

Interne notater

STATISTISK SENTRALBYRÅ

17. februar 1984

84/3

DATABEHANDLING I STATISTISK SENTRALBYRÅ

Av Arne Øien

INNHOOLD

	Side
1. Innledning	2
2. Anskaffelse av utstyr	4
2.1 Datamaskiner og perifert utstyr	4
2.2 Tekstbehandlingsutstyr	4
2.3 Utstyr for grafiske framstillinger	5
2.4 Dataregistreringsutstyr	5
2.5 Optiske lesere	5
2.6 Mikrofiche-utstyr	5
3. Hvor skal utstyret være plassert?	5
4. Systemarbeid og programmering ved Systemkontoret	6
5. Prioritering mellom hovedoppgaver	6
6. Databaser	7
7. Dokumentasjon av data og rutiner	8
8. Opplæring og brukerveiledning	8
9. Organisering av EDB-arbeidet	9

DATABEHANDLING I STATISTISK SENTRALBYRÅ

1. Innledning

Jeg skal her gi en bred oversikt over retningslinjene for databehandlingen i Byrået. I grove trekk er disse trukket opp i Rapporter 82/82 - Perspektiv for 1980-årene som jeg håper de fleste har fått anledning til å se igjennom til i dag. Det er ikke mulig og heller ikke nødvendig å fastlegge detaljer i våre planer på dette området i særlig grad. Heller ikke er det råd å se særlig langt fram i tida. Dertil er vår virksomhet innen EDB for omfattende og den teknologiske utviklingen for rask. Bortsett fra rent prinsipielle retningslinjer må vi nøye oss med konkrete planer bare for noen få år om gangen. Innen denne tidsrammen må det også skje løpende tilpassinger i forhold til planen. Jeg tror dette seminar vil være nyttig både når det gjelder å orientere om retningslinjer og planer og når det gjelder å få reaksjoner på disse. Selv om det ikke blir tatt avgjørelser her i dag, vil de synspunkter som kommer fram påvirke utviklingen i neste runde.

Før vi går løs på de enkelte sider ved vår databehandling kan det være nyttig å se litt på den alminnelige utviklingen på dette feltet. Alle synes å være enige om at teknologien innen EDB og datakommunikasjon stadig er i rask utvikling. De aller siste årene har vi fått dette illustrert spesielt gjennom det varierte tilbudet av mikromaskiner. I dag kan vi for 3000 kr få en datamaskin som på flere områder går utenpå Byråets første datamaskin, og vi kan trygt regne med at det neste år kommer nye modeller med økt kapasitet. Dette har gjort datamaskinen til et leketøy og kanskje et læremiddel for barn og voksne. I en noe høyere prisklasse har det dukket opp maskiner under betegnelsen "personlige datamaskiner" beregnet på yrkeslivet. Bærbare datamaskiner har også blitt en egen spesialitet innen datamaskinområdet. Dette kan tenkes å bli nyttig utstyr for våre intervjuere om noen år. Vi kan tenke oss kontrollerte data overført over linje når intervjueren kommer heim om kvelden. På mini- og stormaskiner har det vært en jevnere utvikling, men det er interessant at det i USA omtales planer om å bygge stormaskiner med hastigheter på 1 milliard regneoperasjoner pr. sek. Disse antas å være ferdige i 1986 og vil da være 50 ganger raskere enn dagens stormaskin, ca 1000 ganger raskere enn Byråets IBM-maskin.

For Byrået er det også av stor betydning at lagringskapasiteten i maskinen øker. Her foregår det stadig en utvikling. For noen år siden var våre

største diskdrives på 200 millioner bytes. Vi regner nå med at vi om 1 eller 2 år vil ha slike enheter med 10 ganger så stor kapasitet. På dette området har det vært mulig for oss hele tiden å ta i bruk det mest effektive på markedet på et ganske tidlig tidspunkt.

På området datakommunikasjon har det hittil ikke skjedd så mye, men det ventes en betydelig utvikling over en lengre periode - 10-20 år. Optiske fibre og satellitter vil øke overføringskapasiteten i de offentlige nett til det mangedobbelte. Utbyggingen av nye offentlige teletjenester basert på disse teknologiene vil ta tid. Men Teledirektoratet har allerede annonsert tjenester som vi ikke har tatt i bruk men som vi sikkert kan ha nytte av, f.eks. teletekst som vil tillate oss å sende over linje aktuell statistikk direkte fra datamaskinen til de viktigste mottakerne. I den siste NOU'en om telematikk er det nevnt som en mulighet at skjemaer til myndigheter kan fylles ut på en dataskjerm og sendes over linje til mottakeren. Tilsvarende kan vi tenke oss automatisk purring som elektronisk post i stedet for brev.

I tillegg til offentlige tjenester utvikles det lokale nett (LAN = local area network) som kan brukes privat innen et avgrenset område. Hensikten med dette er å øke hastigheten på overføringer mellom eget utstyr. Dette gir en bedre utnyttning av databehandlingsutstyret samtidig som det tar seg av intern kommunikasjon, bilder, tale og tekst.

På området systemutvikling og programmering snakkes det om 4. generasjons språk og verktøy. Bestrebelsene tar sikte på å effektivisere arbeidet med utvikling av nye EDB-rutiner, noe som hele tiden synes å være en flaskehals for alle. Utviklingen har to tendenser. Den sterkeste tendensen synes nå å ta sikte på å bringe brukeren mer aktivt inn i løsningen av sitt problem på datamaskinen. Dette fører til språk og programpakker tilpasset bestemte brukergrupper, altså med forholdsvis begrenset generalitet. Samtidig blir de generelle hjelpemidler fra de profesjonelle EDB-folkene bedre. På begge områder, brukeren og profesjonelle, snakkes det om store effektivitetsgevinster ved overgang til neste generasjons verktøy, f.eks. at programmeringsarbeidet kan skjæres ned til 1/10. Det er her ofte upresist hva som menes med programmering og hvor stor andel denne funksjonen utgjør av hele innsatsen i problemløsningen. Det som tross alt synes å være et alminnelig fenomen er mangel på programmering.

Hos oss er det klart at programsystemet TROLL på IBM har bragt oss et betydelig skritt videre i løsning av modeller på datamaskinen. DATSY har hatt en stor betydning og er fortsatt nødvendig for løpende drift av store modeller, men bør nå moderniseres. SPSS er fortsatt et effektivt verktøy for analyse av individualdata. På IBM-maskinen mener vi at vi nå har et enda bedre verktøy i SAS-pakken for statistisk analyse. For dataklargjøring utvikler vi vårt eget system, DOKS. Våre egne standardprogram for tabellproduksjon erstattes

gradvis av TAB-68 på IBM-maskinen samtidig som vi selv utvikler PRINTERTAB for relativt enkle tabeller hvor det viktig å rasjonalisere tekstingen av tabellene.

Etter at IBM-anlegget er bygd ut, regner vi med å kunne skaffe et større databasehåndteringssystem for å lette arbeidet med å utvikle databaser. Dette vil bli vår hittil største investering i ferdige programsystemer og vil trolig koste oss over en million kroner pr. år fra 1985.

2. Anskaffelse av utstyr

2.1. Datamaskiner og perifert utstyr

Det viktigste utstyret er datamaskinen og de dataskjermer, linjeskrivere, magnetbåndstasjoner, platelagre osv. som er knyttet til maskinen. Det har alltid vært slik at dette utstyret har vært en knapphetsfaktor. Det har også vært slik at vi har hatt en høy utnyttingsgrad av utstyret i forhold til andre som vi har hatt muligheter for å sammenlikne oss med. Dette har ført til at prinsippet "mest mulig kapasitet for pengene" har vært lagt til grunn for våre store maskinkjøp. Dette har ført til flere leverandører, manglende standardisering og kjøp av kapasitet hos andre data-sentraler. Noe av det vi har tjent på maskinsiden, har vi tapt i ekstra opplæringskostnader, konvertering av program og datakommunikasjonsproblemer. Det er grunn til å tro at det nå bør legges større vekt på standardisering av utstyr og redusert bruk av eksterne data-sentraler. Det forhold at maskiner stadig blir billigere i forhold til arbeidskraft, tilsier dette. Likevel må vi regne med at det kommer nye leverandører og nye maskiner som vi finner det riktig å ta i bruk, f.eks. innen mikromaskiner.

Vi har i dag fordelt arbeidet på en rekke minimaskiner (Nord) og to stormaskiner (IBM 4341 og Honeywell Bull). Det ville vært enklere, men dyrere, å legge alt arbeidet på en stormaskin. Det kan nå tenkes at mikromaskiner, til visse oppgaver, vil gi oss mer kapasitet for pengene enn det utstyret vi har. Dette må studeres nærmere. Før vi går til noen vesentlig anskaffelse av mikromaskiner, må vi være sikre på hvilke komplikasjoner det fører med seg å fordele arbeidet på enda flere maskiner. IBM-anlegget vil bli bygget ut i 1984. Vi regner med at en av ND 100-maskinene blir bygget ut til ND 550.

2.2. Tekstbehandlingsutstyr

Tekstbehandling har vært og vil fortsatt være lavere prioritert enn databehandling. I 1984 regner vi med å investere 250 000 kr i spesialutstyr for tekstbehandling, samtidig med en fullstendig omlegging til nytt utstyr ved Tekstbehandlingskontoret. Tekstbehandling idatamaskinene vil bli tilgjengelig fra dataskjermen i den grad kapasiteten tillater det.

2.3. Utstyr for grafiske framstillinger

Vi har i dag to grafiske dataskjermer og en plotter. Hittil har vi kommet svært kort når det gjelder å ta i bruk grafisk databehandling. Det utstyret vi har i Kongsvinger må dekke behovet på stedet noen år. I Oslo mangler vi en plotter. Her vil vi få en grafisk skjerm og en fargeprinter i nær framtid.

2.4. Dataregistreringsutstyr

Det kan fortsatt bli nødvendig å investere noe i spesialutstyr for dataregistrering. Dette vil bli gjort snarest mulig. Det kan tenkes at neste modell, XL50, av det utstyret vi har kan konkurrere bra med dataskjermer og DOKS.

2.5. Optiske lesere

Når vi ikke tidligere har funnet grunn til å anskaffe slikt utstyr, synes det nå snart å være for sent. Det kan imidlertid dukke opp en mulighet for samarbeid med Tolldirektoratet om en stor optisk leser som Direktoratet tenker å skaffe til lesing av tolldeklarasjonene.

2.6. Mikrofiche-utstyr

Byrået har et komplett anlegg for framstilling av mikrofiche med tilstrekkelig kapasitet.

3. Hvor skal utstyret være plassert?

Vi må ta sikte på at kontorene i Oslo og Kongsvinger får like gode datamaskintjenester. Dette betyr ikke nødvendigvis at vi må ha datamaskiner begge steder. Våre egne maskiner er i dag fordelt slik:

Oslo: 6 ND-maskiner

Kongsvinger: 1 IBM 4341 og 1 ND-maskin.

I 1984 vil vi forsøke å samle alle ND-maskinene i Oslo for å oppnå en bedre utnyttelse og redusere Driftskontorets arbeid med ND-maskinen. Når IBM-anlegget blir for lite, må vi avgjøre om vi skal ha en maskin også i Oslo, eller om vi skal satse på en stormaskin i Kongsvinger. Dette vil først og fremst være et spørsmål om hvilken løsning som gir oss mest kapasitet for pengene med hensyn tatt til samme krav til svartiden begge steder. Det må imidlertid også tas hensyn til den effekten på det faglige miljøet som en datamaskin på stedet antas å ha.

4. Systemarbeid og programmering ved Systemkontoret

To tendenser gjør seg gjeldende:

- Det blir stadig bedre standardprogram å få kjøpt
- Systemarbeid og programmering blir stadig dyrere i forhold til maskinkapasitet selv om programmering også stadig effektiviseres ved bedre programmeringshjelpemidler.

Det lønner seg derfor etter hvert i større grad å erstatte egen innsats på dette området med kjøpte program. Dette fører til en desentralisering av slikt arbeid fra Systemkontoret til brukerne. Dessuten lønner det seg å bruke også noe av Systemkontorets ressurser til utvikling av standardprogram. Systemkontorets ressurser som kan brukes til spesialprogrammering vil altså øke samtidig som behovet for spesialprogrammering avtar. Dette synes å føre til en sikker nedtrapping ved Systemkontoret. Med den kjennskap vi har til uløste oppgaver og behovet for å utbedre gamle rutiner, må vi regne med bare en beskjeden nedtrapping de nærmeste årene, neppe så sterk som den gjennomsnittlige avgangen de senere årene. Vi har dessuten tidligere hatt den erfaring at løsninger basert på standardprogram etter en tid må erstattes med spesialprogram for å oppnå bedre maskineffektivitet eller tilfredsstillende brukerkravene bedre. Men her må vi være restriktive.

De siste årene har det vist seg mulig å konsentrere innsatsen ved Systemkontoret sterkere om færre og større prosjekter. Dette har gitt gode resultater. Denne tendensen bør vi derfor oppmuntre til.

5. Prioritering mellom hovedoppgaver

I denne sammenheng kan vi dele inn databehandlingen vår slik:

- Registerarbeid
- Dataklargjøring
- Tabellproduksjon
- Analyse
- Tekstbehandling
- Administrativ databehandling
- Grafisk databehandling
- Publisering av dokumentasjon og statistikk (gjennom databaser).

Rekkefølgen her tilsvarer trolig den prioritet vi gir oppgavene i dag. (Vi har ingen klar prioritering og det er ikke mulig å foreta målinger.) Det er liten grunn til å tro at rekkefølgen blir forandret de nærmeste årene. Det skjer imidlertid en langsom utjamning slik at lavt prioriterte oppgaver tildeles større ressurser. Publisering vil trolig om 4-5 år ha rykket noe opp på lista.

6. Databaser

Databaser er en metode for effektivisering av databehandlingen. Utviklingen hos oss har gått langsomt fordi det kreves betydelig mer maskinkapasitet. Databaser ble først tatt i bruk der den var mest lønnsom, på Bedrifts- og foretaksregisteret hvor det ble spart inn 6 årsverk pr. år uten at maskinkostnadene vokste tilsvarende. Vi må fortsatt ta hensyn til lønnsomhet ved overgang til databaser.

Databaseteknikken har for oss disse formål:

- Effektivisere statistikkproduksjonen
- Spre statistikken til brukerne
- Effektivisere programmeringen.

Det er naturlig først å utvikle databaser som vi selv kan ha nytte av. Når basene er der, vil vi vurdere hvilke eksterne brukere vi skal slippe til. Her må vi også ta hensyn til vårt krav om et effektivt datavern. I 1984 og 1985 regner vi ikke med å slippe til eksterne brukere over oppringte linjer til våre databaser. Spesielle løsninger som tillater adgang fra noen departementer til våre databaser kan vi likevel tenke oss i denne perioden.

Vi er i gang med å legge data inn i Teledata som vil bli en åpen database for private brukere. Dette gir oss noe erfaring i hvordan statistikk skal presenteres for et større publikum på dataskjerm, bruk av farger og grafikk. Vi vil forhåpentligvis også få noe føling med hvor stor interesse det er for denne publiseringsformen. Ellers regner vi med at Teledata bare vil fungere som et "annonseorgan" for statistikken med referanse til våre databaser og publikasjoner.

Vi regner med en økende interesse for å lagre vår statistikk sammen med andre data i kommersielle databaser. Vi har ingen betenkeligheter med å levere statistikk til slike formål, men må stille rimelige krav til ajourhold og dokumentasjon av våre data.

7. Dokumentasjon av data og rutiner

I 1983 har vi kommet i gang med arbeid på det praktiske plan med bruk av datamaskinen til dokumentasjon av data. Databaseteknikken fører med seg en sterk integrasjon mellom data og dokumentasjon. Dette er en heldig utvikling. De nødvendige programsystem for lagring og søkning i dokumentasjonen later det til at vi kan skaffe oss, dels ved egen utvikling. Det gjenstår da å fylle systemet med innhold. Dette må gjøres etter fastlagte retningslinjer som vi selv må utvikle og etterleve.

Vi kan ikke regne med å dokumentere på nytt data som befinner seg i arkivet dersom disse ikke reorganiseres for å gå inn i nye beredskapsfiler eller databaser. Slik dokumentasjon i ettertid vil det også være vanskelig å få feilfri. Det er viktig at dokumentasjonen framkommer som et naturlig ledd i arbeidsprosessen og blir brukt ved framstillingen av statistikken.

Generelle retningslinjer for dokumentasjon av data vil foreligge i 1984.

Etterhvert som brukerne deltar mer aktivt i løsningen av sine problemer på datamaskinen, faller et større ansvar for dokumentasjon av EDB-rutinene på disse. Dette er det viktig å være klar over nå mens vi befinner oss i en overgangsfase. Det må tas hensyn til dette i vår opplæring av brukerne.

8. Opplæring og brukerveiledning

Opplæring av EDB-spesialister og brukere må prioriteres. Det er mulig Byråskolens kurstilbud må økes noe. Ellers regner vi med i økende grad å ta i bruk tekniske hjelpemidler i opplæringen. Vi har videoutstyr i Oslo og Kongsvinger som foreløpig ikke er fullt utnyttet. Systemkontoret har begynt å se på datamaskinstøttet undervisning. Det betyr at opplæringen kan foregå fra skjermterminaler rundt omkring i Byrået. Dette gjør det mye lettere å tilby undervisning til rett tid og på rett sted, men vil også stille større krav til planleggingen av opplæringen. Vi bør ta sikte på en integrering av undervisning med lærer, videokurs og datamaskinstøttet undervisning.

9. Organisering av EDB-arbeidet

Vi regner med en desentralisering av arbeid fra Driftskontoret, Systemkontoret og Kontoret for manuell databearbeiding. Dette kan lett føre til tap i effektivitet fordi sentrale organer normalt vil ha personale med større ferdigheter som følge av høyere spesialisering. Desentraliseringen vil imidlertid gi den som er ansvarlig for et prosjekt eller rutine, bedre kontroll over framdriften. Behovet for driftsplanlegging og kommunikasjon vil derfor avta.

Desentraliseringen tar sikte på sterkere integrering av funksjoner. Dette bør føre til mer interessant arbeid og bedre trivsel og mindre belastningssykdommer. Det siste har vært og er fortsatt et problem ved dataregistreringsgruppa. Ved desentralisering av dataklargjøring fra Kontoret for manuell databearbeiding er det derfor viktig at dataregistreringen følger med og ikke blir tilbake som en spesialisert funksjon. Dessuten må vi passe på at desentraliseringen ikke flytter for mye arbeid fra Kongsvinger til Oslo, da dette vil føre med seg flytting av andre oppgaver i motsatt retning.

Det høyt kvalifiserte EDB-arbeidet vil vi fortsatt beholde sentralt. Dette antar vi er nødvendig for å kunne holde på et miljø med tilstrekkelig kompetanse.

De senere årene har Systemkontoret forsøkt å konsentrere innsatsen mer om større prosjekter. Samtidig søker en å finne mer generelle løsninger som fanger opp uforutsette endringer. Denne utviklingen bør fortsette. Slikt arbeid blir organisert i prosjektgrupper hvor brukerne deltar. Prosjekter som spenner over flere brukere, f.eks. DOKS, bør også ha en styringsgruppe. Ledelsen av slike grupper bør normalt ligge hos brukeren, men det kan tenkes andre løsninger.

Byrået har i dag flere hundre selvstendige EDB-rutiner. Det er derfor behov for en konsolidering og forenkling. Vi regner med at databaseteknikken vil ha en slik virkning. Det vil trolig være nyttig å nå fram til rutiner som helt eller delvis dekker statistikkområder i stedet for enkelte statistikker. Dette forutsetter en viss generalisering, større samlet ressursinnsats og mer langsiktig planlegging. Det vil være behov for standardprogram og spesialprogrammering, samarbeid mellom statistikere og EDB-spesialister og et samspill mellom store og små datamaskiner.