NR  : ____ . ____
04 FDTNUMMER: _____

05 KLASS.    : ( ____ . ____ ) ____ . ____ . ____ . ____ . ____
05 DATERING  : ____ . til ____ . _____

07 BRUKSSTED : __ STED { ____ } _____
                ADR _____ GNR/BNR ____ . ____

08 PROD. STED : __ STED _____
                ADR _____

__ NAVN _____ YRKE _____
   FØDT/ETABL ____ . ____ DØD/NEDL ____ . ____
STED _____ ADR _____

__ NAVN _____ YRKE _____
   FØDT/ETABL ____ . ____ DØD/NEDL ____ . ____
STED _____ ADR _____

13 MATERIALE : ( _____ )
/*

```

Fig. 2

- oversynsregistreringa gjev grunnlag for listeprodukt som dekkar dei fleste krava både frå museumspersonalet og publikum.
- oversynsregistreringa tek med klassifikasjonskodar (Outline of Cultural Materials), og det gjer at vi på ein relativt enkelt måte kan avdekka manglar i samlingane våre. Dette gjev oss i neste omgang eit betre grunnlag for prioritering av dei områda der vi bør konsentrera vår innsamlingsinnsats. Edb kan med andre ord hjelpa oss til å få ei meir problemorientert og offensiv innsamlingsverksemd.

## Survey of the use of data processing in museums

*Andrew Roberts*

Five areas of use of automated (computer) systems in museums were discussed:

information management uses (including the storage and processing of collection information as a basis for catalogues, information retrieval and collection control);

- management control uses (including controlling financial, staff, visitor and shop stock information);
- word processing uses (including the production of reports and manuals);
- analytical research uses (including the analysis of statistical results of an excavation or technical examination);
- education uses (including interactive displays in public galleries).

The main emphasis was on information management uses. Influences behind the growing adoption of automated systems for information management include the declining cost and increasing availability of suitable systems, the gradual development of staff expertise and the powerful advantages which they have over a manual approach (such as their support of a higher standard of collections management and their ability to maintain, update and verify records, to produce catalogues and to provide effective retrieval facilities). Against these positive influences must be balanced the need for expertise (rarely available in many museums), the short-term and long-term cost, the shortage of appropriate hardware, the scale of museum collections and the nature of museum records. Some of the hardware and software options available to museums were considered, including the different methods for gaining access to computer facilities and the various approaches to software development. Reference was made to the need for staff for system management, system development and operations and to problems associated with data.

With this background, work on an international scale and within a number of different countries was reviewed. The main coordinating museum committee is CIDOC, the Documentation Committee of the International Council of Museums. This has been developing data standards and exchange standards for museums, and acting as a forum for the various national documentation committees. Work by national organisations and individual museums in different parts of the world were contrasted. The approach in a few countries (Canada, France and Italy) has concentrated on the provision of a national computing centre, funded to assist individual museums and to build up a national inventory. (In Canada, this initial emphasis has now been superseded by the national centre supporting the collections management requirements of its individual users, together with the semi-independent local use of systems within individual museums.) The emphasis elsewhere has been on the requirements of individual museums, with plans for cooperative inventories being considered to be of secondary importance. (The varying roles of the Museum Documentation Association, including advice and training in both manual and computer systems, system development, publications and computing services, was noted.)

Future trends include the likely availability of a growing range of

standard museum applications of computer packages, developed by both commercial and museum organisations; the expanded use of microcomputers for a wide range of functions, including administration, word processing and information management; and the need for more staff with relevant expertise. The effective use of computer systems for information management purposes will depend on the quality of system developments, standards and terminology control. There will be an urgent need for staff training and the availability of advisory support for individual museums. National or regional organisations will play a crucial role. In an area like Scandinavia, a multi-national coordinating approach may prove very effective.

Three sources for further information are available from the MDA:

Microcomputers in museums (1984)

Planning the documentation of museum collections (1985)

Museum documentation systems: developments and applications (1986).

## **Brugen af EDB i museumsarbejdet i Danmark**

*Ole Thamdrup*

I dette indlæg redegøres der hovedsagelig for det arbejde, der foregår i Statens Museumsnævn. Til slut gives en kort omtale af igangværende, praktiske anvendelser af EDB på lokalmuseerne.

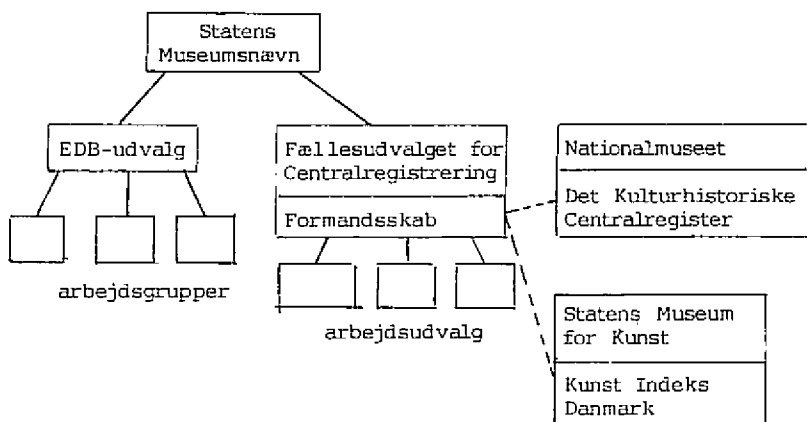
*Statens Museumsnævn* (SMN) er sammensat af valgte repræsentanter hovedsagelig med et museumsfagligt virkeområde.

SMN har to primære opgaver:

- at rådgive og bistå museerne i administrative, økonomiske og faglige spørgsmål
- at rådgive ministeren for kulturelle anliggender

Det daglige arbejde udføres af et sekretariat, beliggende i ministeriet for kulturelle anliggender.

Til løsning af særlige opgaver er der under SMN nedsat en række udvalg, herunder to udvalg vedrørende EDB. Følgende skema skulle gerne tjene til overskuelighed:



Strukturen er afgørende for at forstå det arbejde, der foregår i Danmark.

For det første er der tilstræbt en sikring af indflydelse fra de forskellige faglige grupper inden for museumsverdenen, idet museumsorganisationerne er repræsenteret i udvalgene.

For det andet er der i strukturen indbygget en mulighed for, at EDB-løsninger kan tilpasses både lokalmuseerne og centralregistrene således at en samkøring kan komme til at foregå.

Af hidtidige resultater fra dette arbejde kan nævnes udarbejdelse af rapporter, som er udsendt til alle museerne: «Anskaffelse af EDB-udstyr i de danske museer», hvor der er lagt afgørende vægt på anbefaling af IBM-kompatible mikrodatamater. «Tekstbehandling i museerne» med anbefaling af programmel der også indeholder kartoteksfaciliteter.

Inden længe vil der blive anbefalet programmel til museernes regnskabsfunktioner.

Under EDB-udvalget foregår der et intensivt arbejde med udvikling af et system til registrering af den dokumentation, som museerne indsamler. Planen er, at der inden 1987 foreligger en skitse til systembeskrivelse. Grundprincippet herfor er et system til styring af sager, idet det vurderes, at museerne har behov for andet og mere end et system til registrering af genstande. Også sager vedrørende f.eks., arkæologiske og etnologiske undersøgelser skal indgå i registreringen.

Dette lokalmuseumssystem tænkes også tilpasset, så der bliver mulighed for at indrapportere en delmængde af lokalmuseets oplysninger til de centrale registre: Det Kulturhistoriske Centralregister og Kunst Indeks Danmark.

### *Lokalmuseernes igangværende/praktiske anvendelse af EDB*

Flere museer har inden for den seneste tid anskaffet EDB-udstyr, der først og fremmest anvendes til administrative formål.

Visse museer har dog ved egen indsats erhvervet erfaringer med EDB-baseret registrering:

F.eks.:

Arbejdermuseet, København

Orlogsmuseet, København

Louisiana, Museum for moderne Kunst, Humlebæk

Forhistorisk Museum, Moesgaard, Århus

En række museer i Odense

### **Computerisation in Finnish Museums**

*Jouko Heinonen*

There are about 600 museums in Finland, of which a hundred are run professionally. With the founding of many new museums, the 1970's was a period of rapid development in the museum area, and it was only now that many museums began to increase the numbers of their professionally-trained staff. The 1970's was a time for overhauling premises, collections and exhibitions. Provincial museums were also brought under a central authority during this decade.

The first records to benefit from computerisation were in two art museums (the Alvar Aalto Museum in Jyväskylä and the Pori Art Museum), and they began operating at the end of the 1970's. Information technology became more topical as a subject for discussion at the beginning of the 1980's. In view of the ever-increasing demands made by individual museums on the State, the Ministry of Education set up a committee in 1984 to make a thorough study of the ways in which museums make use of information technology. The chairman of this committee is *Jouko Heinonen*, director of Lahti city museum.

The committee presented its report in the spring of 1985. It made clear that there is already a particularly wide degree of computerisation in museums' financial and administrative matters, but as most museums are under local council ownership, they make use of whatever particular system the council has already adopted. It was noted that there are plenty of ways in which museums can benefit from computerisation, but the Information Technology committee concentrated especially on the use of computers as an aid to documentation in cataloguing. In its

report, the committee focussed on three points: attempts to create national standards for documentation, the establishment of a development centre for computerisation in association with the office of the Finnish Museums Association, and the appropriation of state funds for software purchasing and personnel training.

These proposals were favourably received, but have not resulted in any concrete steps on the part of the state. For this reason several larger museums got together to form standardisation committees of their own that would produce standard documentation models for cultural-history museums on the one hand, and art museums on the other.

To coordinate all the procedures involved in the transition to a computerised system, the Finnish Museums Association set up at the end of 1985 a computerisation committee with representatives both from central museums and from various other types of museums. The chairman of the committee is the general secretary of the Finnish Museums Association, *Anja-Tuulikki Huovinen*. The committee deals with general practical questions relating to the changeover to a computerised system, keeps a check on how things are progressing, and gives its approval to recommendations for general standards. It has recently received the first drafts of proposed standards from the museums' standardisation committees. In addition the committee plans computer training for museum staff, and to look into the need for external financing during the transition period.

With no computer centre forthcoming, many museums have started making preparations for their own computer projects. The most extensive is the project, now in its development stage, at Lahti city museum. The way in which it most noticeably differs from other projects is that it has been aiming right from the start at a command of data that is as all-embracing as possible. Other development plans in the pipeline are based on microcomputers, which makes them necessarily less ambitious. The most noteworthy of these are the project set up at Ateneum, the main Finnish art museum, and a project at the Central Finland Museum in Jyväskylä for the preservation of cultural-historical objects. Ateneum has chosen Norwegian Polydoc software for its Finnish Nokia micro, and the Central Finland Museum is studying the suitability of the same package for a cultural-history museum.

Today about a dozen museums are getting themselves ready for computerisation, and the Board of Antiquities and Historical Monuments has concentrated on an archeological register and a register of buildings. In 1985 the Finnish Museum of Sport in Helsinki together with the Library of Sport set about creating a database of library and archive data, on the basis of which a chargeable information service has been initiated which primarily provides background material for the press. The development of picture processing has been followed with great

interest in Finland. It is hoped that a general project will be got under way during the current year with the purpose of studying the potential of digital picture processing primarily in the referencing of picture archives.

## **Användning av ADB i museiarbetet i Island**

### *Gudmundur Olafsson*

De isländska museerna kan i dag sägas ligga i startgroparna, när det gäller datorisering. Åtminstone 4 museer av ca 30 har redan börjat använda persondatorer i liten skala och jag räknar med att inom de närmaste 2-3 åren kommer ADB-registrering att vara allmänt i bruk bland isländska museer.

Den utveckling som nu är i gång tyder på att museerna kommer att satsa på interna persondatorer heller än nätanslutningar.

#### **Þjóðminjasafn Íslands (Islands Nationalmuseum)**

Den första museidatorn (en Apple IIe 128k) inköptes av Nationalmuseets arkeologiska- och etnologiska avdelningar hösten 1984, för knappt 2 år sedan. Den var delvis tänkt som en introduktion till en framtida datorisering av museet, då det vid den tiden rådde en viss tvekan inför datatekniken bland många av museets personal, men först och främst för att underlätta skrivarbete och avlösa många små kortregister på avdelningarna.

Programmet som användes heter Appleworks, vilket är kombinerad ordbehandlings/D-Base/spreadsheet program. Det passar mycket bra till små arbeten, men klarar tyvärr inte av några större register på grund av mycket begränsad minneskapasitet.

Flera av Nationalmuseets avdelningar har nu planer på att skaffa sig datorer under det närmaste året, och då behövs det ett väl genomtänkt ADB-system för föremålsregistrering, som inte enbart kan användas av Nationalmuseet, utan av alla andra kulturhistoriska museer i landet också. Ännu är det inte bestämt vilket system man skall satsa på, men valet kan mycket väl komma att grundas på de erfarenheter vi gör här.

#### **Listasafn Íslands (Islands konstmuseum)**

Konstmuseet började förbereda EDB-registrering nu i januari med hjälp av datakonsult från Universitetet. De har funderat på några EDB-system men har inte än bestämt sig för något. Eftersom museets samlingar består endast av ca 5000 nummer kommer valet troligen att stå mellan lätt programbara system som dBase II/III eller IBM-Filing assistant.

Konstmuseet har också köpt 2 stycken IBM XT PC-datorer, varav den ena har 20 mb disk. Man räknar med att förarbetet är färdigt i juni/juli och att man då skall ha färdigt program för att påbörja museets ADB-registrering.

Ljósmyndasafnið h/f (Fotografiska museet A/B)

Fotografiska museet har nyligen börjat med ADB-registrering av sina fotografier. De använder sig av ett eget system, komponerat på dBase II/III och som i viss mån bygger på Outlines of Cultural Materials. Deras persondator är av märket Digital.

Minjasafnið að Hjóti (Bygdemuseet i Hnjot)

Ett av de ca 20 bygdemuseerna i Island har satt upp ADB-program på Appleworks program för Apple IIe persondator, men själva registreringsarbetet har inte kommit i gång ännu och de väntar på resultaten från denna kongress, innan de slutgiltigt bestämmer sig.

### *Sammanfattning*

Som synes av det ovensagda, är vi på Island bara i början av datautvecklingen. Jag kan dock säga att intresset är nu mycket stort för att börja utnyttja ADB-tekniken för museernas föremålsregistrering. Det man först och främst väntar på, är att någon tar initiativet till ett bra registreringssystem som alla museer skulle kunna anamma.

I ett land som inte är större än Island skulle det enligt min mening vara ett stort misstag att inte utnyttja det tillfälle vi nu har till att få fram ett enhetligt registreringssystem i hela landet. Hittills har var och en hållit på med att konstruera sitt eget registreringssystem, och faran är att vi inom snar framtid sitter uppe med ett virrvarr av hemsnickrade system, som inte går att köra mellan varandra.

Vi vill naturligtvis också försöka och ta del av de erfarenheter som redan har gjorts i här övriga Norden, och kommer därför att lägga stor vikt vid de resultat som vi får fram på den här kongressen.

## **Användningen av ADB i museiarbete i Sverige i dag**

### *Göran Bergengren*

Automatisk databehandling används för registrering av föremål vid 8 svenska konst- och kulturhistoriska museer eller institutioner.

Verksamhetsformen varierar mellan hyrd tid hos servicebyrå och egen minidatamaskin. De program, som används är oftast egna tillämpningar



av standardprogram. Något för dessa museer gemensamt formulär finns inte.

Av nedanstående tablå framgår vilken maskinvara, programvara, ADB-användning och omfattning, som respektive museum har.

Datautrustning vid Konst- och kulturhistoriska museer i Sverige, i reguljär drift, använd för registrering av föremål maj 1986.

Institution	Maskin	Program	Användning
Armémuseum, Stockholm Box 14 095 104 41 Stockholm 08/611522 Bengt Hermansson	Ej egen. Servicecentral: Medicin- data, Göteborg. Registrerings- central med 4 ABC 800, handkappad personal, disketter.	Se Kulturarvet	För registrering av föremål. 7000 av 100.000 inläs- ta i databas. Ny- förvärv 500-1000 per år
Bohusläns museum Uddevallå 0522/39200 Lars Andersson	Luxor 800 6 st arbets- platser i nät CAT 40 MB win- chestermin- ne	Applikation av standard- vara Magnum	För registrering av föremål. 20.000 (samtliga kultur- historiska) i da- tabas. Nyförvärv ca 500 per år. Ordbehandling redovisning/lön
Kulturarvet, Falun 023/18480 Runc Hermansson	Ej egen. Servicecen- tral: Medicin data, Göte- borg DEC 10. Registrering 7 mikro Bildspelare Philips	Applikation av 3-RIP data- bashanterare	För registrering av föremål. 40.000 föremål i databas 27.000 på band Sammanlagt 135 hembygdsföreningar i Dalarna. Tillväxt ca 5000 per år (15 samlingar större än 1000)
Livruskammaren Slottsbacken 3, 08/209571 Nils Drejholt	Ej egen. Se Nordiska museet. On- line kommu- nikation	Egen applika- av standard- vara databas- hanterare IMAGE	För registrering av föremål. 9000 i databas (30.000 to- talt)
Nationalmuseum Stockholm 08/244200 Ann-Sofi Topelius	IBM PC 12 MB minne 2 st i nät	Applikation av Dataflex	För depositioner och lån ca 10.000 poster, ca 1000 årligen rörliga

Institution	Maskin	Program	Användning
Nordiska museet Stockholm 08/224120 Göran Bergengren	Registrerings- central med Hewlett Packard 1000, 10 ter- minaler, dis- kettminne		Ca 150.000 läns- museiinventarienr per år
	Hewlett Packard 3000 520 MB	Egen applika- tion av stan- dardvara data- bashanterare IMAGE	

#### Användning

Föremålsregistrering. 350.000 egna i databas. 2 läns museer (varierar).  
Ca 1000 nyförvärv per år.

Livrustkammarens databas, se dito Skoklosters slotts databas, se dito  
Förvärvsregister. SAMDOK:s projektregister. Adressregister. Ordbe-  
handling.

Länsmuseidatabaser på band:

Gästriklands museum, Gävle, 22.000 invnr

Halmstads museum, 20.000 invnr

Jämtlands läns museum, Östersund, 20.000 invnr

Jönköpings museum, 35.000 invnr

Löstad slott, 5000 invnr

Länsmuseet Murberget Härnösand, 19.000 invnr

Norrbottnens museum, Luleå, 18.000 invnr

Uppland museum, Uppsala, 20.000 invnr

Varbergs museum, 37.000 invnr

Värmlands museum, Karlstad, 26.000 invnr

Västerbottens museum, Umeå, 18.000 invnr

Välsborgs museum, Vänersborg, 20.000 invnr

Östergötlands läns museum, Linköping, 23.000 invnr

Skoklosters slott se Livrustkammaren	Ej egen Se Nordiska museet	Se Livrustkam- maren	15.000 föremål
Stockholms stadsmuseum Peter mynades backe 6 Stockholm 08/440790 Irene Sigurdsson	Ej egen Kommunens datacentral	Applikations- program	30.000 föremål i databas på på band. List- utskrift. Årlig uppdatering ca 2000 föremål.

Sammanställt av Heidi Henriksson  
Nordiska museet

23.5.1986

# Status for edb-arbeidet i de norske kunst- og kulturhistoriske museer

*Jon Birger Østby*

I samarbeid med NAVFs EDB-senter for humanistisk forskning har museene drevet prøvevirksomhet med edb siden midten av 70-tallet. Dette arbeidet har vesentlig dreiet seg om registrering av gjenstandsdata og katalogproduksjon. Norske Kunst- og Kulturhistoriske Museer (NKKM) har hatt en egen edb-komit e som har hatt ansvar for   tilrettelegge bruk av edb, og komit een utarbeidet et katalogkort med veiledning for registrering av gjenstander. Kortet var opprinnelig beregnet for optisk lesing, men det kan ogs a med fordel benyttes av museer som ikke har planer om   bruke edb.

Ifolge en unders okelse som ble foretatt av Statens Museumsr ad i 1980, ligger 51% eller anslagsvis 4 millioner gjenstander ved norske museer ukatalogisert. Dette inneb erer at ca. halvparten av norske museums-samlinger i praksis er utilgjengelige eller sv ert vanskelig tilgjengelige. Museene st ar her overfor store og viktige arbeidsoppgaver. Katalogisering av denne gjenstandsmassen vil v ere sv ert arbeidskrevende, og det er viktig   dra nytte av ny teknologi.

Databehandlingen har tidligere i vesentlig grad skjedd ved NAVFs EDB-senter for humanistisk forskning. Men etter hvert anskaffet flere museer egne mikromaskiner, og dette ga muligheter for databehandling i egen regi. Med dette ble det behov for   styrke veiledningstjenesten og for   utvikle et felles opplegg for registrering av gjenstander og foto p a mikromaskiner. NAVFs EDB-senter har hatt en helt sentral funksjon i arbeidet med innf oring av edb ved museene, og for museene er det sv ert viktig   bygge videre p a denne kontakten slik at vi ogs a i framtiden kan dra nytte av EDB-senterets kompetanse.

1. april avsluttet vi et to arig edb-prosjekt. Det var et m al   samordne edb-oppleggene slik at museene i st orst mulig grad kan samle seg om felles l osninger b ade med tanke p a selve datateknikken og det maskinlagrede kildematerialet. Med felles edb-opplegg vil vi lettere kunne gjennomf ore felles oppl ering. Det vil v ere enklere   utveksle data, og museene vil ogs a kunne dra st orre nytte av andres erfaringer p a edb-sektoren.

Ved oppstartingen av prosjektet sendte vi ut et sp orreskjema om bruk av edb ved museene. Det kom inn 131 svar. 9 kulturhistoriske museer opplyste at de hadde tatt edb i bruk til gjenstandsregistrering. Disse museene har til sammen lagt inn data for 37.000 gjenstander, mens de hadde en samlet gjenstandsmasse p a ca. 300.000 nummer. 50 museer

oppgå at de hadde planer om å ta i bruk edb til registrering av gjenstander og/eller foto, men for de fleste av disse museene var tidsrammen for gjennomføring av planene noe diffus. Den vesentligste årsaken til dette ser ut til å være problemer med å skaffe midler til innkjøp av utstyr, og noen museer svarte at de ville vente på opplegget fra dette edb-prosjektet. Men 17 museer hadde mer konkrete planer og mente at de skulle komme i gang med registrering med edb i 1984 eller 1985. (De arkeologiske museene er ikke tatt med i denne oversikten.)

### **Anbefaling av utstyr**

I prosjektet har vi vurdert datautstyr som kan egne seg for museer. For vurdering ble det satt opp følgende kriterier:

- Fleksibilitet i utbygging og muligheter for å utvide anlegg uten å måtte skifte ut anskaffet utstyr.

- Kostnader for utstyr, installasjon og vedlikehold.

- Utbygging av leverandørens servicenett.

- Referanser for utstyr og leverandør.

- Skjermkvalitet, utforming av tastatur og støynivå.

Prosjektet ble stående ved Corvus Omninet med personlige datamaskiner fra Scanvest som arbeidsstasjoner i nettverket. Men det har vist seg at mange museer har valgt å kjøpe annet utstyr. En vesentlig årsak ser ut til å være ønske om å handle med lokale firmaer. Dette kan være en fordel når det gjelder oppfølging og service.

Etter denne undersøkelsen ble foretatt, er det kommet en rekke nye produkter på markedet. Anbefalingen er derfor ikke uten videre aktuell i dag. Men det ser fremdeles ut til at personlige datamaskiner i lokalt nettverk vil være en egnet løsning for museer. Dette gir muligheten til å begynne i det små og å utvide etter som det blir endringer i behov eller økonomiske forutsetninger, uten at det er nødvendig å skifte ut utstyr og programvare som allerede er anskaffet.

### **Registreringsprogram**

Hardanger Folkemuseum, Norsk Folkemuseum og Forbundsmuseet i Akershus hadde erfaringer med DataStar. DataStar har den store fordelene at data lagres kompakt. Men søkemulighetene er begrenset, og DataStar er et enbrugerprogram. Det har også vist seg å være problemer med å kjøre DataStar i nyere versjoner av operativsystemet CP/M.

Ved valg av programvare for det nye opplegget var det en forutsetning at programmet måtte gi muligheter for at flere brukere skulle kunne drive registrering, oppretting eller søking i dataene på samme tid. Vi la også vekt på at programmet la data på et format som ville gjøre det enkelt å legge dem over i andre program senere.



*Jon Birger Østby demonstrerer registreringskjemaet for museumsgjenstander for deltakere på den nordiske edb-konferansen for kunst- og kulturhistoriske museer.*

Etter en del undersøkelser og en test på registrering fra flere terminaler samtidig, ble vi stående ved programmet Dataflex. Dataflex er et allsidig databasesystem som gir store muligheter for programmering. Programmet er enkelt i bruk, både ved registrering, søking og utskrift.

Ulempen ved Dataflex er at alle felt lagres i full lengde enten de er utfylt eller ikke. Dette resulterer i krav til stor lagerkapasitet. Dataflex er derfor ikke et egnet program for museer som i overskuelig framtid har basert seg på en mikromaskin med disketter eller et platelager på 10 Mb. For disse museene har vi anbefalt DataStar, og registreringsoppleggene for gjenstander og foto er derfor utarbeidet i begge programmene. For museer som begynner med registrering i DataStar er det mulig å legge dataene over i Dataflex på et senere tidspunkt, og det er laget et program for dette.

Det har vært samarbeid med Seksjonen for kunst- og kunstindustri-museene i NKKM for å prøve å lage et skjema som var tilpasset NKKMs opplegg for gjenstandsregistrering. Men dette arbeidet har ikke lyktes. Gruppen har i steden valgt å lage et eget opplegg for registrering av billedkunst. Opplegget for billedkunst er laget i dBase II. Flere museer har både samlinger av billedkunst og kulturhistorisk materiale. For å

reduere kostnader og opplæring ville det være ønskelig om kunstmuseene og de kulturhistoriske museene i framtiden kunne basere registreringsoppleggene på de samme programmene.

Det er innledet samarbeid med Oslo Bymuseum for å lage et opplegg for registrering av museets kunstsamling. For denne samlingen er det kulturhistoriske aspektet en viktig innfallsvinkel. Her er det utarbeidet et registreringsopplegg i Dataflex som bygger på prosjektets opplegg for foto- og gjenstandsregistrering. Opplegget er under utprøving ved museet.

### **Felles database for museene**

I forbindelse med prosjektet ble det bestemt at det skulle gjøres forsøk med en prøvedatabase for gjenstander fra flere museer. Formålet med dette forsøket var å få erfaringer med organisering og drift av en felles database. Vi ønsket å få erfaringer med overføring av data fra mikro-maskiner til stormaskin, samkjøring, sortering og søking i data fra flere museer.

Til denne databasen har vi brukt tekstsøkesystemet SIFT. SIFT har fleksible søkemuligheter, og erfaringene med søking er svært positive. Men for å få godt utbytte av søking er det en forutsetning at man har opparbeidet et godt kjennskap til innholdet i databasen. Dette kan være en begrensende faktor for generell utnyttelse av slike databaser, og det er et moment som kan tale for at søking i slike baser fortrinnsvis bør utføres i samarbeid med personale ved den institusjon som administrerer basen.

Erfaringer viser at søkeprofilene først og fremst baseres på et lite utvalg av felt. De feltene som er mest aktuelle er: Gjenstand, Brukssted, Produksjonssted, Navn og Materiale. Det ser derfor ut til å være viktig å prioritere arbeidet med ens bruk av termer og navn i disse feltene.

Det viser seg at det ofte er vanskelig å danne seg et godt bilde av gjenstanden på tross av utførlige beskrivelser. Det er derfor grunn til å reise spørsmål om hvor mye arbeid det er riktig å nedlegge på dette området. NKKMs edb-kort var laget på et tidspunkt da det ikke var aktuelt å knytte gjenstandsdata sammen med lagring av bilder, og kortet bærer preg av dette. Gjenstandsregistrering er en tidkrevende og dermed kostbar prosess. Det er derfor nødvendig å vurdere om resultatene svarer til det arbeidet som er nedlagt. Søking i en felles database kan neppe være til annen hjelp enn å finne fram til et relevant gjenstandsmateriale som en ønsker å studere nærmere. I en slik sammenheng ser utførlige beskrivelser ut til å ha en begrenset verdi. De enkelte forskere har forskjellige innfallsvinkler og spørsmål som de arbeider med, og selv en grundig og utførlig gjenstandsbeskrivelse vil sjelden gi svar på de spørsmål den enkelte ønsker å undersøke.

NORSK FOLKEMUSEUMS SAMLINGER			
01	Skjema nr	Asjett, 6	1979
02	Skjema nr	Fruktasjett ?	1079.03 RHJ
03	Skjema nr	Fruktasjett 1	MF.1981-75/80
04	Skjema nr	1935 ca.	264
05	Skjema nr	(01) Dale	
06	Skjema nr	(21) Rosander, Gerd Aarland, museumslektor, HF.	
07	Skjema nr	(11) Vu BD	
08	Skjema nr	(67) Andreassen, Astaug, sykepleier, Alders Hytte, Gjølhøgt. 1, Fredrikstad	
09	Skjema nr	Standby	
10	Skjema nr	Prasset	
11	Skjema nr	DUGER: Blyglasur, gul beghning	
12	Skjema nr	Rund, flat	
13	Skjema nr	Gul med sort dekor	
14	Skjema nr	Sprettet	
15	Skjema nr	Funkis	
16	Skjema nr		
17	Skjema nr	I Vu BD Tur i sirkel, 76026, LRE/6	
18	Skjema nr	DIAM: 16.2, H: 1.8	
19	Skjema nr	god	
20	Skjema nr		1979



*Fra Norsk Folkemuseums fotokatalog. Gjenstandene fotograferes og foto monteres på kort som skrives ut på museets skriver.*

Det viser seg at skjemaene tar svært stor lagerplass. I Dataflex har vi regnet med at det i gjennomsnitt skal kunne lagres 4000 gjenstandsskjemaer på 10 Mb. Med SIFT ser det ut til at vi må regne med at kravene til lagerplass vil bli omtrent det dobbelte. Dette vil variere noe med hvor stor del av skjemaet som er utfyllt. Sammen med erfaringene med at det er et lite utvalg av felt som går igjen i søkeprofilene, kan dette tale for at det ved bruk av SIFT til framtidige felles databaser kan være aktuelt å legge inn bare et utvalg av felt fra registreringsskjemaet, eller at søkemulighetene begrenses til et utvalg av felt.

### Samordning av datatjenester i framtiden

Vi arbeider nå for å etablere en «Museenes datatjeneste» for å ivareta felles arbeidsoppgaver på dette området. «Museenes datatjeneste» skal ha følgende arbeidsoppgaver:

- Fungere som konsultativt organ for bruk av edb ved norske museer.
- Koordinere bruk av edb og følge opp opplegg for registrering, søking og utlising av data.
- Organisere opplæring.

- Gi veiledning ved kjøp av utstyr.
- Forestå drift av felles database for museer.
- Vurdere nye produkter som kan være aktuelle for museer.
- Representere museene i internasjonalt samarbeid på edb-sektoren.

## Konklusjon

Ved avslutningen av prosjektet er oppleggene for foto- og gjenstandsregistrering levert til 13 museer. Tre museer har valgt programmet DataStar, og ni museer har valgt Dataflex. Ett museum har fått begge versjonene, og her skal registrering foregå lokalt med DataStar på diskettmaskiner for senere å legges over i Dataflex på et større anlegg.

Vel 40 museer har hatt deltakere på kurs. I løpet av prosjektperioden har det trolig vært kontakt med ca. 100 museer for generell orientering, demonstrasjon og veiledningstjeneste. I tillegg til dette har det vært en rekke henvendelser fra andre institusjoner og enkeltpersoner.

Ca. 20 kulturhistoriske museer har nå egne dataanlegg. Interessen for investeringer på denne sektoren er meget stor, og det ser ut til at svært mange museer vil komme til i de nærmeste årene.

En satsing på bruk av edb i museumssektoren er i tråd med de retningslinjer som er trukket opp i Stortingsmelding nr. 60, 1984-85. Her er informasjonsteknologi/telematikk og tradisjons- og kulturformidlende forskning to av de fem hovedinnsatsområder som Regjeringen ønsker å prioritere.

Når det gjelder innsamling, registrering og bevaring, står de kulturhistoriske museene i Norge overfor store, uløste arbeidsoppgaver uten at det ser ut til at museene i den nærmeste framtid kan regne med noen tilsvarende økning av arbeidskraften. Det er derfor naturlig å vurdere tiltak som kan effektivisere det nåværende arbeidet, og i denne sammenheng kan bruk av edb være et viktig hjelpemiddel.

For å få en mer rasjonell utnyttelse av museenes ressurser i framtiden, er det også stort behov for å koordinere videre innsamling. Dette er nødvendig både for å avdekke «hull» og for å hindre stor overlappning i museenes innsamling. Det er vanskelig å tenke seg at en slik koordinering vil kunne gjennomføres uten utstrakt bruk av edb. Samtidig vil edb-baserte kataloger gi muligheter for en sterkere aktivisering av gjenstandsmaterialet både i forskning og formidling.

Museene har stor pågang fra eksterne brukere som kommer for å hente kunnskap i museenes samlinger. Det er ofte en tidkrevende oppgave å finne fram til relevant materiale for den enkelte forsker, og slike forespørsler krever stadig omfattende leting i protokoller og mer eller mindre komplette seddelkataloger.

For museene kan katalogdata aldri erstatte selve gjenstanden som primærkilde. Men edb kan nettopp være et godt hjelpemiddel for å søke



fram til det utvalg av gjenstander som kan fortelle noe om de spørsmål man arbeider med. Resultatet av en slik søking vil imidlertid være helt avhengig av kvaliteten på det registreringsarbeidet som er utført.

*Jon Birger Østby har vært forsker-NAVF ved edb-prosjektet for de kunst- og kulturhistoriske museer, og er nå konservator ved Norsk Folkemuseum.*

# CD-ROM – en halv gigabyte i lommen

Øystein Reigem

CD-ROM er en bastard. Den har underholdningsindustrien som mor og databransjen som far. Begge foreldrene vil få stor glede av barnet.

## Compact Disc

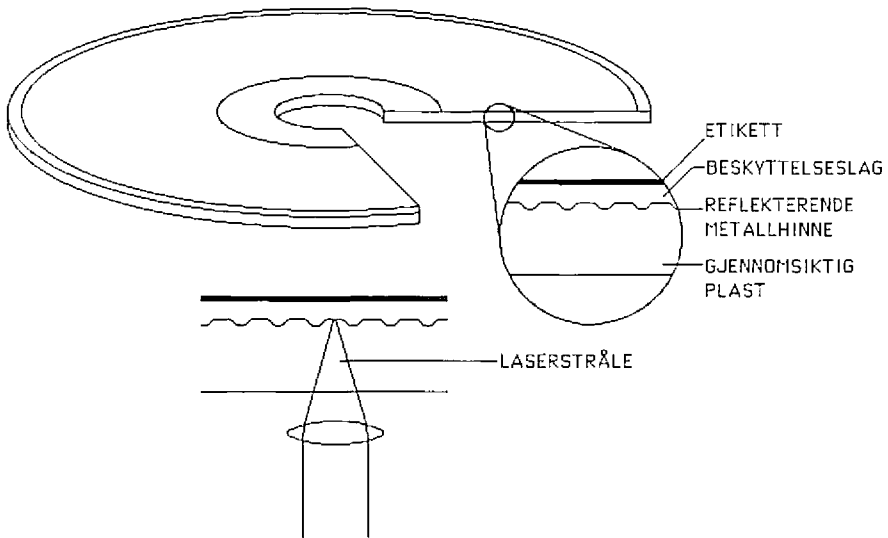
Historien om CD-ROM starter med Compact Disc – den nye grammofonplaten. Seint i 1983 ble de første Compact Disc sluppet ut på markedet i Japan, og tidlig året etter i resten av verden. Nå, to og et halvt år senere, er det klart at intet annet forbrukerelektronisk produkt har hatt en mer vellykket introduksjonsperiode. De to første årene ble det solgt over 65 millioner plater. 10% av alle platetitler er tilgjengelig på Compact Disc.

Compact Disc er en plate med en diameter på bare 12 cm og kun én spillbar side, men med plass til mer enn en hel time krystallklar stereo lyd. Istedenfor at lyden er lagret i en rille preget i vinyl, er den på Compact Disc kodet i et mønster i et tynt metallsjikt. Mønsteret består av fordypninger («pits») som går i et spiralformet spor. Et lag klar plast beskytter mønsteret. (Se illustrasjon 1.) Sporet starter innerst og går utover og har en lengde på 4-5 km. På sin vei gjør det 21.000 omdreininger. Hullene i sporet har en diameter på 0.6 mm (0.6 tusendels millimeter).

De nye platespillerne har istedenfor en stift en liten laser som er i stand til å avlese hullmønsteret. Imidlertid er den viktigste forskjellen i forhold til vinylplaten at lyden ikke lenger er lagret i *analog*, men i *digital* form.

Ikke alle er fortrolige med begrepene analog og digital. Det kan derfor være på sin plass med en forklaring av hva analog og digital lagring av lyd vil si. – Lyd er som kjent svingninger i luften, og kan dermed beskrives vha. av en kurve som i illustrasjon 2. På en gammeldags grammofonplate ligger lyden lagret i en rille med samme form som kurven. Vi sier at rillen er en *analog representasjon* av lyden fordi den *varierer kontinuerlig i samsvar* med lydsvingningene. Det magnetiske signalet på en musikkassett er også en analog representasjon fordi *magnetismen* varierer kontinuerlig i samsvar med lydsvingningene.

*Digital* betyr *på tallform*, og digital lagring av lyd krever at lydsignalet til denne kurven er beskrevet vha. tall. Hvordan lyd *digitaliseres* er vist på illustrasjon 2. Lyden «samples», dvs. utslaget måles med jevne mellomrom (flere tusen ganger i sekundet). Disse målingene gir tallverdier, og resultatet blir dermed en sekvens av tall. På Compact Disc er tallene



Illustrasjon 1. Fysisk oppbygning av Compact Disc.

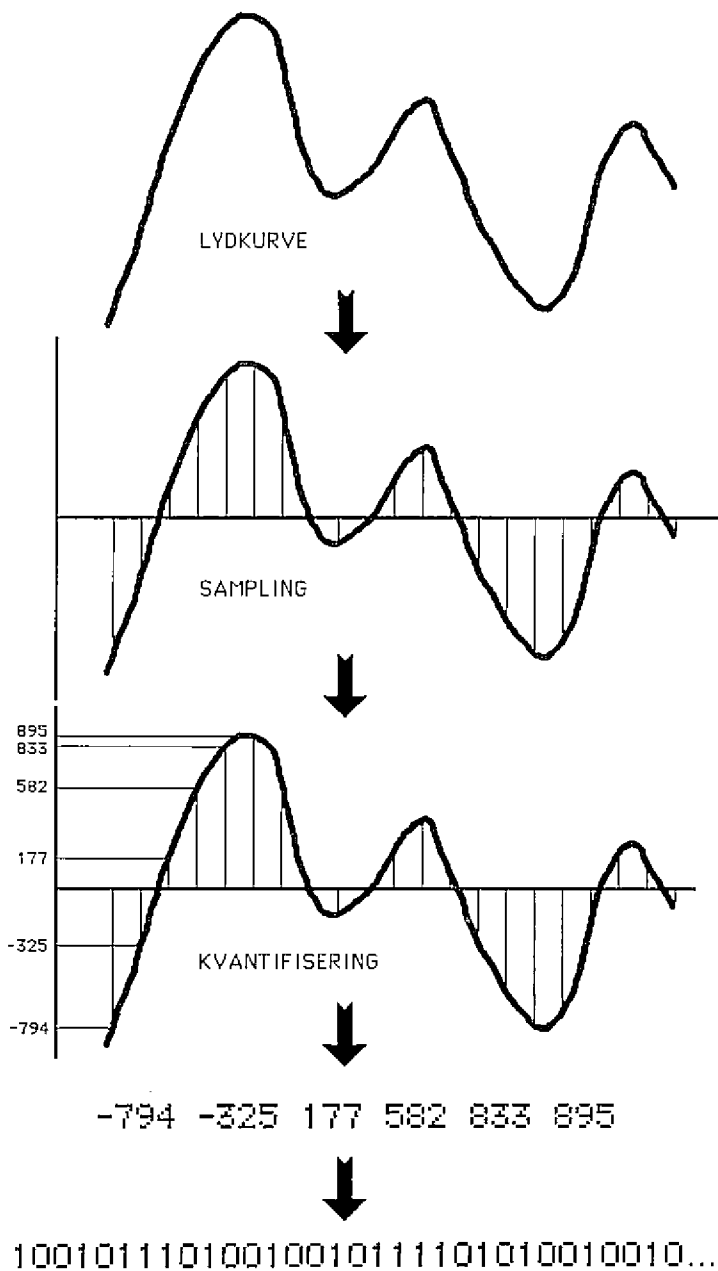
overført til binære tall (0 og 1), akkurat som på et vanlig datalagringsmedium (disk, diskett), og binærsifrene lagres igjen som et hullmønster.

Siden Compact Disc lagrer lyden i digital form, og siden vanlige datalagringsmedier er digitale, er det en nærliggende tanke å ta i bruk platen som et generelt datalagringsmedium. Selv om platen er kun lesbar og ikke på alle måter optimal for datalagring, har den egenskaper som gjør den svært attraktiv. Den har en lagringskapasitet på ca. 550 Mbyte, multimediapmuligheter (lyd, bilde m.m), god holdbarhet, og den kan massekopieres rimelig. (550 megabyte = 550 millioner tegn - tilsvarer 1500 vanlige IBM-disketter.) En kjempefordel i forhold til konkurrerende medier er at CD-ROM kan «ri på ryggen av» Compact Disc-markedet. Dette betyr rimeligere utstyr og potensielt rimeligere og raske produksjon av plater. Datavarianten av Compact Disc kalles CD-ROM, som betyr *Compact Disc Read Only Memory*.

## CD-ROM

CD-ROM er allerede tatt i bruk i en rekke anvendelser, og det er en raskt voksende interesse for mediet til elektronisk publisering av databaser, leksikon og programvare, til lagring av kart, lyd og bilder, og til bruk i underholdning og undervisning. CD-ROM spilles på en egen spiller som kobles til datamaskin som en slags diskstasjon.

En viktig milepel for CD-ROM var den konferansen som det kjente



Illustrasjon 2. Digitalisering av lyd.

amerikanske programvarefirmaet Microsoft arrangerte i Seattle 4.-7.3.86 – «The First International Conference on CD ROM». (Microsoft har bl.a. skrevet MS-DOS – operativsystemet for IBM-kompatible PC-er.) En av foredragsholderne kalte konferansen for «CD-ROM-ens Woodstock», og uten at han ga noen nærmere forklaring, kan en gå ut fra at han tenkte på det store antall deltakere (1172 – mer enn dobbelt så mange som arrangøren regnet med), atmosfæren av fred og fordragelighet (ingen kringel om standardisering), og ikke minst følelsen av å være med på starten av en ny æra. (Forfatteren deltok på konferansen og har hentet mye av informasjonen i artikkelen derfra.)

CD-ROM vil sannsynligvis få to forholdsvis adskilte markeder, og utstyr tilpasset disse. For næringsliv og institusjonsbruk vil CD-ROM-spillere stå som periferutstyr for datamaskin. For underholdning/hjemmemarkedet vil spilleren inngå i et nytt forbrukerelektronikkprodukt, «vieweren». Mer om dette siden.

### Optiske lagringsmedier

Compact Disc/CD-ROM er et såkalt *optisk lagringsmedium* fordi informasjonen avleses optisk, dvs. vha. lys, nærmere bestemt laserlys. På optiske medier ligger informasjonen som regel lagret som hull, fordypninger eller andre endringer i en metallhinne innenfor et gjennomsiktig plastlag. Det finnes i dag en rekke forskjellige optiske medier, og nye typer kommer. De fleste er i form av runde plater, men det finnes også optiske «kredittkort», og optisk bånd er på trappene.

Andre navn som brukes på optiske plater er optiske disk, laserplater, laserdisk, videoplater, videodisk. Videoplater er egentlig en undergruppe av optiske plater (se lenger nede), og enkelte bruker begrepet optiske plater kun om digitale plater (se også lenger nede).

Optiske lagringsmedier byr på en rekke fordeler i forhold til magnetiske medier (dvs. disk, diskett, magnetbånd). Kapasiteten er høy, gjerne i størrelsesorden 1 Gbyte (gigabyte) pr. plateside (1 Gbyte = 1 milliard tegn). På grunn av den høye kapasiteten egner optiske medier seg også for lagring av plasskrevende typer informasjon som bilder og lyd. Optiske plater er robuste og har lang lagringsholdbarhet.

Optiske lagringsmedier faller i tre klasser: *read-only*, *write-once-read-many* og *erasable*. (Forslag til norske oversettelser: «kun lesbar», «skriv-én-gang» og «overskrivbar»). – CD-ROM tilhører den første klassen. Read-only-platene presses omtrent som gramfonplater, og alt materialet må legges til rette på forhånd. Det er ikke mulig å endre eller legge til informasjon på platene siden. Endringer må enten suppleres på et annet medium (f.eks. diskett), eller platene må presses i et nytt opplag. Read-only-plater blir rimelige i store opplag.

Write-once-plater kan man selv legge informasjon inn på. Laseren i

platestasjonen brukes da både til å avlese og skrive inn informasjonen. Informasjonen kan legges inn i flere omganger, men den kan ikke skrives over. Endringer foretas ved å skrive informasjonen på nytt et annet sted på platen. «Write-once-read-many» kalles også «write-once-read-mostly», og forkortes ofte til «WORM».

Overskrivbare plater er ikke på markedet ennå, selv om de i flere år ble spådd å komme «om to år». I det siste er man imidlertid kommet fra laboratoriestadiet til produksjonsprototyper, og platene vil kunne være tilgjengelige tidlig neste år. Overskrivbare plater vil kunne brukes om igjen på samme måten som vanlig disk og diskett.

På tvers av inndelingen i read-only, write-once og overskrivbar, går en inndeling i *analoge* og *digitale* medier. Analoge optiske plater kalles også *videoplater* fordi det er et FM videosignal på platene. (Videosignaler er nemlig analoge og ikke digitale.) På CD-ROM og andre *digitale optiske plater* ligger informasjonen i form av binære siffer (bits – 0 og 1), og de har dermed en viss likhet med magnetiske datalagringsmedier (diskett, disk, magnetbånd).

For å gjøre det hele mer forvirrende finnes det teknikker for å omforme et digitalt signal til et videosignal slik at det kan lagres på en videoplate. Dermed har en også hybridplater med begge muligheter.

### **Viktige typer optiske plater**

For vi vender fullstendig tilbake til CD-ROM, kan det være greit å nevne de vanligste typene optiske plater på markedet i dag. – *LaserVision* er den viktigste typen read-only videoplate. Platen er 12 tommer i diameter, tosidig og rommer pr. side ca. 36 minutter levende video (USA/Japan: 30 minutter) eller ca. 54.000 stillbilder (god TV-kvalitet). Til levende video kan en ha to HiFi lydkanaler (dvs. 1 stereo eller 2 mono). Så langt er platen fullstendig standardisert (dvs. uavhengig av fabrikatet av plate-spilleren), bortsett fra et beklagelig skille mellom amerikansk/japansk og europeisk videosignal (gjelder alle videoplater – ikke bare LaserVision).

Teknikker for omforming av digitale signaler til video gir LaserVision-platen en digital kapasitet på et sted mellom en halv og en gigabyte pr. side. Det er dessverre ingen standard for digitale data på videoplater. Imidlertid regner man med en standardisering med det første. Det finnes amerikanske plater med digital lyd i tillegg til den vanlige analoge, og dette digitale «lydsporet» kan brukes til data isteden. Den digitale kapasiteten er da vel 300 Mbyte. I det europeiske videosignalet er det ikke «ledig plass» til digital lyd, så der vil en antagelig måtte ofre den analoge lyden for å kunne få digital lyd eller digitale data.

På videoplater kan en lagre bilder opptil en viss kvalitet meget kompakt. Lagring av tilsvarende digitale bilder tar mye mer plass. Til gjengjeld mister en med video den fleksibilitet digital lagring gir. Bilder med

tekst (dokumentsider) er uegnet for video. En fordel med videoplater er at bilder og lyd kan vises vha. vanlig forbrukerelektronikk (TV).

Anvendelsesområder for videoplater er bl.a. undervisning, instruksjon, oppslagsverk, presentasjon, salg av varer. På det opprinnelige anvendelsesområdet – spillefilm og annen underholdning – har ikke videoplaten slått gjennom. Markedet består av et lite antall entusiaster. Et unntak er Japan, hvor markedet er i kraftig ekspansjon. (I Japan annonseres det allerede produkter for neste generasjon TV-teknologi – *HDTV – High Definition TV* – med widescreenskjermer og dobbelt så god oppløsning som i dag.) For spillefilm o.l. bruker en som regel et langspillformat som gir opptil 1 time pr. side, men ikke stillbilder.

Det bør nevnes at Pioneer har på markedet en ny spiller for både LaserVision og Compact Disc. Dessverre finnes ikke denne i europeisk versjon.

En plate som ennå bare er under planlegging, er den såkalte *Compact Video Disc* (CVD). CVD er en digital plate, 20 cm i diameter og med 20 minutter Compact Disc lyd, hvorav 5 minutter har digital levende video til.

CD-ROM får også en slektning mer tilpasset data. Fordelen blir raske-  
re aksessid (CAV istedenfor CLV – se Bruk av CD-ROM).

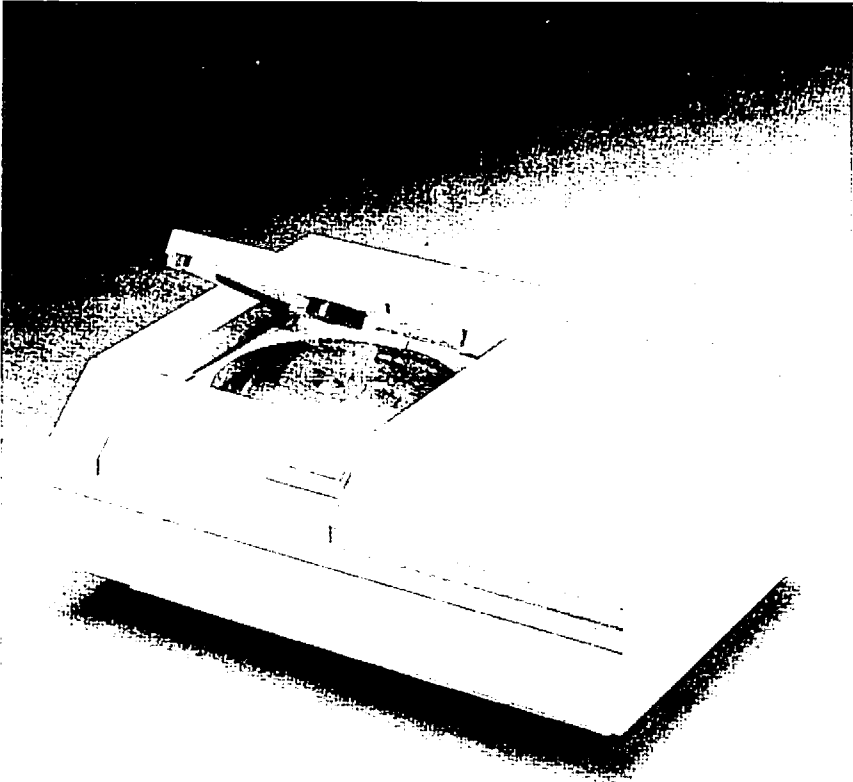
Utstyr med *digitale WORM-plater* leveres av en rekke fabrikanter. Typiske kapasiteter er 200 Mbyte pr. side for 5 1/4" plater og 1 Gbyte for 12" plater. En regner med at enkelte WORM-spillere senere vil kunne spille et helt spekter av digitale plater, både read-only, write-once og overskrivbare.

Anvendelsesområder for digitale WORM-plater er store datamengder som er relativt statiske, backup, databaser med bilder/tegninger/dokumenter, osv. Av de platestasjoner som er levert til i dag, inngår de fleste i nøkkelferdige systemer for lagring av kontordokumenter. Slike systemer har en såkalt *scanner* for innlesing av dokumentene, en høyoppløselig A4-skjerm for visning av dokumenter og en laserskriver for utskrift. Et databasesystem muliggjør registrering av opplysninger for å kunne finne igjen dokumentene.

De som er interessert i en generell oversikt over optisk lagring, kan lese artikkelen «Optisk plate-teknologi» i HD 2-85. HD 2-85 inneholder også to andre artikler med emne optiske plater. Selv om utviklingen går frustrerende fort, er oversiktsartikkelen ennå stort sett up to date. Se også litteraturlisten til slutt i artikkelen.

### **Lagring av forskjellige typer informasjon på CD-ROM**

Hvilke typer informasjon kan en så ha på en CD-ROM, og hva er kapasiteten for de enkelte typer informasjon? – *Lyd* har vi allerede nevnt. Digital lagring av lyd blir en avveining mellom plassbruk og



*Illustrasjon 3. Philips CD-ROM-spiller.*

kvalitet. Kvaliteten stiger med hyppigere sampling. Hvis en vil ha gjengitt lyd opptil en frekvens på  $n$  svingninger pr. sekund, må en sample minst  $2n$  ganger i sekundet. På Compact Disc er lyden samplet over 44.000 ganger i sekundet, og da er hele det lydspekter menneskets øre kan oppfatte dekket. Hvor nøyaktig måleverdiene blir bestemt (hvor mange siffer i tallene) er også avgjørende for kvaliteten. Compact Disc benytter 14 bits til hvert tall, og kan dermed skille mellom over 16.000 verdier.

På neste generasjon av Compact Disc (CD-I – se Standardisering, kompatibilitet), vil en også kunne ha mer lyd av «simplere» kvalitet. Det blir et valg mellom 70 minutter stereo Compact Disc lyd, dobbelt så mye stereo HiFi lyd, 4 ganger så mye stereo MidFi lyd og 8 ganger så mye



lyd i stereo talekvalitet. Går man over til mono, dobles kapasiteten igjen. Kapasiteten for tale i mono blir altså nesten 19 timer. Enkelte nevner også en kapasitet på 32 timer (ikke CD-I).

La oss så se på *bilder*. Med bilder menes her bilder i videste forstand. En dokumentside kan også være et bilde. (I den tidligere nevnte artikkelen «Optisk plate-teknologi» står det at CD-ROM kun egner seg til enklere grafikk og ikke bilder. Artikkelforfatteren trekker herved dette tilbake.)

For bilder blir det også en avveining mellom plass og kvalitet. Den vanlige måten å lagre et bilde digitalt på er som et *rasterbilde*. En kan da tenke seg bildet «rutet opp» i små kvadratiske bildeelementer («pixler»). For hver rute blir det bestemt et tall som forteller hvilken farge/gråtone det er i ruten. Bildet er altså gjort om til en tabell av tallverdier.

Som ved lyd er det to ting som bestemmer kvaliteten av et rasterbilde. Det ene er hvor tett opprutingen går – det som kalles *oppløsningen* av bildet. Det andre er hvor mange farger/gråtoner en kan skille mellom i hver rute. Hvis en for et bilde doubler oppløsningen pr. lengdeenhet, firedobles antall ruter og dermed antall tall i tabellen. Plassbehovet vokser altså raskt med økende oppløsning. For antall gråtoner/farger er plassbehovet avhengig av antall binære siffer i antallet. 2 farger (svart og hvitt) krever 1 siffer, 4 farger krever 2, 8 krever 3, 16 krever 4, 32 krever 5, osv. Gråtoner kan også simuleres med rent svart og hvitt ved å alternere svarte og hvite ruter. Da går imidlertid den virkelige oppløsningen ned.

Bilder på rasterform er vanligvis svært plasskrevende og *komprimeres* derfor ofte før de lagres. Hvor mye et bilde kan komprimeres avhenger mye av innholdet. Komprimering foregår gjerne ved at påfølgende ruter med samme gråtone/farge blir lagret på en mer kompakt form (6 x hvitt istedenfor hvitt hvitt hvitt hvitt hvitt hvitt). Kontordokumenter (rent svart/hvitt) kan som regel komprimeres til ca. 10% på denne måten.

Det er dessverre vanskeligere å komprimere bilder med farger og gråtoner, som f.eks. fotos. Plassbehovet for et *ukomprimert* digitalisert videobilde eller bilde av TV-kvalitet ligger et sted i området 1/4-1 Mbyte. Hvis en regner 1/2 Mbyte, vil en CD-ROM bare få plass til vel 1000 ukomprimerte videobilder. Kapasiteten med vanlig brukte komprimeringsteknikker oppgis som regel til ca. 3000-5000 bilder. Det finnes også i dag (amerikanske) fjernsyn som kan fryse video, digitalisere og komprimere i et lager på 64 Kbyte (64.000 tegn). Med en slik komprimering vil en kunne få 9-10.000 TV-bilder på en CD-ROM, men dette er fortsatt ikke så imponerende sammenlignet med videoplaten. En venter imidlertid at det vil komme helt nye komprimeringsmetoder om ikke så lenge, kanskje med en komprimeringsfaktor på 100:1.

Analog (video) lagring av bilder er altså i dag mindre plasskrevende enn digital. For bilder av skrift og grafikk i rent svart/hvitt kommer

digital lagring mer til sin rett. Skal en skrivemaskinskrevet A4-side gjengis på video, må den stykkes opp så mye at linjene blir delt. En regner at en CD-ROM har plass til minst 10.000 kontordokumenter i A4-format med en oppløsning på 200 dpi (200 punkter pr. tomme = 8 punkter pr. mm). En slik oppløsning er tilfredsstillende for vanlig størrelse trykt skrift. Fotnoter i liten skrift, høyskrift og lavskrift kan kreve f.eks. 300 dpi.

Lagring av trykt skrift på rasterform gir rask og forholdsvis likefram innlesing, men teksten blir ikke søkbar. Søkekriterier må skrives inn separat.

For bilder er et alternativ til rasterform å lagre dem som en samling grafiske opplysninger, dvs. opplysninger om hvordan bildet er bygget opp av linjestykker og manglekanter. En slik lagringsform kan bli adskilleg mer kompakt enn et rasterbilde og egner seg godt til grafisk informasjon som kart, diagrammer, osv. Grafikk laget av datamaskin er ofte allerede i denne formen. En viktig fordel med en grafisk representasjon er at mer «intelligens» kan innebygges. F.eks. kan et kart inneholde opplysninger om hva de enkelte linjer og arealer er (veier, elver, grenser; kommuner, fylker, innsjøer) og hvordan de forholder seg til hverandre (kommune X ligger i fylke Y). Det koster selvfølgelig mer å tilrettelegge bilder på grafisk form enn å lese dem rett inn på rasterform.

(En overføring fra raster til grafisk form vil også kunne brukes i effektive komprimeringsteknikker for mer vanlige bilder. Framtidige komprimeringsteknikker vil f.eks. basere seg på å lagre koordinater og gråtone/farge for et utvalg nøkkelpunkter i bildet. Fargene mellom nøkkelpunktene blir antatt å ha en glidende overgang mellom fargene i nøkkelpunktene.)

En fordel med digital lagring av bilder, enten de er på rasterform eller grafisk, er bedre muligheter for å kunne manipulere bildene videre. Om ikke annet vil en ofte ønske å *zoome* og *panorere* på bilder som er større enn hva skjermen kan gjengi med tilstrekkelig oppløsning.

Uten radikalt forbedrede komprimeringsteknikker kan CD-ROM *ikke* brukes til *levende video*. Levende video er 25 (Europa) eller 30 (USA) bilder («frames») i sekundet, og digital video blir dermed svært plasskrevende. Selv om man ville nøye seg med de ganske få minutter video som er mulig pr. plate, er ikke avspillingshastigheten til en CD-ROM-spiller høy nok (ca. en time pr. plate – tilpasset lyd). Videoplate er det aktuelle alternativet i dag, senere også Compact Video Disc.

Imidlertid har en allerede i eksperimenter med spesielle komprimeringsteknikker lagret opptil 2 timer video på en CD-ROM (E vital Inc., Dallas). Denne enorme kapasitetsforbedringen kan ikke regnes direkte om i stillbilder. Komprimeringen baserer seg på å registrere forskjeller mellom suksessive frames.

I tillegg til bilder, grafikk og lyd, kan en selvsagt lagre «konvensjonelle» typer data som tekst, tall og programmer på CD-ROM. Hvis en regner 2000 tegn pr. maskinskrevet A4-ark, er kapasiteten for tekst 275.000 maskinskrevne A4-sider. En vil riktignok ikke alltid kunne utnytte all denne plassen til tekst. Skal tekst inn i et tekstsøkesystem, må man regne med at systemet trenger minst like mye plass til indekser som til selve teksten.

For å komme opp i denne kapasiteten må selvfølgelig teksten ligge i maskinlesbar form (f.eks. ASCII). Teksten må enten være skrevet inn på tekstbehandling, tatt fra en allerede maskinlagret versjon (f.eks. bånd fra et setteri), eller den må være lest inn fra en trykt utgave vha. en OCR-leser (OCR = «Optical Character Recognition»). OCR-lesing vil kunne få en renessanse siden det nå kommer en ny generasjon av forholdsvis rimelige scannere med OCR-programvare.

Når det gjelder kapasiteten for lagring av *programvare*, vil det neppe være anvendelser der en vil lagre så store mengder at en CD-ROM blir for liten. Programvare til materiale på CD-ROM kommer i dag delvis på samme CD-ROM og delvis på diskett. Fordelen med programvare på diskett er at rettelser og endringer kan gjøres uten å presse en ny plate.

### **Standardisering, kompatibilitet**

For at et nytt edb-produkt skal slå gjennom, må det eksistere mer eller mindre allment aksepterte standarder for å sikre kompatibilitet. (Unntaket er IBM, som har så stort marked at deres produkter automatisk setter standarder.) Standardisering foregår gjennom nasjonale og internasjonale standardiseringsorganisasjoner, og er tradisjonelt en langsom prosess. Det ville sannsynligvis tatt 3-4 år å komme fram til en standard for CD-ROM på denne måten. Samtidig ville CD-ROM bli tatt i bruk i en rekke anvendelser, men en ville ikke få kompatible systemer.

For å unngå anarki dannet derfor industrien i USA i fjor høst sin egen standardiseringsgruppe. Gruppen ble satt sammen av representanter for bedrifter innen maskinvare/elektronikk, programvare og media (media = selve platene). Blant deltakerne var Apple, Digital Equipment Corporation, LaserData, Microsoft, 3M Corporation, Philips, Reference Technology, Sony, TMS, Videotools, Xebec og Yelick Inc. Gruppen kalte seg «the High Sierra Group» siden det første møtet fant sted høyt oppe i Sierra Nevada-fjellene. High Sierra-gruppen hadde sitt forslag til standard ferdig i juni i år. Arbeidet har skjedd i forståelse med de aktuelle standardiseringsorganisasjonene, og forslaget er nå gått videre til disse. En regner imidlertid med at standarden ikke blir offisiell på en stund, kanskje ikke før midten av 1987.

Samtidig utarbeider Philips og Sony en standard mer rettet mot hjemmemarkedet. Denne heter *CD-I* (*CD - Interactive*) og ble annonsert som

konsept i februar. Fullstendige spesifikasjoner kommer ikke før sent i år. En regner heldigvis med at High Sierra-standarden vil bli oppad kompatibel med CD-I-standarden, dvs. at High Sierra-plater vil kunne spilles på CD-I-spillere.

En liten utdyping av standardiseringsspørsmålet kan være på sin plass. CD-ROM bygger direkte på standarden for Compact Disc. På Compact Disc er data, dvs. den digitaliserte lyden, organisert i blokker eller «sektorer» på vel 2 Kbyte. CD-ROM-spillere er i stand til å finne en hvilken som helst sektor, og det er dermed fullt mulig å bruke Compact Disc som datalager. Det finnes allerede en primitiv standard for lagring av data på CD-ROM. Imidlertid må applikasjonsprogramvaren nå selv holde rede på hvilke sektorer data ligger i. På vanlige datalagringsmedier er man vant til at data er organisert i *filer*, og at det er nok å kjenne *navnet* på en fil hvis man vil bruke filen. Filene er også oftest organisert i et eller annet system, f.eks. et hierarki (som filsystemene i MS-DOS og UNIX). Dette ønsker man en standard for også på CD-ROM.

High Sierra-gruppen har ikke tatt med i sin standard opplysninger om hva slags *innhold* en fil har. Standarden har heller ikke opplysninger om datastrukturen i filene (f.eks. hvilket format som skal brukes for lyd og bilde). CD-I, på den annen side, har spesifikasjoner av om filene inneholder lyd, bilde, grafikk, tekst eller programmer, samt flere standardformater for lyd og bilde. Ikke uventet er CD-I-standarden hovedsakelig rettet mot lyd og bilde. CD-I-standarden har også sitt eget operativsystem, CD-RTOS (Real Time Operating System), som skal gå på en Motorola 68000 prosessor.

I tillegg til en standardisering av selve platen kreves også at de enkelte datamaskiners operativsystem utvides til å kunne takle den nye typen lagringenhet. En kan vente at Microsoft kommer med en ny versjon av MS-DOS som gjør IBM-kompatible PC-er i stand til å bruke CD-ROM.

For dem som ønsker å gå i gang med å utvikle CD-ROM-produkter, har det vært en del usikkerhet pga. de to standardene. Konklusjonen ser ut til å være at en ikke bør vente på CD-I-standarden hvis man vil i gang med en anvendelse på det «profesjonelle» markedet (næringsliv, institusjoner). Hvis en senere ønsker å bruke CD-I-utstyr, vil som nevnt High Sierra CD-ROM kunne spilles på CD-I. En bør imidlertid legge opp mot CD-I allerede nå hvis en ønsker å dra nytte av de ferdige mulighetene for større mengder lyd som vil følge med CD-I-spillene.

## Utstyr

Det finnes en mengde produsenter av CD-ROM-spillere – Philips, Sony, Hitachi, Denon, Sanyo, bare for å nevne noen av de mest kjente. I tillegg leverer mange firmaer, spesielt slike som også har sine egne plateprodukter, ferdige pakker hvor spiller inngår.

Skal en selv sette sammen utstyr, kan en ikke uten videre ta en CD-ROM-spiller og koble til en datamaskin som en annen disk- eller diskettstasjon. Det kreves en egen *interface* (overgang), og disse kommer i form av et kretskort som plugges i maskinen. CD-ROM-spillere vil bli mest brukt mot mikromaskiner, og i dag finnes det interfacer for de fleste aktuelle typer: IBM-kompatible PC-er, Macintosh og andre Apple-maskiner, Atari, Commodore, osv. Også for minimaskiner (VAX) fås interface.

Det er flere typer interfacer som er aktuelle for CD-ROM. Ettersom en for magnetiske lagringsmedier til større mikromaskiner nå i stor grad går over til *SCSI (Small Computer System Interface)*, følger CD-ROM etter. Den andre vanlige interfacen i dag er en som kan kobles rett i «bussen» på IBM-kompatible PC-er.

Selve spillerne finnes i flere størrelser. De tidligste var frittstående enheter med dimensjon omtrent som Compact Disc-spillere. Nå får en også modeller som kan monteres i maskinen slik som diskettstasjoner, foreløpig i dobbel høyde. Enkel høyde spillere kommer om ikke så lenge (enkel høyde tilsvarer én 5 1/4" diskettstasjon).

For å spille en CD-ROM med bilder eller grafikk, trengs en bedre skjerm enn en vanlig tekstskjerm, og gjerne et *grafisk kort* som skal plugges i maskinen. Skal man spille av lyd, må man ha øretelefoner eller forsterker og høyttalere. (Se illustrasjon 4.)

Som for write-once-plater vil det komme jukebokser slik at en kan ha flere plater forholdsvis raskt tilgjengelig. På tross av CD-ROM-ens store kapasitet, vil det være mange anvendelser som vil trenge flere plater. Det vil også komme jukeboksutstyr med flere spillere, slik at noen få av et større antall plater vil være tilgjengelige samtidig.

For CD-I tenker man seg en egen type utstyr, gjerne omtalt som en *viewer*. Denne vil ha sin egen innebygde mikromaskin. Vieweren vil kunne kobles mot fjernsynet eller mot en mikromaskins RGB-monitor. Den kan også tenkes ha sin egen skjerm, kanskje en flat skjerm. I utgangspunktet vil den neppe ha noe tastatur, men den vil ha et eller annet til å peke med, slik at en f.eks. kan velge fra menyer på skjermen. CD-I vil være kompatibel med Compact Disc slik at vieweren vil kunne spille «gammeldage» Compact Disc-er. Philips vil sende sin viewer ut på markedet tidlig i 1987. Det vil også komme en CD-I-enhet som gjør nyere modeller av Compact Disc-spillere til CD-I-spillere. Kravet er at Compact Disc-spilleren har digital output.

En CD-ROM-spiller koster i Norge i dag ca. 20.000 kroner. Man spår at prisene snart vil halveres (i høst?). En skjerm for en IBM-kompatibel PC med 16 farger og en oppløsning på 640x400 punkter koster ca. 8-9000 kr., og et grafisk kort omkring 5000 kr.



*Illustrasjon 4. Multimedautstyr for bruk av CD-ROM. Representant for Microsoft demonstrerer «the Multimedia Encyclopedia».*

## **Produksjon**

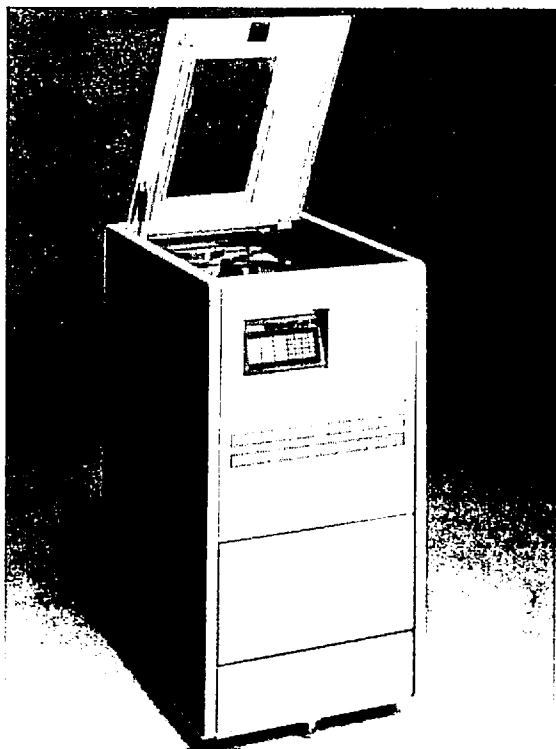
I noen tilfeller vil en flytte eksisterende materiale over på CD-ROM, i andre skal nytt materiale samles inn og bearbejdes. Uansett er det viktig at en under planleggingen tar hensyn til de spesielle egenskaper CD-ROM har (se Bruk av CD-ROM). Innsamling og tilrettelegging kan være et tungt og tidkrevende arbeid pga. den størrelsen materialet gjerne vil ha, og oppgaven kan kompliseres ved at flere typer informasjon inngår. Lyd, bilder og grafikk vil også kreve annen ekspertise og utstyr enn tekst og programmer. Utgiftene ved innsamling og tilrettelegging av materialet for en CD-ROM kan fort komme opp i både 6- og 7-sifrede beløp.

Hvis en selv skal ta hånd om innsamling og tilrettelegging av bilder, grafikk og lyd, trengs spesialutstyr for dette. Bilder tas inn via videokamera eller scanner. Brukes videokamera, må kamera eller datamaskin ha et ekstra kort som digitaliserer videosignalet. Digitaliseringstjenester kan også kjøpes. Grafikk lages best på egne grafiske arbeidsstasjoner med hjelpeutstyr som digitaliseringsbord, mus, programvare for tegning, osv. Til lyd trenger man minimum mikrofon, forsterker, høyttaler og digitalt opptaksutstyr.

Skal materialet inn i en eller annen form for database, må det organiseres og indekser til bruk i søkingen genereres. I de fleste tilfeller må

også ferdige databaser reorganiseres og reindekseres for å fungere på CD-ROM. Så må materialet formateres for å passe til filstrukturen en vil bruke (f.eks. High Sierra-standarden). Det ferdige tilrettelagte materialet vil som regel legges over på vanlig 9-spors magnetbånd, som så sendes til en platefabrikk. Begrepet *premastering* brukes gjerne som en samlebetegnelse på det arbeidet som gjøres med organisering, indeksering og reformatering av et eksisterende materiale, samt overføringen til magnetbånd.

Flere firmaer tilbyr premasteringtjenester, i Norge visstnok bare Data Visual, Byremo, Vest-Agder. Andre firmaer selger utstyr og programvare beregnet som hjelpemiddel i prosessen (se illustrasjon 5). Det finnes både mini- og mikromaskinbasert verktøy. Noen anvendelser vil kreve en minimaskins regnekraft, mens andre vil kunne håndteres av en mikro. De fleste anvendelser vil kreve at utstyret har et stort direkte aksess



*Illustrasjon 5. Utstyr for premastering og simulering av CD-ROM beregnet for IBM XT/AT. Leverandør: Videotools.*

lager, f.eks. en eller flere store diskstasjoner. En magnetbåndstasjon for å kunne skrive ut det endelige resultatet bør også være med. Det mest avanserte utstyret kan simulere en CD-ROMs ytelsesevne (kapasitet, aksestid, overføringshastighet), og en får dermed en realistisk test på hvordan det ferdige produktet vil fungere.

Neste prosess i produksjonen av en CD-ROM er *mastering*. Den består igjen av

- feilkorrigerende koding
- kanalkoding
- innspilling av masterplate

*Feilkorrigerende koding* består i å ta en strøm av bits (i dette tilfellet data fra magnetbåndet), stykke den opp i deler av fast lengde, og for hver del legge på noen ekstra bits. Disse ekstra bitene beregnes matematiske ut fra de foregående, og på en slik måte at selv om en viss mengde data skulle bli ødelagt av støy (f.eks. av riper i platen), vil CD-ROM-spillere være i stand til å rekonstruere data korrekt med høy grad av sannsynlighet.

For CD-ROM kjøres data gjennom minst to feilkorrigerende kodinger, hvorav den siste er den samme som brukes for Compact Disc. F.o.m. denne siste kodingen er produksjonsprosessen for CD-ROM og Compact Disc nøyaktig den samme. Grunnen til at kodingen er mindre grundig for Compact Disc enn CD-ROM er at en for Compact Disc vil kunne akseptere en viss mengde feil. Feil som ikke kan korrigeres, vil nemlig med meget stor sannsynlighet kunne oppdages (det blir en slags paritetsfeil), og dermed kan en Compact Disc-spiller f.eks. blanke ut lyden i et brøkdels sekund uten at lytteren vil merke det. En tilsvarende strategi på CD-ROM er selvsagt helt uholdbar.

En kunne så innbille seg at 1-erne og 0-ene i det nye bitmønsteret gikk ut på platen som et tilsvarende mønster av hull og ikke hull. Så enkelt er det imidlertid ikke. *Kanalkodingen* oversetter nå bitmønsteret til enda et nytt mønster, med spesielle egenskaper tilpasset de fysiske realiteter som spilleren opererer under. Spillere er nemlig i stand til å oppdage fordypninger/mellomrom av en viss minimumslengde, og til å kunne skille mellom forskjellige lengder opp til en viss maksimumslengde.

Det nye mønsteret, altså et mønster av fordypninger av varierende lengde med varierende mellomrom, blir overført til *masterplaten* vha. av en laser. Masterplaten er dekket av et fotosølsomt belegg som påvirkes av laserlyset. De eksponerte områdene vaskes vekk, og platen dekkes til sist med en tynn metallhinne.

Etter *masteringen* lages *stampers*, negative kopier av masterplaten, som igjen brukes til å presse kopier av platen.

Mastering av en *full* CD-ROM koster i dag anslagsvis 75.000 kroner (pris USA \$4000), noe mindre hvis ikke platen fylles opp. Pris pr. kopi



er avhengig av antallet, men i store opplag (over 1000) vil en komme ned i ca. 50 kr. pr. stykk. For opplag på ca. 100 plater må man regne med en hundrelapp eller to.

Det er et begrenset antall fabrikker for mastering og pressing i verden, og det er vanlig med en måneds ventetid fra fabrikken har magnetbåndet i hende. (3M har riktignok annonsert 3 dagers produksjonstid mot et tillegg i prisen.) Årsaken er den store pågangen av Compact Disc-opplag. En måned er selvfølgelig uakseptabelt for mange typer anvendelser, men situasjonen vil bedre seg etter hvert som mange fabrikker øker kapasiteten og flere fabrikker kommer til. Tidsskriftet *Optical Memory News* melder i sitt februarnummer at 16 fabrikker er operative, mens 45 er planlagt i drift i løpet av de første 1 1/2 til 2 år. I marsnummeret blir ytterligere 8 omtalt. I Norge arbeides det med etablering av et firma på Hønefoss - *Scandinavian Laser*.

Mange vil være interessert i å eksperimentere med CD-ROM, men synes utgiftene med en egen plateproduksjon blir for store. For personer/institusjoner innen forskning og undervisning finnes det en løsning på problemet. Professor *Robert Kraft* ved University of Pennsylvania organiserer nå nemlig en tjeneste som har planer om regelmessige utgivelser av plater med testmateriale, muligens hvert halvår. Krafts adresse er: *FCAT, Box 36 College Hall, University of Pennsylvania, Philadelphia, PA 19104, USA*.

NAVF's EDB-senter vil høyst sannsynlig gå i gang med et eget prøveprosjekt, enten alene eller ved å henge seg på et slikt tilbud som nevnt ovenfor.

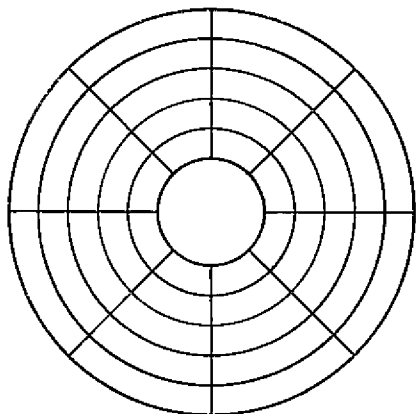
## Bruk av CD-ROM

Fordelene med CD-ROM burde være klare. Men siden CD-ROM ikke ble spesielt utviklet for dataformål, har den karakteristika som må tas hensyn til når man tilrettelegger data og programmer. Det mest kritiske er *aksesstiden*, dvs. tiden det tar for platestasjonen å posisjonere laseren riktig sted og lese det som står der.

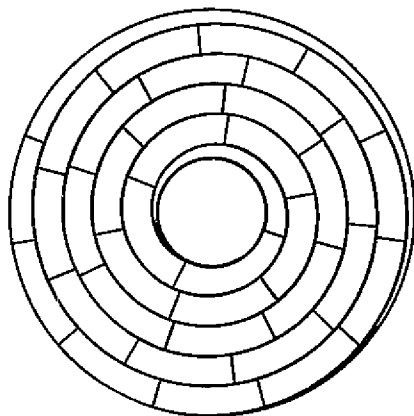
Et vanlig prinsipp for datafagringsmedier i form av plater (disk, disket og mange optiske plater) er at de går rundt med jevn fart, og at ytterste spor eller omdreining inneholder like mye informasjon som det innerste. I tillegg organiseres data i blokker eller «sektorer» som begynner faste steder på platen, dvs. som stykkene i en bløtkake (se illustrasjon 6). Dette er sløsing med plass, da ytterste spor har plass til mer enn innerste. Fordelen er at en raskt finner en hvilken som helst sektor. Aksesstiden blir lav. Betegnelsen som brukes på denne typen plater er *CAV* som er en forkortelse for «constant angular velocity», eller konstant vinkelhastighet på norsk.

Compact Disc, og dermed CD-ROM, bruker et annet prinsipp. Her

CAV



CLV



Illustrasjon 6. Organisering av sektorer på vanlig disk/diskett (CAV) og på CD-ROM (CLV).

roterer platen med konstant *lineær* hastighet, og det vil si at platen roterer saktere når det spilles ytterst (200 omdreininger i minuttet mot vel 500 innerst). Data ligger like tett både ytterst og innerst på platen. Dermed vil sektorene ikke begynne noe fast sted, det vil ta adskillig lengre tid å lete opp en sektor og aksesstiden blir høy. Gjennomsnittlig aksesstid for dagens spillere er over et halvt sekund. Til sammenligning er den gjennomsnittlige aksesstiden for diskstasjoner 30-100 ms (ms = millisekund = tusendels sekund). Heldigvis har mange CD-ROM-spillere meget rask aksess til nærliggende sektorer (1 ms).

Når først en sektor er funnet, er CD-ROM-spilleren i stand til å lese og overføre data i akseptabel fart. Farten er 1,3 Mbit pr. sekund. Overføringshastigheten for en diskettstasjon er typisk 250 Kbit og for en diskstasjon 5 Mbit og oppover. I praksis faller imidlertid overføringshastigheten for magnetiske media pga. at filer vil bestå av sektorer som ligger spredt utover. Dersom man ikke kjører spesielle oppryddingsprogrammer, er mer realistiske overføringshastigheter for diskstasjoner 0,8-2,5 Mbit.

Med CD-ROM er det altså viktig å minimalisere antall aksesser. Dette er ikke noe nytt problem som er dukket opp med CD-ROM. Også for vanlig disk og diskett tar aksessene lang tid i forhold til det som skjer i selve maskinen. Problemet er bare blitt så mye viktigere for CD-ROM, og med de meget store datamengder en gjerne vil ha på CD-ROM forsterkes problemet ytterligere.

Her er noen eksempler på hva en kan gjøre for å holde aksessene nede:

- Prøv å holde data som skal brukes i forbindelse med hverandre fysisk samlet på platen. Dermed kan flere dataenheter leses inn med samme aksess, eller en kan dra fordel av rask aksess til nærliggende sektorer. Et eksempel er trestrukturer, hvor en bør prøve å holde datternodene nær sine mornoder.
- Dupliser opp data for å oppnå nærhet. For de fleste anvendelser er det mer enn nok av plass på platen.
- Bruk ledig plass til å bygge opp ekstra indekser til materialet for dermed å unngå sekvensiell søking.
- Organiser data i store blokker slik at mange dataenheter kan leses inn om gangen istedenfor en og en. Siden CD-ROM-spillere likevel alltid leser hele sektorer på 2 Kbyte (2048 byte, for å være presis), kan det ofte lønne seg med blokker på minst 2 K. Ved å la blokkene ha en størrelse lik et helt multiplum av 2 K vil en av systemtekniske årsaker kunne spare ytterligere tid. Store blokker vil også redusere antall nivåer i trestrukturer (f.eks. de nå så populære B-trærne).
- Behold hyppig brukte blokker i hukommelsen, f.eks. roten i trær.

Både CD-ROM- og CD-I-standardene vil inneholde hjelpemidler for å holde antall aksesser nede. F.eks. vil en anvendelse som skal presentere bilder og lyd-kommentarer til bildene, vanligvis hente bildene fra én fil og lyden fra en annen. Standardene gir en mulighet til å legge filene flettet innimellom hverandre slik at de ikke blir adskilt fysisk.

En vil som regel ha mer datakraft til rådighet ved tilretteleggingen av et materiale for CD-ROM enn ved bruken etterpå. Tilrettelegging vil som regel foregå på mini- og stormaskiner, mens bruken i stor grad vil foregå på mikromaskiner. Tidkrevende oppgaver som leksikalsk analyse og tekstformatering bør altså gjøres på forhånd.

At materialet på en CD-ROM er statisk, er en egenskap som bør utnyttes positivt. F.eks. trenger datastrukturer i databaser ikke lenger være tilpasset både søking og oppdatering. Programvaren kan gjøres enklere fordi oppdatering faller bort og fordi datastrukturene kan gjøres enklere. (Andre faktorer kan selvsagt bidra til å gjøre programvaren mer komplisert igjen.) Blokker kan fylles helt opp da en ikke trenger plass til oppdateringer. Dette sparer både plass og aksesser.

Dette er eksempler på generelle forenklinger, men en vil også kunne gjøre tilpasninger rettet konkret mot det aktuelle materialet. F.eks. kan såkalte hashingfunksjoner skreddersys til de sett av nøkler de bruker. (Hashing er en metode til å plassere og hente dataenheter i en fil. Hashingfunksjonen brukes til å beregne hvor i filen enhetene legges. Beregningen foretas ut fra en nøkkel – en entydig identifikasjon av enhetene. Dersom for mange enheter blir hashet til samme sted i filen,

må noen enheter plasseres andre steder i filen. Dermed blir det flere aksesser for å finne dem igjen.)

I anvendelser hvor en trenger hyppigere oppdateringer enn det er praktisk eller økonomisk overkommelig med CD-ROM, kan en tenke seg at oppdateringene blir publisert på et magnetisk medium. Noen bruker i dag diskett, andre såkalt Bernoulli-disk, som er en slags utskiftbar disk. Et annet alternativ er at oppdateringene gjøres tilgjengelig via en oppringt database. I begge tilfeller er det ønskelig at programvaren er i stand til å presentere materialet fra CD-ROM-en sammen med oppdateringene som et hele, slik at en får inntrykk av å ha en oppdatert CD-ROM.

I mange anvendelser er det bruk for tekstsøking. Det finnes allerede en hel del gjenfinningssystemer for CD-ROM. Her er noen av dem: Turbo Lightning (Borland International), Bluefish (Computer Access Corporation), MicroBASIS (Battelle), Knowledge Retrieval System (KnowledgeSet), Textbank (Group L), D\*CAR\*E (Microforms Trans-Lingual), STA/F Text (Reference Technology), FindIt (Reteaco), Research (TMS), BRS (BRS).

For dem som er interessert i en evaluering av tekstsøkesystemer på CD-ROM, tilbyr U.S. Geological Survey (U.S.G.S.) et interessant produkt, nemlig en plate med samme data lagt inn som base i hele 14 forskjellige søkesystemer. Platen var planlagt ferdig i august. Informasjon kan fås fra *E.J. (Jerry) McFaul, Office of Technology Assessment, U.S.G.S., Reston, VA, USA.*

### **Anvendelsesområder, konkrete anvendelser av CD-ROM**

Den kanskje mest kjente konkrete anvendelsen av CD-ROM i dag er «the Electronic Encyclopedia». Dette er en CD-ROM-utgave av the Academic American Encyclopedia, et 20-binds leksikon utgitt av forlaget Grolier. (Grolier har også eksperimentert med en videoplateversjon av leksikonet.) I den nåværende CD-ROM-utgaven er det kun tekst, men en arbeider med nye versjoner som skal ha både illustrasjoner, tesaurus og annen tilleggsinformasjon. Den nåværende utgaven bruker bare 110 Mbyte av platen (inkludert indekser for søking).

I the Electronic Encyclopedia kan en søke på ethvert ord som forekommer, og på ikke altfor kompliserte kombinasjoner av ord. Søkingen kan også begrenses til artikkeltitler, selve artikkeltekstene, bibliografiske opplysninger, m.m. En kan også sakse materiale og editere videre på disk/diskett. Programvaren er fullstendig menydrivet.

Grolier har priset CD-ROM-leksikonet lavt. The Electronic Encyclopedia har en listepris på \$199 (plate med leksikon + søkeprogram), og årlige oppdaterte versjoner vil koste \$24.95. Til sammenligning koster papirversjonen \$600. En pakke bestående av spiller og plate kan kjøpes over postordre for \$995 fra KnowledgeSet (Activeventure), firmaet som har

skrevet programvaren.

På the First International Conference on CD ROM kunne Microsoft demonstrere en mer avansert prototyp på et leksikon (se illustrasjon 4 og 7). Ved å peke (med mus) på framhevde ord i artikkeltekstene, kunne en



Illustrasjon 7. Skjermbilde fra Microsofts «Multimedia Encyclopedia».

få ytterligere informasjon. Alt etter hvilken farge som var brukt på framhevingen, ble det vist illustrasjoner, illustrasjoner med lyd, animert grafikk (f.eks. kart med bevegelige piler), ordforklaringer fra en dictionary eller beslektede ord fra en tesaurus. «The Multimedia Encyclopedia», som prototypen het, var i tillegg lagt inn under Microsofts «Windows». Windows er en programpakke for IBM-kompatible PC-er som tillater en å kjøre flere programmer i hver sine vinduer. På denne måten kunne leksikonet blitt et verktøy man brukte samtidig med andre programmer, f.eks. tekstbehandling. Microsoft hadde ingen planer om å gjøre the Multimedia Encyclopedia til et komplett leksikon, men det er vel ikke urimelig å tro at noen andre vil gjøre det.

En annen kjent anvendelse er den distribusjon av bibliografiske data som skjer i USA. Firmaet Library Corporation har et produkt de kaller BiblioFile, i sin tid bestående av to CD-ROM med MARC-opplysninger om over 1,4 millioner publikasjoner. Nå er visstnok antall plater lik fem.

(MARC = MACHine Readable Cataloging, et internasjonalt registreringsformat for bibliografiske opplysninger.) Opplysningene på de to første platene favner over alle Library of Congress' engelskspråklige monografier siden 1964, samt hyppig etterspurte titler fra og med århundreskiftet. Distribusjon via CD-ROM er en naturlig oppfølging av virksomheten med distribusjon på magnetbånd. Library Corporation selger CD-ROM-spiller og interface (IBM PC, Atari) for \$680. Med de fem BiblioFile-platene er prisen \$2930. BiblioFile brukes nå i over 250 biblioteker. Det finnes også en rekke andre bibliografiske CD-ROM-produkter og tjenester, både fra Library Corporation og andre.

Distribusjon av store *databaser* via CD-ROM er et alternativ til oppringt søking i basene. Flere databaseoperatører tilbyr nå basene sine på CD-ROM, gjerne gjennom et abonnement på kvartalsmessig eller månedlig oppdaterte plater. I dag er det blant annet tilgjengelig en rekke databaser med opplysninger innen finans/næringsliv, og et helt spekter av baser med abstracts av vitenskapelige artikler (medisin, fiskeriforskning, bioteknologi, kjemi, astronomi, romfart, elektronikk, telekommunikasjon, datamaskiner, helse/sikkerhet/miljø, undervisning, statsvitenskap, jus, for å nevne noen).

Selv om overgang til distribusjon av databaser på CD-ROM ser forlokkende ut for brukerne, kan det være problematisk for baseoperatørene. Hvordan priser man så store datasamlinger? Med oppringte databaser er det enklere, for der tar eierne seg betalt for tid og/eller pr. bit overført materiale. Enkelte databaseoperatører frykter at databaser på CD-ROM skal «kannibalisere» de oppringte basene. Dessverre ser en at noen holder tilbake sine mest verdifulle baser pga. dette.

Et prosjekt solid plantet i humaniora er *Thesaurus Linguae Graecae* (TLG). TLG er en samling tekster som rommer all kjent litteratur av greskspråklige forfattere inntil år 700 f.Kr. Dette inkluderer alle klassiske greske dikt og drama, filosofiske, historiske og medisinske tekster, samt senere religiøse skrifter som det Nye Testamente. Til sammen utgjør dette omtrent 250 Mbyte (250 millioner tegn) tekst. TLG drives av University of California, Irvine. Materialet har hittil vært distribuert på magnetbånd, men er nå også lagt inn på en CD-ROM. På samme plate finnes også flere koptiske og latinske tekster.

Best utbytte av TLG-platen får man ved å bruke Ibycus Systems' «Scholarly Personal Computer». Ibycus Systems er et firma i New Jersey, og Ibycus-maskinen er en spesiallaget maskin for arbeid med manuskripter – tekstkritikk, tekstanalyse og utgivelser. Med maskinen følger tekstsøking og flerspråklig tekstbehandling (bl.a. gresk, koptisk og hebraisk i tillegg til engelsk). Maskinen kan i tillegg programmeres i et eget kraftig høynivåspråk.

*Kart* vil egne seg utmerket for CD-ROM, og da helst såkalte *vektoriserte* kart med *topologiske* opplysninger. Vektorisert vil si at alle streker

(grenser, veier) er lagret som sett av koordinater, altså på det vi har kalt en grafisk form (en vektor er et linjestykke). De topologiske opplysningene er kunnskap om strukturen i kartet, f.eks. opplysninger om hvilke grenser som omgir et areal, opplysninger om forskjellige typer arealer (land, fylker, kommuner), hvordan disse er relatert til hverandre (landet Norge består av fylkene Oslo, Akershus, osv.). Et kart lagret på denne formen kan mer dekkende kalles en *kartografisk database*. En slik lagringsform er kompakt og meget fleksibel i forhold til et «dødt» kart lagret som et rasterbilde («dødt» fordi det ikke inneholder noen strukturelle opplysninger).

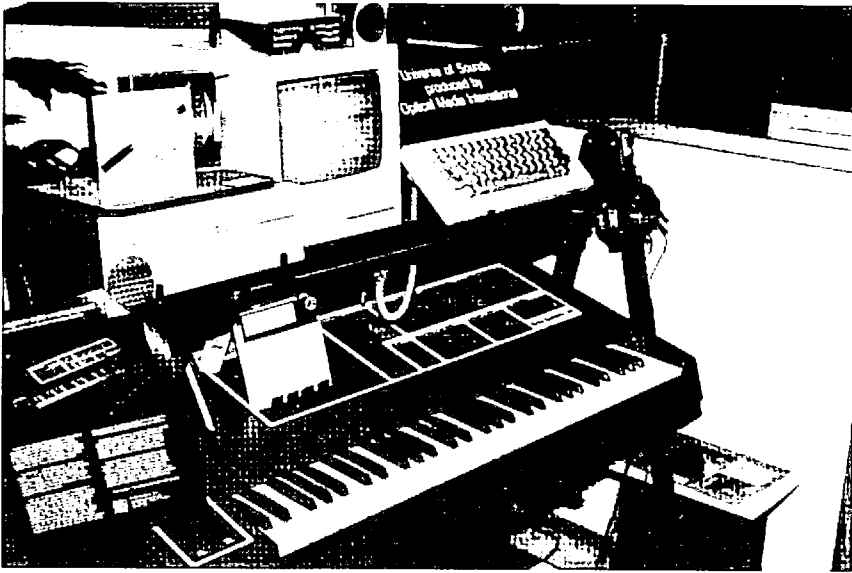
Fleksibiliteten går ut på at et program lettvisnt kan plukke ut enkelttrekk (fylkesgrenser, veinett), kan vise kart i forskjellig skala (noe som krever forskjellig utvalg av detaljer), produsere tematiske kart, foreta beregninger på avstander og areal, osv.

Ett bruksområde for kartografiske databaser er såkalte *navigasjonssystemer* for motorkjøretøyer, skip, fly og raketter. Slike systemer kan være passive, dvs. kun fungere som avanserte kart, eller aktive, dvs. direkte delta i styringen av transportmiddelet. For bilen er det snakk om det første, og bilnavigasjonssystemer anses faktisk for å være et meget stort potensielt marked for CD-ROM. Et system vil typisk bestå av en skjerm for å vise kartet og bilens posisjon, et lager med kart, en mikromaskin og i tillegg et eller annet som kan bestemme bilens posisjon i den virkelige verden og dermed på kartet. Allerede i dag finnes det slike systemer, men med andre lagringsmedier enn CD-ROM.

Verdens første bilnavigasjonssystem ble utviklet av det amerikanske firmaet ETAK. Tidlig i fjor høst gikk de ut på markedet med systemer for San Francisco og Los Angeles, og firmaet arbeider kontinuerlig med tilrettelegging av kart for resten av USA. Senere blir det Europas og Midt-Østens tur. ETAKs system gir i tillegg til bilens posisjon, muligheten for å se på kartet i forskjellig skala og til å finne veien til en bestemt adresse. Avstand og retning til målet blir også kontinuerlig vist. ETAK regner at adekvate kart for hele USA vil kunne få plass på én CD-ROM, men mer interessant er regionale kart sammen med tilleggsinformasjon som f.eks. telefonkatalogens gule sider.

(På CD-konferansen i Seattle ble det dannet en egen interessegruppe for kart på CD-ROM. I Norge er Norsk Regnesentral representert. I Japan er et annet prisverdig initiativ tatt. Der arbeider en gruppe firmaer for standardisering av kartdata for navigasjonssystemer.)

Optical Media International, California, står for den første anvendelsen innen *profesjonell audio* (se illustrasjon 8). Firmaet selger en CD-ROM-spiller som kan styres fra fjernkontroll eller en Macintosh mikromaskin, og en plate med digitalt samplede lyder. Platen er den første i en serie kalt «The Universe of Sounds», og lydene omfatter mer enn 1400 instrumentklanger og effekter. Tidligere var dette lydbiblioteket



*Illustrasjon 8. Utstyr for demonstrasjon av «The Universe of Sounds»: CD-ROM-spiller, Macintosh og Emulator II «digital sampling keyboard», vist på «The First International Conference on CD ROM».*

tilgjengelig på 536 disketter. Lydene kan lastes direkte inn i et Emulator II «digital sampling keyboard», dvs. et elektronisk klaviaturinstrument som lar en spille av vilkårlige lyder fra tangentene. (I motsetning til en synthesizer, som bygger opp lydene fra grunnen av, tar et «sampling keyboard» utgangspunkt i digitaliserte lydprøver. Instrumentet kan selv ta opp og digitalisere «virkelige» lyder via en mikrofon. Ved å manipulere lydprøvene videre digitalt, kan f.eks. instrumentet gjengi samme lyd i forskjellige tonehøyder.)

For å få en anelse om hvilke anvendelser en vil se innen *underholdning/undervisning*, kan det være morsomt å se hvilke planer firmaet The Record Group har. (The Record Group er et underbruk av Warner Communications.) Blant de første demonstrasjonsplatene fra firmaet vil være:

- En interaktiv tur gjennom London. En kan vandre gate for gate og se på de enkelte bygninger, og mange steder kan en også gå tilbake i tid. Materialet vil dekke historisk interessante steder/hendelser helt fra romertiden til i dag.
- En slektshistorisk database over alle verdens overhoder. En vil f.eks. kunne velge to personer og få slektssammenhengen mellom de to.



Databasen kan også brukes som et slags labyrintspill der man selv skal vandre fra den ene til den andre via slektsforbindelsene. Produksjonen er godt i gang og platen skal være ferdig første kvartal 1987. (Spøkefull arbeidstittel: «Princess Di is Related to Chiang Kai-Shek».)

- «Rock Family Trees», som gir en oversikt over sammensetningen av en mengde rockegrupper til forskjellige tider. En kan følge musikerens karrierer gjennom de mer eller mindre kjente grupper de har vært medlemmer av.
- En plate med gammelpop og lokale nostalgiske bilder fra det området Beach Boys og andre surf-grupper holdt til i. Platen har også reklameinnslag og prat av en discjockey. Reklamen og discjockeyen kan man slå av om man vil. Musikken blir avspilt i tilfeldig rekkefølge, noe som vil gi preg av et gammelt radioprogram.
- The CD-I Time Machine. Platen inneholder opplysninger om historiske begivenheter. En kan velge et hvilket som helst år fra 500 f.Kr. til i dag. Platen kan også spilles kronologisk, og da tar presentasjonen ca. en halv time.
- «Adventure», eller «hulespillet» som det ofte kalles på norsk. Den vanlige versjonen av dette klassiske datamaskinspillet baserer seg kun på tekst, men CD-ROM-versjonen vil ha full tale, musikk, tegninger og animasjon. Adventure er et «tenkespill», men det vil også komme et Adventure-aktig spill av reaksjons-typen. Det vil ha adskillig mer realistiske personer enn vanlige videospill, og vi får også vite mer om personenes bakgrunn.

### **CD-ROM i humaniora**

De fleste humanister vil vel se den faglige nytten av CD-ROM. Et spørsmål mange likevel vil stille er om dette mediet kan klare å «slå gjennom» i de humanistiske miljøer. Det er klart at egen produksjon av CD-ROM i mange tilfeller blir for kostbart. Også ren bruk av CD-ROM vil støte på mange hindre i dag. En skranke for noen vil være at CD-ROM krever overgang til datamaskin, for andre at det ikke vil være tilstrekkelig mange interessante CD-ROM-produksjoner til å rettferdiggjøre innkjøp av utstyr, og for andre igjen en så svak økonomi at utstyrskjøp i det hele tatt er utelukket.

Da produksjon er kostbar for små opplag, blir CD-ROM neppe noe rent arkivmedium. (Her vil WORM-plater være en mer aktuell teknologi.) CD-ROM blir først interessant dersom en kan komme opp i et visst antall. I mange tilfeller kan opplaget økes ved samarbeid, enten ved at en legger inn materiale som er interessant for flere anvendelser, eller at man simpelthen deler på plassen.

Det potensielle markedet kan også være større enn man i utgangs-

punktet tenker seg. Mange privatpersoner vil f.eks. være interessert i enkel tilgang til kildehistoriske data, og svært mye humanistisk materiale vil kunne brukes i undervisning i både grunnskole, videregående skole og på høyskolenivå. En CD-ROM-spiller har en overkommelig pris for skole- og høyskolebruk. Dersom CD-I blir en suksess, ligger også hjemmemarkedet åpent.

Det finnes i dag en rekke databaser og tekstsamlinger «bortgjemt» i de humanistiske miljøene i Norge. Mange av disse kan publiseres på CD-ROM uten for mye arbeid. For enkelte fag vil det bli produsert relevante plater utenlands, og i noen tilfeller vil produksjoner fra ikke-humanistiske miljøer være interessante. F.eks. vil en i mange humanistiske anvendelser kunne nyte godt av kart utviklet for andre formål. Et annet eksempel er bruk av musikk på Compact Disc i musikkvitenskap.

Det er allerede nevnt noen konkrete humanistiske anvendelser av CD-ROM, men helt til slutt kan det være morsomt å skissere opp en del andre muligheter:

- Databaser og ordlister med lingvistisk og leksikografisk materiale. Eksempler: Ordbøker, ordlister for orddeling.
- Fonetiske databaser med både taleprøver og transkribert tekst. Bruksområder: Dokumentasjon av dialekter, undervisning i fonetikk, undervisning i fremmedspråk.
- Musikkvitenskapelige databaser (lyd). Eksempel: Samlinger av folke-musikk (med vekt på spillestil, geografiske variasjoner, instrumenter, eller annet).
- Store dataregistrerte tekstsamlinger, f.eks. lagt inn i et tekstsøkesystem. Supplert med scannede versjoner av kildetekster i de tilfeller der det originale skriftbilde også er interessant. Eksempler på materiale: Filosofiske tekster, religionshistoriske tekster, tekstkorpora, enkelte forfatteres produksjon.
- Billedatabaser over arkeologiske gjenstander, museumsgjenstander, faste kulturminner, foto.
- Kart har mange anvendelsesmuligheter i humaniora. En type bruk er visualisering av steds- og tidsrelaterte opplysninger. Eksempler på hva man kan vise med kart + historisk-nominative data (dvs. folketellinger, kirkebøker, emigrantprotokoller osv.): Endringer i bosetningsmønster, endringer i næringsvei, utbredelse av personnavn, slektshistorie, fødselshyppighet, dødelighet, utvandring, osv.
- Kart kan også brukes som inngang til databaser med stedsrelaterte opplysninger (historie, arkeologi, museum). Eksempel på bruk for en arkeologisk funndatabase: Ved å peke på et oversiktskart, får man se et mer detaljert kart over området man er interessert i. Dette kartet har markeringer av fornminner. Ved å peke på en slik markering får man opplysninger om fornminnet, samt et nytt, enda mer

detaljert kart med funnene markert. Ved å peke på en funnmarkering, får man opplysninger om funnet.

- Dokumentdatabaser (scannede dokumenter). Bruksområder: Bibliotek, arkiv, historie, filosofi, litteraturvitenskap, religionshistorie.

#### Litteratur

Steve Lambert and Suzanne Ropiequet (eds.): *The New Papyrus*, Microsoft Press, Redmond, Washington, USA. (Anbefales. Boken finnes naturlig nok også på CD-ROM.)

Flere artikler under samletittelen «Mass Storage» i tidsskriftet *Byte*, mai 1986:

- Bill Zoellick: «CD-ROM Software Development»
- Norman Desmarais: «Laser Libraries»
- Leonard Laub: «The Evolution of Mass Storage»
- Rich Malloy: «A Roundup of Optical Disk Drives»

Oystein Reigem: «Optisk platteteknologi», *Humanistiske Data 2-85*. (Skriv til Senteret etter kopi, da opplaget er «utsolgt».)

#### Tidsskrifter om optisk lagring

*Optical Memory News*, P.O. Box 14817 San Francisco, CA 94114-0817, USA.

*Optical Information Systems*, samt *Optical Information Systems Update* (tidligere *Videodisc and Optical Disk* og *Videodisc and Optical Disk Update*), Meckler Publishing, c/o Eurospan Ltd., 3 Henrietta Street, London, WC2E 8LU, England.

# Educational Uses of Existing Linguistic Software

*Charlotte Toubro*

CALL – Computer Assisted Language Learning – has not yet been fully introduced in the curriculum of language learning, mostly due to lack of courseware sufficiently convincing to justify its presence.

So far, the tendency has been to consider this new tool on a par with other well-known technological tools in (language) teaching, such as audio and video tapes.

On the part of language teachers this view is based on a sound sceptical attitude towards a marketing policy that in many cases has proclaimed this new medium to be new miraculous way of acquiring learning and knowledge, regardless of its actual performance and pedagogical value. By just adding the computer to the technical equipment of the classroom, however, one risks to confine the perception of its potential to that of other already explored tools, as for example the language laboratory.

## **The computer in its own right**

It is obvious that in many ways the computer can be looked upon and used as another tool for a) training purposes – in parallel with written exercises – and b) teaching purposes – in parallel with written material, audio, and video tapes. In these cases the introduction of the computer mainly rests on pedagogical and economic deliberations of the advantages or disadvantages of the computer in terms of its capabilities of infinite repetition, individual speed, personalized, dedicated feed-back (if it is a good program!) as compared to the advantages or disadvantages of other tools or purely human instruction.

Less attention has been paid to the study of the computer «in its own right» in CALL. In other disciplines, in the social and natural sciences, simulation programs provide an example of such a use. Simulation programs exploit some of the computer's capabilities, namely the visual picture combined with the power of fast calculation. However, in language learning the way of acquiring knowledge is different, and the basic elements of the discipline differ completely from those of e.g. simulation programs. And exactly this is the crux of the matter.

By nature the computer is good at manipulating numbers or anything else that can be represented as numbers, because patterns of current switched on or off can be interpreted as numbers and because the sequence of numbers as well as the interaction between them is defined and foreseeable. And the essential feature of a simulation program is just

that: manipulation of numbers and figures. The essential feature of language learning, in contrast, is the manipulation of words. And neither correct words nor correct, i.e. grammatical, structures of word combinations are easily definable and foreseeable. The problem, consequently, lies in the nature of natural language and its inherent incompatibility with the computer's universe of bits and bytes.

This has led to a profusion of gap-filling programs, multiple choice tests, cloze tests, etc., which almost exclusively exploit the computer's capability of comparing a student input to an already given teacher input and deciding whether the two pieces of input are alike or not; if they are alike, the computer outputs an affirmative answer; if not, the student input was incorrect and a negative answer is output. From the teaching/learning point of view it is the latter case that is interesting: in order not merely to test the student, the program must react to an incorrect answer with information that makes it possible for the student to come up with a correct answer next time the same problem arises.

This kind of feed-back can fundamentally be provided in two different ways: a) by furnishing the program with suitable explanations, corrections or answers (and hope that they will fit the situation) or b) by furnishing the program with the means to analyse the student input, deduce the nature of the student's mistake, and on this basis generate remedial output.

The limits of the first solution in terms of acquiring knowledge from one's mistakes will quickly become clear to anybody either designing or working with such programs.

A good example of a tool for courseware construction that provides the means of analysing the student input and of generating suggestions for corrections is a program called FROG (French Robust Grammar Checker) described by Imlah and Boulay (1).

FROG contains a so-called «robust» parser, i.e. a parser that is able to parse ungrammatical phrases and sentences and as a result of the parse to present the student with inconsistencies/errors in the input.

The student types in an arbitrary sentence (in French) and «the program responds by checking the sentence for certain grammatical errors and reports them to the student». The actual program works with the language pair French —> English and in the version described by Imlah and Boulay with a limited grammar (e.g. only declarative sentences). Yet, the examples given reveal quite a comprehensive morphological component and a degree of robustness that seems to make FROG and similar programs a useful teaching tool (see table I). («/» indicates an «accent aigu»). («>» indicates input from the user, the rest is output from the program).

>Le bateaus est sorte/e par mon frere qui portira la pain pour le mangir.

Le bateaus

Unrecognized word "bateaus".

Assumed to be noun (m,p).

Determiner le (m,s) does not agree with noun bateaus (m,p).

Should be "les".

est

Verb "est" does not agree with subject "bateaus".

sorte/e

Unrecognized word "sorte/e".

Assumed to be -ir type (f,s) verb past participle.

Should be ". . . sortie . . ."

Past participle sorte/e (f,s) does not agree with subject bateaus (m,p). Should be "sortis".

par mon frere

Unrecognized word "frere".

Assumed to be noun (m,s).

qui portira

Unrecognized word "portira".

Assumed to be a verb, future tense, (third, s).

La pain

Determiner la (f,s) does not agree with noun pain (m,s).

Should be "le".

pour le mangir

Unrecognized word "mangir".

Assumed to be -ger type verb infinitive "manger".

#### *Table 1*

It is clear, however, that even programs of this type, made to handle natural language, have shortcomings:

- 1) The parser is limited to a subset of natural language. This will usually be the case either because of storage reasons, and/or because of problems with describing the full range of grammatical utterances.
- 2) However robust the parser is, its report on mistakes must be based on some sort of recognizable structure. Certain combinations of mistakes might turn a sentence into garbage, incomprehensible to the program.

- 3) A parser that parses ungrammatical sentences must take some decisions on what the student intended (cf. the example above, where the program assumes that the student wanted to write a sentence in plural on the grounds of the wrong word form «bateaus»).

### Reversing the roles

So far we have focused on the traditional teaching process:

(teacher explains) —> (student absorbs) —> student tries —> teacher/program corrects —> student learns.

The teaching material is provided by the teacher (in one way or another) and the student is «at the receiving end» attempting to master a subject in the cycle of

—> trial —> error —> correction —> trial ...

At the same time we are faced with the problem of how to use this new tool, bearing in mind that programs capable of handling natural language adequately require an analytical component which cannot be produced quickly, nor easily, nor cheaply.

A different approach is offered by the possibility of turning to other, non-teaching oriented natural language computer applications which might bring interesting new perspectives to the field of CALL.

Two such applications are authoring systems and computer assisted translation systems (2). The object of both is to process language, in the former case with a view to designing and producing teaching material, in the latter with a view to making translations.

From a technical point of view both kinds of programs provide a high-level programming language and the borderlines between what is traditionally known as text processing, data base systems, and parsing are often not noticeable for the user.

The student using a teaching program is mainly in the role of the receiver, whereas the user of both an authoring system and a CAT system is a creator, a producer. The idea is then, not only to make use of this already existing language-oriented software but also to «reverse» the roles of teacher and student and let the student take a more active part.

Instead of the learning/teaching algorithm above, we can envisage a different process:

(teacher explains) —> (student absorbs) —> student tries to «explain» via the computer —> student «corrects» himself in interaction with program and teacher —> student learns

which also upgrades the trial-error-correction cycle in the direction of a more active and less tedious working process.

## Authoring systems

An authoring system is a tool mainly aimed at the teacher that is not versed in lower level programming languages; it allows teachers to compose their own teaching programs without prior knowledge of computers or programming. Authoring systems are usually menu-driven and the programs one can produce with them are of the gap-filling type.

A typical example is Brainlearn, an authoring system written by Zettersten and Jacobson (3). The following tables illustrate the format (tables 2 and 3).

LGNERATOR - Edit questions	
Question number: 1/1	
Swedish sentence:	
1	_____
English sentence:	
2	_____
3	Guidance: _____
4	Answer : _____
Feedback:	
5	_____
6	_____
Enter command >	
H=HELP	

Table 2

Most language teachers agree that although exercises on the basis of gap-filling programs, such as structure drills, may be a necessary part of language learning, they hardly constitute the major part of the curriculum. If the student is to achieve communicative competence, a sound basis of linguistic competence is an absolute prerequisite. However, this competence is not obtained through training of isolated grammatical problems only (which is what gap-filling programs offer) but must also rely on exercises practicing the composition of full sentence structures.

In general, the programs that most authoring systems allow teachers to produce are too limited to provide this. On the other hand, the teacher who produces the program uses knowledge of syntactical structures, etc. to single out problems and to decide on both the initial output to the student and the corrective statements needed in case of incorrect student input. If the student, instead of the teacher, tries to design and produce a program, then he or she would have to make use of the same skills.



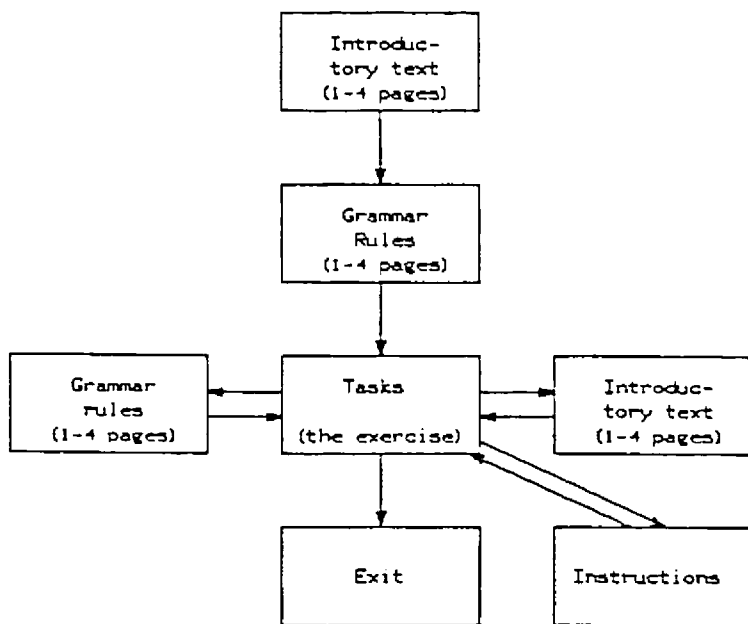


Table 3

At word level one can think of programs dealing with

*paradigms*

They would yield insight into the structure of words

- the delimitation of stems
- the delimitation of flexives
- the structure of a flexive (consisting of several morphemes)
- possible linking elements

and into what it takes to constitute a paradigm.

*word lists*

Producing word lists gives insight into the vocabulary of a language and demands reflection on how to define a class of items and how to describe the items in order to fulfill the purpose of a specific word list. A number of possibilities can be thought of:

- alphabetic lists  
(give rise to questions about capitals, abbreviations, compounds, lemma vs. declined form, etc.)

- frequency lists  
(give insight into the concept «frequency vs. text type»)
- «grammatical» lists  
(requires knowledge of word category, declination patterns, sound  
—> spelling changes)
- collocation lists  
idiom lists  
(give insight into restrictions on collocating words)
- thesaurus lists  
(require knowledge of the organization of word universes and give  
insight into natural relationships between words)
- lists of synonyms and antonyms
- lists of words graded according to the degree of difficulty.

At phrase/sentence level the student can outline and produce programs testing specific syntactic structures:

*grammar instructions*

The focus could be either on the grammatical instructions supporting an exercise, or on the output the program produces in case of mistakes on the part of the student. This requires

- an overview of a grammatical problem
- a terminology in which to express grammatical concepts
- an insight into the grammatical mistakes the particular problem case may lead a student to commit.

**CAT systems**

Computer assisted translation systems are based on interactive use, i.e. the program relies on communication with the user. The degree of communication is generally left at the discretion of the user according to the level of assistance the user expects from the program.

One of the most advanced programs is the ALPS system (4). The ALPS system works on three levels:

- 1) Selective Dictionary Lookup – Allows the translator to consult the automatic dictionary at any point during the translation process on the text processor; the translator creates his or her own dictionary.
- 2) Auto Term (automatic dictionary lookup) – Provides the translator with a word-for-word translation, including idioms and terms. Morphological processing enables the program to associate inflected word forms with dictionary entries (lemmas). Ambiguities and words-not-found are displayed on the screen. Disambiguating is done by the translator, and new dictionary entries can be added during the translation process. The translator has access to KWIC-

lists and frequency lists of the text in question which allows him to check the coverage and the accuracy of the dictionary entry he is adding to the system. The composition of the translation (with correct word order, structural changes, etc.) is the responsibility of the translator.

- 3) Transactive (interactive translation) - Produces a draft translation of the text sentence-by-sentence complete with grammatical processing. The computer still relies on the interactive consultation of the translator, e.g.

user input: X is a big computer user

system output: Is the text talking about

a) big computer

b) big user       ?

For purposes of language teaching, CAT systems exhibit two interesting properties: they focus on the bilingual situation and they include parsing components with regard to both morphology and grammar. Their modular construction (the three levels) seems to make them particularly suitable for teaching purposes - quite apart from their intended uses.

CAT systems offer a number of educational possibilities both at word and at phrase/sentence level.

At word level the main feature of a CAT system is its dictionary. The dictionary is created by the user, whilst the program has the facilities for writing and processing dictionary entries. Moreover, a system like ALPS provides the tools for building dictionaries through special processes like corpus analysis, where a text is segmented into words and (as far as possible) terms.

Most of the instructive uses concerning the word level as mentioned above in connection with authoring systems would also apply to a CAT system. However, it seems more promising to focus on the additional potential of a CAT system, such as

*the organization of a monolingual dictionary*

Requires decisions about the units of the dictionary, the information attached to the units in terms of grammatical information, context, synonyms, idioms, etc.

*the organization of a bilingual dictionary*

Requires additional decisions about target language equivalents, generality vs. specificity of the dictionary, structural differences between source and target language. At phrase/sentence level the program offers at least two possible approaches:

focusing on the production of translations (in particular with the use

of an automatic dictionary lookup program)

The production of translations involves all the well-known problems of choosing the right target language equivalent in terms of text type and style. In as far as the dictionary does not contain correct or sufficiently described entries, choices and decisions must be made with regard to terms, idioms and stock phrases.

focusing on the correction of translations (in particular with the use of an interactive translation program)

Although the translation problems are more or less identical to those just mentioned, the focus could be on the inadequacies of the dictionary and of the grammar. An analysis of the raw translations proposed by the program will point out insufficiencies in the dictionary and in the grammatical component.

When making a dictionary in the void it will not always be clear what information is needed when the process of translation takes place. The interactive dictionary lookup process provides the basis, i.e. a context on which a dictionary can be further refined. The interactive translation program further adds the possibility of investigating what problems arise with regard to the mono-lingual dictionary when looked at in a bi-lingual environment.

Incorrect translations make it necessary for the student to consider what kind of additional information the grammar/parser of the program needs in order to render correct translations. One can single out specific areas such as personal pronouns, word order, (the choice of correct) prepositions, etc.

An example from a translation obtained with the ALPS Transactive component illustrates the idea:

Floppy disk and diskette are generally interchangeable terms. Computer assisted translation is an interesting field to work in. A disk resembles a flexible phonograph record and comes in diameters of 5 and 8 inches.

El disco flexible y diskette son terminos generalmente intercambiables. La traduccion asistida por computadora es un area interesante para trabajar adentro. Un disco se parece a un disco fonografico flexible y puede tener diametros de 5 y 8 pulgadas.

This example reveals the different usage of the definite article in English and Spanish, respectively, and the problems of translating the free infinitive into Spanish.

The following examples stem from another CAT system, Weidner, as demonstrated at the ASLIB conference in 1985:

The man, the girl liked, came            L'homme la fille aimee est venue

< the parser fails to recognize the relative clause >

and traditional tests like

Time flies like arrows

Chronometrer des mouches  
comme les fleches

< the parser does not have correct word class information or does not interpret it correctly >

In both cases better knowledge of sentence structure would allow the system to produce better translations.

### Concluding remarks

This article does not intend to set up a recipe for the design of any particular program. Rather, the aim has been to provide thoughts and suggestions as to how the field of CALL could be advanced further.

At present, one of the dangers is that teachers refuse to use computers in language instruction, arguing that most existing courseware is too limited and too traditional to offer any new perspectives to the teaching situation. Moreover, contrary to the general pedagogical trend, courseware of the gap-filling kind leaves the student in a very passive role.

If the computer is to open up new avenues in language instruction, then the focus has to lie on courseware that is able to

- (1) assign an active role to the student
- (2) handle (a subset of) natural language
- (3) analyse the student input and generate remedial information.

The use of authoring systems or CAT systems as sketched in this article offers one possible line of attack; no doubt, there are many others.

### Notes

- (1) Imlah, W.G., Boulay, J.B.H. du: Robust Natural Language Parsing in Computer-assisted Instruction. in *System*, Vol. 13, No. 2, pp 137-147, 1985.
- (2) Patrick Corness of Coventry Lanchester Polytechnic, Dept. of Language Studies, described at the Aslib Conference, 1985, London, an educational use of one particular CAT system - the ALPS system. Some of his main points were the computer's and the CAT system's ability to expose students to languages and, in particular, to their similarities and contrasts, and the more systematic and conscious decision-making that the computer obliges the student to develop.
- (3) Zettersten, Arne: Authoring Systems - the Key to Future Expansion in CAI. in *System*, Vol. 13, No. 2, pp 133-136, 1985.  
Zettersten, Arne: *New Technologies in Language Learning*. Gyldendal, Cph., 1986.
- (4) ALPS - Automated Language Processing Systems: Multiple Levels of Translator Assistance.

### *Acknowledgement*

My information about the ALPS system is based on visits to the ALPS company; I would like to express my thanks to the ALPS company, in particular to Rebecca Ray, for taking the time to demonstrate their system to me.

*Charlotte Toubro is Research Assistant at Eurotra-DK, University of Copenhagen. Previously having worked on computer-assisted instruction, she is presently engaged in the field of automatic translation.*



# Edb-utbygging og edb-virksomhet i det offentlige arkivverk

*Ivar Fønnes*

I HD nr. 2-84 hadde Hege Brit Randsborg en rapport om «Forskerstipend i automatisert arkivinformasjon». Her ble det redegjort for NAVFs engasjement i arbeidet med å gjøre offentlig arkivmateriale lettere tilgjengelig for forskerne. Det omtalte forskerstipend ble opprettet fra 1.8.1983, og prosjektet ble avsluttet ved utgangen av 1985. Hovedoppgaven for stipendiaten var å utvikle et edb-basert opplegg for produksjon av arkivkataloger og en direkte søkbar database med katalogopplysninger. Ved utgangen av stipendperioden var edb-basert produksjon av kataloger innført som standardssystem i Riksarkivet og enkelte statsarkiver, og det var opprettet en sentral katalogdatabase i Riksarkivet som inneholdt ca. 30.000 kataloginnførsler.

NAVFs forskerstipend inngår som en sentral del i oppbyggingen av en selvstendig edb-virksomhet i det offentlige arkivverk (Riksarkivet og 7 statsarkiver). Stipendiatens arbeid har vært konsentrert om å forbedre og gjøre lettere tilgjengelig informasjon om arkivmateriale som er avlevert til arkivverket, og som i all hovedsak finnes på papir. Like viktig er det imidlertid for arkivverket, og i neste omgang også for forskerne, å ta hånd om det *edb-baserte* materiale som er produsert i den offentlige forvaltning de siste 20 år. Dette materialet skal tas vare på, avleveres til arkivverket og gjøres tilgjengelig for framtidig forskning.

Parallelt med forskerstipendiatens arbeid har det derfor også blitt satt i gang annen edb-aktivitet i arkivverket, særlig knyttet til den offentlige forvaltnings bruk av edb. Internt er det bygget opp en egen edb-avdeling i Riksarkivet og anskaffet edb-utstyr sentralt og lokalt. Opplæring av arkivpersonalet i edb-anvendelser har også vært gitt høy prioritet. Denne utbyggingen danner grunnlaget for å føre videre aktiviteten, bl.a. på det område hvor NAVFs forskerstipendiat har arbeidet.

I det følgende skal jeg gi en oversikt over edb-utviklingen i arkivverket de siste 3 år og antyde noen utviklingslinjer framover.

## **1. Utbygging av edb-ressurser**

Edb-ressurser består først og fremst av kompetent personale, utstyr og programvare. Blant personalet er det nødvendig med godt kvalifiserte edb-spesialister. Men det er også viktig å bygge opp et rimelig kompetansenivå hos den vanlige saksbehandler/arkivar. I det lange løp vil dette

være den viktigste forutsetning for en fornuftig håndtering av edb-oppgaver i etaten.

### **1.1 Riksarkivets edb-avdeling**

Fram til 1983 hadde Riksarkivet en arkivarstilling med oppgaver knyttet til edb, men denne stillingen ble også pålagt en rekke andre oppgaver. Høsten 1983 kom NAVFs forskerstipend i stand, og omtrent samtidig ble det opprettet en edb-avdeling i Riksarkivet med 2 ansatte. Edb-avdelingen hadde først en midlertidig status, men ble gjort permanent fra 1985. Samtidig ble den bygget videre ut, og fra 1.1.1986 er det 4 faste stillinger i avdelingen. Dessuten disponerer avdelingen en av Riksarkivets sivilarbeidere.

Riksarkivets edb-avdeling har inntil videre ansvaret for alt edb-basert arkivmateriale i den sentrale statsforvaltning, og til dels også i lokalforvaltningen. Alle avleveringer av edb-materiale skal foreløpig skje til Riksarkivet v/edb-avdelingen. Det er imidlertid grunn til å anta at deler av dette ansvaret vil bli overført til fagavdelingene og statsarkivene etter hvert som den generelle edb-kompetanse og andre edb-ressurser bygges ut.

Edb-avdelingen har også et overordnet ansvar for etatens interne edb-virksomhet og -utbygging, og avdelingen yter tjenester til andre deler av etaten, særlig i Riksarkivet.

En stor del av det arkivmateriale som produseres i offentlig forvaltning (både på papir og edb-media), er personregistre som krever konsesjon fra Datatilsynet. Det er en rekke berøringspunkter mellom Datatilsynets krav om sletting av personregistre og Riksarkivarens bestemmelser om oppbevaring av arkivmateriale. Disse spørsmålene er også edb-avdelingens ansvar.

### **1.2 Edb-utstyr og programvare**

Arkivverket har i løpet av de 3 siste år investert 3,5-4 mill. kr. i edb-utstyr og programvare. De største investeringene har skjedd i Riksarkivet, begrunnet med edb-avdelingens sentrale og omfattende edb-oppgaver. Men det er også anskaffet utstyr til samtlige statsarkiver.

Riksarkivets hovedanlegg fra 1985 er en ND-530/CX med 4,25 Mb hukommelse, 450 Mb plattelager og 2 magnetbåndstasjoner. Anlegget, som også betjener de andre institusjonene i riksarkivbygningen, har for tiden tilknyttet ca. 30 terminaler. Sentral programvare er tekstbehandlingssystemet NOTIS-WP og databasesystemet FICS-5 som også er et 4. generasjons verktøy. Innenfor FICS er det utviklet en rekke applikasjoner.

I statsarkivene utenfor Oslo blir det i juni 1986 installert et 3Com Ethernet med 3-4 arbeidsplasser (Olivetti PC) og 70 Mb plattelager hvert



sted. Unntatt fra dette er Statsarkivet i Trondheim som benytter en Discovery-maskin, anskaffet i 1984. Inkludert mindre utstyrsenheter som har vært installert tidligere, vil alle statsarkiver fra 1986 ha 3-6 terminal-arbeidsplasser og 60-80 Mb platelager. Grunnleggende programvare er i hovedsak Word Perfect og Dbase II/III.

Kommunikasjon mellom statsarkivene og Riksarkivet er foreløpig ikke etablert, men den nødvendige programvare er anskaffet.

### **1.3 Kompetanseoppbygging: opplæring og praksis**

Oppbygging av edb-kompetanse hos personalet tar dels sikte på at de skal kunne anvende edb i intern virksomhet, men også at de skal kunne ta stilling til edb-spørsmål som knytter seg til deres arkivfaglige virksomhet, bl.a. i forbindelse med feltarbeid overfor offentlig forvaltning.

Ut fra dette er det lagt stor vekt på praktisk opplæring kombinert med øvelser og muligheter for å benytte edb i det daglige arbeid. De fleste har gjennomgått kurs i tekstbehandling, og en stor del av de ansatte benytter dette. Et mindre antall har fått opplæring i databaser og benytter slike. Foreløpig er det imidlertid svært få av de ansatte som har terminal/PC på sitt eget kontor, og dette gjør at edb-anvendelsene blir mindre omfattende enn det oppgavene tilsier.

Men det arkivfaglige personalet må også ha kunnskaper om håndtering av edb-basert arkivmateriale, dels fordi de vil møte slike spørsmål i sin utadrettede virksomhet, dels for at ansvaret for slikt materiale på lengre sikt skal kunne integreres i fagavdelingenes og statsarkivenes virksomhet. Det er derfor gjennomført et større etatskurs i edb for arkivarer, og dette vil bli fulgt opp med lignende og eventuelt også videre opplæring.

## **2. Edb-virksomheten: oppgaver og systemer**

Arkivverkets edb-oppgaver har mange felles trekk med det vi finner i andre offentlige institusjoner, bl.a. når det gjelder kontorautomasjon og administrative registre til bruk i den daglige saksbehandling m.v. Men arkivverket har en omfattende edb-oppgave i tillegg til dette – nemlig det faktum at offentlig forvaltning produserer mer og mer edb-basert arkivmateriale som arkivverket er forpliktet til å ta hånd om på en eller annen måte, enten ved å tillate kassasjon eller ved å få det avlevert og oppbevart for framtidig forskning m.v.

Arkivverkets edb-oppgaver kan ut fra dette deles i 3 hovedområder

- feltarbeid overfor offentlig forvaltning
- depotarbeid og brukertjenester (herunder arkivinformatjon)
- administrative oppgaver, kontorautomasjon

## 2.1 Feltarbeid overfor offentlig forvaltning

Feltarbeidet overfor offentlig forvaltning er et viktig grunnlagsarbeid for å sikre at offentlig arkivmateriale kan bevares for fremtiden. Arbeidet går inn på selve *arkivdanningen* (altså hvordan offentlige institusjoner tilrettelegger, ordner og oppbevarer sitt arkivmateriale samt hva de skal kassere) og på *avlevering* av materialet til arkivverket. Det omfatter både rådgivning og håndheving av regelverk.

*Arkivdanningen* knyttet til edb-materiale er i dag meget omfattende, ikke minst p.g.a. den utstrakte bruk av tekstbehandling og annen kontorautomasjon. Dette gjør at Riksarkivets edb-avdeling ikke har ressurser til å gå inn på arkivdanningen i større bredde. Vi har valgt å konsentrere oss om ett felt, som er særlig sentralt i arkivsammenheng, nemlig utvikling og innføring av *journalssystemer* basert på edb.

I 1984 utarbeidet avdelingen retningslinjer for innføring av edb-basert journalføring i statsforvaltningen. Disse ble gjort gjeldende fra 1.9.1984. Samme år deltok vi, sammen med R-direktoratet, i utarbeidelsen av kravspesifikasjoner for et norsk arkivsystem, NOARK. Rapporten ble utgitt på Aschehougs forlag i 1984, og flere maskinleverandører har utviklet journalssystemer på grunnlag av NOARK-rapporten. En del av disse er nå i ferd med å bli installert i sentrale forvaltningsinstitusjoner.

Edb-avdelingen har også gitt en del konsulentbistand til institusjoner som skal innføre edb-journal, og til leverandører som har arbeidet med NOARK.

*Avlevering* av edb-basert arkivmateriale har i de senere år framstått som et nokså akutt problem. Det finnes materiale som går tilbake til midten av 1960-årene, og de lagringsmedia som benyttes (magnetbånd) egner seg ikke for langtidslagring uten jevnlig ettersyn og fornyelse.

Edb-avdelingen gjennomførte i 1984-85 en kartlegging av hva som finnes av edb-materiale i den sentrale statsadministrasjon, og på dette grunnlag er det lagt planer for gjennomgåelse og avlevering av alt materiale som er eldre enn fra 1980. Under gjennomgåelsen tas det stilling til hva som skal bevares og hva som skal kasseres, og etter planen skal slik kassasjonsbehandling og avlevering være gjennomført for sentraladministrasjonen innen utgangen av 1989. De første avleveringer av magnetbånd fant sted i 1985.

I den lokale forvaltning er det forholdsvis lite gammelt edb-materiale, og dette vil derfor bli behandlet i sammenheng med annet arkivmateriale. Men edb-avdelingen har innledet samarbeid med Kommunedata A/L for å finne fram til depotordninger for kommunalt og fylkeskommunalt edb-materiale, som ikke skal avleveres til arkivverket.

*Forskningsdata*, dvs. edb-materiale som tilrettelegges i tilknytning til forskningsprosjekter, kan det også være av interesse å bevare for ettertiden med sikte på at andre forskere kan ha nytte av det. Dette har nå

vært vurdert av et utvalg nedsatt av NAVF, hvor også Riksarkivaren var representert. Utvalgets innstilling foreligger i disse dager (juni 1986), og det foreslås arkivordninger for slike data i regi av NAVF, men i samarbeid med det offentlige arkivverk.

## **2.2 Depotarbeid og brukertjenester. Arkivinformasjon og framfinning**

*Edb-basert arkivmateriale* som er avlevert til arkivverket, har foreløpig et nokså beskjedent omfang. Totalt er det avlevert knapt 100 magnetbånd, og selv om det er store datamengder på hvert bånd (opptil 160 Mb), er dette bare en spe begynnelse. Vi antar at det finnes minst 50.000 magnetbånd med arkivmateriale ute i forvaltningen. Store deler av dette materialet vil kunne kasseres, men det blir også igjen en del materiale (noen tusen magnetbånd) som skal tas vare på for ettertiden og avleveres til arkivverket.

Ved avlevering av data på magnetbånd må det foretas kontroll både av magnetbåndet (format og datastruktur) og den vedlagte dokumentasjon. Ett av problemene med edb-baserte lagringsmedia, er de mange ulike lagringsformater som benyttes. Arkivverket stiller en del krav til formatet på de magnetbånd som avleveres, og det må kontrolleres (evt. ved stikkprøver) om kravene er oppfylt. Like viktig er det å kontrollere dokumentasjonen og sammenholde denne med datastrukturen på magnetbåndene. Det har vist seg, både i Norge og andre land, at dokumentasjonen ofte er mangelfull og lite informativ. Dette kan gjøre det vanskelig for senere brukere å utnytte materialet.

Et annet problem ved magnetbånd og andre magnetiske lagringsmedia er den forholdsvis korte levetiden. Det antas at data på magnetbånd bør overføres til nye bånd etter ca. 5-10 år dersom man skal være rimelig sikker på å ikke miste informasjon. Dette innebærer at det må legges opp rutiner for jevnlig kontroll og fornyelse av et magnetbåndarkiv.

Foreløpig har naturlig nok de driftsmessige depotoppgavene vært av begrenset omfang. Riksarkivets edb-avdeling har derfor kunnet konsentrere denne del av virksomheten om å bygge opp programvare og rutiner for å behandle et større magnetbåndarkiv. Det er utviklet programvare for å kontrollere og kopiere magnetbånd som er avlevert utenfra. Og det er lagt opp til kontroll og eventuelt kopiering av alle magnetbånd innen 5 år etter avlevering.

Heller ikke brukertjenester på edb-basert arkivmateriale har vært noen omfattende oppgave hittil. Det materialet som er avlevert, er ligningsregistre (hele landet) for 1967, 1970, 1975 og 1980 samt GAB-registrene (grunneiendom-, adresse- og boligregistre) pr. 1.1.1985. Foreløpig har det ikke vært noen brukerforespørsler etter dette. Edb-avdelingen har imidlertid utviklet programvare for å hente ut informasjon fra magnetbånd, og disse programmene skal utvikles videre. Aktuelle brukertjenester på kort



*Fra ett av Riksarkivets magasiner. Lagring av dokumenter på optiske plater istedenfor papir kan radikalt redusere behovet for framtidige arkivmagasiner.*

sikt er å kopiere magnetbånd for brukere som har tilgang til datamaskin på sin arbeidsplass, og å kjøre ut selektive lister fra data.

*Arkivinformasjon og framfinningssystemer* har som formål dels å informere brukerne om hva som finnes i et arkiv, dels å gjøre det mulig å finne fram til de aktuelle arkivsaker. Gode framfinningssystemer er nødvendige for å yte hensiktsmessige brukertjenester både på edb-basert arkivmateriale og på tradisjonelt materiale. Et edb-basert framfinningssystem vil i så måte ha flere fordeler framfor en tradisjonell katalog, bl.a. fordi det er lett å ajourholde til enhver tid, og fordi alle opplysninger er søkbare.

Dette er tankegangen bak den katalogdatabasen som ble opprettet av NAVFs forskerstipendiat (jfr. ovenfor). Databasen har opplysninger på tre nivåer – om arkivskapere, arkivserier og de enkelte arkivstykker (esker, protokoller m.v.) – og det er mulig å søke på alle nivåer. Det gis referanse til arkivinstusjon og arkivstykkets plassering i institusjonens magasiner.

Katalogdatabasen, som ved stipendperiodens utløp hadde ca. 30.000 innførsler, er under stadig utbygging. Alle nye kataloger i Riksarkivet registreres i edb-systemet og legges inn i databasen. Etter hvert vil også kataloger i statsarkivene bli innlemmet. Det gjøres dessuten forsøk med at avleverende institusjoner skriver sine avleveringslister direkte inn i Riksarkivets katalogregistreringssystem. På denne måten kan nye katalogopplysninger innlemmes i databasen uten at Riksarkivet behøver å foreta egen registrering. Dette vil utvilsomt effektivisere arbeidet med å katalogisere arkivmaterialet og dermed medføre at katalogopplysningene i databasen foreligger praktisk talt samtidig med avleveringen.

Et annet spørsmål er hva som skal gjøres med gamle kataloger. Kataloger som skal revideres/justeres/ajourføres, vil det være naturlig å registrere i edb-systemet slik at de også kan innlemmes i basen. Men det er neppe realistisk, iallfall på kort sikt, å skrive om alle gamle kataloger for å få opplysningene inn i databasen. En mulig løsning er å skrive inn opplysningene om arkivskaper og arkivserie (men ikke de enkelte arkivstykker) fra alle kataloger og la arkivserien ha referanse til den gamle katalogen på papir. Dette er en overkommelig oppgave. Fordelen med en slik løsning er at brukeren får direkte tilgang eller referanse til *alle* eksisterende katalogopplysninger ved å gå inn i katalogdatabasen på arkivskaper- eller arkivserienivå.

Foreløpig er katalogdatabasen en forsøksbase, og alle kataloger kjøres også ut på papir i tradisjonelt format. Etter hvert vil basen bli gjort tilgjengelig for saksbehandlerne i arkivverket, og senere tar vi sikte på at også eksterne brukere skal kunne søke i basen (jfr. nedenfor).

### 2.3 Administrative registre, kontorautomasjon

Bruk av elektronisk tekstbehandling har først og fremst kommet i gang det siste året. Men dette hjelpemidlet brukes i stor utstrekning både av saksbehandlere og kontorpersonale for en rekke formål. Det er også bygget opp en del edb-registre som benyttes i administrative sammenheng.

Det neste trinn i arkivverkets kontorautomasjon er innføring av edb-journal. Statsarkivet i Trondheim har benyttet et enkelt system siden 1985. Selv om systemet ikke gir brukbare søkemuligheter, har det vist seg nyttig og arbeidsbesparende i forhold til manuell journal. I Riksarkivet settes det fra 1. juli 1986 i gang prøvedrift med en versjon av NOARK. Det er meningen å gå over til NOARK som fast journalsystem fra 1.1.1987 i Riksarkivet. Lignende ordninger planlegges også for statsarkivene når den nødvendige programvare er anskaffet.

Edb-baserte økonomisystemer og personalforvaltningssystemer er foreløpig ikke vurdert, men spørsmålet vil bli tatt opp som neste punkt i kontorautomatiseringsprosessen.

### 3. Framtidsutsikter

Det er liten tvil om at arkivverket vil bli sterkt berørt av edb-utviklingen i samfunnet. Dette gjelder dels nye muligheter som følge av den teknologiske utvikling, dels nye utfordringer som følge av forvaltningens overgang til edb i arkivsammenheng, og som følge av at brukerne vil stille krav om forbedring og modernisering av tjenestetilbudet.

Ser vi på den teknologiske utvikling, er det særlig optiske lagringsmedia og datakommunikasjon som skaper interessante perspektiver.

Optiske plater gir muligheter for meget kompakt lagring, noe som også gjør det realistisk å kombinere tekst med bilder osv., og de er vesentlig mer arkivbestandige enn magnetbånd. Med tid og stunder vil det trolig bli aktuelt å overføre alle arkivdata fra magnetbånd til optiske plater, og det kan også komme på tale å avfotografere arkivdokumenter fra papir til optisk plate (i stedet for mikrofilm) for å lette framfinningen og for å beskytte originalen. Dersom de arkivskapende institusjoner går over til å lagre dokumenter på optiske plater i stedet for papir, vil behovet for framtidige arkivmagasiner reduseres radikalt.

Utviklingen innen datakommunikasjon vil gjøre det mulig, og vesentlig enklere enn i dag, å koble seg opp mot en database langt borte. Det er grunn til å anta at Riksarkivets katalogdatabase vil bli tilgjengelig for eksterne brukere på denne måten om noen år. Dette vil føre til en revolusjon i arkivverkets brukertjenester. Mange brukere vil kunne hente ut informasjon om arkivsakene direkte på sin egen arbeidsplass, og dersom selve arkivmaterialet ligger på edb-medium, vil de også kunne hente fram innholdet på sin egen dataskjerm.

Slike brukertjenester er imidlertid avhengige av at arkivverket finner det hensiktsmessig å gi et slikt tilbud, og at man har ressurser til å bygge ut tjenester på dette nivå. Men mye tyder på at nettopp katalogdatabasen og en del tilleggsregistre knyttet til denne vil bli gitt høy prioritet. Det er f.eks. planer om å bygge opp et samlet arkivskaperregister med «alle» arkivrelevante opplysninger om de enkelte institusjoner, og et register over hvilken administrativ tilhørighet de enkelte forvaltningsområder har hatt opp gjennom historien (en historisk statskalender). Når to slike registre blir integrert med katalogdatabasen, vil de fleste brukere få et kraftig arbeidsredskap for å finne fram til det ønskede arkivmateriale. Hvor raskt dette kan realiseres er et spørsmål om ressurser.

*Ivar Fonnes har vært avdelingsleder i Riksarkivet fram til 1.8.1986.*

## «Project Emperor-I» – kinesiske skatter på videoplate

*Kristin Natvig*

«The First Emperor of China» er tittelen på to amerikansk-produserte videoplater som presenterer og tolker arkeologiske funn fra perioden 221-206 f.Kr., da Kinas første keiser – Qin Shi Huang Di – regjerte. Mest berømt blant disse funnene er Qins storslåtte gravplass ved Xian, som inneholder 7000 terrakottafigurer av soldater og hester. Utgravingen av denne gravplassen startet i 1973, men det fins foreløpig lite litteratur om funnene.

Leder for «Project Emperor-I» er Dr. *Ching-chih Chen* ved Graduate School of Library & Information Science, Simmons College, Boston. Dr. Chen har arbeidet i mange år med anvendelse av ny teknologi i bibliotek og informasjonsentre og med vitenskapelig, teknisk og medisinsk informasjon. I april besøkte hun NAVFs EDB-senter for humanistisk forskning, hvor hun ga en demonstrasjon av «The First Emperor of China» og holdt foredrag om prosjektet.

### Formål med prosjektet

For Dr. Chen er det viktig at bibliotek samarbeider med informasjonsteknologi-industrien for å formidle informasjon til både forskere og det generelle publikum. Videoplater gir adgang til informasjon som er vanskelig tilgjengelig av økonomiske og praktiske årsaker. På den andre siden åpner videoplateteknologien for nye muligheter for formidling av humanistisk vitenskap.

«Ett av formålene med Project Emperor-I er å fremme publikums forståelse av humaniora generelt, og av en annen kultur», fremholdt prof. Chen under sitt foredrag. «Valget av emne for videoplatene henger også sammen med ønsket om å koble sammen humanistisk vitenskap og ny teknologi, øst og vest, og fortid og framtid.»

Ifølge *Rus Gant*, prosjektets tekniske leder og ansvarlig for videoplatens tilknyttede database, er det meningen at platene skal tjene som en generell prototype på innsamling og lagring av informasjon med mulighet for tilhørende undervisningsprogrammer.

### Videoplatenes innhold

Den første videoplaten inneholder 200 film- og videosekvenser og over 4000 dias, sammen med to 30-minutters lydspor på både engelsk og kinesisk, og musikkinnslag. «Den visuelle informasjonen er lagret i



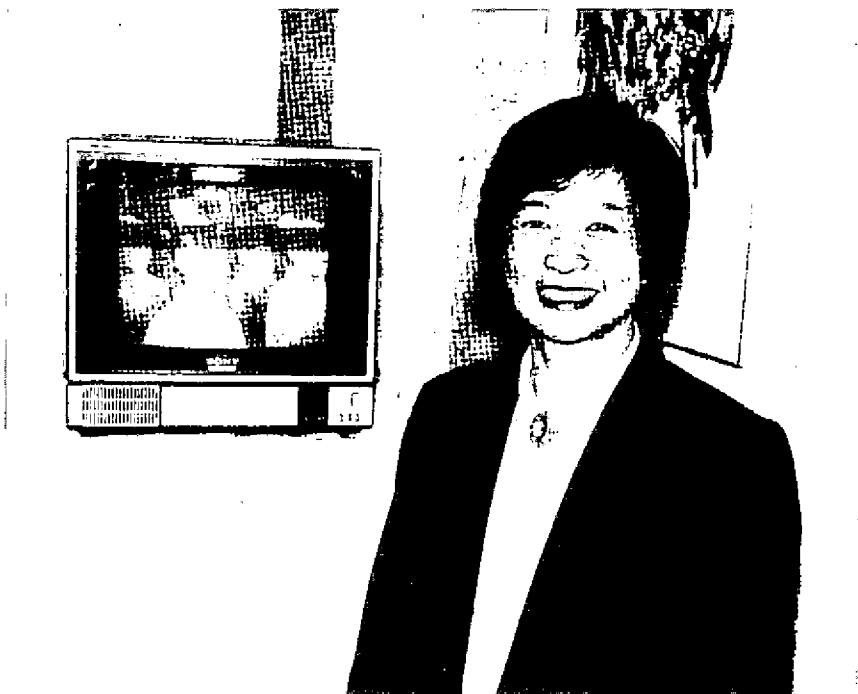
«kapitler» på platen», fortalte Dr. Chen. «Hvert kapittel kan gjenfinnes på tre sekunder – gjenfinning av enkeltbilder skjer også raskt. På den første platen fins dessuten innledninger til keiserens liv og virke, til den kinesiske muren og til selve utgravingen av terrakottafigurene.»

Den andre platen består av intervjuer med verdens 10 fremste eksperter på Qin-dynastiet. Også denne platen er arrangert i kapitler, og intervjuobjektene blir stilt spørsmål av særlig interesse for studenter og forskere, opplyste Dr. Chen.

### **Databasen**

Prosjektets database er fremdeles under utarbeidelse. En applikasjons-generator – sannsynligvis Microcomputer Picture Archive and Communication System fra Artam Ltd. (Israel) – skal brukes i utviklingen av et databasesystem som er dels relasjonell, dels hierarkisk. Data som er innsamlet ved de arkeologiske utgravningene er blitt kodet for å passe inn i hierarkiske strukturer med mulighet for relasjonelle kryssreferanser.

Et eksempel: for hver terrakottafigur er det oppgitt informasjon om



*Dr. Chin-chih Chen demonstrerer «The First Emperor of China».*

eventuell fargelegging, og i så fall, hvilke kroppsdelar som er malt, armstilling osv. Disse opplysningene vil ligge innenfor spesifiserte felt (f.eks. maling på armer), feltene innenfor skjermbilder (f.eks. maling), og skjermbilder innenfor deldatabaser (f.eks. soldater). Deldatabasene vil på sin side ligge innenfor en omfattende database for statuer. Inverterte nøkkelord-søkingar (ved boolske operatorer) vil muliggjøre hierarkisk forgreining gjennom informasjonen.

Databasen for statuer – i likhet med prosjektets øvrige databaser – vil også omfatte deldatabaser for tekst, fotografier og videosekvenser. Informasjonen i disse vil inneholde ulike typer kryss-referansar til andre former for opplysningar, slik at samtidig gjenfinning av tekstlig, visuell og auditiv informasjon vil være mulig.

Til gjenfinning av data skal det benyttes DEC's Interactive Video Information Systems (IVIS) 350 Personal Computer, et 10 Mb platelager for masselagring, en 13" fargemonitor, en videoplatespiller og en IVIS sentralenhet som grensesnitt mellom PC'en og videoplatespilleren. Prosjektmedarbeiderne ønsker også å utvikle systemer for IBM-maskiner og MS-DOS.

### **Pedagogisk programvare**

Dr. Chen opplyste at prosjektmedarbeiderne har planlagt å utvikle flere undervisningsprogrammer for brukere på ulike nivåer, dvs. det generelle publikum, studenter og forskere. Eksempelvis vil ett program hjelpe brukeren til å finne fram i databasene, mens et annet vil gjøre det mulig for arkeologer å overføre informasjon fra flere ulike databaser til sine egne personlege databaser.

Hvert program vil bestå av «leksjoner» på til sammen 8-10 timer. Den enkelte leksjon skal presenteres i flere hundre skjermbilder ved hjelp av menyer, undermenyer, grafikk, vinduer, forgreiningssjanser osv. Tekstlig, visuell, grafisk og auditiv informasjon skal kunne presenteres i kombinasjonar etter brukerens eget valg.

### **Prosjektets organisering**

Foruten Dr. Chen og Rus Gant er Dr. *Robert D. Stueart*, dekanus ved the Graduate School of Library and Information Science, Simmons College, medlem av prosjektgruppa. Stueart er ansvarlig for datainn-samling og -tolkning. Flere konsulenter, bl.a. i kinesisk kunst, er også tilknyttet prosjektet.

Prosjektet har fått bevilget nærmere \$20.000 fra National Endowment for the Humanities, og over \$56.000 fra Simmons College. DEC har donert to IVIS-systemer til prosjektet.

Prosjektet ble starta 1. oktober 1984 og blir avslutta sent på høsten i år. Medarbeiderne tilbrakte to månader i Kina for filmopptak og datainn-



*Soldater og hester i terrakotta i keiser Qins gravplass. (Courtesy of Project Emperor-I).*

samling. Disse oppgavene ble utført i samarbeid med det kinesiske kulturdepartementet og Shansi-provinsens museumsbyrå. Til gjengjeld skal universitetene i Xian og Peking få hvert sitt IVIS-system, ikke bare til utnyttelse av videoplatene og den tilhørende programvaren, men også for å kunne lagre dokumentasjonen av senere funn.

### **Vurdering av prosjektet**

Det er lagt ned et imponerende arbeid i Project Emperor-I. De resulterende videoplatene er av meget høy kvalitet, både teknisk og innholdsmessig, og bør være av interesse for både forskere og det generelle publikum. Prosjektmedarbeiderne har imidlertid måttet arbeide døgnet rundt i perioder for å få produsert platene innen tidsrammen, og prosjektet har vært svært kostbart.

Dr. Chen håper at Project Emperor-I skal bidra til å befeste videoplatens posisjon som ikke bare et underholdningsmedium, men også et seriøst medium for lagring, behandling og gjenfinning av informasjon. En slik status er imidlertid helt avhengig av kvaliteten på de tilknyttede databasene og undervisningsprogrammene, som altså fremdeles er under utarbeidelse.

# RAPPORTER

## Training the Information Researcher for the Future

Anglo-Nordic Research Seminar, Lidingö, Stockholm, 9.-11.4.86

*Jostein H. Hauge*

Hvordan bør en i fremtiden utdanne og dyktiggjøre dem som skal forske i emner med tilknytning til vitenskapelig og faglig informasjon? Dette var en del av temaet for et anglo-nordisk seminar i Stockholm i april i år. Vertskap for seminaret var The British Library Research and Development Department og Det nordiske samarbeidsorgan for vitenskapelig informasjon, NORDINFO. I alt deltok ca. 30 inviterte medarbeidere med tilknytning til biblioteks- og dokumentasjonsfaglige miljøer i Norden og Storbritannia.

Gjennom en rekke foredrag og drøftinger forsøkte man å definere vitenskapelig det forskningsområdet som av flere ble omtalt som «Information Science». Hvordan er avgrensningen til «Library Science» og «Computer Science»? Hvordan vil ny teknologi påvirke hele feltet? Skal en i biblioteks- og dokumentasjonsfaglig virksomhet tilpasse seg eksisterende teknologi eller gjennom samarbeid med og krav til datavitenskapene sikte mot å utvikle nye metoder og teknikker spesielt tilpasset informasjonsbehandling? Sentralt stod også drøftinger om arten av og det akademiske nivået på biblioteksvitenskapelig utdanning i fremtiden. Behovene for å kunne nyttiggjøre seg bedre de økonomiske og personellmessige ressurser i Norden gjennom praktiske samarbeidstiltak ble drøftet i flere av foredragene, og det ble foreslått samarbeidsopplegg som også innbefattet Storbritannia.

I et bredt anlagt forsøk på begrepsavklaring gjennomgikk professor *Marjatta Okko*, Universitetet i Tampere, Finland en del av de forsøk som er gjort på å relatere ulike informasjonsvitenskapelige forskningsområder til hverandre. Det ble bl.a. sagt følgende om «Information Research»: «Information Research deals with problems that arise in connection with academic and research libraries, information science agencies and information resources management. Information Research

is thus tied to two relatively separate professional practices, librarianship and information transfer».

Forfatteren tok for seg IFLAs definisjoner av biblioteksvitenskap og informasjonsvitenskap og konstaterte med undring hvor lite vekt det i definisjonen er lagt på dem som skal nytte informasjonen, dvs. bruker-aspektet. Etter forfatterens mening er library/information science i hovedsak *kommunikasjonsvitenskaper*. Den faglige utdanning bør etter forfatterens mening klart ligge på universitetsnivå slik den alt en tid har gjort i Finland hvor man kan ta både grunnleggende og videregående utdanning i biblioteksvitenskap i Tampere og Abo. Som forfatteren uttrykte det: «There is no plausible reason why practitioners in the library and information field should be deprived of the right to understand with ease what is being achieved in information research».

Dir. *Maurice Line*, Science Technology and Industry Department, British Library tok for seg den forskning som i dag utføres innenfor biblioteks- og dokumentasjonsfaglige emner og fant at svært lite hadde hatt noen direkte virkning på den faglige praksis: «Researchers have often not made clear the practical implications of their research while



To av foredragsholderne, prof. *Marjatta Okko*, Univ. i Tampere, Finland (t.v.) og forsker *Liv A. Holm*, BRODD, Statens bibliotekhøgskole, Oslo (t.h.) sammen med representanter for arrangørene, dir. *Maurice Line*, British Library og generalsekretær *Teodora Oker-Blom*, NORDINFO.

practitioners are either unaware of relevant research or cynical about it – or both». Dette kommer av at forskerne har «an apparent inability to identify real and important issues and to translate them into researchable topics ...». Maurice Line gjorde seg til talsmann for at en i all undervisning først og fremst måtte lære de kommende forskere til å stille de rette spørsmål og utvikle en kreativ og divergent forskerholdning.

Å kunne stille spørsmål som gjelder *hvorfor* og ikke bare *hvordan* er viktig både for praktikere og for forskere sa Line, som ønsket et utdanningssystem som gav samme grunnutdanning for begge grupper. Praktikere må også ta del i biblioteksfaglig forsknings- og utviklingsarbeid mente han og omtalte ulike innovasjonstiltak som er satt i verk i britiske bibliotek. Bl.a. kom foredragsholderen inn på ordningen med ambulerende grupper med biblioteksfaglige ressurspersoner som drar rundt til ulike bibliotek for å analysere og hjelpe det stedlige bibliotek i deres faglige oppgaver. Verdien av operative forskningsmiljøer som utførende bibliotekspersonale kan oppsøke og hospitere ved, ble også understreket.

Sjefsbibliotekar *Stephan Schwarz*, KTH, Stockholm tok opp de mange forsøk som har vært gjort på å etablere en egen forskningsdisiplin som kunne kalles «Library Science». Schwarz mente at avgrensningsproblemer i seg selv er lite faglig interessante. For å komme videre i selve forskningsarbeidet bør en kort og godt innse at Library Science er en administrativ fagenhet som består av en hel rekke fagelementer. Nå er det på tide å legge vekt på de internt-vitenskapelige forutsetninger for denne type forskning, hevdet Schwarz og tok opp sentrale spørsmål knyttet til teori, metode, teknikk og faglig infrastruktur på dette feltet.

I et foredrag kalt «Information technology and training of information researchers» tok forsker *Liv A. Holm*, BRODD, Statens bibliotek- og dokumentasjonsfaglig side bør det i dag bl.a. stilles krav om bedre datamaskiner, nettverk, databasesystemer, programmeringsspråk og søketeknikker. Det er behov for å inngå et samarbeid om en rekke informasjonsvitenskapelige temaer med forskere innenfor datafagene samtidig som det er en oppgave generelt å øke interessen for biblioteksfaglige problemstillinger i slike miljøer. Det er viktig fremover både å øke innslaget av datavitenskapelige emner i bibliotekstudiet og å gi mulighet for å rekruttere kandidater med utdanning i datafag til arbeid med biblioteksspørsmål. Verdien av forskningsmiljøer på feltet ble understreket og av fleksible veilednings- og utviklingsorganer som kan sikre kontinuitet og styrke kvaliteten i forskningsarbeidet.

Forsker *Roland Hjerpe*, LIBLAB, Linköping var også opptatt av å sette forskningsområdene «Library Science» og «Computer Science» i forhold til hverandre, og fant bl.a. at Library Science av flere grunner (bl.a. symbolbehandling) kunne ses på som en del av Computer Science.

Informasjonsteknologien gir i dag helt nye utfordringer til biblioteks- og dokumentasjonssektoren og det gjelder derfor nå først og fremst å utvikle en forskningsstrategi som tar hensyn til de nye mulighetene som åpnes. Hovedbudskapet kan oppsummeres i det følgende: «So far the concern in applying information technology to the tasks in libraries has almost exclusively been the preservation of earlier representatives and modes of doing things; doing the same but with different means. It is high time to start paying increased attention to the impact of information technology, implying i.a.:

- The transformation of the modes of generation of text
- The changes in the forms for making texts public
- The conversion to electronic means for distribution and storage of texts
- The erosion of basic concepts, like document
- The extension of old concepts, e.g. interactive fiction
- The emergence of new concepts, e.g. artificial realities
- The potential for considering what's desirable rather than what's possible, hence
- The feasibility of eliminating constraints imposed by earlier technologies, with
- The need for reassessment of present ways of description and manipulation».

I et foredrag som kom noe på siden av selve hovedtemaet tok førstelektor *Niels Erik Wille*, Institut for kommunikationsstudier, Roskilde universitet, Danmark for seg behovet for informasjons- og dokumentasjonstjenester i de humanistiske fag. Wille understreket at en gjør de humanistiske fag stor urett ved å definere deres behov innenfor rammen av andre fagdisipliner. Mens man f.eks. i mange fag ser på en fast oppbygget faglig terminologi som en forutsetning for effektive datatjenester, er situasjonen, ifølge Wille, i humanistiske fag den at den løse og stadig skiftende terminologi representerer en viktig kommunikasjonsfaktor blant forskerne: «Researchers identify themselves to friends and foes by their language, and this part of the communication of the community of researchers should not be suppressed by the bibliographical information systems, but should be reflected by them».

I tillegg kommer at forskningen i de humanistiske fag i mange tilfeller er avgrenset til nasjonale forhold. Derfor kan utviklingen av internasjonale biblioteksystemer gi mindre hjelp til dette fagområdet enn til mange andre: «The needs of Danish researchers in these fields are not catered to by feeding information about Danish research into the big international networks».

Ulik situasjonen i andre fag er også at biblioteket i seg selv for mange humanistiske forskningsdisipliner ofte representerer selve forskningslaboratoriet gjennom den samling av tekster og primærkilder som der finnes:

«It is far more important for us to have access to ordered collections of the relevant texts, than access to ordered collections of research papers».

Men også Wille ser behovet for moderne informasjonstjenester i de humanistiske fag under forutsetning av at informasjonstjenestene tar tilstrekkelig hensyn til egenarten i disse fagområdene.

I en egen sesjon på seminaret ble erfaringer fra industrirettede informasjonstjenester drøftet med vekt på tjenestenes oppbygging, virkemåte og ønskelige utvikling i fremtiden. Informasjonstjenester for små og mellomstore bedrifter må ta hensyn til en rekke særegne krav enten hos grupper av bedrifter eller hos individuelle foretak.

I sitt foredrag redegjorde dir. *Michael W. Hill* fra Science Reference and Information Service, British Library for engelske erfaringer, og gjennomgikk resultatene fra en rekke studier som har sett på hvilke krav som stilles til denne type formidling. For eksempel må informasjonstjenestene

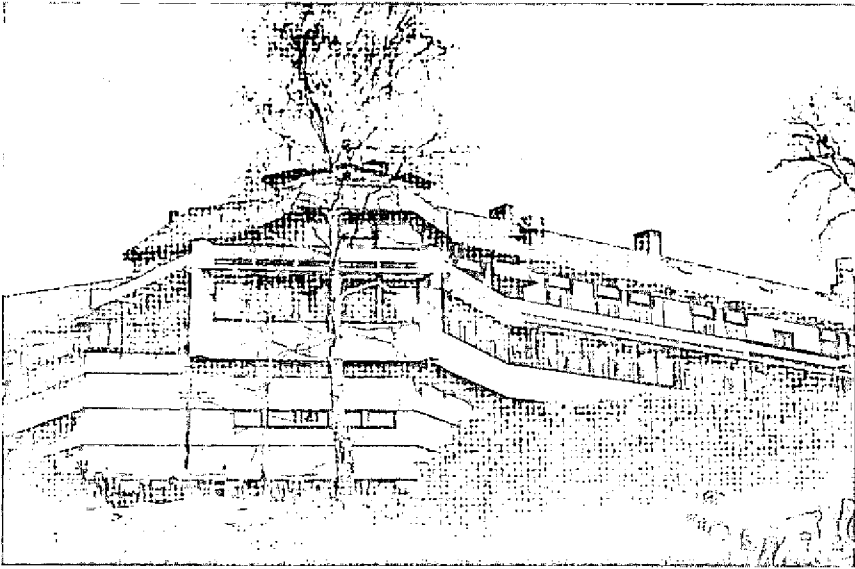
- bare gi den informasjon kunden ønsker - i øyeblikkelig forståelig og anvendbare svar
- være enkle å bruke
- bare involvere kjent teknologi, f.eks. telefon
- kunne nås raskt i det lokale miljø
- baseres på tillitsforhold til dem som driver tjenesten
- være stabil
- ta betaling som kundene kan make.

I slike informasjonstjenester må kunnskapsoverføringen baseres på person til personkontakt og muntlig fremstilling. De kontaktskapende evnene til informasjonsmedarbeiderne blir derfor en helt avgjørende faktor. En omfattende studie fra Suffolk County Library gir data for at markedsinformasjon, teknisk informasjon og næringslivsinformasjon blir høyest prioritert, mens resultater fra forsknings- og utviklingsarbeid var dårligst etterspurt. I andre studier blir imidlertid FoU-virksomhet langt høyere vurdert. I England har en også maktet å bringe de offentlige bibliotekene inn som virkemiddel i informasjonsarbeidet, noe som synes å være vanskelig i andre land.

I et av sesjonens foredrag tok avdelingsbibliotekar *Kirsten Engelstad*, Riksbibliotek-tjenesten, Norge særlig opp spørsmålet om hvordan bibliotekene, og ikke minst de tekniske biblioteker, kan settes i stand til å gi tjenester til næringslivet i samarbeid med de næringslivskonsulenter som i dag finnes i regionene.

I det utredningsarbeidet som frem til i dag er gjort i Norge over informasjonsbehovene i næringslivet, er ikke bibliotekene med i det hele tatt. Engelstad tok til orde for at bibliotekene knyttet til de tekniske høyskoler nå tar selvstendige initiativ til å etablere næringslivskontakt og videreutvikle sine tjenester for nye målgrupper. Engelstad, som selv nylig





*Handelns gård, Lidingö, Stockholm*

har utgitt en rapport om informasjonsbehovene i norske små og mellomstore bedrifter, planlegger sammen med tre tekniske bibliotek prøveprosjekter i informasjonsformidling til næringslivet. Målet er gjennom kontakt med bl.a. bedriftsledere, bransjeforeninger og næringslivskonsulenter å gå inn i utvalgte bedrifter og tilby biblioteksbaserte informasjonstjenester. Provesprosjektene vil bli avsluttet med konferanser for relevante bedrifter i en region der informasjonstjenestenes nytteverdi for næringslivet vil bli evaluert og drøftet.

Fra dansk hold ble det redegjort for det systemet med regionale, tekniske informasjonssentra (13 stk.) som er utviklet siden 1969. Heller ikke her er bibliotekene trukket inn på annen måte enn som sekundære støttemiljøer for den fagstab som bemanner sentrene. Også i Danmark er informasjonsoverføringen basert på muntlig, personlig kontakt, men etter hvert begynner en også å ta i bruk interaktiv søking i informasjonsdatabaser som særlig er tilrettelagt for små og mellomstore bedrifter.

I en avslutningssesjon på seminaret ble ulike samarbeidstiltak mellom de nordiske land og Norden og Storbritannia drøftet. Det ble bl.a. fremsatt forslag om FOU-orienterte sommerskoler hvor forskningsmetoder og forskningsteknikker kunne stå i sentrum. Det ble også foreslått en ordning med utveksling av medarbeidere mellom forskningsinstitusjoner

innenfor BDI-sektoren. Dessuten ble det lagt frem tanker om fellesprosjekter, f.eks. knyttet til spesielle problemfelt som uttesting og utvikling av nye grensesnitt mellom bibliotekssystemer og brukere. Det kan også komme på tale å foreta nasjonale undersøkelser som senere kan sammenlignes, f.eks. om behovene og mulighetene for forskning innenfor biblioteks- og dokumentasjons-sektoren. Konkret ble det her foreslått forskningsoppgaver i tilknytning til det humanistiske og samfunnsvitenskapelige fagområdet.

Kopi av foredragene i de ulike sesjonene er tilgjengelige. NORDINFO vil senere gi ut en rapport fra seminaret. NORDINFOS adresse er:

NORDINFO  
Tekniska högskolans bibliotek  
Otnäsvägen 9  
SF-02150 ESBO  
Finland

## **Utviklingsseminaret på Ustaoset april 1986**

*Asbjørn Brændeland*

14.-16. april ble det på Ustaoset Høyfjellshotell avholdt et utviklingsseminar for edb-personell innenfor humanistisk-faglig arbeid i Norge. (Den noe vage termen «edb-personell ...» refererer til personer beskjeftiget med administrasjon, konsulentvirksomhet, undervisning og/eller forskning innenfor humanistisk edb-virksomhet.) Tilsvarende arrangementer har vært avviklet hvert år siden 1981. Ansvarlig for arrangementet denne gang var EDB-seksjonen ved HF - Universitetet i Bergen og NAVFs EDB-senter for humanistisk forskning. Sistnevnte har for øvrig vært medarrangør ved alle de tidligere seminarene.

Deltakerlisten ved seminaret gjenspeiler stort sett fordelingen av humanistisk edb-virksomhet rundt om i landet, med flest deltakere fra Bergen (NAVFs EDB-senter, Edb-seksjonen ved HF, Nordisk institutt, Engelsk institutt og Norsk Samfunnsvitenskapelig Datatjeneste (NSD)), en relativt stor kontingent fra Oslo (HF-data, Riksarkivet, Germanistisk institutt og Norsk folkemusikksamling), to deltagere fra Institutt for samfunnsvitenskap ved Universitetet i Tromsø, foruten inviterte gjester og foredragsholdere fra andre institusjoner og firmaer (Rådet for humanistisk forskning i NAVF, RUNIT, Digital, Weidner Translation - London, Byremo Telematikk-senter, Salttrød). For første gang var Universitetet i Trondheim uten deltakere.

Seminaret hadde tre hovedtemaer: **nyere lagringsteknologi, data-kommunikasjon og internasjonale datanett og maskinassistert oversettelse.** I tillegg til dette fikk vi rapporter fra de enkelte miljøene om status for forskning, undervisning, personell og utstyr.

### **Nyere lagringsteknologi**

*Tor Thomassen* fra Byremo Telematikk-senter snakket om optiske plater generelt, om integrering av mikromaskiner og drivere for optiske plater spesielt og om forventede trekk ved den videre utviklingen av ny lagringsteknologi.

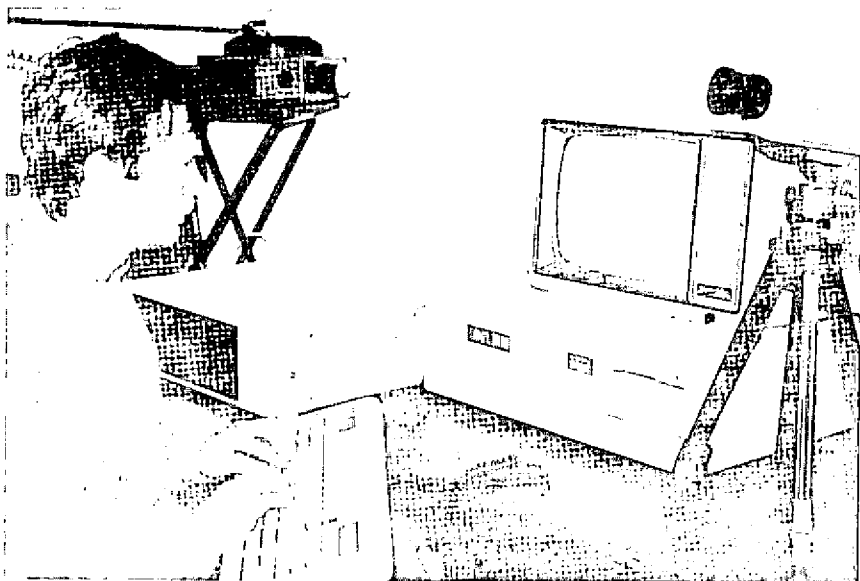
Temaet optiske plater var til dels kjent for mange (og det har da også vært presentert på tidligere utviklingsseminarer og i Humanistiske Data) uten at foredraget dermed ble uinteressant. Foredraget ble for øvrig positivt preget av Thomassens engasjement og entusiasme.

Vi merket oss at Philips i samarbeid med Control Data og Du Pont synes å ha lykkes i å utvikle en slags standard på området (i alle fall foreløpig).

Ved Byremo har man arbeidet – og langt på vei lykkes – med å utvikle et integrert system for mikromaskiner av PC-typen og drivere for optiske plater, noe de faktisk er alene om i verden foreløpig. Et slikt system er interessant av flere grunner – ikke minst fordi det blir vesentlig billigere og mer tilgjengelig enn de systemene man hittil har sett.

Utviklingen fremover vil rimeligvis gå mot pakking av stadig mer data på stadig mer kompakte media. Foruten videreutviklingen av de optiske platene, vil vi formodentlig ganske snart få de lenge bebudede vertikalt magnetiserte platene, som i forhold til dagens horisontalt magnetiserte plater vil få opptil 100 ganger så høy lagringstetthet. (En 5 1/4" PC-diskett som nå tar 360 Kb, skulle dermed kunne få plass til 36 Mb – om vi nå tør tro på dette.) Med en slik utvikling av lagringsmediene, vil flaskehalsene bli å finne i programvaren, i prosessorene og i databus'ene (den interne datatransporten).

En vesentlig del av innsatsen ved database-utvikling legges i å optimalisere søkingen og dermed senke søketiden ved ulike typer søking. Kritiske faktorer her er bl.a. søkealgoritmene, valg av filstruktur, den interne dataoverføringskapasiteten og hastigheten ved ikke-sekvensiell lesing fra ytre lager. En teknikk som vil kunne fjerne de fleste av disse flaskehalsene, består i å sende hele datamengden i form av en bit-seriell datastrøm gjennom en databus, samtidig som man lar en kontrollbrikke «lytte» til datastrømmen og hente ut data etter bestemte kriterier. En slik brikke (som faktisk alt er utviklet av en nordmann) vil kunne programmeres med inntil 120 ulike kriterier, og datastrømmen vil kunne få en hastighet på 100 MB/sekund – eller slik Thommassen så fargerikt uttrykte det: «... en hel bibel i sekundet».



*Jan Telhaug fra Byremo Telematikk-senter demonstrerer en WORM-plate. Bildet på TV-skjermen er lagt inn på platen med videokamera.*

Det var en viss interesse blant tilhørerne for muligheten av å utvikle optiske plater som det kunne skrives på mer enn én gang – såkalte utviskbare (erasable) plater. (De tilgjengelige platene i dag er av typen CD-ROM (Compact Disk – Read Only Memory) og WORM (Write Once Read Many (times).) Det arbeides forsåvidt med teknikker for utviskbare optiske plater, men det er, ifølge Thommassen, til dels misforstått å betrakte de optiske platene som en fremtidig erstatning for dagens magnetiske plater. De optiske platene vil få sin hovedanvendelse i sammenhenger der man ønsker permanent lagring av store datamengder som ikke skal endres.

### **Kommunikasjon og internasjonale data-nett**

*Frithjof Iversen* fra RUNIT snakket om internasjonale data-nett. Iversen arbeider ved den norske EARN-noden, som er plassert i Trondheim, og hadde mest å fortelle om EARN-nettet.

Et nettverk av denne typen består av noder – datamaskiner – fysiske forbindelser mellom nodene i form av kabler eller trådløse forbindelser via satellitt samt programvare for dataoverføring og for en rekke kommunikasjons tjenester. Enhver node i nettet kan fungere både som ut-

gangspunkt og adresse for kommunikasjon. Som utgangspunkt kan node- ne brukes direkte eller kalles opp via lokalt nett eller modem. Kommunikasjonsprogrammets funksjoner kan omfatte bl.a. direktekommunikasjon, meldinger (elektronisk post) og møter. Programmet kan i tillegg være integrert med andre programmer som er installert på noden, så som databaseprogrammer og editorer.

For å få noe særlig utbytte av et slikt nett, bør man være godt kjent med nodens operativsystem og eventuelle editorer og andre programmer som er integrert med kommunikasjonsprogrammet.

«EARN» er navnet på den europeiske delen av et nokså omfattende internasjonalt nett som understøttes av IBM. Dette og en rekke andre nett er primært ment for forskere og skal normalt ikke koste mer å bruke enn det den lokale tilknytningen fra bruker til node koster. Iversen redegjorde for en rekke detaljer, som jeg stort sett ikke skal gjengi her, om bruk av EARN-nettet og andre nett bl.a. angående adressering av noder og brukere.

Av spesielle ting jeg merket meg, var muligheten for tilnærmet direkte kontakt mellom enkeltbrukere. En slik kontakt kan selvsagt ikke konkurrere med telefonkontakt når det gjelder umiddelbar meningsutveksling, men man kan faktisk snakke nokså direkte med hverandre via skjerm og tastatur.

Den mer indirekte meldingstjenesten er på sin side brevvekslingen langt mer overlegen når det gjelder utveksling av skriftlige meddelelser.

En form for tilnærmet direkte meningsutveksling lar seg for øvrig også etablere gruppevis, ved at én enkelt node kjører et slags åpent-forum-program, der flere brukere fra ulike noder kan slippe til samtidig.

Ellers fikk vi inntrykk av at meldingstjenestene og de mer asynkrone møtene var det som ble mest brukt.

Mer eller mindre i tråd med Iversens foredrag redegjorde *Knut Hofland* fra NAVFs EDB-senter og *Mette Cathrine Jahr* fra Digital for sine erfaringer med bruk av data-nett - henholdsvis offentlige nett og interne nett i Digital's organisasjon.

Ifølge Hofland varierer de offentlige nettene i Norge i kvalitet, og de kan også være mer eller mindre kostbare å kjøre. Her fremhever Televerkets DATAPAK seg som driftssikkert og relativt rimelig. Universitetenes eget UNINET har vist seg både ustabil og lite brukervennlig. Det burde ifølge Hofland videreutvikles og følges opp.

Begge disse nettene er rene dataoverføringsnett uten de tilleggstenestene som f.eks. EARN-nettet tilbyr. En motsatt type kommunikasjonsmedium er programmer som tilbyr bl.a. meldingstjenester og møter, men som mangler dataoverføringsfunksjonen. Disse programmene er dermed bare tilgjengelige for brukere av én og samme maskin, men en bruker kan selvsagt befinne seg et helt annet sted enn den aktuelle maskinen og være innlogget via f.eks. UNINET. Programmet KOM (eng. COM) som går

på DEC-maskiner og som er installert på maskiner bl.a. i Oslo og Stockholm, er et eksempel på et velutviklet program av denne typen.

Jahr hadde mye pent å si om Digital's interne nett. Dette er av typen DEC-NET og tilbyr alle de typer tjenester som har vært omtalt hittil. For Jahr har nettet blitt helt uunnværlig i hennes daglige arbeid som bl.a. omfatter utstrakt kontakt med hennes medarbeidere utenfor Norge - i Europa og til dels i USA.

*Odd de Presno* som arbeider som freelance program/systemutvikler, journalist og foredragsholder, snakket om kommersielle data-nett for nyhetspublikasjon. Dette er nett der kommunikasjonen stort sett går bare én vei: fra nettets databaser til brukeren. Til gjengjeld inneholder databasene gjerne en svært mangeartet og omfattende informasjonsmengde - om enn mest myntet på næringslivet.

Når disse nettene i tillegg er nokså kostbare i bruk, er det om å gjøre for brukeren å finne raskest mulig frem til den informasjonen han eller hun ønsker. Til dette formålet er det utviklet et eget program QMODE for automatisk nett-tilkobling og informasjonssøking. Programmet unngår kostbare feiltastinger, foretar effektive unnvikelser ved linjebrudd eller andre feil ved forbindelsen, søker etter informasjon i ulike databaser etter kriterier som er satt opp på forhånd og er selvsagt betydelig raskere enn en menneskelig bruker. Programmet er nokså omfattende og delvis slutningsbasert.

### **Maskinassistert oversettelse**

*John Newton* som er ansatt ved Weidner Translation i London, snakket hovedsakelig om et programsystem for maskinassistert oversettelse - MicroCAT - som hans firma har utviklet.

«CAT» står for «Computer Assisted Translation», og det er altså ikke snakk om et fullautomatisert oversettelsessystem. Systemet gir et forslag til råoversettelse som brukeren senere korrigerer. Et slikt system egner seg godt for oversettelse av f.eks. teknisk dokumentasjon, brukerveiledninger, salgsdokumenter, forretningskorrespondanse og visse typer vitenskapelige tekster, men er sannsynligvis helt ubrukelig for oversettelse av normalprosa og skjønnlitteratur. Ifølge Newton er det vanskelig å tenke seg at man skal kunne utvikle systemer for fullautomatisert oversettelse.

Systemet er nokså omfattende og inneholder moduler bl.a. for syntaktisk analyse og parsing, redigering av leksikon og korrigering av programets rå-oversettelser. Til det siste formålet inneholder systemet en avansert editor med flere spesialfunksjoner så som ordombytting, tegnombytting og parallell visning av kilde teksten og oversettelsen på en delt skjerm.

I tillegg til de enkelte programmodulene inneholder systemet et kjerneleksikon med fra 9000 til 15.000 termer (avhengig av kilde språket). Dette

kan utvides og redigeres av bruker, og systemet gir også muligheten for opprettelse av et brukerdefinert spesialleksikon. Redigeringen av leksikon – innleggelsen av nye termer og regler eller endring av gamle – vil normalt utføres av oversetteren – gjerne i forbindelse med korrigeringen av de enkelte oversettelsene. Leksikonredigeringen er imidlertid en nokså komplisert prosess og stiller krav til brukerens lingvistiske kompetanse. Grunntermene i leksikon er lemmer, og homografvariantene genereres på grunnlag av regler. Leksikonet kan dessuten inneholde idiomer, dvs. faste uttrykk. Materiale av den typen systemet benyttes mest på – salgsdokumenter, teknisk dokumentasjon etc. – vil ofte inneholde mange, og til dels nokså lange, faste uttrykk.

Ifølge Newton vil systemet neppe være særlig egnet direkte som oversettelsessystem innenfor humanistisk forskning. Det vil derimot med fordel kunne brukes som pedagogisk verktøy innenfor språkundervisningen.

Newton's foredrag bar preg av at foredragsholderen passet godt til forsamlingen og var etter min mening seminarets klart beste foredrag.

### **Statusrapporter fra miljøene**

Jeg skal ikke gå nevneverdig inn på rapportene fra miljøene. Disse var nokså spekket med detaljert informasjon om forsknings- og undervisningsaktiviteter, personellrekruttering og utstyrsinvesteringer. Det generelle inntrykket var at aktivitetsnivået var nokså høyt – og økende – på alle disse områdene, og at humanistisk informasjonsteknologi er i framgang både når det gjelder kvalitet, omfang og anerkjennelse i og utenfor de humanistiske miljøer.

### **Generelt om det faglige utbyttet**

De utviklingsseminarene jeg selv har deltatt på, har variert både når det gjelder faglig kvalitet, tematisk bredde og arbeidsintensitet. I disse henseender må vel årets seminar sies å ligge omtrent midt på treet. Som nevnt var foredraget om datamaskinassistert oversettelse etter min mening det beste. Jeg hadde også mye utbytte av foredraget om nyere lagringsteknologi, mens de to hovedinnslagene om kommunikasjon og data-nett ble noe mindre utbytterikt.

*Asbjørn Brændeland er amanuensis ved HF-data, Universitetet i Oslo.*

## XVIII International Congress of Papyrology

*Espen S. Ore*

Kongressen ble arrangert i Aten 25.-31. mai 1986. Det var flere parallelle sesjoner, og de spente fra religionshistorie til bruk av edb innen papyrologi. Det var samlet ca. 250 deltagere fra 21 nasjoner, blant dem tre fra Norge. *Knut Kleve* og *Brynjulf Fosse* vakte spesielt stor oppmerksomhet da de presenterte de siste forfininger i sin metode for å rulle opp forkullede papyri (se HD 1-85). De kunne også fortelle at de nå hadde begynt å undervise personalet ved Nasjonalbiblioteket i Napoli, der de herculanensiske papyri oppbevares, i å bruke metoden.

*John F. Oates* og *William H. Willis* fra Duke University, North Carolina, opplyste om status for prosjektet *The Duke Data Bank of Documentary Papyri* (DDBDP). Formålet er å maskinregistrere *alle* utgitte dokumentpapyri (kontrakter, testamenter o.l.) og ostraka (f.eks. påskrevne potteskår) på gresk og latin. Til nå er alle aktuelle tekster utgitt etter 1966 (unntatt *Oxyrhynchus Papyri* som blir registrert i en databank ved Oxford University) skrevet inn. Omtrent halvparten er ferdig korrekturest. Fase 2, registrering av tekster utgitt i perioden 1925-66 er påbegynt, og den siste planlagte fasen vil bestå i å registrere tekster utgitt mellom 1891 og 1924.

De registrerte tekstene er nå tilgjengelige på magnetbånd. De er også tatt med i materialet som ligger på en CD-ROM-plate (se Øystein Reigems artikkel i dette nummer av HD) som i løpet av sommeren 1986 blir utgitt av *Thesaurus Linguae Graecae* (TLG, ved University of California, Irvine). På denne platen vil man også finne TLGs komplette materiale. Tekstene på platen har koder som er beregnet på å kunne utnyttes av den nye *Ibycus Personal Computer*. Denne maskinen er spesielt egnet for bruk innen klassisk filologi og beslektede fag. Den tillater flerspråklig tekstbehandling (gresk, hebraisk og koptisk) og viser de aktuelle alfabeter med komplette sett av diakritiske tegn.

*Deborah Hobson* fra York University, Ontario, demonstrerte et pilotprosjekt der data fra et dokumentarkiv fra 1. årh. ble analysert ved hjelp av *dBase II*. Hensikten med prosjektet var å vise hvorledes alminnelig tilgjengelig programvare kunne brukes til å utvide det potensialet for sosio-økonomiske historiestudier av gresk og romersk Egypt som finnes i dokumentpapyrussamlinger. I dette prosjektet ble det foretatt krysskoblinger mellom fordelingen av greske kontra egyptiske navn og hvilke roller de hadde i avtalene papyrene dokumenterte (f.eks. *jordutleier* kontra *leietaker*). *Hobson* innrømmet at dokumentsamlingen var relativt liten (ca. 100 dokumenter), og at det derfor var mulig at de samme



\*\*\*\*\* Papyri Data Base Report \*\*\*\*\*

File Transaction Type S E First Party #1 S E 2nd Party #1 Money Transact  
ed Document E T First Party #2 E T 2nd Party #2 Grammatikon  
Contract Type I H I H

REC. #	35							
Tebtunis	P.Mich.123r ii 43	m g Didumos	m e Harpaeus	24 dr.				
6,Theoth,19	daneion		f wife	6 ob.				
-----								
REC. #	70	arguriou dr. 90						
Tebtunis	P.Mich.123r iii 33	? e Tesenophis	m e Harepseus	90 dr.				
6,Theoth,26	daneion		f e wife (Theubastios)	4 dr.				
-----								
REC. #	83							
Tebtunis	P.Mich.123r iv 2	m e Papebubnis	m e Kleus	100 dr.				
6,Phaaphi,2	daneion		f wife	14 ob.				
-----								
REC. #	92	arguriou						
Tebtunis	P.Mich.123r iv 12	m e Orseus	m g Ptolemaios	26 dr.				
6,Phaaphi,4	daneion		f wife	8 ob.				
-----								
REC. #	94	arguriou						
Tebtunis	P.Mich.123r iv 14	m g Herodion	m e Psuphis	500 dr.				
6,Phaaphi,4	daneion		f e Theubastis	8 dr.				
-----								
REC. #	99	arguriou						
Tebtunis	P.Mich.123r iv 19	m e Mustharion	m e Harpaeus	100 dr.				
6,Phaaphi,5	daneion		f e wife (Thaesis)	6 dr.				
-----								
REC. #	108	arguriou						
Tebtunis	P.Mich.123r iv 31	m g Lusiachos	m e Peleesis	28 dr.				
6,Phaaphi,8	daneion		f wife	7 ob.				
-----								
REC. #	191	arguriou						
Tebtunis	P.Mich.123r vi 35	m e Harsisouchos	m e Onnaphis	60 dr.				
6,Hathyr,9	daneion		f wife	7 ob.				
-----								
REC. #	217	meriteas						
Tebtunis	P.Mich.123r vii 21	m e Kalloulhos	m e his sons	0				
6,Hathyr,19	homologia		f his wife	8 dr.				
-----								
REC. #	265	arguriou						
Tebtunis	P.Mich.123r viii 29	m e Petouris	m e Papatos	60 dr.				
6,Chorak,16	daneion		f wife	12 ob.				

Transaksjonsrapport generert i dBase II i pilotprosjektet beskrevet av Deborah Hobson.



*Under kongressen ble det arrangert en utflykt til Kap Sounion. Bildet viser Poseidon-tempelet.*

resultatene kunne oppnåes ved hjelp av tradisjonelle verktøy som pappesker og kartotekkort, men hun var selv overbevist om at verktøy a la dBase II representerte fremtiden.

*Giuseppe Tibiletti* fra Università Cattolica, Milano presenterte et bibliografisk prosjekt. Formålet er å lage en bibliografisk databank som spenner fra 1920 til i dag. Databanken skal i fremtiden bli løpende oppdatert. I dette prosjektet så han mange av de tradisjonelle problemene i bibliografisk databehandling: bruk av emneord og i så fall åpent kontra lukket vokabular, evt. automatisk uttrekking av emneord osv. Siden prosjektet er i en planleggingsfase, består det nåværende arbeidet først og fremst i å løse disse problemene.

Undertegnede la frem status for edb-arbeidet i det norske papyrusprosjektet (se HD 1-85). Som en del av Senterets videodiskprosjekt (se samme nr.) er det lagt inn fotografier av papyrustekster på platen. Det ble lagt frem en skisse som viste hvorledes vi tenker oss et system der fotografier, tegnede omriss av bokstaver, utgitte tekster og tekster under bearbeidelse kan lagres på de medier som er best egnet for de forskjellige behov. Det ble også poengtert at det antydde systemet avvek grunnleggende fra f.eks. Ibycus-prosjektets. Vårt system vil særlig bygge på programvare, mens Ibycus-maskinen er spesialutviklet for formålet.

## ICAME 7th

*Knut Hofland*

Den 7. «Conference on English Language Research on Computerized Corpora» ble holdt i Amsterdam 8.-11. juni 1986 med over 40 deltakere fra 8 land. Det var i år ingen deltakere fra USA, muligens pga. frykt for terror og radioaktivt nedfall. Deltakerne kom fra Nederland, Sverige, Canada, Storbritannia, Israel, Norge, Finland og Kina. Som ved tidligere konferanser var det en blanding av enkeltstående undersøkelser basert på korpusmateriale og foredrag som presenterte prosjekt eller delprosjekt innen analyse og behandling av korpusmateriale. I tillegg var det i år to «workshops» og en sesjon med demonstrasjoner.

*Anna-Brita Stenström* og *Karin Aijmer*, Lund, hadde studert bruk av hhv. *right* og *all right* og *Oh* og *Ah* i talespråkkorpus (London-Lund). *Nina Devons*, Jerusalem, hadde gjort en undersøkelse i bruk av *one's* og *oneself* i Brown Corpus. *Göran Kjellmer*, Göteborg, hadde studert kollokasjoner i Brown Corpus. Han delte disse inn i grupper etter frekvens, kompleksitet og lengde og undersøkte om det var forskjeller i bruk innen de 15 tekstkategorier som Brown Corpus består av.

I Lund fortsetter prosjektet TESS (Text Segmentation for Speech). En studerer her segmentering av engelsk tale i toneenheter og har som langsiktig mål å forbedre kvaliteten til tekst-til-tale konvertering. Siste år har *Jan Svartvik* revidert settet med ordklasseklassifikasjoner og bl.a. tilpasset disse til bruk av dBase III. Klassifikasjonene er delt inn i 3 nivåer og består av ca. 200 kategorier. I Lund har noen datastudenter laget program for grafisk presentasjon av prosodisk informasjon. *Mats Eeg-Olofson* har tilpasset programmet for statistisk merking av tekst til det nye kodesystemet. Dette er blitt testet på tekster fra Brown Corpus med en feilrate på 3-6 prosent. *Bengt Altenberg* har gjort studier av hvorledes en kan predikere tekstsegmentering til toneenheter basert på forskjellige former for skille mellom setningstyper og setningsledd.

*Gerry Knowles*, Lancaster, presenterte foredraget «Tone-grouping by Numbers». Prosjektet er et samarbeid med IBM i Winchester, og som i Lund er målet syntetisk tale fra tekst. En har bygget opp et lite korpus på ca. 30.000 ord. Hoveddelen av prosjektet går ut på å definere intonasjonsregler. Tittelen henspiller på at prosodisk informasjon kodes som deler (bit) i et binært tall. Ordene grupperes deretter trinnvis (3 nivåer) fra bunnen opp og fra venstre mot høyre.

*John Sinclair*, Birmingham, tok i foredraget «Upward and Downward Collocation» utgangspunkt i arbeid som er gjort ved York University i Toronto. For et gitt ord vil en, når en søker etter ord som er brukt

sammen med dette ordet og som har en lavere frekvens, finne ordsammenstillinger som har semantiske trekk (downward collocation, *eat food, eat meat*). Tilsvarende vil hyppigere forekommende ord enn utgangsordet gi kombinasjoner av syntaktisk karakter (upward collocation, *to eat, you eat*).

Fra miljøet i Nijmegen var det fire foredrag. *Jan Aarts* og *Nelleke Oostdijk* innledet med «Grammars in Corpus Analysis» som ga en oversikt over metodiske erfaringer som en har høstet i løpet av 10 års arbeid med korpuslingvistikk. *Theo van der Heuvel* fortsatte med «Interaction in Syntactic Corpus Analysis», der han diskuterte skillelinjen mellom databehandlerens og lingvistens domene innen korpuslingvistikk. *Hans van Halteren* presenterte LDB - Linguistic DataBase, som er et programsystem for å lagre og søke i et analysert korpus. Setningene lagres som trær med informasjon tilknyttet hver node. Programmet kjører på VAX utstyr og er tilgjengelig for vitenskapelig bruk. Hittil er det eksportert til 7 universiteter, bl.a. Oslo. *Pieter de Haan* hadde brukt LDB for å undersøke bruk av nominalfraser i forskjellige typer tekst. Frasene var delt inn i 4 grupper etter kompleksitet og det ble brukt log lineær analyse for å finne eventuelle signifikante forskjeller i kompleksitet mellom teksttypene.



*Deltakere i panelet «Corpus Linguistics in Relation to Other Areas of Research and Application.»*

*Lou Burnard*, Oxford, presenterte CAFS, «Content Adressable File System», som er en integrert del av utstyr fra den britiske maskinleverandøren ICL. CAFS gjør det mulig å søke sekvensielt gjennom en tekst på platelager med en hastighet på 1 Mb pr. sek. Søkeargumentene kan bygges opp som boolske kombinasjoner av tekst og merker som kan være lagt inn i teksten. For data i størrelse fra 10-30 Mb vil dette være et alternativ til tekstsøking. Andre firma har lignende system, men gjennombruddet vil først komme dersom dette kan implementeres i integrerte kretser på leverandøruavhengig utstyr, f.eks. i forbindelse med optisk plateteknologi.

*Nancy Behmore*, Canada, holdt foredraget «Word Classification Systems for Language Understanding». Hun hadde her gjort bruk av QUERY programmet (se nedenfor) for å trekke ut eksempler på *-ed* ord i Brown Corpus for videre klassifikasjon.

*Jeremy Clear*, Birmingham, introduserte en «teksttrål» i foredraget «Monitor Corpora». For å bygge opp store eksempelsamlinger til leksikologisk arbeid, filtrerer en ut belegg med lav dekning. Høyfrekvente ord registreres kun med frekvens. Føreløpig er mer enn 20 mill. løpende ord registrert.

*Ted Briscoe*, Lancaster, er tilknyttet Alvey-prosjektet «Natural Language Tools», et samarbeid mellom universitetene i Edinburgh, Cambridge og Lancaster. Briscoes prosjekt bygger opp en grammatikk og bruker LOB korpus som en test på dekningen av grammatikken. Av 63 eksempler er 58 analysert riktig pr. i dag.

*Stig Johansson*, Oslo, beskrev det grammatisk merkede LOB corpus som nå er tilgjengelig. Han gikk gjennom tekstformatene og konkordansen og presenterte en del analyser som er blitt gjort av tagfrekvenser, ordfrekvenser, tagkombinasjoner og ordkombinasjoner. Foredraget ble utarbeidet i samarbeid med undertegnede.

Fra Helsinki presenterte *Ossi Ihalainen*, *Merja Kytö* og *Matti Rissanen* to korpora som er under oppbygging der. Det ene er et engelsk diakronisk korpus fra år 800 til 1800. Hovedkorpuset skal bli ca. 1.5 mill. ord med et tilhørende tilleggskorpus, dvs. tekster som er til overs fra hovedkorpuset. Korpuset skal være ferdig til distribusjon i 1988. Det andre korpuset består av engelske dialekter basert på lydbåndopptak i 1970-årene. I øyeblikket er det ca. 130.000 ord, mens målet er 500.000. En del prøvetekster er analysert og lagt inn i dBase III.

*Jan de Jong* og *Pieter Masereeuw*, Amsterdam, holdt foredraget «A new implementation of the LSP Grammar». Dette er Naomi Sagers «Linguistic String Parser» som implementeres i språket PARSPAT, som er en parser-generator. *Eric Akkerman*, også fra Amsterdam, gav en statusrapport for ASCOT, «Automatic Scanning System for Corpus Oriented Tasks». En har her bl. a. bygget opp et codesystem for leksikon med utgangspunkt i LDOCE (Longman Dictionary).

En sesjon (workshop) var avsatt til foredrag/diskusjon som skulle belyse korpuslingvistikk i forhold til andre felt. *S. C. Dik*, Amsterdam, innledet om «Functional Grammar», *T. van Dijk*, Amsterdam, om «Discourse Analysis», *H. Kerkman*, Nijmegen, om «Experimental Psycholinguistics» og *L. Pols*, Amsterdam om «Speech Technology». Denne sesjonen ble ikke helt vellykket. De to første innledningene ble kun en presentasjon av fagfeltet uten at koblingen til korpuslingvistikk stod i fokus. De ble her for mye stoff i forhold til den avsatte tid. Innlegget om psykolingvistikk understreket hvor sentralt *ord* står i denne type forskning. I Nederland brukes en stor database, CELEX, som inneholder 110.000 lemma, 600.000 ordformer og 32 millioner løpende ord. Pols presenterte prosjekter innenfor taleprossesering i Nederland og EF (Esprit/Alvey). I Amsterdam brukes bl.a. parsergeneratoren PARSPAT til overføring fra tekst til fonemisk representasjon. Et tilsvarende prosjekt fins også i Nijmegen.

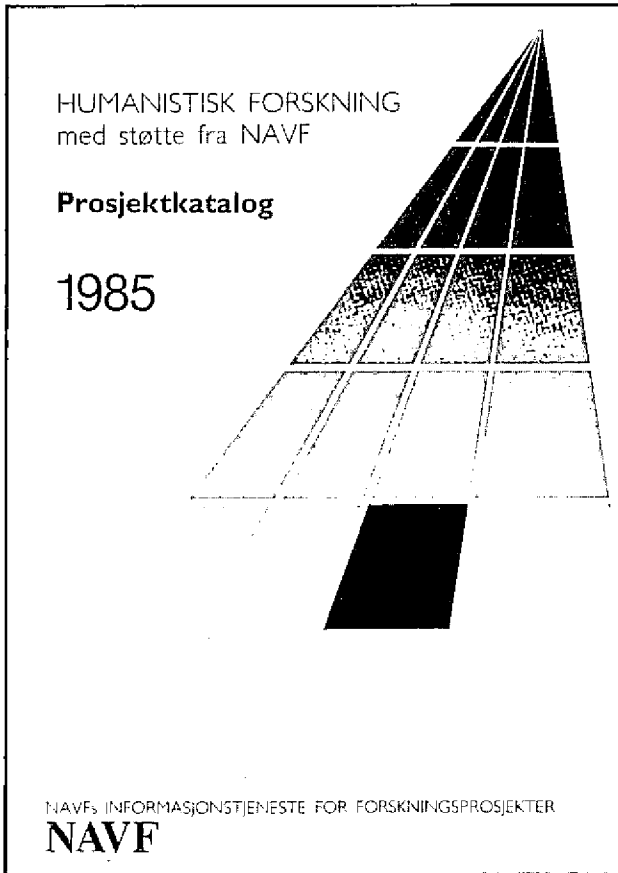
En kveldssesjon var avsatt til diskusjon om koordinering og utveksling av program. *Gert van der Steen*, Amsterdam, hadde laget et notat og innledet. Deltakerne presenterte deretter det maskin- og programutstyr som ble brukt og de ønsker en hadde, særlig på programvaresiden. Det ble diskutert hvorledes en utveksling kunne organiseres, hvilke medium som skulle brukes til transport (nettverk, diskett, magnetbånd) og typer program som skulle inngå i tjenesten. Videre ble det gjennomgått en standard for beskrivelse. Det ble også diskutert programmeringsspråk og operativsystemer uten at det ble satt noen restriksjoner her. Felles produksjon av en CD-ROM var også drøftet, men dette ble utsatt til neste år. Resultatet av sesjonen ble at NAVFs EDB-senter for humanistisk forskning tar på seg et koordinerings- og distribusjonsansvar for programvare for bruk i korpuslingvistikk i tillegg til den nåværende distribusjon av tekster. En vil i stor grad bruke nettverk som overføringsmedium og som et første skritt er det satt i gang en egen filtjener ved EARN noden i Bergen (FAFSRV at NOBERGEN), se egen melding i dette HD.

Under demonstrasjonene ble *Gert van der Steens* program QUERY presentert på forskjellige typer materiale. Dette er en svært avansert mønstergjenkjenner som kan brukes både på umerkede og grammatisk merkete korpus. Videre ble LDB i Nijmegen demonstrert. Det ble også demonstrert et søkeprogram for LDOCE og programmet for fonetisk transkripsjon av tekst.

Organisasjonkomiteen med *Willem Meijs* i spissen stod for en velorganisert konferanse. Den eneste innvending måtte være at programmet var svært tett og varte utover kvelden. Foredragene vil bli utgitt i bokform i høst. På konferansen ble det også nedsatt et rådgivende styre for ICAME. Neste konferanse vil bli holdt i Helsinki i mai eller august/september 1987.

# Nytt fra RHF/NAVF

Humanistisk forskning med støtte fra NAVF  
Prosjektkatalog 1985



NAVF's informasjonstjeneste for forskningsprosjekter har gitt ut en katalog (152 sider), som inneholder opplysninger om 188 igangværende forskningsprosjekter med støtte fra RHF/NAVF i 1985. Det gis også opplysninger om de publikasjoner som er skrevet på grunnlag av prosjektene.

Katalogen koster kr 120 og kan bestilles fra: *NAVF's informasjonstjeneste for forskningsprosjekter, c/o NAVF's EDB-senter for humanistisk forskning, Postboks 53, 5014 Bergen-Universitetet.*

## Utredning om humanistisk informasjonsteknologiforskning

### *Bakgrunn*

På bakgrunn av at regjeringen har utpekt informasjonsteknologi til et hovedinnsatsområde, har det fra flere hold (bl.a. fra Datapolitisk råd) blitt påpekt nødvendigheten av at anvendelser innen andre fagområder enn teknisk-naturvitenskapelig sektor får oppmerksomhet og tilstrekkelige midler. Betydningen av en fokusering på konsekvenser av kulturell og samfunnsmessig art er også blitt påpekt.

NAVF har et særlig ansvar for denne delen. Da området ikke ble tilgodesett i budsjettproposisjonen for 1986 besluttet NAVFs styre, innen egen ramme, å øremerke 1 mill. kr. til utredning og forskningsstøtte innen temaet «Konsekvenser og anvendelser av informasjonsteknologi innenfor humanistiske, medisinske og samfunnsvitenskapelige områder». Styret vedtok 10.12.85 at midlene skulle fordeles med 1/3 på hvert av de tre aktuelle fagrådene i NAVF for 1986.

Rådet for humanistisk forskning (RHF) ønsker å stimulere alle relevante humanistiske fagmiljøer til integrert innsats på området innenfor et program for humanistisk informasjonsteknologiforskning. Det ble i utgangspunktet ansett at *datalingvistikk*, *akustisk og perseptorisk fonetikk*, og *filosofi og logikk* med vekt på anvendt logikk og kunnskapsteoretiske og språkfilosofiske problemer kunne være aktuelle satsingsområder innenfor et slikt program. En mente videre at det burde være aktuelt å prioritere forskning om *kulturelle endringer* som følger av innføring av ny informasjonsteknologi.

### *Om utvalget*

RHF vedtok derfor på møte 13.12.85 at det skulle nedsettes et utvalg med representanter fra RHF og fra de fire universitetene som skulle kartlegge nåværende status innenfor programområdet og utforme strategier for det videre arbeidet. Det ble videre avsatt midler til oppstartning av ett prosjekt i 1986. Utvalget fikk følgende mandat:

«Å foreta en kartlegging av status innenfor humanistisk informasjonsteknologisk forskning i Norge ved universiteter og høyskoler og å legge opp en nasjonal handlingsplan for videre satsing».

Utvalget har hatt følgende sammensetning:

Professor *Helge Dyvik*, Institutt for fonetikk og lingvistikk, Universitetet i Bergen, formann

Amanuensis *Asbjørn Brændeland*, HF-data, Universitetet i Oslo

Førsteamanuensis *Cathrine Fabricius Hansen*, Germanistisk institutt Universitetet i Oslo (RHF)

Direktør *Jostein H. Hauge*, NAVFs EDB-senter for humanistisk forskning, Bergen (RHF)



Førsteamanuensis *Lars Hellan*, Lingvistisk institutt, Universitetet i Trondheim

Førsteamanuensis *Ove Lorentz*, Institutt for språk og litteratur, Universitetet i Tromsø

Førstekonsulent *Tore Ellingsen*, RHF, har vært sekretær for utvalget.

### *Faglig avgrensning*

Informasjonsteknologi er teknologi for lagring, overføring og bearbeidelse av informasjon; særlig aktuell er teknologi basert på elektronisk databehandling.

Det finnes tre hovedtyper av koblinger mellom informasjonsteknologi og humanistiske fag, med ulik grad av relevans for utvalgets mandat:

1. Humanistiske fag som leverandører av innsikt ved utviklingen av informasjonsteknologiske produkter;
2. Humanistiske studier av kulturelle konsekvenser av informasjonsteknologien i samfunnet;
3. Bruk av informasjonsteknologi innenfor humanistiske fag.

Utvalget så det slik at det nå var viktigst å klarlegge punktene 1 og 2 ovenfor.

Innføringsteknologiske hjelpemiddel har siden slutten av 60-årene vært brukt til å løse humanistiske fag- og forskningsoppgaver og denne virksomhet er etter hvert vel innarbeidet i en rekke fagmiljøer og dels godt dokumentert, bl.a. gjennom tidsskriftet *Humanistiske Data*.

Den etter hvert omfattende og til dels spesialiserte edb-erfaring som er opparbeidet gjennom datastøttet forskning, har i dag skapt grunnlaget for at de humanistiske fag nå også ønsker å markere sine selvstendige bidrag til den generelle informasjonsteknologien.

Sentralt i innstillingen står redegjørelser for hvordan humanistiske fag kan bli leverandører av innsikt til informasjonsvitenskapelig utviklingsarbeid og for forhold som tilsier at vi må ruste opp forskning omkring de kulturelle konsekvenser av informasjonsteknologien.

### *Humanistiske fag som leverandører av innsikt*

Mange av de sentrale problemene innenfor informasjonsteknologien, særlig innenfor det stadig mer aktuelle delområdet kunnskapsteknologi («kunstig intelligens»), er av humanistisk art. Ifølge utvalget blir informasjonsteknologien stadig mer integrert i kommunikasjon mellom mennesker, også ikke-kyndige brukere, og er i økende grad aktuell ved formidling av lagret informasjon på en brukervennlig måte. Mer avanserte systemer kan sies å simulere menneskelige kommunikasjonspartnere på ulike måter. Det er følgelig generelt et formål at informasjonsteknologien tilpasses en menneskelig kommunikasjonsform, en menneskelig måte å omgås symboler på.

Mer spesielt kan vi for det første tale om et humanistisk bidrag ved at

teorier og forskningsresultater fra humanistiske fag direkte legges til grunn for utviklingen av programvare. For det annet kan en tenke seg et bidrag i form av kritiske undersøkelser av den kunnskapsteknologiske forskningens forutsetninger om menneskelig kunnskap, forståelse og kommunikasjon, og dermed av grensene for dens muligheter.

Den første typen bidrag er aktuell ved automatisk analyse og syntese av naturlige språk, ved representasjon av kunnskap og ved modellering av slutningsprosesser. Språkrelaterte forskningsdisipliner står sentralt her. Formell syntaktisk, morfologisk og semantisk teori og prosessmodeller av språkanalyse (parsere) og språksyntese er aktuelle ved utvikling av systemer for språkforståelse og språkproduksjon og faller inn under emnet data lingvistikk. (Anvendelsesområder for slike systemer omfatter naturlig-språklig kommunikasjon mellom menneske og maskin i spørsmål-svar-systemer og ved oppbygging og bruk av ekspertsystemer; helt eller delvis automatisert oversettelse; systemer for avansert, evt. flerspråklig, tekstbehandling, informasjonssøking i naturlig-språklige tekstmasser, språkpedagogiske hjelpemidler, språkkontrollerende hjelpemidler for skribenter, m.m.).

Akustisk og perseptorisk fonetikk i kombinasjon med teknikker for signalanalyse og signalprosessering benyttes ved utvikling av systemer for talesyntese og talegjenkjenning. (Anvendelsesområdene omfatter bl.a. robotkontroll, hjelpemidler for språkhemmede, automatisk prosesskontroll i industrien, telefoniske informasjonstjenester og kontorautomasjon, f.eks. hjelpemidler ved korrekturlesning, og «talk-writers» som omformer muntlig diktat til skriftlig form).

Videre er formell logikk og logisk semantikk grunnlag for mye av betydningskomponenten i naturligspråklige systemer, og for resonnerende komponenter f.eks. i ekspertsystemer. Kunnskapsteoretiske arbeider er av betydning ved fastleggelsen av prinsippene for kunnskapsrepresentasjon i ekspertsystemer og databaser. Ferdigheter i humanistisk kontekst- og meningsanalyse er sentrale ved innsamling og fortolkning av selve kunnskapen.

I innstillingen blir det også omtalt hvordan forskning omkring billedspråk og billedforståelse er av relevans for forskning innen visuell kommunikasjon. Kunnskapsteori og teorier om forståelse, tenkning, kommunikasjon og kontekst kan gi bidrag til og kritikk av utviklingen av kunnskapsbaserte systemer.

### *Kulturelle konsekvenser av informasjonsteknologi*

Det heter bl.a. i innstillingen at utvikling av ny teknologi har alltid ført med seg endringer som går ut over teknologiens umiddelbare bruksammenhenger. Særlig iøynefallende er de kulturelle og samfunnsmessige endringer som oppstod ved overgangen fra manuell til maskinell produksjon i den industrielle revolusjon. Med den dominerende rolle informasjonsteknologien vil spille på mange livsområder, er det grunn til å tro at



*Filosofi kan være et aktuelt satsingsområde innenfor et program for humanistisk informasjonsteknologiforskning. Edvard Munch: Portrett av Friedrich Nietzsche. Thielska galleriet, Stockholm.*

også denne vil føre med seg slike endringer. Disse endringene utgjør et viktig forskningsfelt for en rekke humanistiske fag. Dette er dessuten et felt der humanister bør kunne samarbeide både med samfunnsvitere, jurister og informatikere.

Utvalget peker på at man innenfor samfunnsforskningen, historieforskningen og filosofien i nyere tid bl.a. har vært opptatt av hvorledes den teknologiske utviklingen påvirker enkeltmenneskets opplevelse av seg selv – særlig i forhold til produksjonslivet. Man har bl.a. ment at den teknologiske utviklingen fører til økende *fremmedgjøring og atomisering av mellommenneskelige relasjoner*. Hvorvidt og i hvilken grad informasjonsteknologien vil bidra til å forsterke eller motvirke denne tendensen, er imidlertid et ubesvart spørsmål.

Et spørsmål som kan sees i sammenheng med dette, er etter utvalgets mening hvorvidt den teknologiske utviklingen også påvirker enkeltmenneskets opplevelse av eget og andres *ansvar* for de prosesser det deltar i. Etter hvert som arbeidsprosessene i økende grad kompliseres og automatiseres, vil den enkelte deltagers innsikt i prosessene lett bli mangelfull. Dette aksentueres særlig tydelig i forhold til bruken av kunnskapsbaserte systemer. Disse skal simulere beslutningstagere, og kan bli tillagt en uberettiget selvstendighet i beslutningsprosessene.

Bruk av informasjonsteknologi har alt hatt en åpenbar innvirkning på vår språkbruk hevder utvalget – i alle fall hva angår vokabular. Det er en viktig oppgave for humanistisk forskning å studere hvor dypt denne virkningen går, bl.a. når det gjelder vokabularet i de enkelte nasjonale språk, hos oss i første rekke i norsk språk.

Menneskelig kommunikasjon er styrt av normer. Informasjonsteknologien vil sannsynligvis påvirke disse normene etter hvert som den veves inn i menneskelig kommunikasjon, og det er liten grunn til å tro at påvirkningen vil bli positiv. Et datasystem som simulerer deler av menneskelig kommunikasjon, vil alltid være mindreverdig i forhold til menneskelige kommunikasjonspartnere, og bør derfor ikke være normdannende. Ikke desto mindre vil ulike systemer lett kunne bli det – det vil f.eks. være nærliggende å be medarbeidere å holde seg til et syntaktisk repertoar og et ordforråd i sine skriftlige produkter som firmaets automatiserte oversettelsessystem klarer å tolke.

Utvalget kommer også inn på hvordan informasjonsteknologien også får en stadig viktigere rolle som metafor: Dens kompleksitet og generelle anvendelighet gjør den nærliggende som metafor ved nesten ethvert fenomen som involverer kompliserte prosesser. Dette betyr at den samtidig blir styrende for vår forståelse av disse fenomenene. Den datamaskin-analøge forståelsesmåten er ofte effektiv, og utelukker derfor desto mer effektivt alternative forståelsesmåter. Eksempler på dette opptrer ikke minst i ulike vitenskaper, kanskje særlig slike som savner veletablerte paradigmer. Således søker viktige retninger innenfor psykologi og lingvi-

stikk å forstå menneskelige evner ved hjelp av begreper hentet fra informasjonsprosesseringsdomene.

Generelt reiser utviklingen av ny teknologi et spørsmål om forholdet mellom det mulige og det ønskelige: Hva *kan* vi bruke våre kunnskaper til, og hva *bør* vi bruke dem til? I humaniora må dette spørsmålet stilles på en mer grunnleggende måte enn f.eks. i informatikken og i samfunnsvitenskapen, dvs. det ønskelige må relateres til de mest grunnleggende menneskelige verdier. Hvordan konsekvensene av vår kunnskapsanvendelse kan bestemmes, og hva som er de mest grunnleggende menneskelige verdier er et moralfilosofisk spørsmål. I og med at informasjonsteknologien griper såvidt direkte inn i mange sider ved menneskets liv, synes nettopp denne å utgjøre et særlig interessant forskningsfelt for moralfilosofien, mener utvalget.

### *Virksomheten i dag*

I innstillingen gis det en detaljert oversikt (m/vedlegg) over relevant forsknings- og utviklingsarbeid ved instituttene innenfor datalingvistikk, fonetikk og taleprosesserings, anvendt logikk, bildeanalyse og grunnlagsproblemer i kunnskapsteknologien. Prosjekter som i dag bruker kunnskapsteknologiske metoder i humanistisk forskning blir også omtalt sammen med de undervisningstilbud som i dag tilbys innenfor de enkelte temaområder.

Av oversikten fremgår det at de fire universitetene har ulike tyngdepunkter i forskningsvirksomheten. På områdene datalingvistikk og taleprosesserings er Bergen og Trondheim kommet lengst, mens Oslo dominerer på området anvendt logikk, og vel også i forskningen omkring grunnlagsproblemer i kunnskapsteknologien. Det er ikke uten videre rimelig å la denne fordelingen være normgivende for de videre utbyggingssplanene. Det er tale om små miljøer og store forskningsfelt, og de ligger i den naturlige forlengelse av utviklingen innenfor veletablerte fag som lingvistikk, fonetikk og epistemologi. Alt dette taler imot en hardhendt geografisk fordeling av innsatsområder på dette stadium: Det er ønskelig at flere miljøer arbeider med beslektede problemer.

Alle fire universiteter har, eller er i ferd med å få, studietilbud på grunnnivå innenfor datalingvistikk, riktignok av svært ulikt omfang, og til dels også innenfor andre humaniora-relevante edb-emner og logikk-relevante emner. I hvert fall i Bergen kan man regne med et hovedfagstilbud i datalingvistikk i overskuelig fremtid, med det nye professoratet i faget, og med datalingvistikk som et av universitetets innsatsområder. Men der er et klart behov for å utbygge studietilbudene på viderekomment nivå ved flere av universitetene.

Skal det være tilrådelig med forsterket satsing på edb-relatert forskning og undervisning, må imidlertid de respektive fagmiljøer rustes kraftig opp utstyrmessig, både med basisutstyr og spesialisert edb-utstyr for ulike faglige anvendelser.

Det mest løynefallende problem ved utbyggingen av humanistisk informasjonsteknologiforskning i Norge er mangelen på kvalifiserte personer. Siden mye av denne forskningen kan være industrielt interessant, kan finansiering av forskningsprosjekter i enkelte tilfelle være et mindre problem. Dette tilsier at kompetanseoppbygging prioriteres foran stort anlagte forskningsprosjekter i første fase.

#### *Innsatsområder og virkemidler*

Som en konklusjon på sin behovsanalyse og sin virksomhetsbeskrivelse foreslår utvalget et fire-leddet nasjonalt forskerprogram for humanistisk informasjonsteknologi innenfor en 7-års innsatsramme.

Utvalget foreslår følgende innsatsområder:

- Tekstforståelse og informasjonssøking i naturligspråklig tekst
- Akustisk og perseptorisk fonetikk og taleprosessering
- Billedanalyse
- Kunnskapsteknologiens grunnlagsproblemer
- Kulturelle og verdimeslige sider ved informasjonsteknologi

I innstillingen blir innholdet i et forskerprogram for de respektive felt gjennomgått.

For å nå de oppstilte mål forutsetter utvalget at det alt fra 1986 av blir iverksatt en rekke ulike, men gjensidig kompletterende, tiltak. Disse sikter i en sum mot å bygge ut undervisningstilbudet innenfor de oppstilte felt, øke rekrutteringen, styrke samarbeid og internasjonal kontakt, skaffe til veie høvelig vitenskapelig utstyr og stimulere forskningen gjennom direkte økonomisk støtte.

Utvalget har sett det som en viktig oppgave å legge frem så klare tilrådninger på hvert av punktene at de direkte kan legges til grunn ved en forskningspolitisk drøfting av en nasjonal handlingsplan.

Utvalget er imidlertid klar over at enkelte av forslagene forutsetter ytterligere utredning eller bør ses i sammenheng med hverandre. For å få en effektiv nasjonal koordinering foreslår utvalget at det opprettes en styringsgruppe for et program for humanistisk informasjonsteknologi.

Av de virkemidler som utvalget foreslår nevnes:

- opprettelse av nye stillinger
- forfinansiering fra NAVF av amanuensisstillinger
- videreutdanning av vitenskapelig tilsatte
- bruk av limelærere
- koordinert bruk av ulike institusjoners fagtilbud
- student/utdannings-/dr.art. stipend
- internasjonalt samarbeid
- investering i vitenskapelig utstyr
- øremerkede prosjektbevilgninger

Utvalget har utarbeidet en økonomisk plan for investeringer innenfor

de ulike tiltak med en innsatsramme på ca. 25 mill. over en sju-års periode.

#### *Videre arbeid med innstillingen*

Rådet for humanistisk forskning behandlet saken på sitt møte i mai i år og vedtok bl.a. å holde et felles kontakt- og informasjonsmøte med aktuelle institusjoner om innstillingen i august.

I neste nummer av *Humanistiske Data* vil vi redegjøre nærmere for de drøftinger som nå føres mellom RHF og universitetene med tanke på å legge grunnlaget for et koordinert handlingsprogram for humanistisk informasjonsteknologi.

*Jostein H. Hauge*

## **Fra bokhylla**

**Danell, Christina (red.): Demografiska databasen vid Umeå universitet. En presentation av verksamheten. Umeå, 1985. 223 s.**

Demografiska databasen har utgitt en bearbejdet og oppdatert utgave av beskrivelsen av virksomheten som kom ut for første gang i 1980. Innholdet favner om bl.a. databasens etablering og organisasjon; det svenske kirkebokmaterialet; databasens ulike samlinger; systemløsninger; arbeidsprosessen fra kilde til database; uttaksmuligheter, standardprodukter og dokumentasjon; forsøk med innsamling av statistikk og såkalt parallelt materiale; forskertjenesten; og eksempler på forskningsmuligheter i databasens materiale.

**Walter B. Hewlett/Eleanor Selfridge-Field (eds.): Directory of Computer Assisted Research in Musicology. 86 s.**

Årets utgave av *Directory of Computer Assisted Research in Musicology* fokuserer på pågående aktiviteter innenfor programevareutvikling og -anvendelser innen musikologi. Oversikten omfatter dels nyoppstartede prosjekter, dels prosjekter omtalt i 1985-utgaven (se HD 3-85). I tillegg inneholder publikasjonen bl.a. en artikkel om trykking av musikk via edb.

Hefteet koster \$5 og fås ved henvendelse til: *Center for Computer Assisted Research in the Humanities, 525 Middlefield Road, Suite 120, Menlo Park, California 94025, USA.*

**Madsen, Torsten: Numerisk dataanalyse for arkæologer. Institutt for forhistorisk arkæologi, Aarhus Universitet, 1985. 216 s.**

Ifølge Torsten Madsen er formålet med *Numerisk dataanalyse for arkæologer* å gi en introduksjon til den beskrivende og struktursøkende statistikk, med spesielt henblikk på anvendelse innenfor arkeologi. Presentasjonen spenner fra beskrivelser av en enkelt variabel til kompliserte multivariable datareducerende metoder. Vekten legges på det praktiske bruksaspektet, hvordan de statistiske metodene fungerer, hva de kan og ikke kan brukes til, hvordan de tolkes, og hvordan man kan unngå feiltolkninger.

Boka gir bl.a. også en gjennomgang av noen konkrete analyseprogrammer og de utskriftene som er resultatet av disse, i tillegg til ulike arkeologiske eksempler under anvendelse av de aktuelle metodene.

**Price, Anne Jamieson: A Survey of Computer Use in Projects Sponsored by the National Endowment for the Humanities. The Office of Scholarly Communication and Technology, Washington D.C., 1986. 36 s.**

The Office of Scholarly Communication and Technology har foretatt en spørreundersøkelse om bruk av edb blant 53 humanistiske forskere som fikk økonomisk støtte fra the National Endowment for the Humanities i 1983 og 1984. Den resulterende rapporten viser at humanister i høy grad benytter seg av edb-metoder – og at de er stort sett fornøyd med den tilgjengelige maskin- og programvaren. Noen hovedtrekk fra rapporten:

«Bruk av edb» omfatter en rekke anvendelser fra tekstbehandling til en kombinasjon av PD'er og stormaskin med kompliserte spesialskrevne programmer. Omtrent halvparten av prosjektene hadde samlinger av brev og tekster som forskningsgjenstander – resten dreier seg om ordbøker, leksikon, indekser og databaser. En rekke ulike språk ble brukt i prosjektene, inkludert gresk, hebraisk, arabisk og indianerspråk.

De fleste forskere brukte datamaskiner for tre eller fire formål, som tekstbehandling, analyse av språk, temaer eller semantikk, indeksering av tekst og registrering/gjenfinning av informasjon.

Forskerne benyttet seg av en rekke ulike typer maskiner. Mer enn halvparten fikk edb-assistanse fra edb-sentrene ved sine respektive institusjoner og fra mer erfarne kollegaer. Dokumentasjonen av både maskin- og programvare ble ansett av et flertall for å være god.

Flertallet av forskerne hadde planer om å ta i bruk elektronisk publisering, mens bare 1/3 hadde tilgang til et eksternt datanettverk.

Rapporten fås ved henvendelse til: *Office of Scholarly Communication and Technology, 1717 Massachusetts Avenue NW, Suite 401, Washington D.C. 20036, USA.*



# MELDINGER

## Artikkelsamling i humanistisk databehandling

NAVFs EDB-senter for humanistisk forskning har gitt ut en artikkelsamling som gir en innføring i ulike emner med tilknytning til edb-bruk i humaniora. Artikkelsamlingen er resultat av et faglig og redaksjonelt samarbeid med edb-tjenestene ved HF-fakultetene ved universitetene i Bergen, Oslo og Trondheim.

Samlingen inneholder følgende artikler:

Gunnar Thorvaldsen: Edb-anvendelser i historisk forskning

Arvid O. Vollsnes: Norsk musikkvitenskap og edb

Stig Welinder: Edb og arkeologi

Jon Birger Østby: Bruk av edb ved kulturhistoriske museer

Espen S. Ore: Edb i idéfag og klassisk filologi

Hege Brit Randsborg: Edb i arkivverket

Eirik Lien: De egenskapene edb har som gjør at det kan være et tjenlig hjelpemiddel for humanistisk forskning

Ivar Fønnes: Den humanistiske forskers møte med edb: Krav, problemer, tilpasninger

Roald Skarsten: Hvilken sammenheng er der mellom tradisjonelle metoder og metoder basert på bruk av edb

Knut Hofland: Et programeksempel

Div. forfattere: Nasjonal virksomhet

Kristin Natvig/Espen S. Ore: Internasjonal virksomhet Espen S. Ore: Juridiske aspekter

Del 2 av artikkelsamlingen skal utgis i løpet av høsten.

Artikkelsamlingen koster kr 50,- og kan kjøpes i universitetsbokhandlene i Bergen og Oslo. Den kan også bestilles direkte fra NAVFs EDB-senter (porto kommer i tillegg).

## Telemøte for humanistisk databehandling

Det er nå opprettet et møte, *HUMKOM*, på konferansesystemet KOM ved UiO. Møtet er i første omgang tenkt som et åpent forum der deltakerne kan ta opp problemer, komme med gode råd osv. Dersom det kommer frem ønske om det, kan det senere opprettes møter for særinteresser/spesielle fag.

For å delta må man være registrert som bruker på DEC-10 ved UiO. De som er tilknyttet UiO, registrerer seg gjennom sitt institutt. Personer fra andre institusjoner kan kontakte *Baard M. Johannesen*, tlf. (02) 45 58 49, eller henvende seg skriftlig til *USE (Universitetets Sentrale EDB-tjeneste)*, P.b. 1059 Blindern, 0316 Oslo 3.

Når man har logget seg inn på DEC-10, starter man konferansesystemet med kommandoen R PUB:KOM. Man blir så bedt om å oppgi navn, og som ny bruker, et personlig passord. Deretter gir man kommandoen GÅ HUMKOM. For øvrig vises det til den hjelp man kan få fra programmet og til brukerveiledningen for KOM-systemet. (Før man går inn i KOM kan man gi kommandoen HELP KOM.)

## Filtjener ved UiB

Som et resultat av ICAME 7th konferansen (se rapport i dette HD) er det blitt laget en filtjener ved EARN datanettnoden i Bergen. Denne skal være en hjelp i formidlingsarbeidet av programvare til bruk innen korpuslingvistikk og tilgrensende områder. Brukere som er tilknyttet EARN kan sende en interaktiv melding HELP til FAFSRV at NOBERGEN. Brukere tilknyttet andre nettverk kan sende brev med *Subject: Help* til FAFSRVNOBERGEN.EARN (blir foreløpig lest manuelt). Bidrag til programvarebiblioteket tas imot med takk!

## Klassisk gresk litteratur på CD-ROM

Forskere ved Brown University, USA er i ferd med å overføre ca. 250 Mb klassisk gresk litteratur til én enkelt CD-ROM. Platen skal også inneholde en indeks på ca. 150 Mb. Tekstsamlingen omfatter alle kjente verk av 178 forfattere, fra Homers tid til ca. 600 e.Kr.

Gjenfinningssystemet for platen blir laget for IBM RT PC, og den greske teksten skal behandles av Interleaf dokumentbehandlingssystem for RT'en, som skal modifiseres for det gammelgreske alfabetet.

For å lette tilgjengeligheten til systemet, blir det skrevet et terminal-program for Macintosh. Med dette programmet kan Macintosh-maskiner fungere som terminaler på IBM RT og gi en nøyaktig gjengivelse av den greske teksten. CD-ROM-systemet skal ferdigstilles i løpet av sommeren 1986.

## Elektronisk publisering for nybegynnere

Ved Rochester Institute of Technology (RIT) i USA samarbeider eksperter i edb og typografi om et elektronisk publiseringssystem for hjemme- og kontorbruk. Utgangspunktet på maskinsiden er Xerox arbeidsstasjoner, og systemet skal bestå av menyer som muliggjør valg av sidetype (tittelside, kapittelstart, brødtekst osv.) og format (bok, brosjyre osv.). Systemet blir tilgjengelig fra 1987, og vil kunne brukes av alle som har tilgang til en PC.

(ACH Newsletter, Summer 1986)

## Aktuelle tidsskrifter

### Archaeological Computing Newsletter

*Archaeological Computing Newsletter* inneholder artikler og rapporter av relevans for edb-anvendelser i arkeologi og arkeologiske museer, inkludert stoff om program- og maskinvare. I tillegg fins det i hvert nummer en «oppslagstavle» for meldinger om kommende begivenheter, kommunikasjon mellom leserne osv.

Meldingsbladet utgis hvert kvartal og koster £4 i abonnement. Adresse: *Archaeological Computing Newsletter, Department of Computing, North Staffs Polytechnic, Blackheath Lane, Stafford ST18 0AD, United Kingdom.*

## Nytt i biblioteket

Abercrombie, John R.: Computer programs for literary analysis. Philadelphia, Pa: University of Pennsylvania Press, 1984. 203 s.

Hellberg, Staffan: The morphology of present-day Swedish. Word-inflection, word-formation. Basic dictionary. Stockholm: Almqvist & Wiksell International, 1978. 130 s.

Höglund, Lars/Persson, Olle: Information och kunskap. Informationsförsörjning – forskning och policyfrågor. Umeå: INUM, 1985. 164 s.

Jahr, Ernst Håkon/Lorentz, Ove: Morfologi/Morphology. Oslo: Novus, 1985. 182 s.

Thorvaldsen, Gunnar: Lenking av historiske individdata. Eksemplar nr. 1. Tromsø: 1986. 48 s.

Tottie, Gunnel/Bäcklund, Ingegerd: English in speech and writing. A symposium. October 5, 1984. Uppsala universitet, 1986. 204 s.

Vurdering av data-administrasjonssystemer for PC-maskiner. Institutt for informasjonsvitenskap, Universitetet i Bergen, 1985. 87 s.

## Konferanser

### Konferanse om nordisk forskning i tekstforståelse og informasjonssøking

Denne konferansen arrangeres av Institutionen för Lingvistik ved Stockholms Universitet 8.-9. desember 1986 på Säby ved Stockholm. Konferansens formål er i første rekke å planlegge framtidige nordiske forskningsprosjekter innenfor områdene datamaskinbasert tekstforståelse og informasjonssøking. Konferansen er først og fremst beregnet på representanter fra universiteter i Norden som er i gang med eller planlegger forskning på disse og nærliggende områder. Det forventes at representantene gir en 20 minutters presentasjon av sin forskning. Også andre interesserte kan delta i konferansen.

Påmelding innen 1. oktober til: *Kari Fraurud, Institutionen för Lingvistik, Stockholms Universitet, 106 91 Stockholm, Sverige.*

### Computers and Teaching in the Humanities

En konferanse om edb og undervisning i humanistiske fag skal holdes ved University of Southampton, 10.-11. april 1987. Spørsmål som skal belyses ved konferansen, er: hvilke aspekter ved edb passer for undervisning i humaniora under hovedfagsnivå? hva bør være de underliggende prinsippene for undervisning i edb? hvilke praktiske hensyn må tas ved utvelgelse og anskaffelse av maskin- og programvare?

Konferansen skal bestå av en blanding av plenumssesjoner om spørsmål av generell interesse, og mindre formelle, spesialiserte sesjoner om enkelte emneområder og edb-verktøy og -teknikker. Det skal i tillegg holdes fem sesjoner for deltakere som ikke har praktisk edb-erfaring.

De spesialiserte sesjonene skal omhandle de fleste humanistiske fag i tillegg til maskin- og programvare, databaser, programmering, kunstig intelligens og logisk programmering, biblioteksfag, og integrerte pakker.

Det endelige programmet fås ved henvendelse til konferansearrangøren: *May Katzen, Office for Humanities Communication, University of Leicester, Leicester LE1 7RH, England.*

# SUMMARY

## **Nordisk edb-konferanse for kunst- og kulturhistoriske museer**

### **Nordic computing conference for museums of art and cultural history**

In June the Centre arranged a Nordic computing conference for museums of art and cultural history. According to Karl Ragnar Gjertsen, Manager of Nordmøre Museum, the conference demonstrated the importance of co-operation between museums, particularly at the national level, to the preparation of functional computing systems. In addition, the conference revealed the potential for museums of the technical possibilities for combining data and images.

The first day of the conference dealt with the current status of computing in museums, the second day was reserved for technical issues and demonstrations, and the subjects of the third day's program were future projects and strategies.

Along with Gjertsen's report, abstracts of each speaker's papers are provided. Göran Bergengren presented computing systems at the Nordic Museum, Stockholm, with particular attention to the transferral of large amounts of data from a manual to an automatic system. Jon Birger Østby of the Norwegian Folk Museum gave an account of experiences of work on nomenclature, which poses a particular problem when preparing data for computer entry. Rune Hermanson's paper dealt with computing activities at The Cultural Heritage in Falun, Sweden, with emphasis on a combined information system for object data and pictures. Knud Holm and Carsten U. Larsen spoke on the use of computing at the National Museum in Copenhagen - Holm on present and planned applications under the auspices of the Department of Conservation, Larsen on the museum's future documentation project. Magne Velure, Director of the Sandvig Collection, discussed the gap between ideal and reality in connection with object data entry.

A general overview of computing in Danish museums was given by Ole Thamdrup from the Danish Museum Association. Icelandic museums have just started to take computing into use, reported Gudmundur Olafsson of the National Museum of Iceland. The state of the art in Sweden was presented by Göran Bergengren, who gave details of computer applications at 8 museums of art and cultural history.

## **Status for edb-arbeidet i de norske kunst- og kulturhistoriske museer**

### **Computing in the Norwegian museums of art and cultural history**

On April 1st a two-year computing project for museums of art and cultural history was concluded. The project was carried out by Jon Birger Østby of the Norwegian Folk Museum in collaboration with the Norwegian Computing Centre for the Humanities. Østby reports that the project aimed at creating a joint

system for the entry and storage of source data at all museums. Previously, the association of Norwegian Museums of Art and Cultural History had designed a catalog card for the entry of such data on computers.

In the course of the project, computing equipment suitable to museum use has been assessed. Personal computers in a network are recommended. Data entry systems for objects and photographs have been created in the data base programs DataStar and DataFlex. A program has been written for conversion from DataStar to DataFlex.

The Norwegian art museums have created their own system for data entry in dBaseII, apart from the City Museum of Oslo, which is currently testing an adaptation of the data entry form used by the museums of cultural history.

In connection with the project a trial data base for objects from several museums has been established, in order to gain experience with the organization and management of such a data base. The text retrieval system SIFT has been used, with good results. However, a fair knowledge of the contents of the data base is necessary and SIFT requires much storage space.

An initiative has now been taken to obtain funding for a joint computing service for museums. Such a service will co-ordinate computer applications, organize training, give advice on the purchase of equipment, manage a joint data base, assess new products, and represent the museums in international co-operative ventures.

### **CD-ROM – en halv gigabyte i lommen**

#### **CD-ROM – half a gigabyte in one pocket**

Senior Computing Officer Øystein Reigem of the Centre starts this article by explaining the principles of Compact Disc technology and giving a general introduction to media for optical storage, including a rundown of the various types of optical disks.

The storage of different kinds of information on CD-ROM is dealt with in the article. Reigem also presents attempts at standardizing the technology and discusses the issue of compatibility. Equipment necessary for CD-ROM technology is outlined, along with the production of a disk and necessary considerations when preparing data and programs for a CD-ROM. In conclusion, Reigem gives examples of CD-ROM applications, with emphasis on the technology's potential for the humanities.

### **Edb-utbygging og edb-virksomhet i det offentlige arkivverk**

#### **Computer activities in public archives**

Ivar Fønnes, previously Section Manager at the National Archives of Norway, gives an overview of the development of the use of computers in public archives in the course of the last three years, and reviews future developments.

The Research Council's research fellowship in automated archive information (see HD 1-86) formed a central part of the development of independent computer applications in Norwegian public archives (the National Archives and seven

regional state archives). Parallel with this work activities have been initiated in connection with the public administration's use of computers. A separate Computing Section in the National Archives was set up in 1983, and computing equipment has been purchased by each institution.

Fonnes describes the general tasks of the National Archives' Computing Section and outlines the hardware and software taken into use at the public archives. Much emphasis has been put on the training of archives staff in computer applications.

Fonnes gives a detailed description of the computer activities of the public archives, which consist of developing systems for securing and storing public archive material for posterity, user services (including archive information) and administrative tasks.

In Fonnes' view, optical storage media and new developments in computer communication give rise to interesting perspectives for the future of the public archives sector.

### **Project Emperor-I – kinesiske skatter på videoplate**

### **Project Emperor-I – Chinese treasures on videodisc**

«The First Emperor of China» is the title of a pair of videodiscs which present and interpret archaeological finds from the Qin Dynasty. These finds include the famous terracotta warriors and horses excavated from Qin's burial grounds.

The Director of «Project Emperor-I» is Dr. Ching-chih Chen of the Graduate School of Library & Information Science, Simmons College, Boston. Dr. Chen visited the Centre in April, where she gave a demonstration of the videodiscs and spoke on the project.

The first videodisc contains 200 sequences of film and video and 4000 slides, together with two 80 minute soundtracks in English and Chinese and musical interludes. The other videodisc consists of interviews with the world's ten foremost experts on the Qin dynasty. The project's data base – still under development – will be partly hierarchical, and partly relational. The project staff will develop courseware at several levels: for the general public, students, and researchers.

The project has been awarded \$20.000 from the National Endowment for the Humanities and more than \$56.000 from Simmons College. It was started on October 1st, 1984 and will be concluded later this year.

### **Training the Information Researcher for the Future**

This is the title of an Anglo-Nordic seminar arranged in Stockholm in April by The British Library Research and Development Department and the Nordic Council for Scientific Information and Research Libraries.

Director Jostein H. Hauge of the Centre reports that in a number of papers and discussions a definition of the research field «Information Science» was sought. Another central theme was the nature and academic level of future training in library science. The need for making better use of funds and personell

in the Nordic countries was also debated in several papers. Hauge gives a detailed account of the major speakers' papers.

In a separate session experiences of information services aimed at private industry were discussed. In the closing session several concrete suggestions were made for collaboration between both the Nordic countries and these countries and the UK.

#### **Utviklingsseminaret på Ustaoset april 1986**

##### **Developmental seminar at Ustaoset April 1986**

This year's developmental seminar for computing staff within the humanities was arranged jointly by the Centre and the computing service for the humanities at the University of Bergen. Senior Lecturer Asbjørn Brændeland of the University of Oslo reports that the seminar had three main topics: new storage technology, computer communication and international networks, and computer-assisted translation. In addition, status reports were given on research, teaching, personnel and equipment at each institution represented at the seminar.

Papers were presented on the following topics: optical disks, the integration of microcomputers and drivers for optical disks, and future developments within storage technology; the EARN network; experiences with using public networks and the internal network of the Digital Corp.; commercial networks for news publication; and Weidner Translation's computer-assisted translation system MicroCAT.

#### **XVIII International Congress of Papyrology**

Senior Computing Officer Espen S. Ore of the Centre reports that this congress was held in Athens in May for 250 participants from 21 countries. Papers given in the session on computing and papyrology are summerized.

Knut Kleve and Brynjulf Fosse, Norway, presented the latest refinements of their method for unrolling documentary papyri (see HD1-85). John F. Oates and William H. Willis, USA, gave a status report on the Duke Data Bank of Documentary Papyri. Deborah Hobson, Canada, demonstrated a pilot project in which data from a document archive of the first century is analyzed with the aid of dBasell. Guiseppe Tibiletti, Italy, presented a project aiming at creating a bibliographic data base from 1920 to the present.

Ore himself described the Centre's plans for experimenting with a system for optical storage of papyrus texts treated in various ways for research purposes.

#### **ICAME 7th**

The 7th «Conference on English Language Research on Computerized Corpora» was held in Amsterdam in June for 40 participants from 8 countries. Senior Computing Officer Knut Hoffland of the Centre reports that as previously the conference consisted of papers presenting both investigations based on corpus material and projects dealing with the analysis of such material. Hoffland gives details of each paper.



This year's conference also featured a demonstration session and two workshops. One workshop dealt with corpus linguistics in relation to other fields of research - in the other a discussion took place of the co-ordination and exchange of computer programs.

The proceedings of the conference will be published this autumn. The next ICAME conference will be held in Helsinki in 1987.

## **Nytt fra RHF/NAVF**

### **News from the Council for Research in the Humanities**

The Council's information service for research projects has published a catalog of data on 188 on-going projects in the humanities funded by the Council in 1985.

Director Jostein H. Hauge of the Centre reports that the Council has appointed a committee in connection with the initiation of a program on humanistic information technology research. The committee's mandate has been to chart the status of research in this field and to suggest a plan of further action. The results of the committee's work have recently been published in a report.

The committee has concentrated on clarifying two issues: the humanities as suppliers of knowledge in connection with the development of information technology products, and humanistic studies of cultural consequences of information technology in society.

In the committee's view, information technology should be adapted to a human form of communication. The humanities can contribute to this goal in two ways: Theories and research in the humanities can directly form the basis of the development of software, and critical investigations can be conducted of the prerequisites of research in knowledge technology. Relevant disciplines are linguistics, phonetics, logics and image analysis.

There are many important research fields for the humanities in connection with cultural and social consequences of information technology. These include studies of how information technology influences the individual person, use of vocabulary, and general norms and values - issues of particular interest to linguists and philosophers.

In the report a detailed overview is given of relevant Norwegian research and training in the disciplines mentioned above. In order to strengthen these activities, more computing equipment and competence is necessary. The committee concludes its report with a detailed plan for a seven-year national research program on humanistic information technology.

## **Meldinger**

### **News**

The Centre has published a collection of articles that give an introduction to computing in the humanities. A second collection will appear this autumn.

A national teleconferencing system for the humanities has been established at the University of Oslo.

Researchers at Brown University, USA, are currently transferring 250 MB of classical Greek literature to a single CD-ROM disk. The collection includes all known works by 178 authors, from the time of Homer to approximately 600 A.D.

At the Rochester Institute of Technology, USA, computing and typesetting experts are developing an electronic publishing system for home and office use. The system will be available from 1987.

*Archaeological Computing Newsletter* contains articles, reports and news of relevance to computer applications in archaeology and at archaeological museums. The newsletter is published quarterly.

A Nordic conference on text comprehension and information retrieval will be held in Sweden in December.

A conference on computers and teaching in the humanities will be held at the University of Southampton in April 1987.

*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. Samlet pris kr. 180. (Rapportene kan også kjøpes enkeltvis).
- 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. ISBN 82-7283-039-6. Pris kr. 60.
- RAPPORT nr. 35. *Ole Lauvskar: Bruk av statistiske metoder 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.

Av innholdet:

Div. forfattere: Nordisk edb-konferanse for kunst- og kulturhistoriske museer

Jon Birger Østby: Status for edb-arbeidet i de norske kunst- og kulturhistoriske museer

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

Charlotte Toubro: Educational Uses of Existing Linguistic Software

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

Kristin Natvig: «Project Emperor-I» – kinesiske skatter på videoplate

Returadresse:

**NAVFs EDB-senter for humanistisk forskning**

Boks 53

5014 Bergen – Universitetet