

Datalogi – en ny hjælpevidenskab i humanioras tjeneste

Af Lene Rold

1. Den fagopdelte verden

Siden oplysningstiden er menneskets viden om sig selv og naturen vokset kolossalt. Så sent som i højmiddelalderen kunne et velinformeret menneske med en central geografisk placering godt have et overblik over sin tids videnskabelige stade, men efterhånden som viden ved stadig forskningsindsats er vokset, er det blevet umuligt at overskue hele spektret.

Ved begyndelsen af det 20. århundrede var det stadig muligt at opretholde et overblik over et enkelt, eller måske endda flere fagområder. Inden for de yngre videnskaber, som f.eks. arkæologien, har det været muligt at bevare dette overblik længere end i de gamle videnskaber.

Så sent som i 1957-60 udkom Johannes Brøndsteds reviderede og ajourførte »Danmarks Oldtid«. Et digert værk på tre store bind, som med nogen ret kunne hævdes at indeholde alt om Danmarks oldtid. Det kunststykke har ingen arkæolog siden kunnet gøre efter, og årsagen hertil er ikke, at arkæologerne er blevet dovne og uduelige, men simpelthen det faktum, at den arkæologiske viden er blevet så stor, at det idag endog er svært at holde sig orienteret i blot en enkelt af de traditionelle »aldre«: sten-, bronze- og jernalder.

2. Større viden – mindre visdom

Vi oplever derfor det dilemma, at jo større vor viden bliver, desto mere må vi opsplitte verden i ekspertom-

råder, og jo flere ekspertområder desto mindre mulighed for at bevare overblikket. Med denne udvikling står det os desværre også klart, at vi ikke kan anse en



Fig. 1. Mekaniske metoder til hurtige udregninger har været brugt til forskellige tider. Kuglerammer blev f.eks. brugt i middelalderen både i Europa og i Kina. Den første kendte mekaniske regnemaskine blev bygget af Blaise Pascal (1623-1662). Efter W. de Beauclair: *Rechnen mit Maschinen*. Braunschweig 1968.

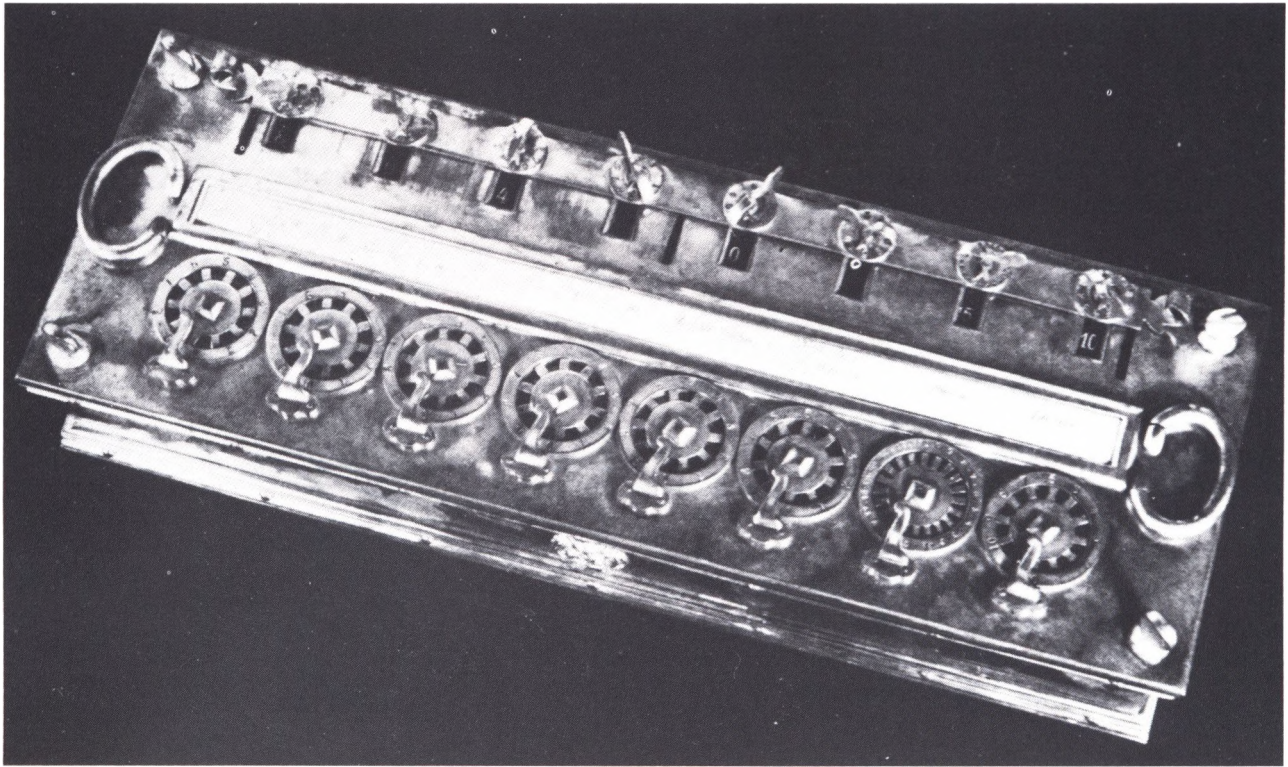


Fig. 2. Maskinen, som Pascal byggede i en alder af kun 20 år, blev vistnok brugt af Pascals far, som var skatteopkræver. Efter W. de Beauclair 1968.

stor ekspertviden inden for et lille ekspertområde for en anvendelig og naturlig viden om verden – vi bliver ikke »vise« ved at blive eksperter.

En videnskabspllosion synes kun at føre til, at det enkelte menneske vil blive reduceret til en biologisk »chip«, hvis eneste formål er at vedligeholde et lillebitte vidensområde – med en enorm stor viden – i et kæmpemæssigt vidensunivers.

Som tankeeksperiment kunne man godt forestille sig, at man indlæste al arkæologisk viden i en datama-

skine og lod maskinen skrive »Danmarks Oldtid«, men resultatet ville være så omfattende, at værket kun kunne »læses« af en anden datamaskine! Alle andre udgaver af værket måtte nødvendigvis forkortes og bearbejdes til versioner, der kunne læses i hvert fald inden for et menneskes levetid, med alle de problemer dét så ville give med at sørge for, at værket hele tiden var opdateret med den seneste viden!

Mange har forestillet sig, at man med datalogiens hjælp kan genskabe det tabte overblik. Med datalogien

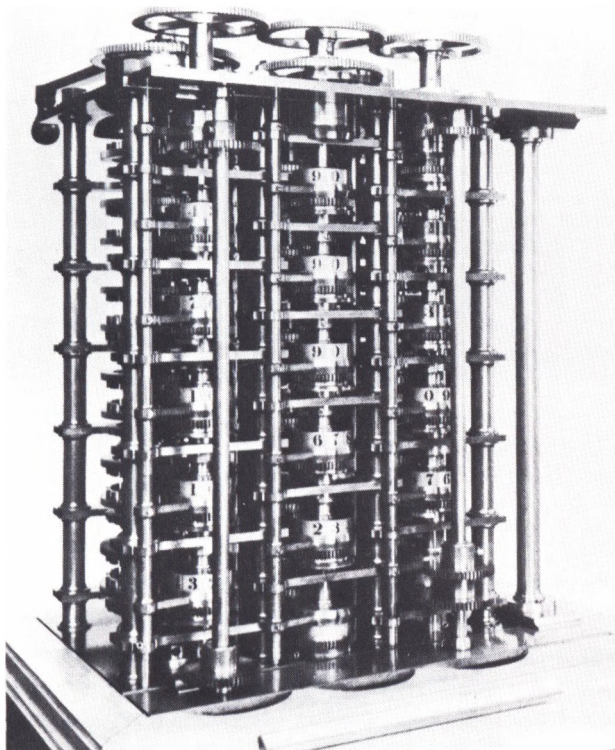


Fig. 3. Med forbedringer, som blev indført fra begyndelsen af 1800-årene, blev mekaniske maskiner en virkelig hjælp ved store beregninger. Englænderen Charles Babbage's beregningsmaskiner blev især berømte, selvom ingen af hans maskiner nogensinde blev bygget helt færdige. Men Charles Babbage er med rette indskrevet i datalogiens historie, fordi han som den første forestillede sig en maskine, der kunne »programmeres«. Efter D. Stein: Ada. A life and a legacy. London 1985.

er der skabt mulighed for hurtigt at få overblik over meget store stofområder, hvilket i sig selv vil have en sammenfattende funktion – for en tid i hvert fald – ligesom datalogien også kan være med til at skabe en tværfaglig sammenhæng. Hvis man imidlertid nøjedes



Fig. 4. Charles Babbage blev protegeret og delvist finansieret af en kreds af det engelske aristokrati, som interesserede sig for matematik – en højst aparte interesse på den tid – som vist for de flestes vedkommende mest blev anvendt til beregninger af favorable forbindelser i hesteavlen på baggrund af opnåede resultater på væddeløbsbanerne. Til denne kreds hørte også Ada, countess of Lovelace. Hun var lord Byrons datter og kunne være blevet en habil matematiker, om hun havde levet i en anden tid. Efter D. Stein 1985.

med at anvende datalogien til disse simple funktioner, ville man næppe løse det ovenfor skitserede eksistentielle problem i det lange løb. Tværtimod ville datalogien snarere sætte endnu mere skub i vidensproduktionen og dermed i udskillelsen af ekspertområder.



Fig. 5. Kort før 2. Verdenskrig startede den udvikling, som førte til dagens datamaskiner. De tidligste maskiner var kolossalt store, men kunne ikke klare udregninger af samme omfang og kompleksitet som nutidens programmerbare lommeregner. Her den berømte Mark I, som blev fremstillet i årene op til 1939 af IBM i samarbejde med Havard University. Mark I var i funktion indtil 1944, hvorefter den blev ombygget og forbedret. Maskinen optog et gulvareal på ca. 150 m². Efter W. de Beauclair 1968.

Løsningen på dilemmaet synes i øjeblikket at findes i den tankegang, at man *med maskinens hjælp* kan skabe versioner af den samlede viden. Det kan dreje sig om mere eller mindre detaljerede versioner af ekspertområder eller bredere tværfaglige redegørelser – versioner som også kan være tilpasset bestemte formål og bestemte målgrupper.

3. *Dataintensitet og vidensintensitet*

De ovenfor skitserede problemer og løsningsmuligheder, som involverer en »Analytisk Maskine«, kan stadig forekomme som science fiction. Ikke desto mindre er problemerne reelle nok, og den »Analytiske Maskine« er allerede i anvendelse inden for flere områder.

På museerne og inden for de museale fagområder er den maskinelle udvikling også forlængst startet – omend ikke i stort omfang til udførelse af opgaver, som tangerer »Kunstig Intelligens«. Gennem indsamling

af genstande og oplysninger – ikke blot oplysninger om de indsamlede genstande, men også i langt bredere forstand – har museerne efterhånden oparbejdet en kolossal mængde data. Dertil kommer en rig tradition i registrering og organisering af disse data. En sådan mængde af organiserbare data er med en datalogisk terminologi et *dataintensivt* område.

Den maskinelle behandling af store mængder data er idag en rutinepræget arbejdsopgave, som kort går ud på at lagre data og at give hurtig adgang til disse data for at anvende dem til forskellige formål så som a-jourføring, organisering på forskellig måde, beregninger osv.

Gennem forskning i bl.a. disse data produceres viden såvel på museerne som på universiteterne og alle andre steder, hvor de museale fag er repræsenteret. I den datalogiske terminologi er vore fag *vidensintensive*.

Udviklingen er kun lige startet, men mange mener, at en realisering inden for dette område vil få endnu

større og mere indgribende betydning i vort samfund, end hele den hidtidige udvikling inden for EDB-teknologien allerede har haft.

Det, man i den yderste konsekvens ønsker af et system baseret på vidensintensitet, er at kunne simulere en menneskelig hjernes tankeproces – f.eks. få en maskine til at fortolke og til at tage ved lære af sine fejltager. Nogle filosoffer, her iblandt Alan Turing, som var en af pionererne ved datamaskinens udvikling, ville gå så langt som til at påstå, at en maskine vel også kunne bringes til at forelske sig eller føle stolthed/skam ved et veludført/dårligt udført stykke arbejde!

4. Datalogiens anvendelsesområder på museer og i den humanistiske forskning

Fordele og ulemper ved at overføre museale data til elektroniske medier er et indviklet problemkompleks, som dog efterhånden er blevet diskuteret og beskrevet ofte nok. De fleste er derfor bekendt med problematikken såvel som perspektiverne – specielt når det drejer sig om dataintensive funktioner, såvel i administrative som i forskningsmæssige situationer.

Det følgende er derfor blot et supplement til den aktuelle debat med hovedvægt på de områder, hvor anvendelse af datalogiens teori og teknik synes at kunne true eller vanskeliggøre en bevaring af den humanistiske forskningstradition.

Mangfoldighed og standardisering

Det er et faktum, at humanistiske data er svære at registrere – i forhold til naturvidenskabelige data som f.eks. mål og beregninger af planeters bane. Ikke desto mindre registreres der på livet løs på museerne, og der stilles sjældent spørgsmålstejn ved værdien af al denne registreringsflid, som omfatter alt fra meget konkrete oplysninger vedrørende genstandes accession, proveni-

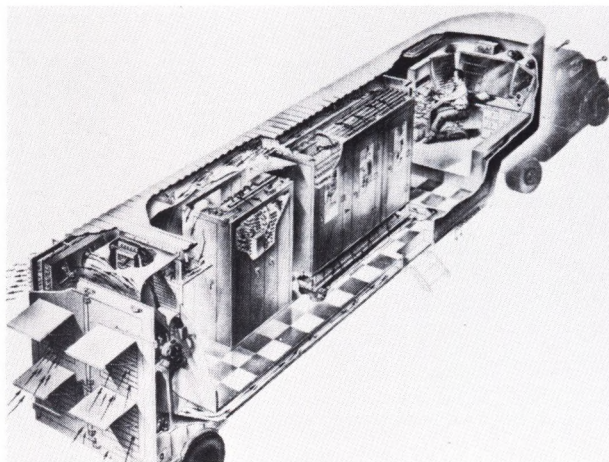


Fig. 6. På grund af størrelsen var datamaskinerne ikke flytbare, men i USA, hvor livet jo foregår på hjul, byggedes allerede i 1953 en mobil datamaskine. Lommeregnerens idémæssige forfader? Her en tegning af denne maskine i sit »mobile home«. Efter W. de Beauclair 1968.

ens og administration til mere eller mindre sikre beskrivelser, bestemmelser og »rørende historier«.

Følelserne kan imidlertid komme i kog, når det foreslås at omlægge registreringen til EDB, og årsagen hertil er en udbredt myte om, at EDB-registrering kræver standardiseringer, der i bund og grund er i modsætning til den museale registrerings formål.

Hvordan kan man f.eks. kræve en éntydig betegnelse på en genstand, når det jo er alment kendt, at mennesker nødigt kalder en skovl en skovl? Og hvordan kan man bibeholde en længere beskrivelse af en ceremoni i forbindelse med indsamling af de genstande, som indgik i ceremonien?

Det er korrekt, at etableringen af databaser stillede ret vidtgående krav netop til standardisering for ganske få år siden. Efterhånden har erfaringerne fra brugen af databaser imidlertid ført til, at baserne i langt

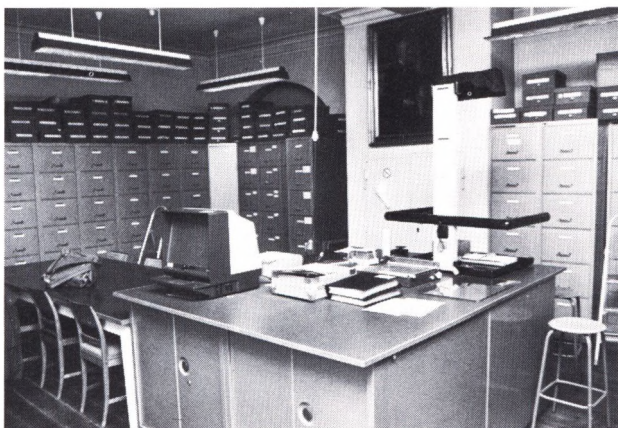


Fig. 7. Beretningsarkivet på 1. Afdeling, Nationalmuseet. De omfattende manuelle arkiver langs væggene bliver i øjeblikket nedfotograferet på mikrofilm til fremvisning på læseapparatet på bordet midt i billedet. Til højre for maskinen ses to sorte bøger, som indeholder mikrofilm af samtlige 1. Afdelings accessionsprotokoller helt fra omkring 1807 til i dag. Protokollerne selv fylder et skab på ca. 2 m i højden og 80 cm i bredden. Foto Lars Kann-Rasmussen. Dokumentationsenheden, Nationalmuseet.

højere grad kan fungere på samme varierede måde som manuelle registre, uden af den grund at overtage ulemperne fra de manuelle registre.

Brugen af databaser er nu betydeligt mere varieret, bl.a. fordi man anvender en nuanceret strategi ved etablering af en base, lader maskinen udføre så meget arbejde som muligt, og drager nytte af, at elektroniske medier omfattende billede, lyd og tekst kan bringes til at arbejde sammen. Desuden øjnes så småt muligheder for, at maskinerne selv kan tolke traditionelle medier som f.eks. maskinskrevne papirer, tale og billeder/genstande og omsætte disse til maskinlæsbarhed på elektroniske medier.

Ved etableringen af en database er man idag mere omhyggelig med at undgå for megen standardisering

end det modsatte, og man vil fraråde at udføre en arbejdskrævende overførsel, hvis der anes muligheder for, at en sådan overførsel kan ske maskinelt inden for en overskuelig tid.

Kvantitet og kvalitet

Ved hjælp af databaser kan man udføre visse forskningsopgaver, som før ville tage dage og år, på ganske få sekunder. Det er imponerende at se et udbredelseskort over en bestemt genstandstype blive skrevet ud, eller på en brøkdal af et sekund at få at vide, i hvor høj grad officerer i handelsflåden var involveret i indsamlingen til den Etnografiske Samling på Nationalmuseet før 1864. Ved at søge i udenlandske baser kan man desuden hurtigt sammenligne danske forhold med forhold i resten af verden.

Det er faktisk inden for rækkevidde at indkredse og afprøve sit kildemateriale til endog store videnskabelige projekter på en eftermiddag, hvis man er en ihærdig, på-hit-som og dreven databasebruger. Det er derfor sandsynligt, at humaniora vil opleve en bølge af projekter med baggrund i store kvantitative undersøgelser, og det kan måske virke skræmmende.

Der er dog næppe grund til at frygte, at en sådan bølge vil være langvarig, eller at den vil kunne ændre forskningen på længere sigt. Snarere vil man kunne undgå, at kvantitative undersøgelser, som er anerkendte og nødvendige dele af de fleste projekter, får mere betydning end de reelt har krav på. Det vil således næppe kunne ske, at folk kan få accepteret specialafhandlinger eller artikler, hvis aspekt kun er indsamling og organisering af kildemateriale – sådan som det kan forekomme med de manuelle metoder, der benyttes idag. Der vil med god grund kunne stilles større krav til vurderingen af den indholdsmæssige betydning i kildematerialets vidnesbyrd og tolkningen heraf.



Fig. 8. Undervisningslokalet i Dokumentationsenheden, Nationalmuseet. Der gives undervisning i alt det programmel, som museet anvender lige fra tekstbehandling over regneark til brug af det store databasesystem, som anvendes til museets forskellige registreringsprojekter. Foto Lars Kann-Rasmussen. Dokumentationsenheden, Nationalmuseet.

Økonomi og humanistisk forskning

En væsentlig indvending mod indførelse af EDB er simpelthen, at det er for dyrt!

Datamaskinerne er, såvel på maskin- som programmelssiden, blevet mange gange bedre og billigere, siden udviklingen tog fart i slutningen af 70-erne. Det ser ud til, at der også fremover kun kan vindes økonomisk ved at vente med anskaffelsen af EDB-udstyr.

Imidlertid er det hverken maskinerne eller kommercielt udviklet programmel, der er den største økonomiske belastning. De virkelig store udgifter er på længere sigt dels at uddanne museernes personale og forskerne i brugen af datamaskinen, og dels de udgifter, der må påregnes til konsulenthjælp udefra. Sidst – men af størst betydning – vil de udgifter gøre sig gældende med stigende styrke, som skal afsættes til *udvikling af programmel*, der kan dække humanioras behov.

Til at afhjælpe problemet har Statens Museums-

nævn nedsat et EDB-udvalg, som i stigende grad beskæftiger sig med udvikling og udvælgelse af egnet programmel til museerne, og i flere år har både de kulturhistoriske og kunsthistoriske museer haft centrale EDB-registre, DKC (Det kulturhistoriske Centralregister) og KID (Kunst Indeks Danmark).

Nationalmuseets nyoprettede Dokumentationsenhed anser også dette område for meget vigtigt. De to største enkeltopgaver, som vil optage staben de første par år, er design og anvendelse af en database til registrering af genstande, såvel tekst som billede, og udvikling af formidlingsprojekter med anvendelse af EDB-teknologi, bl.a. interaktiv video. I de kommende par år vil der dog også blive lagt stor vægt på undervisning til alle museets ansatte i brug af almindeligt kommercielt programmel som f.eks. tekstbehandling, bogholderiprogrammer osv. Museets ansatte gennemgår således dét, man i erhvervskredse kalder »Det Teknologiske Løft«.

Da Nationalmuseets forskellige afdelinger selv undersøger undervisningsbehov og tilrettelægger undervisningsplaner sammen med Dokumentationsenheden, opnås dels at uddannelsen styres og tilskrives efter afdelingernes eget ønske. Desuden koster undervisningen kun en brøkdel af, hvad en tilsvarende undervisning ville koste »ude i byen«. Som eksempel kan nævnes at et kommercielt 4-dages kursus, som afvikles uden hotelophold og fordyrende luksusmåltider nemt kan koste kr. 10.000 pr. person, mens et tilsvarende kursus i Dokumentationsenheden ikke koster noget.

En sådan udregning er selvfølgelig ikke acceptabel, da Dokumentationsenheden koster en hel del i etablering og drift, men der er sikkert god økonomi i en sådan »Enhed«, og man kunne forestille sig, at noget lignende på amtsligt plan til service for landets museer ville være en god investering.

Humanioras største økonomiske problem ved indførelse af EDB-teknologi er derfor slet ikke tangeret endnu. Den fremtidige udvikling inden for vidensbaserede systemer vil stille helt anderledes økonomiske krav, som humaniora ikke vil kunne klare med de ressourcer, som allokeres idag – og her tænkes på såvel penge som organisation. Resultatet kan derfor blive, at humaniora nødsages til at anvende programmeler, som egentlig er designet til erhvervslivet, og dette vil mere end noget andet kunne true den humanistiske forskningstradition.

Humanist og datalog

Indtil videre har den datalogiske teori og teknik ikke haft nogen større betydning i den humanistiske forskning. Kun ganske få humanistiske forskere har taget teknikken i anvendelse, og endnu færre har ladet sig inspirere af datalogiens teorier.

Årsagen hertil er sandsynligvis, at den tekniske side – bl.a. brugen af datamaskinerne – faktisk har været meget besværlig at lære og dertil krævede en form for teknisk viden, som de fleste humanister følte sig meget fremmede overfor.

En anden årsag er ganske givet, at EDB-teknologien er behersket af en specielt ondartet og uforståelig fagterminologi, hvor det svirrer med CD-ROM'er, sekventielle ASCII-filer, bits, bytes og meget andet af samme skuffe. Faktisk er denne fagterminologi ikke altid nødvendig, og den anvendes oftest af halvstuderede røvere, som på bedste Erasmus Montanus-vis forsøger at dupere omverdenen med halvfordøjet viden eller forsøg på at dække over helt manglende viden.

Heldigvis involverer indsigt i datalogi ikke længere nødvendigheden af at tilegne sig al denne tekniske viden. Det er således helt unødvendigt at lære programmering. Kommunikationen med datamaskinen er alle-

rede forholdsvis let, og i løbet af de næste par år vil kommunikationen blive lige så gnidningsløs som brugen af almindeligt TV, selvom en forskningsmæssig brug af datamaskinen ikke tænkes at foregå som envejskommunikation.

Den sværeste del ved at inddrage datalogien i den humanistiske forskning er dog stadig nødvendigheden af at forstå og mestre datalogiens muligheder og begrænsninger på det teoretiske plan. Humanisterne skal, samtidig med bibeholdelse af egne forskningsparadigmer, lære dels at stille krav til datalogien og dels at lade sig inspirere til nye forskningsmetoder.

Dette arbejde på det teoretiske plan vil dog næppe afskrække humanister, som da også vil komme i selskab med folk som Ludwig Wittgenstein, Bertrand Russell, Kurt Gödel og Alvin Toffler for blot at nævne nogle få forskere, som har leveret ammunition til såvel datalogien som de humanistiske videnskaber.

5. Humaniora og samfundet

Det er i det foregående ofte blevet påpeget, at anvendelse af EDB-teknologien kun lige er på begyndelsesstadiet. Selvom samfundet tilsyneladende er gennemsyret af EDB-teknologi, gælder dette kun på de *dataintensive* områder, dvs. at EDB-teknologien anvendes til behandling af store datamængder. De *vidensintensive* områder, hvor en stor maskinlagret viden anvendes til at opnå resultater på baggrund af ganske få data, er det næste skridt og et skridt, hvor humaniora ikke alene bør være repræsenteret, men endda måske bør gå forskningsmæssigt i forvejen.

Med de vidensbaserede systemer nærmer man sig et yderst kontroversielt emne: kunstig intelligens.

Menneskelig intelligens kontra kunstig intelligens er blevet diskuteret i århundreder, og med datamaskinen

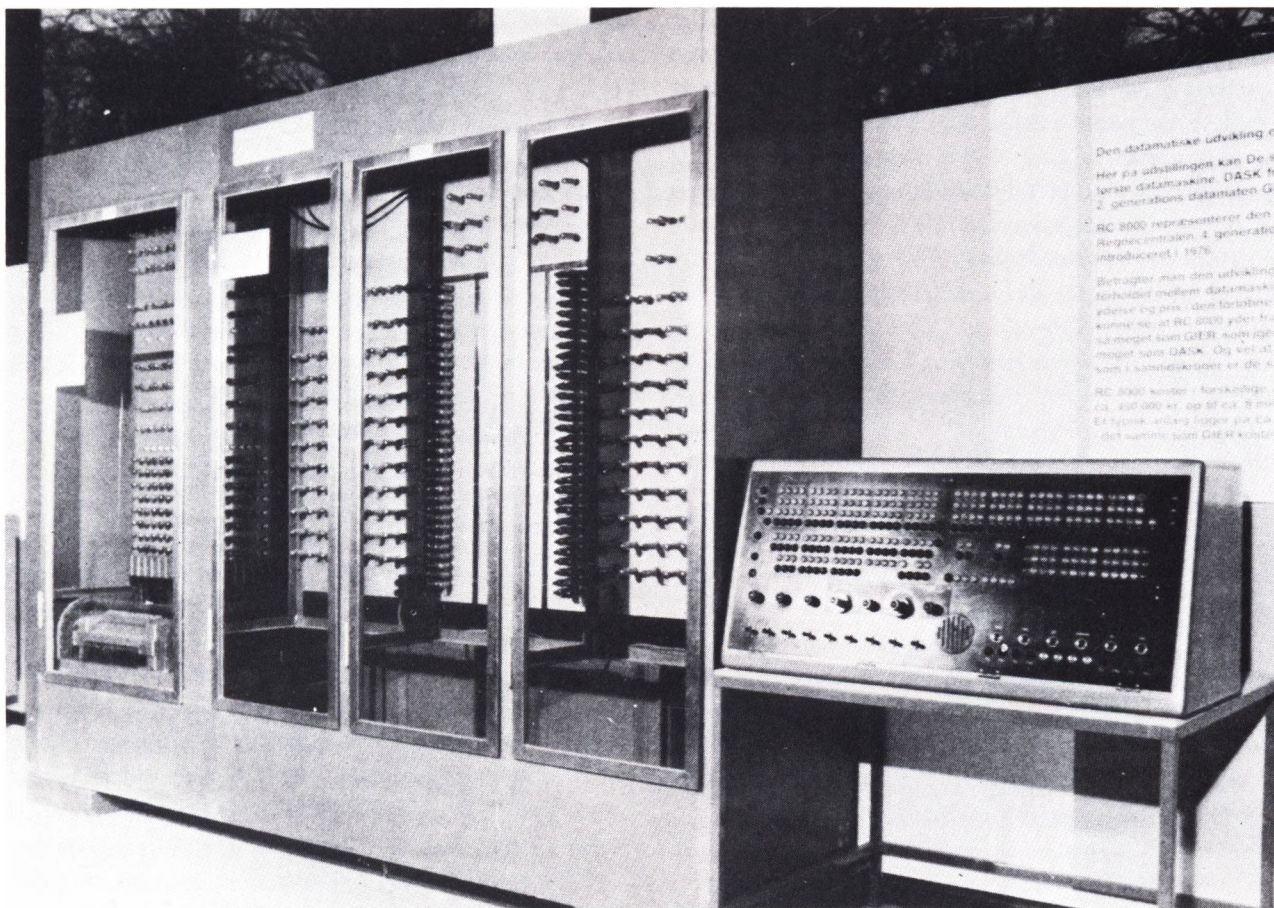


Fig. 9. Den første danske datamaskine, DASK (Dansk sekvens Kalkulator), blev bygget 1955-57 og fungerede med udbygger, især til hjælp for maskinens hukommelse, i 10 år. Maskinen var meget indviklet at bruge og det krævede derfor en særlig uddannelse. Efter T. Blankholm Nielsen: *Teknologi – Menneske*. København 1987.

har diskussionen fået fornyet relevans. Ofte stilles to aspekter af intelligens op mod hinanden. På den ene side den rene tanke: den matematiske, logiske verden, som anses for at kunne programmeres. På den anden

side står den rene natur: den kunstneriske, uforudsigelige, følelesesbetonede verden, som anses for ikke at kunne programmeres.

Forsøget på at bringe »den rene tanke« og »den

rene natur« sammen er af gammel dato. Som regel er moralen, at mennesker kan blive klogere og bedre ved at anvende lidt mere af »den rene tanke«, mens det at bygge et menneske, altså at skabe natur gennem videnskabelig teknik, er katastrofalt. Berømte afskrækkende eksempler er middelalderens Homunculus og Frankensteins Monster fra nyere tid.

Ud fra denne grundholdning vil man formentlig også vige tilbage, hvis det blev foreslået at anvende vidensbaserede systemer til at skabe en sammenhæng i den fragmenterede og fagopdelte verden, som vi idag lever i. Ikke desto mindre er det nok værd at overveje, om datalogien kan ændre det faktum, at politiske be-

slutninger i høj grad foretages alene ud fra tekniske og økonomiske overvejelser.

Det er sikkert ikke så meget ond vilje, men snarere mangel på overblik, som idag næsten udelukker anvendelsen af humanistiske forskningsresultater og synsvinkler i samfundets beslutningsprocesser. Men man kan næppe forestille sig at situationen kan ændres, hvis humaniora ikke selv tager initiativ til en ændring. Der påhviler derfor humaniora et stort ansvar for, at samfundsudviklingen ikke i fremtiden skal styres af et Frankensteinuhyre uden humanistiske kredsløb. Og en af måderne at sikre dette på kan være deltagelse og engagement i den teknologiske udvikling.