

Interne notater

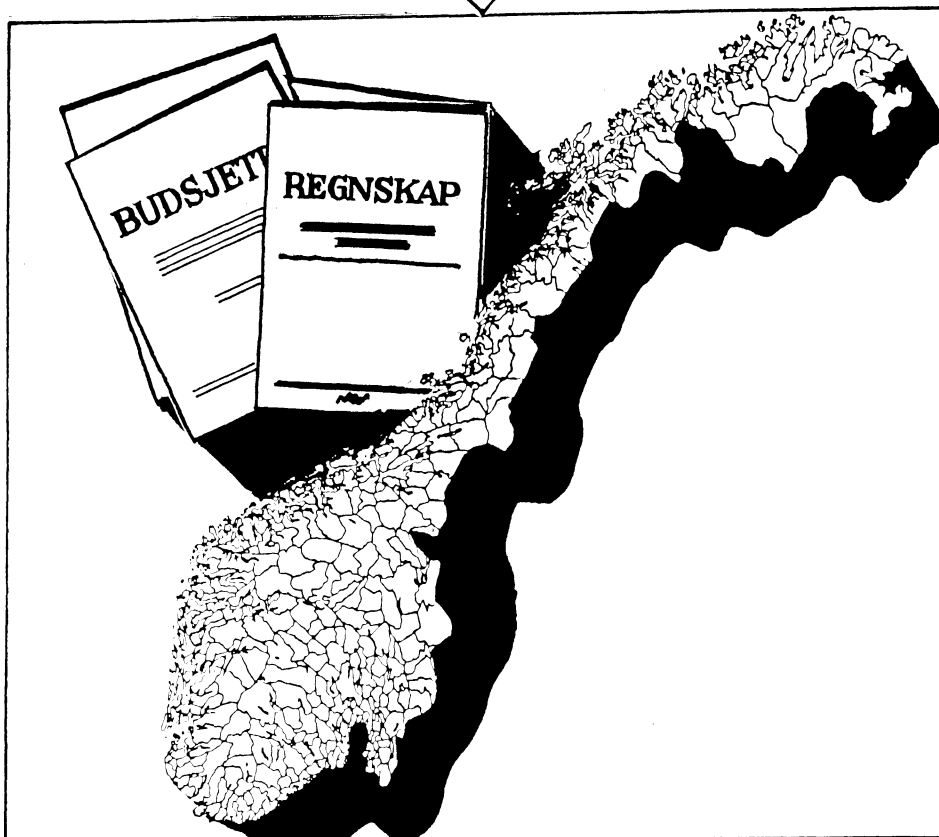
STATISTISK SENTRALBYRÅ

81/33

3. november 1981

FORPROSJEKT: DATABASE FOR KOMMUNALØKONOMI

Mandat: "Forprosjektgruppen skal kartlegge oppgaver, omfang, egnede systemer og kostnader ved en kommunaløkonomisk databank. Gruppen skal også sette opp en tidsplan for gjennomføring av hovedprosjektet."



Forslag fra prosjektgruppen

En database for kommunaløkonomi etableres innen 1. januar 1983, og ordinær tabellproduksjon kjøres fra databasen fra samme tidspunkt. Dette vil kreve ca. 3,6 årsverk system- og driftsarbeid og 0,6 årsverk ved 3. kontor i 1982. I 1983 og 1984 fullføres rutiner for integrert dataregistrering, revisjon, kontroll og feilretting via dataskjerm, det tas i bruk generelle program-pakker for analyseformål m.v. Tekniske løsninger (og planene for hovedprosjektet) baseres stort sett på bruk av maskin- og programvare som Byrået har fra før:

IBM 4341, OS/VS1	under VM/370
COBOL	(programmeringsspråk)
CMS	(for interaktiv programmering/testing)
VSAM	(aksessmetoder)
CICS	(kommunikasjonsinterface)

FORPROSJEKT

3 1 3 7 D A T A B A S E F O R K O M M U N A L Ø K O N O M I

Prosjektdeltakere:

Fra KDØ, Hamar: Tor Svelle
Fra 3. kontor: Jan-Tore Pedersen
Fra Systemkontoret: Tormod Solvin
 Reidar Waaler (prosjektledelse, rapport)
 (Halvor Strømme har deltatt i vurderingen
 av programpakker)

Referansegruppe: Liv Bjørnland og Bjørn Bleskestad, 3. kontor

Prosjektarbeidet har foregått i tidsrommet 17. august - 15. oktober 1981.

I N N H O L D

	Side
1. Innledning	6
1.1. Bakgrunn for forprosjektet	6
1.2. Mandat for forprosjektgruppen	6
1.3. Deltakere og arbeidsdeling i forprosjektgruppen	7
1.4. Gjennomføringen av arbeidet i forprosjektgruppen	7
2. Sammendrag av rapporten	7
2.1. Datagrunnlaget	8
2.2. 3. kontors ønsker om funksjoner i tilknytning til data- basen	8
2.3. Tekniske/generelle krav	8
2.4. Konkretisering og prioritering av hovedfunksjoner	9
2.5. Valg av tekniske løsninger	9
2.6. Gjennomføring av hovedprosjektet	10
2.7. Drift av databasen	10
2.8. Budsjett, finansiering	10
2.9. Driftskostnader og innsparing i forhold til nåværende rutiner	11
3. Ramme for hovedprosjektet	11
3.1. Formål, oppgaveramme	11
3.2. Tids- og kostnadsramme	12
4. Datagrunnlag	12
4.1. Omfang	12
4.1.1. Foreløpige og reviderte tall	12
4.1.2. Ikke regnskapsmessige data	13
4.1.3. Flexibilitet med hensyn til antall variable og enheter	13
4.1.4. Trinnsvis gjennomføring	13
4.2. Datainnngang	14
4.2.1. Regnskapsmessige data	14
4.2.2. Ikke regnskapsmessige data	15
5. Funksjonsbeskrivelse for forventet ny rutine	15
5.1. Integrert dataregistrering, revisjon, kontroll og feil- retting	15
5.1.1. Kontrollrutiner	16
5.1.2. Oppretting	16
5.2. Sikring av databasen	17
5.3. Tabeller	17
5.3.1. Ferdig redigerte tabeller	17
5.3.2. Tabeller som produseres årlig	18
5.3.3. Andre tabeller	18
5.4. Omregning til nasjonalregnskapets standard	18
5.5. Dataanalyse	18
5.5.1. Statistiske analyseteknikker	18
5.5.2. Grafisk presentasjon	19

5.6.	Tidsserieanalyse	19
5.7.	Sortering, seleksjon og utlisting	19
5.7.1.	Sortering	20
5.7.2.	Seleksjon	20
5.7.3.	Horisontal og vertikal utlisting av datamatriksen .	20
5.8.	Fleksibilitet, endringer i tabell- og analysebehov	20
5.9.	Automatisk omregning	20
5.9.1.	Kommuneendringer	21
5.9.2.	Endringer i kapittel- og postgrupperingen	21
5.9.3.	Omregning til enhetskostnader	22
5.9.4.	Omregning til faste priser	22
5.9.5.	Automatisk skille mellom fylkes- og kommune- regnskapstall	22
5.9.6.	Prosent-, prosentil- og gjennomsnittsberegninger .	22
5.10.	Brukervennlig uttakssystem	23
5.10.1.	Tilknytning av eksterne brukere	23
5.10.2.	Overføring av data fra databasen	23
5.10.3.	Konvergerende	23
5.10.4.	Interaktiv	23
5.11.	Datadokumentasjon	24
6.	Framtidsperspektiver	25
6.1.	Kommunikasjon med andre databasesystemer	25
6.2.	Kartproduksjon	25
6.3.	Modeller	25
6.4.	Disaggregering av budsjettall	25
7.	Maskin- og programvarehensyn	26
7.1.	Maskinkonfigurasjon IBM 4341	26
7.2.	Programutrustning	26
7.3.	Andre databaser i drift og under planlegging i Byrået	26
7.4.	Krav til filsystem/databasesoftware	27
7.4.1.	Kriterier for valg av DBMS	27
7.4.2.	Konklusjon	28
7.4.2.1.	Primær målsetting	28
7.4.2.2.	Sikkerhet	28
7.4.2.3.	Datastruktur	28
7.4.2.4.	Mulig filsystem	28
7.4.2.5.	Kommunikasjon	28
8.	Andre krav	30
8.1.	Lagringsstrukturer for statistiske databaser	30
8.2.	Dokumentasjon (spesifikasjon)	32

	Side
9. Konkretisering av hovedfunksjoner i systemet	32
9.1. Prioritert rekkefølge	32
9.2. Beskrivelse av funksjonene	33
9.2.1. Etablering av databasen med data for eksisterende magnetbåndregistre	33
9.2.2. Dokumentasjon av variablene i systemet	34
9.2.3. Bruken av variablene i systemet	35
9.2.4. Integrrert databearbeiding	36
10. Valg av tekniske løsninger	37
10.1. Forslag til registre i systemet	37
10.1.1. Hovedregistre (bare kommuneavhengige)	37
10.1.1.1. Regnskapsdata	38
10.1.1.2. Budsjettdata	39
10.1.1.3. Balansedata	39
10.1.1.4. Produkt/mengdedata	39
10.1.1.5. Sosioøkonomiske data	40
10.1.2. Hjelperegistre, kommuneavhengige	40
10.1.2.1. Kontoplan	41
10.1.2.2. Avledede data	41
10.1.2.3. Lønns- og prisindekser	41
10.1.3. Hjelperegistre, kommuneuavhengige	42
10.1.3.1. Kommunetabell	42
10.1.3.2. Kode/tekst	42
10.2. Valg av databasehåndteringssystem	42
10.2.1. Forprosjektgruppens forslag	42
10.2.2. Vurdering av alternative løsninger	43
10.3. Valg av dokumentasjonssystem	45
10.4. Valg av programpakker	46
11. Gjennomføring av hovedprosjektet	46
11.1. Systemanalyse/-design	47
11.2. Detaljert systemdesign	50
11.3. Programutvikling	51
11.4. Systemtest	52
11.5. Implementering	53
11.6. Vedlikehold/videreutvikling	54
12. Drift av databasen	54
12.1. Satsvise kjøring	54
12.2. Interaktiv databehandling	55
13. Forslag til budsjett for hovedprosjektet og for drift av databasen	55
13.1. Kostnader for hovedprosjektet	55
13.2. Finansiering av hovedprosjektet	56

	Side
13.3. Driftskostnader og innsparing i forhold til nåværende rutine	57
13.3.1. Bruk av program- og maskinvare som er finansiert på annen måte	57
13.3.2. Driftskostnader og besparelser pr. år i for- hold til nåværende rutine	58

V e d l e g g:

1. Prosjektskriv for "Database for kommunaløkonomisk statistikk"	59
2. Tidsplan for forprosjektet	67
3. Variabelliste	71
4. Dokumentasjon (spesifikasjon)	76
5. Kriterier for valg av databasesystem. Datamodell for kommunaløkonomi	89
6. Vurdering av programpakker	94
7. Fotosats ved trykking av tabeller mv. grafisk databehandling for FOB-80	96
8. Maskinkonfigurasjon IBM 4341	100
9. Beskrivelse av databaser i drift og under planlegging i Byrådet	101
10. Tids- og ressursplan for hovedprosjektet	105
11. Søknad til Kommunaldepartementet	106
Søknad til Norges allmennvitenskapelig forskningsråd	121
12. System- og databegreper	124

1. INNLEDNING

1.1. BAKGRUNN FOR FORPROSJEKTET

3. kontor tok initiativet til å starte opp arbeidet med forprosjektet i prosjektskriv om etablering av database for kommunaløkonomisk statistikk (LBj/HaR, 23/3-81 (vedlegg 1)). I prosjektmøte i Byrådet 13. mai 1981 ble det besluttet å nedsette en forprosjektgruppe for å få en nærmere konkretisering og utredning av prosjektet.

Etter kontakt med Kommunedatasentralen for Øst-Norge A/L (KDØ), Hamar, om konsulentbistand (Byråets brev av 8. juli 1981 og KDØs svarbrev av 14. juli 1981) ble konsulent Tor Svelle, KDØ, avsatt til forprosjektet på halv tid i tidsrommet 15. august til 15. oktober. Byrådet viste til avtalen med KDØ i en søknad av 8. juli 1981 til Kommunaldepartementet om støtte til forprosjektet, og departementet ga 20. juli 1981 tilsagn om inntil kr 50 000 til slik konsulentbistand.

1.2. MANDAT FOR FORPROSJEKTGRUPPEN

I Byråets brev av 8. juli 1981 til KDØ spesifiseres mandatet for forprosjektgruppen slik:

- a) Kartlegge brukernes behov ut fra det materialet som foreligger, slik at en får konkretisert hvilke funksjoner databasen skal ha.
- b) Finne fram til det databasehåndteringssystem som vil være best egnet til å dekke disse funksjonene.
- c) Vurdere dette databaseprosjektet i relasjon til andre planlagte databaseprosjekter i Byrådet, og ta kontakt med tilsvarende prosjekter som er i gang.
- d) Konkretisere hvordan hver av hovedfunksjonene til databasen bør løses.
- e) Gi et kostnadsoverslag spesifisert på utviklingskostnader, årlig vedlikehold av programmene og ajourhold av dataene. Kostnadene for leie av ferdige datapakker som ønskes tilsluttet dette systemet, skal det også gis et kostnadsoverslag over.
- f) Utarbeide en tidsplan for gjennomføringen av hovedprosjektet.

I et møte hos Aurbakken 17. august 1981 ble det i tillegg bestemt at:

- pkt. a) skal ha en tilføyelse om at hovedfunksjoner og data skal graderes slik at det blir mulig å ta standpunkt til en trinnvis gjennomføring.
- pkt. e) også skal omfatte inntekter og finansiering.

Forprosjektgruppen skulle også formulere et utkast til søknad til Kommunaldepartementet om støtte til hovedprosjektet.

1.3. DELTAKERE OG ARBEIDSDELING I FORPROSJEKTGRUPPEN

Følgende har deltatt i forprosjektarbeidet (i parentes er angitt hva de hovedsakelig har arbeidet med, bokstavene er henvisninger til spesifikasjonene i pkt. 1.2. ovenfor):

- Tor Svelle, KDØ (b, d, e, f)
- Jan Tore Pedersen, 3. kontor (a, vurdere programpakker)
- Tormod Solvin, Systemkontoret (b, c, d)
- Reidar Waaler, " (e, f, prosjektledelse, rapport)

Halvor Strømme, Systemkontoret, har deltatt i vurderingen av programpakker.

1.4. GJENNOMFØRINGEN AV ARBEIDET I FORPROSJEKTGRUPPEN

En del rapportredigering er gjort etter 15. oktober, ellers har prosjekt-deltakerne (i Hamar, Oslo og på Kongsvinger) arbeidet ca. halv tid i tidsrommet 17. august til 15. oktober.

Prosjekt møter og samordning har krevd litt ekstra tid på grunn av avstanden Hamar - Oslo - Kongsvinger, og data om ressursbehov til budsjettet for 1982-86 og til søknaden til Kommunaldepartementet måtte framskaffes på et litt ubeleilig tidspunkt i forhold til en naturlig prosjektframdrift, men arbeidet har gått stort sett etter planen (vedlegg 2).

I tilknytning til prosjektarbeidet har Solvin og Waaler besvart spørsmål om bruk av databaser i Byrået (fra Aarno Soivio, Finnland, Nordiskt kontaktmøte om informasjonsfrågor).

Utbetalingene for konsulentbestand til KDØ (Tor Svelle) vil være godt innenfor den økonomiske rammen på kr 50 000.

2. SAMMENDRAG AV RAPPORTEN

Formålet med hovedprosjektet er å etablere en database for kommunal-økonomi ved Byråets IBM-maskin på Kongsvinger innen 1. januar 1983. Systemet skal bygges videre ut med nye funksjoner i 1983 og 1984. Dette vil effektivisere databehandlingsrutinene innenfor det aktuelle statistikkområdet, samtidig med at det åpner for nye og mer avanserte muligheter.

Det forutsettes eksternt finansiering til deler av utviklingsarbeidet, og at det benyttes eksterne konsulenter til en del av systemerings- og programmeringsarbeidet i 1982.

2.1. DATAGRUNNLAGET

Databasen skal inneholde data fra regnskapsoppgaver fra hver enkelt fylkeskommune og kommune, og i tillegg noen utvalgte ikke-regnskapsmessige variable, som produktenheter som kan knyttes til fylkene og kommunenes produksjon, mengdedata som kan knyttes til produksjonsfaktorene, lønns- og prisindekser og sosioøkonomiske data.

2.2. 3. KONTORS ØNSKER OM FUNKSJONER I TILKNYTNING TIL DATABASEN

Databasen skal oppdateres med data f.o.m. 1974 (evt. -72). I første omgang benyttes data som er registrert, revidert og rettet ved hjelp av Byråets nåværende rutiner, men senere (i 1983) skal også disse rutinene legges om, slik at arbeidet kan foregå interaktivt mot databasen.

Redigering og utlisting av tabeller bør kunne foregå slik at 3. kontor enkelt skal kunne ta ut løpende tabeller, foreta endringer i tabelloppsettet og ta ut nye tabeller. Rutinene i forbindelse med redigering og trykking av tabeller/hefter bør effektiviseres.

Ved kommuneendringer, endringer i kapittel- og postgrupperinger og i en rekke andre sammenhenger er det behov for at ofte brukte omregninger skjer automatisk.

Det bør legges opp til at dataene er tilgjengelig for statistiske analyser ved hjelp av programpakker som SPSS/SCSS.

På lengre sikt bør det legges opp til muligheter for grafisk framstilling av statistikk og analyseresultater, oppbygging av modeller for konsekvensanalyser, disaggregering av budsjettall, og overføring av data fra/til andre databaser.

En forutsetning for å kunne få til dette er at det benyttes et dokumentasjonssystem som gjør det mulig å finne ut raskt (og ha kontroll med) hva dataene i databasen står for, endringer over tid mv.

2.3. TEKNISKE/GENERELLE KRAV

Byråets IBM 4341 på Kongsvinger er en forholdsvis liten edb-maskin (se vedlegg 8), men det er likevel stor nok kapasitet til kommunedatabasen. Det er også mulig å benytte operativsystemer og aksessmetoder som er i bruk i forbindelse med folketellingsdatabasen.

Kriteriene for valg av databasehåndteringssystem og for dimensjonering av teknisk utstyr gir imidlertid ett svar hvis kommunaløkonomidatabasen er eneste begrunnelse, men et annet hvis det vurderes ut fra allmenne behov i Byrådet.

Det samme gjelder for datadokumentasjonssystemene. En enkel database-løsning er godt tjent med et enkelt, selvlaget system, mens det totale behovet dekkes best ved å velge et av de som finnes på markedet.

2.4. KONKRETISERING OG PRIORITERING AV HOVEDFUNKSJONER

Forprosjektgruppen foreslår at hovedfunksjonene prioriteres slik:

1. Oppdatering av data for 1974 (-72) - 1981
2. Dokumentasjon av variablene i systemet
3. Løpende tabellproduksjon
4. Interaktiv registrering, revisjon, kontroll og feilretting
5. Statistiske analyser ved hjelp av programpakker som SPSS/SCSS
6. Tidsserieanalyser

På lengre sikt:

7. Grafisk databehandling
8. Modeller for konsekvensanalyse
9. Disaggregering av budsjettall
10. Kommunikasjon med andre databasesystemer

2.5. VALG AV TEKNISKE LØSNINGER

Prosjektgruppen foreslår at operativsystemer og aksessmetoder som brukes til folketellingsdatabasen (OS/VS1 under VM/370, CICS, VSAM) også benyttes til kommunaløkonomidatabasen. Det gir tilfredsstillende løsninger i forhold til 3. kontors kravspesifikasjon.

Det er imidlertid mye som taler for å velge et mer omfattende database-håndteringssystem, men pris, krav til maskinkapasitet mv. tilsier at mer generelle vurderinger og en allmenn bruk i Byrået må legges til grunn for en slik avgjørelse.

Ut fra de kriteriene forprosjektgruppen har funnet fram til kan en prioritert liste over aktuelle databasehåndteringssystemer i tilfelle se slik ut:

1. RAPID (utviklet av Statistics Canada)
2. Subject (effektive løsninger mht. å få tak i deler av databasen)
3. AXIS (effektivt mht. lagring og framhenting av "ferdige tabeller")
4. CS (som er stort/omfattende)

Valget av dokumentasjonssystem er, på samme måten som for databasehåndteringssystemet, avhengig av om kommunaløkonomidatabasen er eneste begrunnelse, eller om generelle behov i Byrået skal legges til grunn. Systemer som SYSDOC (norsk) og DATAMANAGER kan være aktuelle, men ut fra de forutsetningene forprosjektgruppen har foreslår vi at det utvikles et eget enkelt system som er tilpasset kommunaløkonomidatabasen.

Bruk av programpakker er aktuelt i forbindelse med

- ekstrahering av data fra databasen (filbehandlings- og omkodingsprogrammer)
- tabellhåndtering
- statistiske beregninger og analyse

Programmer som er utviklet i Byrået kan danne utgangspunkt, men må tilpasses interaktiv bruk mot IBM-maskinen. Tabellhåndteringssystemet TAB-68 kan også brukes i enkelte sammenheng. Statistiske beregninger og analyse er godt dekket med SPSS (SCSS), som Byrået kjøper inn uavhengig av kommunaløkonomidatabasen.

2.6. GJENNOMFØRING AV HOVEDPROSJEKTET

Prosjektgruppen har planlagt (se vedlegg 10) med sikte på at databasen skal etableres innen 1. januar 1983 og oppdateres med data f.o.m. 1974 (-72) som er registrert, revidert, kontrollert og rettet ved hjelp av Byråets nåværende rutiner. Løpende tabellproduksjon kjøres fra databasen fra samme tidspunkt.

Dette vil kreve 900 timeverk ved 3. kontor, 150 ved Driftskontoret, 1300 ved Systemkontoret og 2840 timeverk fra eksterne konsulenter.

I 1983 og 1984 bygges databasesystemet ut med rutiner for integrert registrering, revidering, kontroll og feilretting, og det tas i bruk programpakker for statistiske beregninger og analyser. 3. kontor har satt av 600, og Systemkontoret regner med å bruke 2 600 timeverk til dette.

2.7. DRIFT AV DATABASEN

Satsvise kjøring vil dominere den første tiden etter at systemet er tatt i bruk, senere vil interaktiv bruk ta mer over.

Databasen vil kreve diskplass til ca. 100 mill. tegn): 1/30 av diskkapasiteten pr. i dag.

Uttaket av CPU-tid til satsvis databehandling er beregnet til ca. 15 timer pr. år. Interaktiv databehandling vil etter hvert kreve 2-3 ganger så mye.

2.8. BUDSJETT, FINANSIERING

Utviklingsarbeidet i hovedprosjektet i 1982 er beregnet til å koste 275 000 kr for arbeid utført i Byrået og 640 000 kr til eksterne konsulenter, tilsammen 915 000 kr. I tillegg kommer utgifter til teknisk utstyr, 75 000 kr i Byrået og 25 000 kr utenfor Byrået, slik at de totale kostnader i 1982

blir 1 015 000 kr. Til deler av dette søkes finansieringsbistand fra Norges almenvitenskapelige forskningsråd og Kommunaldepartementet.

I 1983 regner 3. kontor med å utføre arbeid for 60 000 kr og Systemkontoret for 320 000 kr, tilsammen 380 000 kr. Det er aktuelt å søke om støtte også til dette.

2.9. DRIFTSKOSTNADER OG INNSPARING I FORHOLD TIL NÅVÆRENDE RUTINE

De største inntektene og utgiftene i forbindelse med driften av databasen lar seg ikke tallfeste innenfor rammene av forprosjektet. Det gjelder mulighetene for framtidige besparelser, fordi det blir mulig å få bedre grunnlag for beslutninger, økende inntekter til Byrådet for betalte oppdrag, og andeler av utgifter ved bruk av program- og maskinvare som brukes av mange i Byrådet og som er finansiert på annen måte.

Leie av teknisk utstyr vil utgjøre 150 000 kr pr. år eller mer, avhengig av hva en økende interaktiv bruk vil kreve av utstyr. Innsparing av 600 timeverk ved Kontoret for manuell databearbeiding og noe ved Systemkontoret og Driftskontoret, og redusert maskinleie ved Statens driftssentral, vil imidlertid føre til at Byrådets direkte utgifter bare øker med 20000 kr pr. år.

3. RAMME FOR HOVEDPROSJEKTET

3.1. FORMÅL, OPPGAVERAMME

Det skal etableres en database for kommunaløkonomi ved Byrådets IBM 4341, med sikte på å effektivisere produksjonsrutinene, bedre datatilgjengeligheten og aktualisere statistikken innenfor dette statistikkområdet. Årlige budsjett- og regnskapsdata som Byrådet samler inn for fylker og kommuner, og noen utvalgte ikke-regnskapsmessige variable, skal inn i basen.

I første omgang skal løpende tabellproduksjon kjøres fra databasen, som oppdateres med data som er behandlet i Byrådets nåværende rutiner for registrering, revisjon, kontroll og feilretting. Det skal tas i bruk et maskinstøttet dokumentasjonssystem som gjør det mulig å identifisere alle dataene i basen, også mht. endringer over tid.

Senere skal databasesystemet bygges ut med nye funksjoner. Opplegget for registrering, revisjon, kontroll og feilretting legges om (foreløpig definert som eget prosjekt), slik at dette arbeidet kan utføres interaktivt mot databasen. Programpakker som muliggjør statistiske beregninger og analyser (f.eks. SPSS/SCSS) og grafisk databehandling skal kunne brukes i tilknytning til systemet.

3.2. TIDS- OG KOSTNADSRAMME

Innen 1. januar 1983 skal databasen være etablert, og oppdateres med data for 1974 (-72) - 1981. Disse dataene er registrert, revidert, kontrollert og rettet ved hjelp av Byråets nåværende rutiner. Løpende tabellproduksjon kjøres fra databasen fra samme tidspunkt.

I 1983 og 1984 skal databasesystemet bygges ut med rutiner for interaktiv registrering, revisjon, kontroll og feilretting, og det skal tas i bruk program-pakker som muliggjør statistiske beregninger og analyser.

Det er for 1982 søkt om finansieringsbistand fra Kommunaldepartementet og Norges almentvitenskapelige forskningsråd/RFSP. Under forutsetning av at den nødvendige finansiering kan skaffes til veie i 1982, skal Systemkontoret avgi 1 årsverk pr. år i 1982-83 og -84, det øvrige system- og programmeringsarbeidet må i tilfelle utføres av eksterne konsulenter.

Ved 3. kontor vil det bli brukt 0,6 årsverk i 1982, og et nødvendig antall timer (mindre enn i -82) i 1983 og -84.

Det er forutsatt at mesteparten av programmeringen og testingen vil skje mot maskiner utenfor Byrådet (evt. ved KDØ), slik at Driftskontoret, Kongsvinger bare vil bli belastet med et lite antall timeverk i forbindelse med utviklingen av systemet.

4. DATAGRUNNLAG

4.1. OMFANG

Databasen skal inneholde regnskapsdata og i tillegg noen variable av ikke-regnskapsmessig art fra hver enkelt kommune og fylkeskommune. For tiden har vi 454 kommuner og 18 fylkeskommuner. Databasen vil dermed bestå av 472 enheter. Hver enhet vil være definert ved en 4-sifret kode.

I databasen ønsker 3. kontor å få lagt inn de regnskaps- og budsjett-data som blir innhentet. Dette omfatter skjema 2 "Nasjonaløkonomisk gruppering av utgiftene og inntektene på kommuneregnskapet", skjema 3 "Kommunens balanse-konto" og skjema 1 "Nasjonaløkonomisk gruppering av utgiftene og inntektene på kommunebudsjettet". Tallene er i 1 000 kr.

4.1.1. Foreløpige og reviderte tall

Brukerne av kontorets statistikk har sterkt behov for å få denne på et tidlig tidspunkt. Det er derfor viktig at det blir mulig å legge inn foreløpige tall som senere skiftes ut med endelige tall. Brukerne vil på denne måten få muligheten til å få en grov oversikt over den kommunaløkonomiske situasjonen på et meget tidlig tidspunkt.

4.1.2. Ikke-regnskapsmessige data

For hver enkelt fylkeskommune og kommune bør det legges inn i basen noen utvalgte fysiske variable som totalt antall innbyggere og antall innbyggere i ulike aldersgrupper, antall elever og antall klasser i grunnskolen osv. Slike data ligger lagret på forskjellige magnetbånd i Byrået. Det vil senere være aktuelt å supplere denne variabellisten med flere variable som kan knyttes direkte til kommunens utgifter til ulike formål. Foreløpige data for antall innbyggere osv. skal enkelt kunne skiftes ut med endelige tall.

For å komme fram til et bedre mål for utgifts- og inntektsendringene i kommunesektoren, er det viktig å få innarbeidet lønns- og prisindekser i systemet. Lønnsindeksene vil bli utarbeidet innenfor de enkelte kapitler. Felles for alle kapitler vil det bli lagt inn prisindekser for utstyr, nybygg og nyanlegg, vedlikehold av bygg og anlegg og for andre driftsutgifter.

4.1.3. Fleksibilitet med hensyn til antall variable og enheter

For at systemet skal kunne bli tilpasningsdyktig til utviklingen over tid, bør det bygges opp slik at det er mulig å innpasse endringer i antall enheter, kapitler og poster. Kommuner kan bli delt eller slått sammen, derfor vil det kunne være noe forskjell i antall enheter på to ulike tidspunkt. Det har skjedd noen få slike endringer i perioden 1974-80. I kapittel- og postgrupperingen vil det også bli endringer. Disse vil i første omgang være av lite omfang, men på lang sikt vil vi få en hovedrevisjon av denne inndelingen.

Også for variablene som er av ikke-regnskapsmessige karakter, vil det bli behov for utvidelser. Spesiell interesse vil det være for å få inn flere variable som er direkte knyttet til de enkelte kapitlene, slik at en etterhvert kan begynne å regne ut bl.a. enhetskostnader.

Det vil også bli behov for å legge midlertidig inn spesielle variable i basen, som skal analyseres i sammenheng med andre variable i basen.

4.1.4. Trinnvis gjennomføring

1. Trinn. Data for årene 1974(-72)-1981 legges inn fra skjema 2. For disse årene vil regnskapsdataene ligge ferdig revidert og kontrollert på magnetbånd i Byrået. I tilknytning til innleggelsen av regnskapsdataene ønsker 3. kontor at ikke-regnskapsmessige data blir lagt inn for de samme årene, disse ligger også på magnetbånd i Byrået.

2. Trinn. Årlig ajourhold av databasen fra og med årgang 1982. Inntil det er etablert et opplegg for integrert databearbeiding via skjermterminaler ved 3. kontor, vil regnskapsdataene overføres ferdig revidert og kontrollert fra magnetbånd i Byrådet til databasen (jfr. avsnitt 3).
3. Trinn. Innleggelse av budsjettall, første årgang 1983 og balanse-tall fra og med 1980 (1983).

4.2. DATAINNGANG

4.2.1. Regnskapsmessige data

Regnskapsmessige data fra kommunene kommer inn på forskjellig måte. Noen kommer på magnetbånd, andre på maskinlister, og noen er også håndskrevet.

Kommuneregnskap (skjema 2)

Av de 472 kommuner/fylkeskommuner vi har, er det 194 kommuner som er tilknyttet de interkommunale datasentralene, det er kun disse som sender magnetbånd til Byrådet. De øvrige sender enten maskinlister (100 kommuner) eller håndskrevne skjemaer (178 kommuner) som punches i Byrådet (Kongsvinger).

Antall kommuner	EDB-tilknytting	Dataoverføring 1981
194	Interkommunale datasentraler	Magnetbånd
70	Kienzle	Maskinlister som punches i Byrådet
7	NCR	" " " " "
6	Olivetti	" " " " "
6	Computer Systemer	" " " " "
4	IBM (usikkert:Oslo, Drammen, Sandefjord)	" " " " "
2	Nordlandsdata	" " " " "
5	Tandberg	" " " " "
178	Ikke EDB-kommuner	Håndskrevne lister som punches i Byrådet

Kommunebudsjett (skjema nr. 1)

I de siste år har det foregått et samarbeid mellom Kommunaldepartementet, Kommunenes sentralforbund og Byrådet når det gjelder en hurtiginnsamling av data fra kommunenes budsjett. Dataene blir revidert i Sentralforbundet og punchet i Byrådet. Alle data kommer her inn på enkle skjemaer som punches i Byrådet

(Kongsvinger). Byrådet skal i samarbeid med Kommunenes sentralforbund fra budsjettåret 1983 innhente budsjettdata fra kommunene på nye skjemaer. Det er først fra dette tidspunkt at 3. kontor ønsker å få lagt inn budsjettall i databasen.

Kommunenes balansekonto (skjema nr. 3)

Tall for kommunenes balanse innhentes av Byrådet. Disse data kommer inn på skjema og blir punchet i Byrådet (Kongsvinger). Dette skjemaet fikk sin hovedform i 1980, men vil gis noen mindre endringer i 1982. 3. kontor ønsker å legge inn disse dataene fra 1980, evt. 1983.

Konklusjon

Det er kun for utgifts-/inntektsoppgavene fra regnskapene at vi får endel av dataene inn på magnetbånd for i alt 194 kommuner. For de resterende 278 kommunene punches disse oppgavene i Byrådet. Samtlige budsjett og balanseoppstillinger blir også punchet i Byrådet.

Av de oppgaver som Byrådet innsamler, ønsker en å få lagt inn budsjettdata fra og med budsjettet for 1983, balansen fra 1980 (83) og utgifter/inntekter fra 1974.

4.2.2. Ikke-regnskapsmessige data

De ikke-regnskapsmessige data som skal legges inn i databasen, vil i første omgang i sin helhet bli hentet fra magnetbånd i Byrådet. Men basen bør bli egnet til å kunne ta inn variable som kontoret innhenter også fra andre insitusjoner. Se variabelliste i vedlegg 3.

5. FUNKSJONSBESKRIVELSE FOR FORVENTET NY RUTINE

Databasen bør organiseres på en slik måte at dataene til enhver tid er eller kan bli tilgjengelige. For de fleste formål er det ønskelig å ha interaktiv adgang til databasen, men det kan også være aktuelt med satsvise kjøringer ved større arbeidsoperasjoner.

5.1. INTEGRERT DATAREGISTRERING, REVISJON, KONTROLL OG FEILRETNING

Det er ønskelig med framskynding og forbedring av statistikken utarbeidet på grunnlag av kommunenes regnskaper. Den endelige statistikk er nå ferdig først 14-16 måneder etter utløpet av regnskapsåret. Dette skyldes dels at oppgavene har kommet sent til Byrådet, dels at bearbeidingsrutinene er meget tidkrevende. Da det er store feil og mangler i de innsendte statistikkskjemaer, er revisjonen i Byrådet meget omfattende.

Under revisjonen skal Byrådet rette opp feil som har oppstått ved overføring til statistikkskjemaene og større avvik fra regnskapsforskriftene. Målet er at detaljerte regnskapsdata på kommunenivå som publiseres av Byrådet, skal være sammenlignbare kommunene imellom.

For å kunne gjennomføre en effektivisering av produksjonsrutinene, er det nødvendig å gå over til ny teknikk. 3. kontor mener det vil kunne oppnås både framskynding og kvalitetsmessige fordeler ved bruk av data-skjerm til integrert dataklargjøring og bearbeiding av denne statistikken.

For å kunne foreta bedre revisjon, er det nødvendig å innføre maskinelle kontroller knyttet til regnskapsdata for foregående år. Under dette arbeidet og ved framskrivning og estimering av foreløpige tall, ville det være effektivt å hente fram tidsserier fra den database som foreslås etablert i 1982.

Det er også behov for å forbedre rutinene for datainnsamling og purring ved at Byrådet i større grad tar direkte kontakt med kommunene, men dette forutsetter et bedre register for forvaltning og foretak. Videre er det ønskelig å påvise feil og mangler overfor kommunene, bl.a. ved et enkelt system for utkjøring av feillister som kan sendes til kommunene. Jfr. godkjent prosjektskriv "Opplagg for integrert dataregistrering, revisjon, kontroll og feilretting via dataskjerm" (LBj/MeS, 1/4-81).

5.1.1. Kontrollrutiner

3. kontor ønsker følgende hovedtyper av kontroller:

1. Prosentavviket fra året før for hvert tall som skal legges inn
2. Sumkontroller
3. Kontroll av at det er tall i de viktigste rutene og at det ikke forekommer i ruter som ikke er i bruk. Skille mellom fylkeskommuner og kommuner i denne kontrollen
4. Kontroll av at enkeltposter ikke er unormalt høye i relasjon til til andre poster
5. Gi melding om store avvik fra gjennomsnittet pr. innbygger for henholdsvis gj. fylker og landet som helhet

5.1.2. Oppretting

I opprettingsarbeidet bør en ta i betraktning at systemet bør bygges ut slik at følgende punkter blir tatt hensyn til:

1. Når en retter et tall, får dette innvirkning på mange sumtall som automatisk bør rettes opp

2. Ved feilmelding bør det komme ut en enkel beskjed. Dersom den som reviderer ikke forstår feilen, bør det være lagt inn en detaljert forklaring på hvordan en skal gå frem for å rette opp denne feilen
3. For hver retting bør det legges inn kommentarer om hvorfor og hvordan rettelsen er foretatt. Neste årgang gis automatisk beskjed om denne rettelsen, slik at det blir sammenheng i revisjonen over tid

5.2. SIKRING AV DATABASEN

Datainnleggelse og endringer i variabler og verdier på basen skal kun bli foretatt av de personer som er ansvarlige for basen. Men dette ansvaret bør kunne delegeres for bestemte områder. 3. kontor ønsker derfor at dette sikringssystemet blir gjort så fleksibelt at ulike medarbeidere kan arbeide på ulike felt i basen, og de skal kun ha mulighet til å gjøre opprettinger innenfor sine avgrensede felt.

For eksterne brukere bør det kunne opprettes hjelpefiler utenfor databasen som de kan benytte til å legge inn data som skal analyseres mot databasen.

Men for å forsikre oss mot at noe går galt, bør en minst en gang pr. måned ta en kopi av basen, slik at det blir mulig å oppdatere basen, dersom noe skulle gå galt. Dette er spesielt viktig i innkjøringsfasen.

5.3. TABELLER

Databasen bør gjøre det mulig å redigere tabellutkjøringene på en slik måte at de kan benyttes som offsetoriginal direkte.

Dette forutsetter:

1. At utlistingen fra databasen skal kunne skje på en printer
2. At tabellteksten kan settes inn og eventuelt endres etter behov
3. At basen inneholder navnene på kommunene og fylkeskommunene. Disse skal skrives ut i tabellenes forspalte i tillegg til det 4-sifrede kommunenr.
4. At en skal kunne omredigere tabeller på en enkel måte. F.eks. bytte om på tabellkolonnenes rekkefølge

5.3.1. Ferdig redigerte tabeller

Basen må kunne lagre tabeller som Byrået gir ut, og som kan være av interesse for viktige brukergrupper. I tilknytning til disse tabellene lages det et tabellregister som fortløpende, automatisk blir oppdatert.

5.3.2. Tabeller som produseres årlig

Det skal være mulig å sette opp tabellprogrammer som benyttes f.eks. i strukturtallpublikasjoner, en gang, slik at uttaksprogrammet ikke må skrives hvert år. Når det blir flere årganger tilstede, vil en måtte gi beskjed om årgang. Uttaksprogrammet må derfor automatisk kunne endres til å gi data for sisteår. I tillegg til uttaksprogram for hver deltabell bør det også være mulig å lage et uttaksprogram som styrer utkjøringen av alle tabeller til en publikasjon, slik at tabellene kommer ut ferdig sortert og nummerert. Det må også produseres tidsserietabeller.

5.3.3. Andre tabeller

Databasen vil i økende grad bli benyttet til tabelluttak ut fra interne og eksterne brukeres egne spesifikasjoner. Det er viktig at det blir lagt opp til et brukervennlig og konverserende system for å få dette til uten for store problemer. Ved større tabellkjøringer som det ikke er rasjonelt å kjøre ut direkte, bør brukeren få første del av tabellen ut, for å se om tabellen er i overensstemmelse med behovet. Det bør for slik bruk være lagt opp til et system som gjør det enkelt å rette opp, og endre de spesifikasjoner som alt er gjort.

5.4. OMRREGNING TIL NASJONALREGNSKAPETS STANDARD

Det skal til kapittel og post knyttes koder som tilsvarer kodene i nasjonalregnskapet for produksjonssektor, art og formål, slik at tallene i kommuneregnskapene skal kunne omgrupperes til nasjonalregnskapets standard. Konkret vil det si at de fleste av variablene skal tilordnes koder som gjør det mulig å omkode kommuneregnskapet til nasjonalregnskapets standard.

Dette vil muliggjøre en utvikling der kommunenes og fylkeskommunenes økonomi på et mer detaljert nivå kan bli innarbeidet i planleggingsmodeller, f.eks. i en delmodell til Modis.

5.5. DATAANALYSE

5.5.1. Statistiske analyseteknikker

Systemet bør kunne gi beregninger som er fullt på høyde med de mest benyttede samfunnsvitenskapelige datapakker. En bør her benytte SPSS/SCSS som et sammenligningsgrunnlag.

I systemet må det ikke være noen begrensninger med hensyn til hvilke variable som kan kjøres mot hverandre. Slike begrensninger er heller ikke nødvendige, da basen ikke vil inneholde sensitive data.

Omkodinger og konstruksjoner av nye variable ut fra datamateriale bør det legges vekt på å få til så brukervennlig som mulig.

5.5.2. Grafisk presentasjon

For visuelt å få fram sammenhengen mellom variable, er det behov for grafiske løsninger. I denne forbindelse er det spesielt viktig å få innarbeidet plottdiagrammer. I flere av de nyere safmunnsvitenskapelige programpakkene er det inkludert grafisk presentasjon i tillegg til statistiske analyseteknikker.

5.6. TIDSSERIEANALYSE

Det er behov for at basen gir muligheter for å studere utviklingen i kommunaløkonomien over tid. Systemet bør bygges opp så brukervennlig at når brukeren har spesifisert de årgangene som skal analyseres samt variabelspesifikasjonen for ett av årene, finner programmet automatisk fram samme variabel for de øvrige år.

Det er i forbindelse med tidsserieanalyse at omregning til verdi-beløp regnet i faste kr og omregninger for endringer i kapittel- og post-grupperingen er sentral. Også seleksjonsrutiner for kommuner som ikke finnes i hele perioden, kommer inn her (se pkt. 5.9.).

Databasen bør legges opp slik at det er mulig å lage tidsserier for enkeltkommuner eller grupper av kommuner. Det vil være ønskelig med tids-seriediagram, slik at en grafisk vil se utviklingen over tid for ulike kommuner.

5.7. SORTERING, SELEKSJON OG UTLISTING

5.7.1. Sortering

I dette systemet ønsker vi en sorteringsrutine som gjør det mulig å sortere enhetene (kommune) i tabellutkjøringer etter andre variable enn kommunenr. Men i de overveiende fleste tabeller vil det være sortering etter kommunenr. som er det ønskelige.

5.7.2. Seleksjon

Det er behov for å trekke ut kommuner som skal analyseres for seg. F.eks. kan en ønske å lage analyser for kun en kommune, en gruppe av kommuner, f.eks. kommuner med færre enn 1 000 innbyggere eller for alle kommuner med og uten Oslo. Det vil ofte bli behov for å kjøre den samme analysen med ulik seleksjonsorden. Programmet bør derfor stille spørsmål til brukeren om denne ønsker å foreta samme utkjøring med alternative seleksjonsspesifikasjoner.

5.7.3. Horisontal og vertikal utlistering av datamatriksen

I databasen (matriksen) bør det være enkelt å få laget utlisteringer både vertikalt og horisontalt. Ved dette vil en kunne få utlistet verdiene for alle kommuner på en variabel eller alle variabelverdiene for en kommune. I siste tilfelle bør vi arbeide oss fram til en "statistisk årbok" for det kommunaløkonomiske felt i hver kommune. En ønsker her å få ut enkeltkommunenes variabelverdier for hver variabel i banken sammenholdt med variabelens gjennomsnitt for fylker og landet under ett, samt en del andre mål på fordelinger med vekt på tidsserier av de ulike variable.

Beregninger av statistiske mål for den enkelte variabel bør lagres for at slike utkjøringer som her er beskrevet, ikke skal bli unødig kostbare.

5.8. FLEKSIBILITET, ENDRINGER I TABELL- OG ANALYSEBEHOV

Over tid utvikles nye teknikker for tabell og analyse. Vi må derfor legge opp systemet på en slik måte at det blir relativt enkelt å gjøre bruk av nye programpakker eller deler av disse etterhvert som de blir utviklet. Dersom vi ikke legger stor vekt på dette, vil systemet bli foreldet på få år. Vi må derfor legge opp til et system som kan utvikles etter behov i form av nye tilleggsprogrammer.

5.9. AUTOMATISK OMREGNING

Den mest tidkrevende og kompliserte operasjon i dataarbeidet er ofte ikke selve analysen av dataene, men bearbeidingen av enhetene og variablene, slik at de blir i overensstemmelse med brukernes behov. Det er i denne fasen store muligheter for feil som senere er vanskelige å oppdage. For å forenkle denne delen av dataarbeidet, er det behov for å innføre et system slik at ofte brukte omregninger skjer automatisk.

5.9.1. Kommuneendringer

I dataanalysen har en behov for at enhetene over tid er stabile. For vår sektor er ikke dette alltid tilfelle, vi har kommunedelinger og sammenslåinger, noe som kan gjøre det problematisk å studere utviklingen over tid. De to vanligste måtene å løse dette problemet på, er enten å utelukke de kommuner som har vært berørt av kommuneendringer eller å regne om dataene til de kommunene som har hatt endringer, slik at de blir i overensstemmelse med kommunestrukturen i basisåret. Begge disse metodene har svakheter.

Omregningsmetodene er på det kommunaløkonomiske området lite egnet på grunn av at en kommunes regnskap er en helhet som vanskelig kan deles opp i overensstemmelse med en kommunedeling. Det er ikke rimelig å forvente at nyoppdelte kommuner vil ha den samme prioritering som den opprinnelige storkommune. Det vil derfor ofte være urimelig å omregne storkommunens disposisjoner til de nye kommuner.

Elimineringsmetoden gir, i motsetning til omregningsmetoden, sikre resultater for de kommuner som blir med, til gjengjeld kan det i enkelte tidsperioder bli svært få kommuner igjen i datamaterialet. Men siden det har skjedd svært få kommuneendringer i den tidsperioden som det er aktuelt å legge inn i databasen, samtidig som det heller ikke er ventet dramatiske endringer, fremtrer elimineringsmetoden som den eneste aktuelle løsning.

Konkret ønsker vi at løsningen blir slik at når brukeren har oppgitt tidsrommet som skal studeres og et basisår, vil brukeren få spørsmål om han ønsker bruk av elimineringsmetoden. Det kan forekomme situasjoner f.eks. utkjøring av tabeller, der en ikke ønsker å bruke elimineringsmetoden, for derved å få ut alle kommuner. I slike sammenhenger skal alle kommuner som har eksistert på ett eller annet tidspunkt i perioden, med som enheter. For de årene de ikke eksisterer, gis det et symbol for dette i tabellen.

Ved eliminering skal programmet automatisk sortere ut de kommuner som ikke finnes for alle årene, i den perioden brukeren ønsker å studere.

5.9.2. Endringer i kapittel- og postgrupperingen

Vi har under pkt. 1.1.3. beskrevet de endringer som er aktuelle i kapittel- og postgrupperingen. Det er ønskelig at det legges inn et omkodingsprogram, slik at det vil bli mulig å analysere kommunaløkonomiske tall over tid, selv om det skjer mindre endringer i kapittel- og postgrupperingen.

Konkret ønsker vi at systemet gir muligheter til å omkode kapittel- og postgrupperingene i det enkelte år til det basisår brukerne til enhver tid ønsker.

5.9.3. Omregning til enhetskostnader

Det bør legges opp til at systemet automatisk skal kunne omregne alle tall til kr pr. fysisk størrelse. F.eks. pr. innbygger, pr. familie, pr. elev osv.

Det er ønskelig at brukerne etter å ha spesifisert de regnskapsdata de ønsker, får spørsmål om det er behov for å kombinere disse med ikke-regnskapsmessige variable, f.eks. folketall. Dersom brukeren ønsker å benytte alle variable i tall pr. innbygger, bør programmet automatisk gjøre denne operasjonen. Ønsker brukerne å kombinere regnskapsdataene med ulike fysiske størrelser, gis det også anledning til dette på en enkel måte.

5.9.4. Omregning til faste priser

Systemet skal inneholde lønns- og prisindekser. Ved tidsserieanalyse skal brukerne gis tilbud om automatisk omregning av regnskaps-tall fra ulike år til faste priser. Brukerne skal selv ha mulighet til å velge basisår.

5.9.5. Automatisk skille mellom fylkes- og kommuneregnskapstall

Det bør gis fire valgmuligheter i uttaksprosessen:

1. Kommuneregnskap med og uten Oslo
2. Fylkeskommuneregnskap
3. Både fylkes- og kommuneregnskap, men i hver sin tabell
4. Både fylkes- og kommuneregnskap i samme tabell

5.9.6. Prosent-, prosentil- og gjennomsnittsberegninger

Gjennomsnitt skal automatisk bli beregnet for kommunene fylkesvis i tabellutkjøringen. På dette området bør systemet muliggjøre to typer av gjennomsnittsberegninger. For det første gjennomsnitt av råtallene som vi i dag benytter, for det andre gjennomsnittsberegninger av gjennomsnittene på kommunenivå.

Basen bør gi mulighet for brukerne til å velge absolutte tall, relative tall (evt. prosentueringsretning) og/eller prosentiler.

5.10. BRUKERVENNLIG UTTAKSSYSTEM

Det er behov for at systemet blir så enkelt at brukere i og utenfor Byrået selv på en enkel måte kan få ut de data de til enhver tid har behov for. For å få dette til, er det viktig at systemet er konverserende og interaktivt.

5.10.1. Tilknytning til eksterne brukere

Et siktemål må være at større brukere utenfor Byrået etterhvert skal kunne betjene seg selv via terminaltilknytning til databasen. Da det ikke skal legges inn beskyttede eller sikkerhetsmessig graderte data i basen, er det trolig ikke nødvendig å innføre noen sperringer for eksterne brukere i databasen. Dette punkt må tas opp med Sikkerhetsutvalget.

5.10.2. Overføring av data fra basen

Til forskningsprosjekter der data fra denne databasen kun skal utgjøre en mindre del av datatilfanget, vil det være hensiktsmessig at det er mulig å overføre de data som er aktuelle.

I praksis vil det være den beste løsning dersom eksterne brukere kunne koble seg til databasen, og ta ut, og overføre de data de selv ønsket. Brukerne av disse dataene vil i stor grad ha behov for å få ut dataene i bl.a. SPSS-form.

5.10.3. Konverserende

For at brukerne ikke skal måtte bruke mye tid på å lære seg et komplisert program, vil det være riktig å gjøre systemet konverserende. Et slikt system vil ved hjelp av spørsmål og veiledning i alternative svarmuligheter, føre nye og ukyndige brukere fram. Men for at de som etterhvert blir kjent med programmet ikke skal bruke mer tid enn nødvendig ved terminalen, bør systemet innledningsvis få svar på hvor spesifisert veiledning brukeren har behov for. For brukere som kjenner systemet, er det ikke nødvendig med veiledning, de vil kun ha behov for spørsmål uten forklaring.

5.10.4. Interaktiv

Ved at det er mulig å gjøre interaktive kjøring mot basen, å få resultatene fortløpende ut på egen terminal, vil brukerne fort se sine eventuelle feil. De vil da uten å kaste bort tid, kunne få endret utkjøringen.

Ved kommunaløkonomisk analyse vil en også fortløpende kunne nyansere de modeller en bygger opp, slik at de blir mest mulig overensstemmende med datagrunnlaget.

Et eksempel på en brukervennlig løsning er den konverserende og interaktive utgave av SPSS som har fått navnet SCSS. Dette programmet er lagt opp på universitetets datamaskin i Oslo.

5.11. DATADOKUMENTASJON

De variable som legges inn i databasen, må være så godt dokumentert at bruken av data blir riktig. Spesielt er det viktig å dokumentere endringer i lover og forskrifter som fører til brudd i tidsserier.

Datadokumentasjonen til hver variabel (rute) i skjema nr. 2 for kommunenes inntekter og utgifter, gis ut fra en nøyaktig kapittel- og postbeskrivelse. Ved å gi hvert kapittel og hver post en god dokumentasjon, vil en kunne få ut dokumentasjon for hver variabel (rute) i skjema 2.

	000-599	000-099	100-149
1.00			
1.07			
1.08			
1.09			
1.0			
1.10			
1.11			

Dokumentasjonen av databasen bør gis ut som egen publikasjon, som bør inneholde:

1. Variabelliste med variabelnr. og variabelbetegnelse
2. Variabeldefinisjoner og endringer i disse over tid
3. Enkel statistikk for hver variabel, 2 typer gjennomsnitt, max-min., gjennomsnittsavvik, totalsum etc. (Dette bør være en liste på samme måte som SPSS gir.) Denne listen vil være fin som en første tilnærming til materiale

6. FRAMTIDSPERSPEKTIVER

6.1. KOMMUNIKASJON MED ANDRE DATABASESYSTEMER

3. kontor ønsker at basen vil kunne tilknyttes andre databaser som kommer på kommunenivå, slik at det blir mulig, på en relativt enkel måte, å kjøre variable i en base mot variable i andre baser. Eller i det minste å overføre variable fra en base til en annen. I denne forbindelse planlegger 3. kontor å samle data bl.a. i en database for finansregnskap (jfr. prosjektskisse 30/8-81).

6.2. KARTPRODUKSJON

Kommunegrensene i Norge er i et samarbeid mellom Byrået og NSD blitt koordinatfestet for hele perioden 1837-1980. Program er anskaffet av Byrået som kan gjøre bruk av dette grunnlagsarbeidet. For å bedre presentasjonen av kommunaløkonomiske data, ønsker 3. kontor å få benytte dette kartskraveringsprogrammet i tilknytning til databasen.

6.3. MODELLER

I tilknytning til databasen ønsker vi å bygge opp delmodeller for å beregne konsekvenser av f.eks. endringer i overføringsregler, lønns- og prisnivå, lovfestede pålegg fra staten etc.

Vi ønsker bl.a. å bygge opp modeller som sammenholder kommunenes inntektsendring med en beregnet kostnadsendring som følger av lovfestede pålegg og lønns- og prisendringer. Vi ønsker også å bygge opp modeller for kommunalt behov og sammenlikne dette med den nåværende kommunale inntektsfordeling.

6.4. DISAGGREGERING AV BUDSJETTALL

For de aggregerte budsjettall kontoret innhenter, ønsker vi å lage en modell slik at disse kan disaggregeres. Vi ønsker å fordele de budsjetterte beløp på hvert kapittel, etter den innbyrdes relative sammensetning innen hvert kapittel i kommunenes regnskap, på underkapitler og poster i det disaggregerte budsjettet.

7. MASKIN- OG PROGRAMVAREHENSYN

7.1. MASKINKONFIGURASJON IBM 4341

Det er forutsatt at Byråets IBM-maskin på Kongsvinger skal benyttes til kommunaløkonomidatabasen.

Den er nå utstyrt med (se vedlegg 8):

- 60 skjermterminaler
- 5 diskstasjoner, hver med plass til ca. 650 millioner tegn
- 2 magnetbåndstasjoner
- 2 linjeskrivere

7.2. PROGRAMUTRUSTNING

IBM 4341 har operativsystemene OS/VS1 under VM/370, og CMS benyttes for interaktiv programmering/testing.

Dessuten er CICS (kommunikasjonsinterface) og VSAM (aksessmetoder) installert, og brukes i forbindelse med folketellingsdatabasen.

7.3. ANDRE DATABASER I DRIFT OG UNDER PLANLEGGING I BYRÅET

Databehandlingen i Byrådet har tradisjonelt vært og er fremdeles sterkt dominert av sekvensiell prosessering av datafiler som er lagret på magnetiske bånd. Nå vil mye av den statistiske bearbeidingen av data nettopp være av en slik sekvensiell natur, men det er først og fremst størrelsesorden på data-mengdene som har forhindret en mer utstrakt bruk av mer moderne databaseteknologi. Denne teknologien er jo basert på hurtige, direkte aksesserbare lagringsmedia som har vært en knapp og kostbar ressurs. I de senere år har tilgangen på slik lagerplass økt betraktelig, og på flere områder er dataene organisert og lagret slik at de kan søkes frem og behandles på forskjellige måter og av flere brukere og applikasjoner (samtidig). I vedlegg 9 er det gitt en kort beskrivelse av de viktigste av disse:

I drift:

- Bedrifts- og foretaksregisteret
- Personregister (del av -)
- GAB-register
- Database for Folketellingen 1980
- " " industristatistikk
- " " abonnenter (utenrikshandelstatistikk)
- " " inngivere av korttidsstatistikk
- " " utførselsdeklarasjoner
- " " personaladministrasjon
- " " innsjekking av intervjumateriell
- Systemer for statistisk analyse

Under planlegging:

- Database for makroøkonomiske tidsserier
- " " tidsskriftsirkulasjon
- " " energistatistikk

Det foreligger ønsker om å etablere:

- Database for stats- og trygderegnskap og budsjett
- " " bankregnskap
- " " andre kredittinstitusjoner og forsikring
- " " aggregerte finansregnskap

For de databasene som er laget, er det i liten grad benyttet eller utviklet generell software for definering, lagring, fremhenting, fjerning, endring, kontroll og sikring av data. Pga. at mye av dataene og databehandlingen i Byrået er så likeartet i struktur, bør det være mulig å komme langt i en standardisering på dette området. Dette vil i så fall kunne effektivisere systemutviklingen betydelig.

7.4. KRAV TIL FILSYSTEM/DATABASESOFTWARE

7.4.1. Kriterier for valg av DBMS

Nedenfor er de viktigste punktene som er listet opp i vedlegg 5 "Kriterier for valg av DBMS" trukket ut og kommentert:

Økonomiske:

- Oppfyllelse av kravspesifikasjonen medfører investering i spesialsoftware for bruk og vedlikehold av databasene (tabellgenerator, analyseverktøy, datakatalog-system)
- Økonomi i forbindelse med bruk av systemet bør være et vesentlig moment for å vurdere om systemløsningen er vellykket.

Operative:

- Primært skal dataene raskest mulig ut til bruker (online), og sekundært skal vi ha akseptable behandlingstider i batch
- Sikkerhet mot ødeleggelse, misbruk osv. må være tilstede i både online og batch
- Det må være stabilt i bruk (tilgjengelighet)

Strategiske:

- Prosjektskrivene tilsier en utstrakt bruk av databaseteknikk for å løse oppgavene
- Bruken av 4341 må bli sett i sammenheng med både 3. kontors behov og for SSB totalt

- Valg av et DBMS kan binde bruken av 4341 for andre systemer, og til en viss grad binde fremtidig utvikling i SSB (3. kontor)

Det er derfor naturlig å avvente en grundigere vurdering av SSBs behov for et DBMS.

Vårt valg av filsystem og aksessmetoder bør derfor være utifra det som eksisterer på 4341 pr. dd, eller et system/metode som ikke kommer i konflikt med de ovennevnte vurderinger.

Funksjonelle/tekniske

- For kommunikasjon bør kommunikasjonsinterfacet ikke være bundet av et bestemt DBMS
- Vertsspråk/brukerspråk må være kjent eller enkelt å ta i bruk for programmerere (COBOL, PL/1)
- Minst mulig krav til opplæring, for spesielt å utnytte ekstern bistand til utvikling av prosjektet og senere vedlikehold/videreutvikling internt (systemerere, operatører, programmerere)
- Gi mulighet for en fleksibel videreutvikling av både dette system og kjent fremtidig bruk av Byråets EDB-anlegg for andre systemer
- Systemets spesielle datastruktur må enkelt innpasses i dette filsystemet
- Endringer i datainnhold må kunne gjøres uten endringer i eksisterende program (logisk datauavhengighet) og uten endringer i filene (fysisk datauavhengighet)
- Man må ta hensyn til bruk av programpakker for fremstilling av statistikk og analysing av data (direkte mot systemets registre eller via egne filsystem)
- Bruk av høynivå brukerspråk er meget aktuelt, sammen med programpakker som nevnt foran
- Datakatalog-system (data dictionary) bør tas i bruk tidligst mulig i prosjektet. Dette er en investering, men inntjenes på sikt ved å skape oversiktlig, ryddig og dokumentert system

7.4.2. Konklusjon

7.4.2.1. Primær målsetting

Data raskest mulig ut til bruker betyr at "veien" fra der dataene er fysisk lagret og til brukeren ser dataene, på skjerm bilde eller rapport, må gjøres "kortest" mulig.

Brukerens (3. kontor) kravspesifikasjon må tilfredsstilles.

7.4.2.2. Sikkerhet

Bruk av DBMS har også noe med sikkerhet å gjøre, og det er klart at man ved ikke å benytte et DBMS må gi avkall på noe.

Vi kan her snakke om to sider ved datasikkerhet:

- sikkerhet innen dette systemet
- sikkerhet mot ødeleggelse/misbruk av data i andre systemer

Under online vil dette ikke være noe problem, selv med eksterne brukere, da all tilgang til systemet skjer mot identifikasjon av autoriserte brukere. Tvil om tilstrekkelig datasikkerhet kan oppstå i batch, der passord alltid vil være kjent av noen. Dette er et allment problem, men da slike kjøringar ikke skjer av eksterne brukere, ser vi ikke dette som noe problem.

Vi har her med både prosjektavhengige og -uavhengige aktiviteter å gjøre, for å skape et "idiotsikkert" system, og det må være rutiner rundt ethvert anlegg og system for å forebygge og oppdage misbruk av dataene.

Innen dette systemet er ikke konfidensielle data lagret, slik at våre krav til sikkerhet er begrenset.

7.4.2.3. Datastruktur

Når vi strukturerer dataene og lager forslag til systemløsning, må vi legge oss på en linje som ivaretar kravet til fleksibilitet i datainnhold og bruk som 3. kontor skisserer.

Dette mener vi til en viss grad kan oppnås ved å strukturere dataene og fysisk lagre de på en slik måte at aksessmetoder, som aktuelle DBMS bygger på, kan tas i bruk.

En slik oppstart gir systemet fleksibilitet med tanke på videreutvikling, og binder ikke utvikling av andre systemer til en bestemt DBMS.

Vi kan slå fast at dette systemets datastruktur er forholdsvis enkel, samt at systemets krav til datauavhengighet kan tilfredsstilles ved en enkel datastruktur (ref. eget punkt om datastruktur og registerbeskrivelse).

7.4.2.4. Mulig filsystem

Vi tenker her spesielt på VSAM som innehar 3 typer filer (ESDS, KSDS, RRDS) med mulighet for alternativ indeksering. VSAM er dessuten godt kjent i Byrået (Kongsvinger), og er enkelt å ta i bruk.

Standard software inneholder dessuten de nødvendige utilities.

VSAM har vært og er stabilt i bruk. Det benyttes av flere DBMS, og vil oftest være enkelt å konvertere til et DBMS med andre aksessmetoder.

7.4.2.5. Kommunikasjon

Valg av telekommunikasjonsinterface bør for vår del følge samme linje som valg av filsystem: Ikke binde systemet til et bestemt DBMS, men gi mulighet for senere valg av databasesystem som kan tilpasses Byråets utnyttelse av 4341 når en totalvurdering er gjort.

CICS som interface er valgt i de mest aktuelle DBMS, samt at CICS, som VSAM, er kjent i Byrået. Dette gjør at en anbefaling av CICS, sammen med VSAM, vil gi oss god nok fleksibilitet med tanke på overgang til et DBMS og for videreutvikling av systemet.

I CICS er dessuten de nødvendige rutiner for logging og restart ved systemutfall innebygget, og erfaringsmessig tilfredsstillende CICS de vanlige krav til datasikkerhet.

8. ANDRE KRAV

8.1. LAGRINGSSTRUKTURER FOR STATISTISKE DATABASER

Den vanligste lagringsformen i databaser er å gruppere sammen i en rekord (post) alle dataelementer (kjennetegn, variable) som hører til en enhet (objekt) eller en relasjon mellom enheter. Poster av samme type lagres oftest i sekvens i en fil. Sekvensen kan være vilkårlig eller ordnet på en bestemt måte (sortering). Filen behandles ved sekvensiell gjennomløping av alle postene eller direkte oppslag på enkeltposter. For sistnevnte metode må det opprettes søkemekanismer for de aktuelle dataelementene (primærnøkler, sekundærnøkler). Mest brukt er indekser (B-trær, inverterte lister), randomisering (hashing) og relativ adressering.

Statistiske databaser er særlig karakterisert ved at de inneholder mange dataelementer og ofte et stort antall forekomster av de viktigste posttypene. Typiske statistiske beregninger (antall, summer, gjennomsnitt) bruker som regel forholdsvis få dataelementer, men omfatter gjerne mange poster. Selekttering skjer via identifikatorer, men mest ved kategorivariabel (klassifiserende kjennetegn) som f.eks. alder, kjønn, næringsgren. Verdimengdene til disse er små og stabile. Generelt er det få oppdateringer av allerede inngitte data, mens det er en forholdsvis jevn (årlig, månedlig) tilgang på nye data (poster). Data vil sjelden bli direkte slettet da historiske data er en viktig statistikkilde. I en del statistiske databaser blir det ofte opprettet nye variable pga. endrede statistikk- og analysebehov.

Den foran omtalte rekordorienterte lagringsstrukturen er også mest benyttet i statistikk. I de tradisjonelle tapebaserte systemene er slike store rekorder med mye redundante data mest praktisk. I online databaser som i det alt vesentlige ligger på direkte aksesserbare media kan mesteparten av dobbeltlagringen (redundansen) unngås slik at rekordene blir såkalt normaliserte. I tillegg kan det benyttes et par andre teknikker på den fysiske lagringsstrukturen for bl.a. å effektivisere på plassforbruk og søking.

Ved såkalt transponering av filer splittes rekorden i flere deler. Den opprinnelig filen blir på denne måten delt opp i flere delfiler. Ved fullstendig transponering inneholder rekordene kun et dataelement. De splittede rekordene vil ha samme posisjon i sine respektive filer slik at den opprinnelige rekorden kan rekonstrueres. Denne teknikken effektiviserer den sekvensielle behandlingen hvor kun et fåtall dataelementer brukes, da det blir færre fysiske aksesser og mindre overføring av unødvendige data pga. de korte rekordene. Indekser kan brukes på vanlig måte. Å legge til nye variable vil ikke påvirke de eksisterende filene da det opprettes en ny delfil. Denne teknikken vil selvfølgelig kreve endel ekstra søketid ved operasjoner som omfatter mange dataelementer i den logiske rekorden, da disse må hentes fra flere steder (en ekstra aksess pr. dataelement ved fullstendig transponering). Databasesystemet RAPID (Statistisk Sentrabyrå i Canada, Sverige) bruker fullstendig transponerte filer. I SCSS benyttes også denne teknikken.

En annen teknikk kan kalles for invertert lagring. Dette betyr at dataelementer "trekkes ut" av posten. Verdiene til slike inverterte dataelementer lagres i egne filer sammen med referanser til postene hvor de logisk forekommer. En bestemt verdi vil derfor kun lagres en gang. Postreferansene kan være fysiske eller symbolske pekere. Dette er altså analogt med indekser (for søking) implementert som inverterte lister. For indekser er det derimot vanlig at dataelementer også er representert i posten slik at plassforbruket ikke reduseres men økes. Referanselistene vil ta en del plass og vil kunne være kostbare å administrere og søke i. Referanser kan derimot unngås ved å ordne postene på en bestemt måte. En måte å gjøre dette på, er å organisere filen som en matrise hvor hver dimensjon svarer til en invertert variabel og matriseelementene inneholder verdier til de gjenværende variablene. Denne måten gjør det også lettere å finne verdier til inverterte variable gitt at en "står på" en bestemt rekord (element i matrisen) enn ved inverterte lister. Det er kun dataelementer som benyttes for søking (kategorivariabel), som er aktuelle å invertere. Pga. at matrisene må lagres på en linearisert form, vil de søkemønstre som "matcher" denne formen være mye mer effektive enn de som krever uttak av elementer som ligger i stor fysisk avstand fra hverandre. Ved lineariseringen må en derfor preferere de dataelementer som det søkes på mest. Generelt vil invertert lagring egne seg dårlig for vanlig sekvensiell behandling hvor inverterte dataelementer inngår. Det kan derimot lett genereres standard sekvensielle filer særlig fra de nevnte matriser. De statistiske databasesystemene AXIS (Sverige) og SUBJECT (California Univ.) bruker matriseorganisering.

En kombinasjon av de nevnte metoder vil i praksis være mest aktuelt i statistiske databaser. Som en hovedregel kan en si at transponering vil være aktuelt for dataelementer som inneholder statistiske tall og der hvor sekvensiell behandling er dominerende. Kategorivariable, som antar en relativt liten og fast verdimengde, bør inverteres ved den nevnte matrisemetoden e.l. for å minske plassforbruket og for å oppnå hurtig aksessering ved en del høyt prioriterte søkemønstre. Disse metodene vil også gjøre det enklere å redefinere (legge til nye variable i) databasene, og kanskje også i større grad gjøre det mulig å utvikle en mer generell og datakatalogstyrt databasesoftware (fordi databasene vil inneholde mange, men enkle filer).

Et annet prinsipp som må benyttes ved store statistiske databaser, er å lagre de på et flernivå hierarki av lagringsmedia. Data som brukes ofte ligger på de hurtigste media, mens de som er sjelden i bruk har sitt permanente opphold på langsomme (billige) media, men bringes automatisk til et hurtig medium når de refereres. På denne måten kan selv de største databaser fungere i et online-miljø.

8.2. DOKUMENTASJON

Et dokumentasjonssystem bør ta utgangspunkt i begreper og terminologi som brukerne (3. kontor) kjenner. Samtidig er det en fordel hvis systemet dekker edb- og systemeringsbehov for dokumentasjon, og at det kan være et analyse- og struktureringshjelpemiddel. Vedlegg 4 går nærmere inn på dette.

"Data" er en særdeles viktig ressurs i Byrået. Opplysninger om dem og om hvordan de oppstår og brukes må være lett tilgjengelig. Effektiv bruk av denne ressursen er bare mulig hvis kvaliteten kan kontrolleres.

Et maskinstøttet dokumentasjonssystem er det viktigste hjelpemiddel i arbeidet med å effektivisere datadokumentasjonen. I vedlegg 4 finnes også eksempler på bruk av blanketter i SYSDOC (norsk system). Det er utviklet et avansert edb-system for dokumentasjon etter "SYSDOC-metoden" (3. kontor får benytte blankettene gratis til utvikling av kommunaløkonomidatabasen, men får da ikke tilgang til programmene som benyttes i behandlingen av dem).

9. KONKRETISERING AV HOVEDFUNKSJONER I SYSTEMET

9.1. PRIORITERT REKKEFØLGE

1. Oppdatering av data for 1974(-72) - 1981, som er registrert, revidert, kontrollert og rettet ved hjelp av Byråets nåværende rutiner
2. Dokumentasjon av variablene i systemet
3. Løpende tabellproduksjon
4. Interaktiv registrering, revisjon, kontroll og feilretting

5. Statistiske analyser ved hjelp av programpakker som SPSS/SCSS
6. Tidsserieanalyser

På lengre sikt:

7. Grafisk fremstilling av statistikk og analyseresultater
8. Oppbygging av modeller for konsekvensanalyser
9. Modell for disaggregering av budsjettall
10. Kommunikasjon med andre databasesystemer for overføring av data

9.2. BESKRIVELSE AV FUNKSJONENE

Med utgangspunkt i kravspesifikasjonen kan vi gruppere systemets hovedfunksjoner i følgende hovedpunkter:

1. Etablering av databasen med data fra eksisterende magnetbåndregistre
2. Dokumentasjon av variablene i systemet
3. Bruk av variablene i systemet
4. Integrert databearbeiding

Funksjonenes innhold og mål vil ikke være permanent, men vil over tid bli avløst av nye eller bli endret. Dette er godt beskrevet i kravspesifikasjonen, og vi vil her ikke sitere denne altfor mye, men forsøke å beskrive hvordan funksjonene er tenkt gjennomført.

9.2.1. Etablering av databasen med data fra eksisterende magnetbåndregistre

Grunnlaget for databasene som skal etableres med data fra 1972, er for årene fram til 1982 eksisterende registre i Byrået.

Ikke-regnskapsdata, fysiske data, vil bli plukket ut fra forskjellige registre avhengig av type data (produkt/mengde/sosioøkonomiske/indekser).

Dette medfører at grunnlagsregistrene må konverteres til en form og et innhold som tilsvarer det som er beskrevet i "Forslag til registre i systemet". Den endelige utformingen av registrene og beskrivelse av behandlingsregler for konverteringen må gjøres i fasene "Systemanalyse/-design" og i "Detaljert system-design".

Denne rutinen for etablering av regnskapsdatabasene for årene 1972 til 1982, vil være aktuell i sin tenkte form inntil systemet for "Integrert data-registrering, revisjon, kontroll, og feilretting via dataskjerm", er satt i drift. Dette er planlagt i 1983.

Inntil den tid må også en midlertidig rutine for ajourhold og oppretting av databasene etableres. Det er naturlig å ha ovenfornevnte system i tankene når en slik rutine for ajourhold og feilretting utformes, slik at programmene ikke er "midlertidige", men kan benyttes også i det nye systemet.

Volumet som databasene krever vil være avhengig av hvor mange år dataene skal lagres for online bruk. Vi anser det som sannsynlig at ikke alle år fra 1972 er aktuelle for direkte online bruk. Dette vil vi ta hensyn til ved systemdesign, slik at vi får databaser med innhold som rullerer på år. Er de siste 5 år en sannsynlig tidsserie, må likevel de øvrige år raskt være tilgjengelig, enten på magnetbånd eller som en egen database over "historiske data".

Bruk av Byråets registre vil vanligvis bli vurdert også med tanke på sikkerhet mot misbruk og ødeleggelse av data. Grunnlagsregistrene som skal benyttes i disse registrene er også vurdert av 3. kontor med tanke på dette. Ingen av disse registrene inneholder konfidensielle data, og 3. kontor har heller ikke funnet det nødvendig å få bruken av nåværende magnetbåndregistre i dette systemet godkjent av Byråets interne sikkerhetsutvalg.

9.2.2. Dokumentasjon av variablene i systemet

Det markedsføres ganske mange dokumentasjonssystemer. De går som regel under betegnelsen "Data Dictionary" (Datakatalog). Flesteparten av disse systemene er laget i forbindelse med et bestemt databasehåndteringssystem (DBHS) som f.eks. TOTAL, ADABAS, IDMS, DATACOM. De brukes primært for å dokumentere databaser implementert vha. tilhørende DBHS, men er også nødvendige for å utnytte andre viktige verktøy som er knyttet til et databasesystem (rapportgeneratorer, spørrespråk, høynivå programmeringsspråk, testdatageneratorer etc.). Valg av dokumentasjonssystem kan derfor ikke betraktes isolert, men må særlig ses i sammenheng med hvilket DBHS som (skal) benyttes. Nå finnes det systemer som ikke er bundet til et bestemt DBHS, men gir en viss "support" til flere databasesystemer. Det mest utbredte online-systemet av denne typen er DATAMANGER. Et annet er det norskutviklede systemet SYSDOC, som brukes for å dokumentere kravspesifikasjoner (problemorienterte, konseptuelle beskrivelser). Dataene beskrives her ved hjelp av begreper som entitetsklasse (enhet), relasjon, dataelement (variabel), rapport o.l. Eksempel på bruk er vedlagt. Også i de mer datateknisk orienterte dokumentasjonssystemene er det mulig å bruke mer konseptuelle termer, og en del dokumentasjonstyper kan defineres av brukerne selv s.a. f.eks. egen fagterminologi kan benyttes.

Systemene nevnt foran er mest brukt på administrative systemer. Nå er det ingen særlige prinsipielle skillelinjer mellom administrative databaser og statistiske databaser som inneholder individualdata (mikrodata). Ofte er de sammenfallende mens det er bruken av databasene som er helt forskjellig. Så i motsetning til hva som ofte er tilfellet med databasehåndteringssystem, så kan et datadictionary-system være like (eller vel så) velegnet i statistiske systemer som i administrative.

Som tidligere nevnt er det spesielt på makronivå mest aktuelt å integrere database og dokumentasjon pga. at slike statistiske databaser ofte er så generelle og omfattende. Den dokumentasjonen som derfor må legges inn i systemet bør derfor være så omfattende at det ikke blir behov for noen ekstra datadictionary ved siden av. Eksempler på statistiske databasesystemer med delvis integrert dokumentasjon er AXIS (RSDB), RAPID og SUBJECT.

For den overordnede dokumentasjonen, som ikke gjelder en bestemt database eller applikasjon, men totalen, trenger en et system som er mer skreddersydd for Byrået. I Sverige (SCB) har de et system (VARKAT) som muligens kan benyttes. Et generelt dokumentasjonssystem kan alternativt brukes ved at en selv definerer de dertil egnede dokumentasjonstyper (variabel, database, statistikkområdet, produsert etc.).

9.2.3. Bruken av variablene i systemet

Kravspesifikasjonen setter store krav til fleksibilitet i systemet.

Dette med hensyn på

- bruken av dataene til de formål 3. kontor har beskrevet
- de betingelsene som settes til endringer i datainnhold (midlertidige og faste endringer av antall variable)
- brukervennlighet (veiledende menyer og dokumentasjon), og det generelle krav til sikkerhet mot misbruk og ødeleggelse av data, og autorisasjon for bruk

Ut i fra kravspesifikasjonen kan vi også definere et ambisjonsnivå for systemet: Vi skal etablere registre og rutiner rundt disse som på kort sikt skal dekke 3. kontors behov for

- fremstilling av faste og spesialbestilte tabeller ved hjelp av tabellgenerator
- gjennomføring av statistiske analyser ved hjelp av DMS (programpakker) som f.eks. SPSS/SCSS, og
- gjennomføring av tidsserieanalyser

Dette betinger bruk av faste beregningsrutiner, at datene må kunne selekteres og sorteres etter forskjellige kriterier, og at utlistingen av en matrise må kunne skje både vertikalt og horisontalt.

På lang sikt skal databasene være grunnlag for

- grafisk fremstilling av statistikk og analyseresultat
- oppbygging av modeller for konsekvensanalyser
- modell for disaggregering av budsjetttall, og
- kommunikasjon med andre databasesystemer for overføring av data

Med utgangspunkt i ambisjonsnivå, tilgjengelige ressurser og det som er praktisk gjennomførbart, vil vi foreslå at utvidelsen av omfanget gjennomføres i etapper som nødvendigvis må gå over flere år. Dette er vist i planen for gjennomføring av systemet.

Vi vil ta hensyn til de forannevnte krav og ønsker i våre løsningsforslag, i den utstrekning dette er mulig i et forprosjekt.

Dette vil vi gjøre ved å foreslå registre med et innhold og en datastruktur som enkelt kan tilfredsstille kravene til fleksibilitet i f.eks. endring av antall variable, seleksjonskriterier, sorteringsrekkefølge, online - kontra batch - bruk osv.

En effektiv utnyttelse av databasene betinger bruk av både spesielle programpakker (tabellgenerator) og generelle DMS for statistikkfremstilling og analyse av dataene (SPSS(SCSS)).

Vi vil også se det som viktig å binde løsningen til bestemte DBMS, men legge mulighetene til rette for en enkel overgang til et databasehåndterings-system når dette kan ses i sammenheng med andre systemer, etter en helhetsvurdering for utnyttelse av Byråets maskin.

Dette vil vi gjøre ved å foreslå bruk av aksessmetoder og kommunikasjonsinterface som aktuelle DBMS bygger på. Konkret er dette VSAM og CICS. Dessuten vil vi kun bruke kjente høynivå programmeringsspråk (COBOL, PL/1) og interaktiv programutvikling og testing under CMS. Dette er stabile og velkjente verktøy både i SSB og KD.

Krav til sikkerhet kan ivaretas i vårt valg av software. Dette er beskrevet i pkt. 7.4.

Vi vil understreke betydningen av datasikkerhet, da denne skal gi bruker tillit til at dataene han bruker er korrekte, gi et driftssikkert system og gi bruker adgang kun til sine aktuelle data.

Forøvrig er det hensiktsmessig med et dokumentasjonsverktøy for styring og kontroll med bruken av dataene.

Den praktiske gjennomføringen av prosjektet som bygger på ovenfornevnte krav og forutsetninger, er beskrevet i kapitlet "Gjennomføring av hovedprosjektet".

9.2.4. Integrert databearbeiding

Denne funksjonen er aktuell for trinn 2 i datainnleggelsen som nevnes i pkt. 4.1.4., og vil bli gjennomført når systemet for online kontroll og revisjon av regnskapsdataene er i drift. Dette er planlagt f.o.m. 1983.

En forløper for dette systemet vil være en midlertidig ajourholds-/korreksjonsrutine for dataene som legges inn i databasene frem til dette tidspunkt.

10. VALG AV TEKNISKE LØSNINGER

10.1. FORSLAG TIL REGISTRE I SYSTEMET

Vi har funnet en hensiktsmessig inndeling i

- hovedregistre og
- hjelperegistre

og på grunnlag av datainnholdet, delt disse to gruppene i

- (fylkes)kommune-avhengige og
- (fylkes)kommune-uavhengige

Grunnlagsregistre for både hoved- og hjelperegistre er eksisterende registre i Byrådet. Dette vil være tilfelle for årene fra 1972 og frem til systemet for "Integrert dataregistrering, revisjon, kontroll og feilretting via dataskjerm" kan implementeres (1983?).

Denne inndelingen av dataene følger naturlig av datastrukturen og av den bruken av dataene som beskrives i kravspesifikasjonen fra 3. kontor.

I denne registerinndelingen er logisk struktur lik fysisk struktur, og det er i første omgang ikke aktuelt med logiske koblinger mellom registrene, men alternative indekser på hovedregistrene bør vi ha muligheten for å implementere etter behov.

Dataene i systemet har ingen mening for bruker før budsjett-/regnskaps-/balansedata er tilgjengelig, dvs. ikke-regnskapstall har **liten betydning før** de kan knyttes til regnskapsdata.

Regnskapsstatistikken for en kommune omfatter følgende regnskapsførende enheter:

- by- og herredskommuner
- fylkeskommuner

Grovt regnet vil hoved- og hjelperegistrene kreve ca. 40 mill. bytes i lagringskapasitet for 5 år.

5 år anses som ønskelig lagringstid for data som skal brukes online. Nevnte behov for lagringskapasitet gjelder alle registrene, men her er hjelperegistrene forholdsvis små. Plassbehovet for hjelperegistrene kan anslås til ca. 5 millioner bytes. Eventuelle indekser kommer også i tillegg.

10.1.1. Hovedregistre (bare kommuneavhengige)

Generelt om budsjett-, regnskaps- og balansedata:

Inndelingen i registre etter type variabel er en konsekvens av bruken av dataene og måten registrene vil bli ajourført på.

Registrene vil inneholde data for alle landets kommuner og fylkeskommuner, dvs. 454 + 18 enheter.

Regnskap og budsjett vil inneholde et variabelt antall posteringer pr. enhet. Gjennomsnittlig vil de tre registrene omfatte ca. 1 700 variabler pr. enhet pr. år.

Registrene er planlagt etablert med data fra 1972, og input til registrene vil komme fra Byrådet, senere dels fra Byrådet dels fra kommuner og dels fra KD.

Etablering av registrene for årene 1972 til systemet er i drift, vil medføre en konvertering av eksisterende registre. Denne konverteringen må inneholde en del kontrollfunksjoner, f.eks.:

- kontroll mot kontoplan
- kontroll av regnskap mot budsjett
- kontroll mot forrige års budsjett/regnskap
- logiske kontroller

Det må også opprettes ajourholdsrutiner, og disse sammen med kontrollfunksjonene, må ses i sammenheng med rutinene for kontroll og revisjon som skal etableres på et senere tidspunkt.

Nedenfor beskrives hvert av registrene:

10.1.1.1. Regnskapsdata

Input til dette register kommer fra "skjema nr. 2". Gjennomsnittlig vil det inneholde ca. 1 000 variabler pr. kommune pr. år, og være det største registeret i systemet.

Et foreløpig forslag til variable i registeret:

- år (2 pos - 2 bytes)
- (fylkes)kommunenummer (4 pos - 3 bytes)
- kapittel (5 pos . 3 bytes)
- felt for angivelse av brukte postnr. (36 pos - 36 bytes)
- variabel pr. post max. (9x36 pos - 5x36 bytes)

Lengden på hver variabel er satt til 9 posisjoner, dvs. beløp på 999 milliarder kan lagres.

Formålet med "felt for angivelse av postnr." er å spare plass, samt gjøre det mulig å lage alternative indekser på spesielle postnr. Denne løsningen medfører variabel rekordlengde.

Hvis vi regner gjennomsnittlig 12 posteringer pr. kapittel, får vi 104 bytes pr. record (dvs. pr. kapittel). Totalt pr. år for alle 472 enhetene utgjør dette: 4,7 millioner bytes.

Hvor mange år som skal være tilgjengelig online vil være avgjørende for diskplassbehovet. Anslår vi 5 år som lagringstid, betyr dette at dette registeret trenger ca. 24 millioner bytes til dataområdet. I tillegg kommer eventuelle indekser.

Data for de øvrige år kan lagres på et annet medium, f.eks. magnetbånd og derfra være lett tilgjengelig ved behov.

10.1.1.2. Budsjettdata

Input til dette register kommer fra "skjema nr. 1". Budsjettdata er mer aggregert enn regnskapsdata, og vil gjennomsnittlig inneholde ca. 400 variabler pr. kommune pr. år.

Registeret vil få tilsvarende layout som for regnskapsdata, og det foreløpig forslaget til variable i registeret vil være det samme.

Volumet til budsjettdata tilsvarer ca. 10 prosent av det regnskapsdatene. Dvs. ca. 3 millioner bytes for 5 år for alle enhetene.

10.1.1.3. Balansedata

Input til dette registeret kommer fra "skjema nr. 3". Registeret vil inneholde ca. 300 variabler pr. kommune pr. år, og vil oppta ca. 2 millioner bytes for 5 år. Variablene fremkommer ikke direkte ved regnskapsmessige posteringer, men ved å trekke sammen beløpene fra de forskjellige balansekonti.

Dataene i registeret vil være:

- år
- (fylkes)kommune
- kontor for aktiva/passiva
- post
- variabel

Disse dataene vil være av samme størrelsesorden som i de to registrene foran, men det er ikke gitt et tilsvarende løsningsforslag for rekordlayout. Etableringsår er ikke oppgitt.

10.1.1.4. Produkt/mengdedata

Dette registeret vil inneholde data som kan knyttes til produksjonen av tjenester og til produksjonsfaktorene i den enkelte kommune/fylkeskommune.

Input til registeret vil komme fra Byråets interne registre, og omfang og hyppighet i ajourhold vil sannsynligvis følge de etablerte rutiner for grunnlagsregistrene.

Eksempel på variable i registeret:

- antall klasser i forskjellige type skoler
- antall elever i forskjellige type skoler
- antall pasienter i forskjellige type institusjoner
osv.
- antall ansatte i sentraladministrasjonen
- antall ansatte i forretningsdrift
osv.

Dette gir en variabel pr. kommune pr. gruppering eller enhet.

Nødvendige data i registeret vil være:

- år
- (fylkes)kommunenr.
- variabel-type/-kode
- variabel

Antall variabeltyper vil variere, og registeret må være slik oppbygget at dette behovet for fleksibilitet dekkes. Pr. dd er antall typer ca. 25 pr. kommune pr. år. Dette gir et diskplassbehov på ca. 0,9 millioner bytes for alle enhetene for 5 år.

10.1.1.5. Sosioøkonomiske data

Dette registeret vil inneholde data om grupper av personer i forskjellige sammenheng. Input og omfang/hyppighet i ajourhold som for "produkt/mengdedata".

Eksempel på variabler i registeret:

- folketall pr. 31/12
- antall personer i forskjellige aldersgrupper (0-15 år, 16-19 år osv.)
- andel av sysselsatte ansatt i forskjellige næringer
- antall arbeidsløse
- osv.

Diverse koder vil bli lagt i et eget hjelperegister for koder o.a.

Registeret vil inneholde ca. 200 variabler pr. kommune pr. år, og vil kreve ca. 7 millioner bytes diskplass for lagring av data for 5 år.

Øvrige data i registeret som for "produkt/mengdedata", og det bør vurderes om disse to registrene kan slås sammen.

10.1.2. Hjelperegistre, kommuneavhengige

Disse registrene inneholder data som er knyttet til en spesiell enhet. Opprettes samtidig med regnskap, budsjett og balanseregistrene.

10.1.2.1. Kontoplan

Registeret skal inneholde gjeldende kontoplan for alle (fylke-) kommuner, og med link til eventuelle tidligere utgaver av kontoplanen. Registeret vil bli brukt til verifisering av kontonr./post og fremhenting av navn til rapporter.

Data i registeret:

- år
- kommune/fylkeskommune
- kontonr. (formål)
- post (art)
- tidligere kontonr./post
- kontonavn
- postnavn
- andre

En rekord som antydnet ovenfor vil oppta ca. 60 bytes, avhengig av lengden navnefeltene. Totalt for 1 år ca. 3 millioner bytes.

Ajourholdsmessig vil registeret bli forholdsvis stabilt.

10.1.2.2. Avledede data

Det vil være hensiktsmessig lagre bearbeidede data som er av en slik art at de kan brukes til generering av rapporter eller skjermbilder, og som grunnlag for analyse av dataene, på et vilkårlig tidspunkt. Betingelsen er at "avledede data" samtidig genereres på nytt hvis datagrunnlaget endres.

Hensikten med slike registre er å spare bearbeidingstid og gjentakelser av samme type beregninger. Spesielt er dette viktig i et onlinesystem, med tanke på svarstider, men også utifra økonomi er dette er "sunt" prinsipp.

Sannsynligvis vil det være mest hensiktsmessig å opprette mange slike registre etter som behovet melder seg, og at slike "avledede data" kan inngå som en del av tilbudet til bruker.

Det vil være naturlig at detaljer omkring dette behovet først kommer frem i systemdesignfasen.

10.1.2.3. Lønns- og prisindekser

Inneholder indekser for forskjellige postgrupper (f.eks. nybygg/nyanlegg, utstyr) og pr. kapittel. Indeksene genererer pr. år.

Data i registeret:

- år
- indekstype
- variabel (indeks)

Pr. år vil registeret kreve bare ca. 160 bytes. Det er mulig at denne type "register" mer hensiktsmessig bør lagres som en tabell, og tas inn i programmet som et "copy-member".

10.1.3. Hjelperegistre, kommuneuavhengige

Disse registrene inneholder data som er generelle for alle enheter.

10.1.3.1. Kommunetabell

Dette hjelperegisteret skal brukes til verifisering av kommunenr. og fremhenting av kommunenavn og forskjellige koder.

Data i registeret:

- gjeldende kommunenr.
- link til eventuelt tidligere nr.
- kommunenavn
- kommuneadresse og postnr.
- landsdelskode
- kommunetypekode
- arbeidsgiveravgiftssone
- målføre

En rekord som antydnet ovenfor vil kreve ca. 60 bytes, avhengig av lengden på navn/adressefelt. Ajourholdsmessig er også dette registeret stabilt.

10.1.3.2. Kode/tekst

Dette registeret vil inneholde samtlige koder, meldinger, headinger og kommentarer som vil forekomme i systemet: på skjermbilder i online og på rapporter i batch.

Det er viktig at dette samles ett sted, slik at koder og meldinger kan standardiseres og ikke bli "programavhengig".

Rekorden vil sannsynligvis inneholde kun kode og tekstfelt, og bli på ca.60 bytes.

10.2. VALG AV DATABASEHÅNTERINGSSYSTEM

10.2.1. Forprosjektgruppens forslag

Forprosjektgruppen foreslår at programvare som allerede finnes i Byrået (OS/VS1 under VM 370, CICS, VSAM) brukes til kommunedatabasen, selv om det kanskje ikke er riktig å bruke betegnelsen "databasehåndteringssystem" om den. Det er fullt mulig å lage en databaseløsning som tilfredsstillende de kravene brukerne (3. kontor) har ved hjelp av denne programvaren. Kjente og anerkjente aksessmetoder mv., og en enkel oppbygging av registre (filer), gjør også at det kan bli en overkommelig oppgave å ta i bruk et mer omfattende databasehåndteringssystem senere, når/hvis Byrået finner fram til et som er hensiktsmessig (se også pkt. 10.2.2. nedenfor).

Det teller med i vurderingen at folketellingsdatabasen er bygd opp på denne måten og fungerer tilfredsstillende, at mange ved Systemkontoret på Kongsvinger kjenner systemene, og at KDØ har (og bruker til daglig) de samme systemene. (Det kan være aktuelt å leie både programmerings- og maskin-kapasitet hos KDØ under utviklingen av databasesystemet).

Databaser som er i bruk eller under utvikling i Byrået pr. i dag gir, med unntak av folketellingsdatabasen, få holdepunkter for et valg av formåls-tjenlig databasehåndteringssystem til IBM-maskinen.

10.2.2. Vurdering av alternative forslag

Forprosjektgruppen innser at det hadde vært en fordel hvis tids- og ressursrammene og mandatet for forprosjektet hadde gitt anledning til å velge og å finansiere et "virkelig" databasehåndteringssystem (i motsetning til CICS, VSAM mv.). Det hadde løst en rekke problemer i forbindelse med lagring av, tilgjengelighet til og dokumentasjon og sikring av data i databasen. Databasen for kommunaløkonomi alene gir ikke tilstrekkelig grunnlag for et slik valg. Mange forhold (pris, krav til maskinkapasitet, behov for ensartede rutiner for og kommunikasjon mellom flere databaser, opplæring, oppdatering av nye funksjoner mv.) tilsier at mer generelle vurderinger må legges til grunn, og at kostnadene må vurderes i forhold til en allmenn bruk i Byrået.

Forprosjektgruppen har skaffet seg nødvendig oversikt over problemstillingen ved hjelp av en forholdsvis overflatisk vurdering av noen få systemer som er "skreddersydd" for statistiske databaser. I pkt. 7.4. og 8.1. og i vedlegg 5 finnes det en oppsummering av kriterier for valg mellom slike systemer, og noen betraktninger om generelle krav til lagringsstrukturer for statistiske databaser og om det som karakteriserer en datamodell for kommunaløkonomi. "Vanlig" administrativ databehandling er nokså forskjellig fra Byråets, derfor har vi ikke tatt standpunkt til f.eks. TOTAL (som har svært stor utbredelse).

"Statistiske databaser" kjennetegnes bl.a. ved at

- datagrunnlaget vanligvis består av et stort antall dataelementer (der tid ofte er et viktig element) og store forekomster av hver posttype.
- etter klargjøring, kontroll og korreksjon er det få eller ingen endringer i det som ligger i databasen.
- opplysningene må være lett tilgjengelige, men samtidig beskyttet mot misbruk.
- dataene må lagres på detaljnivå og det må brukes et standardisert dokumentasjonssystem, av hensyn til en variert bruk som er vanskelig å forutsi.
- ordinær statistikkproduksjon krever et effektivt opplegg som tar hensyn til at "alle" dataene i databasen berøres.

- bruken av dataene i tabellproduksjon og til statistiske beregninger/analyser krever effektiv programvare for uttak av deler av databasen (eventuelt aggregerte).

Prosjektgruppen konkluderte på et tidlig tidspunkt med at kommunedatabasen ikke gir tilstrekkelig grunnlag for en vurdering av databasehåndteringssystemer, men ut fra generelle kriterier vil en prioritert liste kunne se slik ut:

1. RAPID er utviklet av Statistics Canada for å løse problemer som tilsvarer de som Byrået har (men i større målestokk). Det ser ut til å være effektivt når det gjelder å lagre store datamengder, tilgjengelighet til enkelte variable og til deler av databasen, muligheter for å utføre analysefunksjoner mv. SCB i Sverige har tatt i bruk dette systemet, og deltar i arbeidet med å videreutvikle det.
2. SUBJECT kjennetegnes først og fremst ved sine effektive rutiner for lagring av store datamengder og ekstrahering av del- og sumdatabaser for "hovedbasen".
3. AXIS egner seg best til lagring og framhenting av "ferdige tabeller".
4. CS (Consistent System) er svært omfattende, og bygges videre ut. Krever store maskinressurser (ville kanskje ellers vært det beste alternativet). Et avansert databasehåndteringssystem (JANUS) er "bygd ut" med store mengder programvare for analyseformål mv.

Arbeidet med å velge databasehåndteringssystem til IBM 4341 kan videreføres ved å vurdere og prioritere kriteriene for valg, å undersøke hvordan aktuelle systemer tilfredsstiller kravene, og å finne ut hvordan finansieringen kan ordnes. Det kan også bli nødvendig å ta standpunkt til en eventuell utvidelse av maskinkapasiteten, hvordan brukerne helst skal koples til IBM-maskinen (via NORD?), og hvordan arbeidsdelingen mellom maskiner bør være. Den raske utviklingen tilsier imidlertid at det har lite for seg å gå for detaljert til verks. Det er langt viktigere å få en rask avklaring, tilpasset situasjonen pr. i dag, enn en som er 100 prosent riktig også på lang sikt.

Forprosjektgruppen har ikke laget alternative planer for hovedprosjektet, fordi mye avhenger av valget av databasehåndteringssystem, men regner med at det alternativet som er planlagt er det mest arbeidskrevende når det gjelder det som bør belastes kommunedatabasen (selv om det er det rimeligste totalt sett hvis kommunedatabasen er eneste begrunnelse). Det beste ville være om det kan foreligge en avklaring før hovedprosjektet startes opp, slik at de tekniske løsningene for kommunedatabasen og planene for hovedprosjektet kan endres i samsvar med dette. Uten en slik avklaring vil det være mest hensiktsmessig å benytte program- og maskinvare som allerede er i bruk i Byrået.

10.3. VALG AV DOKUMENTASJONSSYSTEM

Behovet for å styrke datadokumentasjonen og dataadministrasjonsfunksjonen i det hele er klart erkjent i Byrået (jfr. EDB-plan), da det å bedre både kvaliteten og tilgjengeligheten av statistikken i stor grad avhenger av dette. Arbeidet med standarder og retningslinjer på dette området må selvfølgelig gjelde all databehandling i Byrået, slik at en kommer fram til én standard og et dokumentasjonssystem. Det vil kreve alt for mye ressurser å dokumentere alle eksisterende systemer fullstendig ved hjelp av slike nye metoder. I praksis er det derfor i forbindelse med utvikling av nye systemer og utvidelse/omarbeiding av eksisterende systemer at en kan ta i bruk nye dokumentasjonsmetoder og verktøy. I disse tilfellene vil nytteverdien for selve utviklingsprosessen være så stor at det ikke vil kreve ekstra kostnader, bortsett fra der hvor en første gang prøver ut nye teknikker. Denne utprøvingen må for en stor del skje i prosjektene for å utvikle nye databaser (tidsseriedatabase, kommunaløkonomisk database, energistatistikk m.m.).

Databasen for økonomiske tidsserier er så generell at den i prinsippet må inneholde all dokumentasjon om dataene (metadata). Det som så skal til for å dokumentere selve denne fysiske databasen (metametadata), er såpass begrenset at det kan gjøres manuelt. Et slikt generelt databasekonsept vil kunne brukes på flere databaser (spesielt på makronivå).

I vedlegg 4 drøftes dokumentasjon (spesifikasjon) i forhold til et vel-etablert begrepsapparat innenfor statistikk, hva dokumentasjonen kan omfatte og hva den kan brukes til.

Kommunaløkonomidatabasen vil være en mer spesialisert database for primærdata. Den vil derfor ligne mer på tradisjonelt registeroppbygde databaser. Hvis det ikke blir aktuelt å bruke et spesielt databasehåndteringssystem bør en vurdere å benytte enten SYSDOC eller DATAMANAGER som dokumentasjonssystem. Alternativt kan en utvikle et eget enkelt system. I første omgang er det viktigst å benytte og prøve ut metoder for dokumentasjon. Også i denne basen vil det være aktuelt med en viss form for integrering av data og metadata. En bør derfor koordinere aktivitetene på dette området med det foran nevnte tidsseriedatabaseprosjektet.

Hvis det brukes et "virkelig" databasehåndteringssystem vil et dokumentasjonssystem normalt være tilgjengelig i forbindelse med dette. Ellers bør en beslutning om valg av system (SYSDOC, DATAMANAGER eller et eget enkelt system) falles etter en behandling i arbeidsgruppen (Møglestue) som utreder dette problemet.

Hvis ingen slik avklaring foreligger, anbefaler forprosjektgruppen at det utvikles et eget enkelt system som er tilpasset kommunaløkonomidatabasen.

10.4. VALG AV PROGRAMPAKKER

Vedlegg 6 gir en vurdering av programpakker som er aktuelle i de forskjellige aktivitetene.

Bruk av programpakker/standardprogrammer er aktuelt i forbindelse med

- ekstrahering av data fra databasen (filbehandlings- og omkodingsprogrammer)
- tabellhåndteringsprogrammer
- statistiske beregninger og analyse

Byråets egenproduserte programmer (SFB og SOP) kan danne utgangspunkt for ekstraherings- og omkodingsfunksjonen, forutsatt at de tilpasses interaktiv bruk mot IBM-maskinen.

Programmer for tabellhåndtering er allerede tilgjengelig (Tab-68), og Intervjukontorets tabellprogram kan konverteres til IBM og "gjøres interaktivt". Vi anser det som viktig at det utvikles en programmodul som forbereder bruk av fotosats i forbindelse med trykking. Dataprint A/S har laget en prøve (se vedlegg 7), og er villig til å samarbeide om en løsning som er fornuftig for begge parter.

Statistiske beregninger og analyse ser ut til å være konsentrert om SPSS (evt. SCSS). Dette vil dekke det behovet som finnes pr. i dag. (Vi regner med at dette kjøpes inn uavhengig av kommunaløkonomidatabasen.) SAS kan være et aktuelt alternativ senere.

Maskin- og programvare for grafisk databehandling i tilknytning til IBM 4341 (folketellingen) vurderes av en arbeidsgruppe, og brukerne ved 3. kontor og andre bør avvente resultatene av dette arbeidet.

11. GJENNOMFØRING AV HOVEDPROSJEKTET

Prosjektgruppen forslår at hovedprosjektet gjennomføres slik (se tids- og ressursplan for 1982 i vedlegg 10):

	Timeverk ved			
	3. kontor	Drifts- kontor	System- kontor	Eksterne kons.
1982				
Problemorientert systemering, dokumentasjon, prosjektledelse mv. ved 3. kontor	900	-	-	-
Prosjektledelse (maskinorientert del), systemanalyse, systemdesign ..	-	-	250	300
Programutvikling og testing				
- Programmer for å opprette databasen med ferdig reviderte data	-	-	-	600
- Ajourhold av filer	-	-	-	550
- Ekstraheringsprogrammer for del- og sumdatabaser	-	-	-	700
- Sortering, backup, restore	-	-	-	200
- Datadokumentasjonssystem	-	-	150	150
- Tilpasse standard tabellprogrammer	-	-	520	-
- Opplegg for faste tabeller	-	-	150	-
Systemtest, dokumentasjon	-	75	165	170
Implementering	-	75	65	170
Sum timeverk i 1982	900	150	1 300	2 840
Sum for 1983 og 1984				
Etablere rutiner for integrert revisjon, kontroll og feilretting (840 tv. system- og progr.arb.), ta i bruk generelle programpakker for statistisk analyse, vedlikeholde og videreutvikling	600	-	2 600	-
Sum for hovedprosjektet	1 500	150	3 900	2 840

De forskjellige fasene i prosjektframdriften er beskrevet nedenfor.

11.1. SYSTEMANALYSE/-DESIGN

Hovedprosjektet må starte med systemanalyse og -design, og basere seg på rapporten fra forprosjektet og brukerens kravspesifikasjon. I dette tilfellet har vi hatt en uvanlig klar kravspesifikasjon som raskt kunne legges frem i forprosjektfasen, og som muliggjorde en større grad av detaljplanlegging i forprosjektet enn det vi i utgangspunktet hadde trodd var mulig, på den forholdsvis korte tiden forprosjektet har vart.

Målet med dette aktiviteten er å komme frem til en systemløsning som er klar for detaljplanlegging og gjennomføring. Dette betinger at følgende må gjennomføres:

Detaljplanlegge systemanalysen

Enhver hovedaktivitet må planlegges med angivelse av tidsforbruk, start-/ferdigtidspunkt, ressursbehov og kostnader pr. aktivitet.

Mål-/krav-analyse

Uavhengig av hvor detaljert forprosjektrapporten er, bør den input den gir studeres:

- Er målsettingene forståelige, entydige og relevante?
- Er muligheten for måloppfyllelse tilstede?
- Er nye momenter kommet til siden forprosjektrapporten ble skrevet og etter at problemstillinger/løsningsforslag er "modnet"?

Vi må her ha muligheten til en ytterligere målpressisering, både hva angår informasjonsbehov, ambisjonsnivå og systemets samspill med omgivelsene.

Fastlegge teknisk profil

Vi må konkretisere krav til maskin- og programutstyr. Angående programutstyr har vi i utgangspunktet at sentralanlegget er IBM 4341 med eksisterende programutstyr, og vi har i budsjettforslaget for hovedprosjektet satt opp at vi skal anskaffe hensiktsmessige verktøy for programutvikling og bruk av databasene. Med programutstyr forstår vi databasesoftware/aksessmetoder, programmeringsspråk, verktøy for online programutvikling og testing, testdatagenerator, utilities, rapportgenerator og pakker for brukervennlig uttrekk, behandling og redigering av data fra registrene.

For at vi skal nyttiggjøre oss av disse verktøy må vi dessuten utarbeide nye, eller ta i bruk eksisterende metoder, standarder og prosedyrer.

Hensiktsmessige verktøy for dette prosjektet vil være:

- COBOL som programmeringsspråk
- CMS for interaktiv programutvikling og testing
- CICS som kommunikasjonsinterface
- VSAM som aksessmetode for spesielt hovedregistrene

For programutvikling kan forøvrig JSP (Jacksons Structure Programming) og METACOBOL tas i bruk som metode/verktøy sammen med COBOL som nevnt foran. Dette gir automatisk god dokumentasjon, forenkler/muliggjør programendringer av andre programmer og er dessuten kostnadsbesparende. KDØ bruker denne metoden og verktøyet i sine prosjekter og anbefaler dette.

Et datakatalogsystem bør tas i bruk så tidlig som mulig i prosjektet. Dette er en investering, men vil gi inntjening på sikt ved å skape et oversiktlig og godt dokumentert system. Deler av systemet SYSDOC kan tas i bruk spesielt for dette prosjektet, eller et system som DATAMANAGER kan installeres med tanke på bruk i flere prosjekter i SSB. Dette må imidlertid ses i sammenheng med en annen utredning i SSB om katalogsystemer.

Informasjon-/behandlingsanalyse

Dette er den viktigste aktiviteten i denne fasen, fordi den skal gi input til den detaljerte systemdesignfasen.

Vi må her kartlegge informasjonsflyten i systemet:

- hvilke data kommer inn
- hvilke data går ut av systemet
- hvilke behandlingsprosesser må foretas av inn-/utdata
- til hvilke perioder får vi inn-/utdata
- hvilke krav sikkerhet og kontroll er minimum
- har systemet naturlige "delsystem"

Resultatet av kartleggingen skal være

- detaljert kravspesifikasjon, eventuelt pr. delsystem
- spesifisering av avhengighet mellom delsystem, med tanke på funksjon, registre og tid
- vurdering av gjennomførbarhet (pr. delsystem og totalt)

Input til denne aktiviteten er kravspesifikasjonen, forprosjektrapporten og den kjennskap vi har til dagens system.

Fastlegge dokumentasjonsgrad

I løpet av prosjektperioden skal system-, bruker- og driftsdokumentasjon lages. Innhold og form av denne bør godkjennes av styringsgruppen. Brukerdokumentasjonen bør vies spesiell oppmerksomhet da dette er et onlinesystem, og formen på dokumentasjonen må være både skriftlig og være knyttet til onlinebruken av systemet. Datadokumentasjonen må ses adskilt fra dette. En hensiktsmessig form for denne er et "data dictionary"-system som nevnt foran under metode/verktøy.

Fastlegge krav til opplæring/kompetanse

For at systemet skal utvikles, implementeres, settes i drift og brukes etter intensjonene, må personell inneha en viss kompetanse eller gis opplæring.

Beslutning om videre arbeid

Alternative utfall av systemanalysen vil være:

- Prosjektet stoppes
- Systemanalysen må bearbeides videre
- Klarsignal for detaljert systemdesign gis

Når det videre arbeid startes, bør kravspesifikasjonen "fryses". Ved eventuelle endringer må det vurderes om disse kan utsettes til etter implementeringen, dvs. overføres til vedlikeholdsfasen. Er endringen av en slik natur at de er vesentlige for systemløsningen, eller innvirker på tids-, økonomi- og ressursplaner, må en eventuell umiddelbar implementering av endringene godkjennes i styringsgruppen, med de nødvendige endringer i planverket.

11.2. DETALJERT SYSTEMDESIGN

Denne fasen er en direkte fortsettelse av "Systemanalyse/-design", og baserer seg på den detaljerte kravspesifikasjonen og at klarsignal er gitt for fortsettelse av prosjektet.

Resultatet av denne fasen skal først og være input til programutviklingsfasen, men her startes også aktiviteter som vil gå parallelt til systemet er i drift.

Følgende aktiviteter skal gjennomføres i denne fasen:

- Planlegging/rapportering
Dette omfatter estimering av tidsforbruk og start/ferdig pr. aktivitet, ressursbehov og detaljkostnader for gjennomføring av aktivitetene, samt etablere de nødvendige rapporteringsprosedyrer. Vi må her ta i betraktning at flere og nye personer begynner i prosjektet.
- Valg av løsning
Alternative løsninger pr. delsystem burde ideelt sett vært vurdert og avgjort før denne fasen. I praksis opplever vi at man her kan oppleve nok en "runde" på de enkelte delsystem, etter som graden av detaljering øker.
- Detaljert registerbeskrivelse
Med utgangspunkt i registerbeskrivelsen i forprosjektrapporten, skal hvert register beskrives i detalj:
 - type (hoved/hjelpe-register, kommuneavhengig/-uavhengig)
 - informasjonsinnhold
 - form og størrelse
 - ajourholdsregler
 - nøkkelbegrep (primært og sekundært)
 - organisasjon og aksessmetoder
 - sikkerhetskrav og autorisasjon

- Program-/modul-beskrivelse
Innhold av denne aktiviteten er delvis anhengig de metoder og verktøy som skal anvendes: JSP, METACOBOL og Tabellgenerator etc. Vi må her beskrive funksjon pr. program, beskrive input/output på detaljnivå, og estimere tid pr. program/modul.
- Påbegynne dokumentasjonen (drift og bruker)
For at dette arbeidet ikke skal bli en forsinkende faktor senere i hovedprosjektet, bør dokumentasjonen påbegynnes så tidlig som mulig. Ideelt burde brukerdokumentasjonen være ferdig før selve programmeringsarbeidet startes.
- Konsekvensanalyse
Det bør foregå en kontinuerlig analyse av de følger systemutformingen får med hensyn på:
 - måloppfyllelse
 - tid og ressursbruk
 - drift
 - besparelser/inntjening
- Revurdering/beslutning om det videre arbeid
Før programutviklingsfasen og den der på følgende implementering, må det foretas en fullstendig gjennomgang av prosjektet. Dette angår den tekniske EDB-løsningen, prosjektets totale målsetting inklusiv tidsrammer og økonomi. Dette bør gjøres med deltakelse av "friske øyne" og med tanke på:
 - teknisk og applikasjonsmessig systemløsning
 - driftsmessig løsning
 - dokumentasjon (system, drift og bruker)
 Dette bør vurderes mot det mandat som gruppen har fått.
 Prosjektets videre skjebne må her avgjøres, og alternative beslutninger vil være:
 - prosjektet stoppes
 - ytterligere bearbeiding av systemløsningen
 - klarsignal gis for fortsettelse med programutvikling

11.3. PROGRAMUTVIKLING

De enkelte aktiviteter vil i stor grad være avhengig av de verktøy og metoder som man skal benytte seg av.

Følgende aktiviteter er inkludert i denne fasen:

- Planlegging
Som for de andre fasene må aktivitetene planlegges med estimat på tid og start/ferdigtidspunkt. Erfaringsmessig er det her at avvik hyppigst oppstår. Det er derfor viktig at rapporteringen fungerer og at avvik straks blir tatt opp med prosjektleder.
- Etablere prosedyreverk og hjelpemidler for programutviklingen.
Eksempelvis: Navnestandarder, pre-prosorsors (Meta-Cobol), bibliotek, CMS-rutiner, Data Dictionary, Testdata-generator etc.

- Programmering/testmateriell/testing
Innholdet vil være preget av verktøy og metoder som benyttes. Det er viktig at testingen kan foregå under en viss kontroll og med fasit slik at delsystem- og systemtesten er forberedt, men den enkelte programmerer skal selv ha ansvaret for testing av sine programmer.
- Dokumentasjon
I tillegg til dokumentasjonen programmereren får, må han detaljdokumentere hver modul og program. Måten dette gjøres på er helt avhengig av metode som benyttes, og det tilstrebes at alle benytter samme metode. Dette er av betydning for senere rettelser og utvidelser som skal kunne gjøres av andre personer.

11.4. SYSTEMTEST

Hovedansvaret for systemtesten bør ligge hos prosjektgruppen, men med sterk deltakelse fra bruker. Systemtesten vil bestå av følgende aktiviteter:

- Planlegging/rapportering
Gjennomføringen av systemtesten må planlegges med tanke på innhold, tidsplaner, ressurser og økonomi. Systemtesten vil vanligvis gå over en kort periode med deltakelse av alle i prosjektgruppen samt representant fra 3. kontor (bruker). Rapporteringen får derfor et annet preg enn under de foregående fasene, og vil ofte medføre at feil/mangler/misforståelser umiddelbart må oppklares og korrigeres. Vanligvis får dette størst innvirkning på behandlingsreglene i programmene.
- Testdata og fasit
Utarbeidelsen av testdata med fasit bør startes umiddelbart etter den detaljerte systemdesignfasen. Det er vesentlig at fasit er utarbeidet, slik at feil, misforståelser o.l. kan korrigeres umiddelbart. Testdataene bør være laget slik at de kan brukes til uttesting av både delsystem og det totale system.
- Ferdigstillelse av brukerdokumentasjonen
Brukerdokumentasjonen er viktig i systemtestfasen. Den skal selv testes og dessuten være grunnlag for integrasjonstesten.
- Testing av delsystem
Resultatet av systemtesten skal avdekke om det enkelte delsystem fungerer som spesifisert i kravspesifikasjonene.
- Integrasjonstest
Dette er den egentlige systemtesten, og skal utføres kun v.h.a. bruker- og eventuelt driftsdokumentasjonen. Her skal systemet "kjøres" og testes mot de krav til samspill mellom delsystemene som er satt i kravspesifikasjonen. Det er viktig at 3. kontor som bruker, her har stor del i testen.
Spørsmål som må besvares er:
 - Er dokumentasjonen forståelig?
 - Fungerer systemet som spesifisert?

Ved negative svar må de nødvendige rettelser utføres og hele systemtesten gjennomkjøres på nytt inntil svarene er positive.

- Utfall av systemtest
En godkjent systemtest bør være skriftlig, spesielt når man benytter konsulenter utenom SSB, og være underskrevet av de berørte parter (styringsgruppe/bruker og prosjektleder). En slik godkjenning bør også inneholde kommentarer til det videre arbeid, spesielt hvis de opprinnelige planene for implementering er forskjøvet.

11.5. IMPLEMENTERING

Her gis en oversikt over aktiviteter i forbindelse med implementeringen av EDB-systemet. I tillegg kan aktiviteter som fysiske endringer av lokaler for plassering av teknisk utstyr (skjermer, printere o.l.), omorganisering av personell og eventuelle nyansettelser komme inn i bildet. Dette bør avklares på et så tidlig tidspunkt som mulig.

Det praktiske implementeringsarbeidet omfatter følgende aktiviteter:

- Planlegging/rapportering
Nye personer kommer her inn i bildet. Det gjelder driftsavdelingen og brukeren av systemet, og disse må tildeles delansvar for gjennomføringen av fasen. Rapporteringen kan her få et mer uformelt preg, men for å sikre en planmessig gjennomføring, bør de formelle rapporteringsprosedyrer benyttes som for de andre fasene i prosjektet.
- Slutføring av dokumentasjonen
Dette gjelder spesielt driftsdokumentasjonen. Denne bør "gjennomkjøres" før den godtas.
- Opprette driftsrutiner
Dette medfører følgende aktiviteter:
 - vurdere behovet for maskin- og mannetid for daglig drift og for spesielle perioder i året hvor behovet erfaringsmessig er høyere
 - opprette programbibliotek
 - opprette lager med blanketter/formular
 - lage JCL for daglig drift av systemet i batch og online
 - lage backup/restart-rutiner (batch og online)
 - fastsette driftstid og hyppighet på batchkjøringer
 - vurdere behovet for tape/disk
- Konvertering
Her skal den endelige overgangen til ferdig system skje. For dette må regler for konvertering, aksept og godkjenningssprøve være fastlagt. Hvis det må kjøres parallelle systemer, må varigheten av dette bestemmes. Erfaring tilsier at alle er tjent med et testsystem. Dette bør man forsøke å implementere samtidig, og med data som kan gi et reelt bilde av alle funksjonene i systemet.
- Igangkjøring/overlevering
Systemet anses som godkjent når kriteriene for konvertering er oppfylt og godkjent, og systemet er i drift hos bruker.

11.6. VEDLIKEHOLD/VIDEREUTVIKLING

Etter at implementeringen er over tas forslag til systemendringer opp til vurdering. Slike endringsforslag kan eventuelt ha blitt lagret siden systemløsningen ble endelig bestemt, være planlagt videreutvikling av systemet eller komme først når systemet er satt i drift.

Dette tilsier at det avsettes personer og ressurser med tanke på en vedlikeholdsfase.

12. DRIFT AV DATABASEN

12.1. SATSVISE KJØRINGER

Det meste av kjøringene mot databasen blir i utgangspunktet satsvise kjøring: Oppdatering og kontroll av data fra kommuner og fylker, utkjøring av kontroll- og feillister, løpende tabellproduksjon mv. Slike kjøring bør likevel kunne startes opp fra 3. kontors terminal, og 3. kontor bør også ha mulighet til å ta ut resultatene på egen skjerm eller printer. En tilkopling til IBM 4341 via NORD-utstyr i Oslo, som Byrået tenker seg som en generell ordning, gir muligheter til at mer tekstbehandlingspreget databehandling kan gjøres lokalt i Oslo, og til at overføring av "jobber" og data mellom Oslo og Kongsvinger blir rasjonell.

Fra det tidspunktet systemtesten starter (planlagt til november 1982) vil det være behov for:

I Oslo:

- 1 printer
- 1 skjermterminal
- Tilgang til modem/utstyr for tilkopling til IBM 4341 via NORD

På Kongsvinger

- Diskplass til ca. 100 mill. tegn,): 1/30 av diskkapasiteten pr. i dag (databasen, 50 mill. tegn, og nødvendig plass til andre filer i forbindelse med statistikkproduksjon mv.)
- Magnetbånd til backup mv., som etter hvert vil kreve arkivplass til ca. 40 bånd.
- Uttak av CPU-tid ved IBM 4341, foreløpig beregnet til ca. 15 timer pr. år.

12.2. INTERAKTIV DATABEHANDLING

"Integrert dataregistrering, revisjon, kontroll og feilretting" er definert som et eget prosjekt, men har sterk tilknytning til databaseprosjektet. Forprosjektgruppen har ikke hatt muligheter til å vurdere i detalj hva denne funksjonen vil kreve av maskiner og utstyr eller hva den betyr arbeidsmessig. Det dreier seg om ca. 600 timeverk pr. år ved Kontoret for manuell databasearbeiding i nåværende rutine. Arbeidsomfanget kan reduseres ved å effektivisere innhenting av data fra kommuner og fylker, og ved å bedre tilbakerapportering om vanlig forekommende feil og mangler slik at kvaliteten på dataene øker.

3. kontor regner med å overta dette arbeidet uten å øke bemanningen ved kontoret.

Bruken av dataene i basen til analyseformål o.l. vil kreve små ressurser i 1983, men kravene vil øke etter hvert som programvaretilbudet blir bedre og brukerne får kjennskap til mulighetene.

Det blir behov for 2 skjermterminal^{er} til ved 3. kontor fra det tidspunkt rutine for integrert dataregistrering, revisjon, kontroll og feilretting implementeres. Senere vil antall terminaler avhenge av utviklingen mht. statistisk analyse o.l.

Forbruket av CPU-tid som en følge av interaktiv bruk av dataene i basen vil være ubetydelig fra starten, men etter hvert som mulighetene og kunnskapen om dem øker vil forbruket bli større enn for satsvis behandling (15 timer), sannsynligvis 2-3 ganger større.

13. FORSLAG TIL BUDSJETT FOR HOVEDPROSJEKTET OG FOR DRIFT AV DATABASEN

13.1. KOSTNADER FOR HOVEDPROSJEKTET

	Kroner		Sum
	I Byrået	Utenfor Byrået	
1982			
Prosjektledelse, systemering, programmering, testing, implementering:			
Ved 3. kontor	100 000	-	100 000
" Driftskontoret	15 000	-	15 000
" Systemkontoret*	160 000	-	160 000
Bruk av eksterne konsulenter	-	640 000	640 000
Sum systemutvikling i -82	275 000	640 000	915 000

Kostnader for hovedprosjektet (forts.)

	Kroner		Sum
	I Byrået	Utenfor Byrået	
Installasjon. leie, bruk av utstyr:			
1 skjermterminal til 3. kontor (2. halvår)			
1 skjermterminal til Systemkontor (hele året)			
1 printer til 3. kontor (2. halvår)			
Utstyr til programutv. hos eksterne konsulenter	75 000	25 000	100 000
Sum kostnader til hovedprosjektet i 1982	350 000	665 000	1 015 000
1983 og 1984			
Videreutvikling (rutiner for integrert dataregistrering, kontroll og feil- retting, analysefunksjoner mv.			
Ved 3. kontor.....	60 000	-	60 000
Ved Systemkontoret (herav kr 95 000 for rutine for integrert data- registrering mv.)	320 000	-	320 000
Sum for hele prosjektet	730 000	665 000	1 395 000

13.2. FINANSIERING AV HOVEDPROSJEKTET

1982

Sum kostnader for hovedprosjektet		kr 1 015 000
- Søknad til Norges almenvitenskapelige forskningsråd/RFSP om:	kr 288 000	
Søknad til Kommunaldepartementet om:	" 610 000	" 898 000
Finansieres av Byrået i 1982		kr 117 000
1983 og 1984		
Sum kostnader for hovedprosjektet		" 380 000
Finansieres av Byrået, hele prosjektet		kr 497 000

Spørsmålet om støtte til prosjektet til dekning av Byråets utgifter i 1983 og 1984 bør utredes i hovedprosjektet, når mer detaljerte planer for siste del av prosjektet foreligger.

13.3. DRIFTSKOSTNADER OG INNSPARING I FORHOLD TIL NÅVÆRENDE RUTINE

Databaseløsningen gir 3. kontor større muligheter til å yte rask og riktig service til Byråets kunder. Betalte oppdrag kan utføres raskere, og det vil være muligheter for at eksterne brukere kan få tilgang til databasen, og aktuelle programpakker, fra egen terminal. Dette vil gi offentlige myndigheter et bedre grunnlag for å treffe riktigere beslutninger raskere, men selv om det er mye å tjene, så er det vanskelig å forutsi nøyaktig de økonomiske konsekvensene av dette. Det vil også gi større inntekter til Byrådet enn nå, men det er heller ikke mulig å tallfeste dette på det nåværende tidspunkt.

3. kontor vil bli i stand til å utføre databehandling på egen hånd som de tidligere måtte ha hjelp til fra Produksjonsavdelingen. Dermed unngås dobbeltarbeid, kommunikasjonsvanskeligheter, venting på ledige russurser mv. Det regnes også med at det er mulig å få til mindre arbeidskrevende rutiner for innhenting av data fra kommuner og fylker.

Som nevnt er de virkelig store inntektene, totalt sett, ikke med, men nedenfor gis det likevel en oversikt over ressursbruk, driftskostnader og besparelser i Byrådet i forhold til nåværende rutine.

13.3.1. Bruk av maskin- og programvare som er finansiert på annen måte

IBM 4341:

- 15 timer CPU-tid pr. år til satsvis databehandling
- 10 timer, økende til ca. 50 timer, CPU-tid pr. år til interaktiv databehandling
- Diskplass til ca. 100 mill. tegn (1/30 av total kapasitet)
- Behovet for magnetbånd øker gradvis til ca. 50 Brukes til backup-formål mv. i forbindelse med databasen
- Div. programvare (grafisk databehandling, SPSS/SCSS, SAS o.l.)

Utstyr ellers, hvis tilkopling til IBM 4341 via NORD-maskin:

- Kapasitet på NORD-utstyr, sambandsutstyr og overføringslinjer
- Programvare for tekstbehandlingsfunksjoner

13.3.2. Driftskostnader og besparelser pr. år i forhold til nåværende rutine

Leie av teknisk utstyr mv.

(3 skjermterminaler og 1 printer,
sambandsutstyr, linjer) kr 150 000

Antall skjermterminaler vil øke i takt med behovet.

Innsparing:

- Maskinleie mv., Statens Driftssentral	ca.kr	80 000
- Timeverk ved Systemkontoret og Driftskontoret (usikkert)	" "	5 000
- 600 timeverk ved Kontoret for manuell databearbeiding	" "	<u>45 000</u>

Sum innsparing pr. år ca.kr 130 000

Kostnader for leie av teknisk utstyr (beregnet av Driftskontoret) er basert på bruk av IBM-utstyr og at faste kostnader ved tilkopling av dette utstyret belastes kommunedatabasen.

Direkte utgifter i forbindelse med løpende kommunaløkonomistatistikk vil m.a.o. øke med 20 000 kroner pr. år.

PROSJEKTSKRIV FOR 1982 FRA 3. KONIOR

DATABASE FOR KOMMUNALØKONOMISK STATISTIKK (STAT.NR. 3131-3136)

I. Bakgrunn og formål

En viktig forutsetning for planlegging og forskning på området kommunaløkonomi er en god og lett tilgjengelig statistikk. Det er blant annet behov for et bedre informasjonsgrunnlag for å kunne analysere utviklingen i kommunale utgifter og inntekter. På denne bakgrunn foreslås det å opprette en database for kommunaløkonomisk statistikk i Byrådet.

De årlige regnskapsoppgaver for hver enkelt fylkeskommune og kommune bør legges inn i denne databasen. Disse detaljerte regnskapsdata bearbeides og lagres på magnetbånd i Byrådet. I tillegg bør også noen utvalgte fysiske data legges inn i basen slik at det kan foretas direkte beregninger av strukturtall som inntekter/utgifter pr. innbygger, pr. gruppe av innbyggere, pr. elev osv. Det vil også være aktuelt å legge inn implisitte prisindekser for driftsutgifter og investeringer (hentet fra nasjonalregnskapet) og spesielle lønnsindekser for å kunne foreta beregninger i faste priser.

Ved å opprette en slik database i Byrådet vil det kunne oppnås effektivisering og kvalitetsforbedring både ved bearbeiding av løpende regnskapsstatistikk og ved produksjon av statistikk- og analysetabeller. En database i Byrådet vil kunne være løpende ajourholdt og godt dokumentert. Det må videre utvikles hensiktsmessige brukerprogram. For Byrådet og enkelte sentrale brukere vil det være behov for interaktiv adgang via terminal til databasen. Byrådet vil også mot bestilling raskt kunne kjøre ut forskjellige analysetabeller (tidsserie- eller tverrsnittsdata).

II. ProsjektresultatDatabasens omfang

1. Databasen for kommunaløkonomisk statistikk bør etableres for årene fra 1974 da vi har god dokumentasjon av data fra dette år. (En eventuell utvidelse bakover til 1972 kan overveies, da vi har regnskapstall på magnetbånd fra dette år.) Databasen skal senere holdes løpende ajour. Det må utvikles gode rutiner for oppdatering slik at foreløpige data som legges inn i databasen, lett skal kunne skiftes ut med endelige data.

2. Følgende informasjon skal overføres til den kommunaløkonomiske database:

(i) For hver enkelt av de 18 fylkeskommuner og de 454 kommuner skal det legges inn detaljerte regnskapsdata.

Regnskapsoppgaver ligger lagret samlet for hver enkelt årgang, på magnetbånd i Byrådet. I regnskapsoppgavene er det en maksimal inndeling i 85 kapitler med en underoppdeling i 30 poster (inntekts-/utgiftsarter). En rekke kapittel x post kombinasjoner kan ikke forekomme eller forekommer meget sjelden. Det maksimale antall mulige kombinasjoner av kapittel og post (antall variable) i et regnskap er 2 250.

[Det er mulig det er hensiktsmessig å begrense antall variable i databasen til f.eks. 1 500 og foreta aggregeringer over mindre viktige kapitler.]

Verditallene er i 1 000 kr. [Tallene bør kanskje avrundes til mill. kr med 1 desimal.]

Det vil kunne være behov for å legge inn i databasen flere versjoner av regnskapsoppgaver for samme årgang, f.eks. for å foreta analyser av sammenheng mellom budsjett og regnskapstall.

Foreløpige (ureviderte) regnskapstall skal enkelt kunne skiftes ut med reviderte og endelige tall.

(ii) For hver enkelt fylkeskommune og kommune bør det legges inn i basen noen utvalgte fysiske data som totalt antall innbyggere, antall innbyggere i forskjellige aldersgrupper, antall elever og antall klasser i grunnskolen osv. Slike data ligger lagret på forskjellige magnetbånd i Byrådet. Det vil senere være aktuelt å supplere denne variabelliste med flere fysiske størrelser som kan knyttes direkte til kapitler i kommuneregnskapet. Som eksempel på slike data kan nevnes liggedager i forskjellige helseinstitusjoner, antall kilometer vei o.l. eller antall sysselsatte.

Foreløpige data for antall innbyggere osv, skal enkelt kunne skiftes ut med endelige data.

(iii) For å komme fram til bedre mål for utgiftsendingene kan det være aktuelt å legge inn enkelte lønns- og prisindekser i denne database. Det kan utarbeides spesielle lønnsindekser for lønnsutgifter innenfor de enkelte kapitler. Ved å overføre et sett med implisitte prisindekser for vareinnsats og investeringsutgifter fra nasjonalregnskapet, kan det foretas fastprisberegninger for drifts- og investeringsutgifter i de enkelte kommuner.

Implisitte prisindekser hentet fra foreløpige nasjonalregnskap skal skiftes ut med indekser fra endelige nasjonalregnskap.

I vedlegget er det en foreløpig oversikt over variabellisten. Det vil senere bli aktuelt å legge inn hovedtall fra kommunenes balanseoppgaver.

Krav til databasen

1. Dokumentasjon av data

De variable som legges inn i databasen må være godt dokumentert slik at bruken av data blir riktig. Spesielt er det viktig å dokumentere endringer i lover og forskrifter som fører til brudd i tidsserier. Det er mulig det i databasen etterhvert kan bygges inn et system for en fullstendig og ajourholdt dokumentasjon som kan utnyttes ved produksjon og publisering av statistikken. 3. kontor må også bygge opp en viss beredskap til veiledningstjeneste og beskrivelse av data for brukere utenfor Byrådet.

2. Datatilgjengelighet

Databasen bør organiseres på en slik måte at dataene til enhver tid er eller kan bli tilgjengelige. For de fleste formål er det ønskelig å ha interaktiv adgang til databasen, men det kan også være aktuelt med satsvis kommunikasjon ved større arbeidsoperasjoner.

Hvis 3. kontor får gjennomført et prosjekt for overgang til integrert databearbeiding av kommuneregnskapene, er det nødvendig å ha direkte toveis samband med de siste årganger av databasen under revisjon og maskinkontroll.

Til databasen må det være knyttet egnede beregnings- og tabellprogram. Det vil også bli behov for programpakker for statistisk og grafisk analyse. Det vil bli behov for å skrive ut statistikk- og analysetabeller av tverrsnittsdata, kombinerte tidsserie-/tverrsnittsdata og aggregerte tidsseriedata.

Det må være beregningsrutiner for aggregeringer (over kommuner og/eller kapittel og/eller poster), for beregning av undersummer og hovedsummer og for beregning av prosentvise endringer m.v. Videre skal det beregnes forskjellige forholdstall som verditall satt i forhold til fysiske størrelser eller verditall satt i forhold til pris- og lønnsindekser.

De vanligste bearbejdede tall kan lagres i databasen hvis dette gir kostnadmessige fordeler.

Hvis 3. kontor selv kan foreta utkjøringer fra terminal, vil dette gi store muligheter til å vurdere tallmaterialet og foreta alternative beregninger og analyser. Dette vil gi bedre kunnskap om dataene. 3. kontor vil ha behov for minst en skjermterminal med printer til direkte utskrift av mindre statistikk- og analysetabeller. Etter opplæring vil 3. kontor derved selv kunne produsere forskjellige spesialtabeller til bruk i og utenfor Byrådet.

Et siktemål må være at større brukere utenfor Byrådet etterhvert skal kunne betjene seg fra terminalen. Da det ikke skal legges inn beskyttede eller sikkerhetsmessig graderte data i denne basen, er det ikke nødvendig å innføre noen sperringer ved eksterne brukeres kontakt med basen.

Oversikt over viktige brukere utenfor Byrådet

1. Kommunaldepartementet, Finansdepartementet bl.a. ved arbeidet med skatteutjæmningsproposisjon og nasjonalbudsjett/revidert nasjonalbudsjett.
2. Kirke- og undervisningsdepartementet og flere andre departement som bl.a. administrerer store overføringer til kommunesektoren.
3. Forskjellige forskningsinstitusjoner og regionale planleggere.
4. Norske Kommuners Sentralforbund.
5. Kommuner og fylkeskommuner.
6. Offentlig oppnevnte utvalg.

III. Tilknytning til andre prosjekter

1. Effektivisering ved produksjon av løpende statistikk

En framskynding av foreløpig og endelig kommuneregnskapsstatistikk vil gjennomføres hvis Byrådet utvikler et system for interaktiv databehandling via dataskjerm ved produksjon av den løpende kommuneregnskapsstatistikken. For å kunne foreta forskjellige maskinelle kontroller er det nødvendig å ha direkte aksess til regnskapsdata for foregående år, data for innbyggertall m.v. Under dette arbeid og ved framskriving og estimering av foreløpige tall ville det være nødvendig å hente fram tidsserier fra en database. Ved hjelp av databasen vil en også kunne studere avvik for enkeltkommuner mellom budsjettall, foreløpige regnskapstall beregnet ved hjelp av utvalgsundersøkelser og endelige tall.

2. Revisjon og utbygging av publikasjonen NOS: Struktur tall for kommunenes økonomi

Struktur tallspublikasjonen har blitt gitt ut fra 1974. Formålet med publikasjonen er å gi statistikk som beskriver kommunenes økonomiske stilling og som kan nyttes ved sammenlikninger mellom kommunene. I hver publikasjon er det bare gitt tall for en enkelt årgang, hovedsaklig enhetskostnader som driftsutgifter pr. innbygger, pr. elev eller pr. seng. Fra brukerhold er det blitt uttrykt sterke ønsker om revisjon og utbygging av publikasjonen. I dette arbeidet vil en database være svært nyttig og på enkelte områder også en forutsetning. Det vil da kunne kjøres ut tabeller som gir endringer i kommunenes utgifter og inntekter over tid og fordelinger mellom kommunene. Videre vil vi søke å beregne andre utgiftstall pr. enheter enn de utgiftstall pr. innbygger osv. som nå beregnes.

3. Kommunaløkonomisk analysepublikasjon

Etter å ha etablert en database i Byrådet knyttet til egnede program-pakker, ville det være enkelt å utarbeide forskjellige analysetabeller og figurer til en analysepublikasjon.

3. kontor vil så snart det er avklart om det kan etableres en database i 1982, legge fram en prosjektbeskrivelse for en slik analysepublikasjon.

IV. Ressursplan

3. kontor i Byrådet har de nødvendige ressurser til å utføre fagkontorets del av planleggingen av dette prosjekt. For å gjennomføre det, er det antagelig nødvendig å få ekstern finansiering til utgifter til systemarbeid, teknisk utstyr m.v. Et helt foreløpig utgiftsanslag vil være ca. kr 200 000 - kr 300 000. Kostnadene vil avhenge av databasens størrelse og anvendelighet (interaktiv med mulighet til å trekke ut og kombinere alle variable i databasen eller bare med mulighet for spesielle programmerte kombinasjoner av elementer).

Da de fleste data som i første omgang skal overføres til databasen, ligger lagret på magnetbånd, vil det bli svært små utgifter til dataoverføring. Etter at databasen er etablert vil det spares ressurser ved Driftskontor og Trykningskontor.

V. Tidsplan

Det ville være ønskelig å starte planlegging av dette databaseprosjekt i 1981 med det siktemål at databasen kan bli operativ sommeren 1982.

VI. Prosjektledelse

Ved 3. kontor vil prosjektet bli ledet av Jan Tore Pedersen med støtte fra en prosjektgruppe (Bjørn Bleskestad, Jan Larsen og Liv Bjørnland).

UTKAST TIL VARIABELLISTE

1. Regnskapsdata for den enkelte kommune og fylkeskommuneKommunennummer eller fylkeskommunennummer:

Kapittel: 100	Kommunestyre	Post: 09	Lønn
Kapittel: 100	Kommunestyre	Post: 14	Utstyr
Kapittel: 100	Kommunestyre	Post: 19	Vedlikehold
...	
Kapittel: 100	Kommunestyre	Post: 67	Salgsinntekter
...	
Kapittel: 800	Politi og rettspleie	Post: 09	Lønn
...	

Objekt: 700 Kasse, postgiro, bankinnskott

Objekt: 810 Ihendehaverobligasjonslån

... ..

2. Produktenheter som kan knyttes til den enkelte kommunes eller fylkeskommunes produksjon av tjenester:

Antall klasser i grunnskolen

Antall klasser i videregående skole

.....

Antall plasser ved aldershjem

Antall plasser ved kombinerte alders- og sykehjem*

Antall plasser ved somatiske sykehjem

.....

Antall barnehageplasser

.....

Km kommunale veier

Km fylkesvei

3. Mengdedata som kan knyttes til produksjonsfaktorene i den enkelte kommune eller fylkeskommune:

Kommunal sysselsetting, sentraladm.
 " , undervisning

4. Lønns- og prisindekser som kan knyttes til produksjonsfaktorene (Lønns- og prisindeksene kan være felles for alle kommuner):

Lønnsindeks
 Prisindeks, utstyr
 Prisindeks, vedlikehold av bygninger og anlegg

5. Sosioøkonomiske data

Byråets kommunetypekode
 Alternative typekoder

Folketall pr. 31/12
 Antall personer 0-15 år
 Antall personer 16-19 år

Antall personer 67 år og over
 Andel av sysselsatte ansatt i primærnæringene

Personlige skattyteres nettoinntekt pr. innbygger
 Antall arbeidsløse

VARIABELLISTE

1. Regnskapsvariable

Variabel nr.	(1) Kommuneregnskap inntekter og utgifter (skjema 2)			
1	Kapittel: 1.07	Maskinanskaffelse		Post: 000-599 Bruttoutgifter
2	" "	" "		" 000-099 Lønn
3	" "	" "		" 100-149 Utstyr
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
3600	Kapittel: 1.99	Årets regnskapsover-/underskott		Post: 900-999 Finansieringsinntekter

Variabel nr.	(2) Kommunenes balansekonto (skjema 3)			
3601	Kapittel: 2.70	Kasse, postgiro, bankinnskott		1. Kommunekassen
3602	" "	" "		12. Total sum
3603	" "	" "		13. Havnekasse
3604	" "	" "		14. Pensjonskasse
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
3964	Kapittel: 2.84	Annen gjeld		14. Pensjonskasse

Variabel nr.	(3) Kommunebudsjett (skjema 1)			
1	Kapittel: 1.07	Maskinanskaffelser		Post: 000-599 Bruttoutgifter
2	" "	" "		" 000-099 Lønn
3	" "	" "		" 100-149 Utstyr
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
3600	Kapittel: 1.99	Årets regnskapsover-/underskott		Post: 900-999 Finansieringsinntekter

Fotnoter se neste side

FOTNOTER

(1) Alle rutene i skjema 2 er benyttet som variable. I realiteten er det ikke tall i alle disse, men situasjonen kan endres og noen av disse rutene kan bli tatt i bruk, for at det ikke skal bli nødvendig å forandre variabelnummereringen. Av den grunn vil det være fordelaktig om alle rutene i skjemaet er nummerert.

(2) De skjemaene (skjema 1 og 3) som kommunenes budsjett og balanse blir innhentet på, skal omarbeides. På nåværende tidspunkt er det vanskelig å gi noen helt konkret beskrivelse av forandringene. Antall variable for kommunenes balansekonto vil ikke forandres nevneverdig. Når det gjelder kommunenes budsjett, vil det bli innhentet opplysninger sortert på kapittel og post, men mer aggregert enn i regnskapet. Det antall variable kontoret vil innhente om kommunenes budsjett, vil derfor fortsatt bli relativt ubetydelig sammenlignet med kommuneregnskapet.

(3) For de aggregerte budsjettall kontoret innhenter, ønsker en å lage en modell, slik at disse kan disaggregeres. En planlegger her å fordele de budsjetterte beløp på hvert kapittel etter den innbyrdes relative sammensetning innen hvert kapittel i kommunenes regnskap på underkapitler og poster i budsjettet. Selv om en i første omgang kommer til å få et lite antall budsjettvariable, vil det være ønskelig, for å lette oversikten for brukerne, at regnskap- og budsjettvariable har samme variabelnr.

2. IKKE-REGNSKAPSMESSIGE VARIABLE

2.1. Fysiske variable 3. kontor benytter i dag

Folketall

Elever pr. klasse

Antall senger og kurdøgn ved helseinstitusjoner

Antall elever ved grunnskoler og gymnas

Antall klasser ved grunnskoler og gymnas

Elever pr. klasse, gjennomsnitt

2.2. Variable 3. kontor har behov for av ikke-regnskapsmessig art⁽¹⁾2.2.1. Produktenheter som kan knyttes til fylkenes/kommunenenes produksjon

Antall klasser i grunnskolen K

Antall elever i grunnskolen K

Antall klasser i videregående skoler F

Antall elever i videregående skoler F

Antall aldershjem

Antall plasser ved aldershjem

Antall pasienter ved aldershjem

Km kommunal vei

Km fylkesvei

Km riksvei

Antall kombinerte alders- og sykehjem

Antall plasser ved alders- og sykehjem

Antall pasienter ved alders- og sykehjem

Antall somatiske sykehus F

Antall plasser ved somatiske sykehus F

Antall pasienter gjennomsnittlig ved somatiske sykehus F

Antall barnehager K

Antall barnehageplasser K

Antall barn som benytter barnehageplassene K

(1) Denne listen er foreløpig, og må kunne utvides ettersom behovet for å kombinere regnskapsdata med andre variable øker.

2.2.2. Mengdedata som kanknytttes til produksjonsfaktorene (disse data kan legges inn fra 1981)

Kommunalt ansatte i sentraladministrasjon (antall)

"	"	" undervisning	
"	"	" helsevern	"
"	"	" sosialomsorg	"
"	"	" kirke, kultur	"
"	"	" utb. og boligformål	"
"	"	" forretningsdrift	"
"	"	" ymse	"

2.2.3. Lønns- og prisindekser (disse indekstallene vil være felles for alle kommuner. Muligens vil de kunne beregnes for kommune og fylkeskommune uavhengig)

Prisindeks andre driftsutgifter

Prisindeks nybygg og nyanlegg (1 stk.)

Prisindeks vedlikehold av bygg og anlegg (1 stk.)

Prisindeks utstyr (1 stk.)

Lønnsindekser (18 stk.) forskjellig for hvert kapittel og for endel underkap.

Veid indeks til å beregne kommunenes inntekter i faste priser

2.2.4. Sosioøkonomiske variable

Landsdel

Befolknings tetthet

Folketall

Kommuner gruppert i 13 grupper etter folketall

Kommunetypologier (9-delt)

Aldersfordeling antall personer i ulike aldersgrupper

0-5	år
5-7	"
7-14	"
14-16	"
16-20	"
20-30	"
30-40	"
40-50	"
50-60	"
60-67	"
67-70	"
70-75	"
75-80	"
80-85	"
85-90	"
over 90	år

Antall personer i de 16 aldersgruppene, endring fra året før

Kjønnfordeling: Antall menn
Antall kvinner

Andel av innbyggere som er skattytere

Inntekt gjennomsnitt pr. skattyter

Inntekt gjennomsnitt pr. innbygger

Antall arbeidsløse

Antall uføretrygdede

Antall med sosial stønad

Skilsmisefrekvens

Antall enslige forsørgere

Yrkesfrekvens kvinner

Yrkesfrekvens menn

Antall enslige over 67 år

Antall enslige 30-67 år

Dødelighet

Sykepengedager pr. innb. gjennomsnitt

Innpendling

Utpendling

Antall husholdninger

Andel av sysselsatte ansatt i primærnæringer

Andel av sysselsatte ansatt i sekundærnæringer

Andel av sysselsatte ansatt i tertiærnæringer

Boligmassens aldersstruktur

Andel av boliger bygd før 1920

" " " " 1921-1945

" " " " 1945-1970

" " " " etter 1970

Kommunestyrets politiske sammensetning (ca. 10 variable)

DOKUMENTASJON (SPESIFIKASJON)

1. Hva skal dokumenteres?
2. Hvilke opplysninger om de forskjellige dokumentasjonsenhetene skal inngå?
3. Hvordan bruke dokumentasjon?

Datadokumentasjon (spesifikasjon)

En dokumentasjon (spesifikasjon) skal først og fremst være en brukerorientert (konseptuell) beskrivelse av problemet/virksomheten som betraktes. En bør derfor ta utgangspunkt i de begreper og den terminologi som benyttes innen det fagfeltet problemet faller inn under. Hvis man på denne bakgrunn kan finne fram til begreper som også egner seg godt for å strukturere dataene, vil en slik brukerorientert beskrivelse også være en god innfallsport til den data-tekniske dokumentasjonen. Sistnevnte kan da ofte avledes fra den konseptuelle modellen ved hjelp av enkle regler.

Innenfor statistikken finnes det et veletablert begrepsapparat for å beskrive og analysere data.

Viktige begreper i praktisk statistikk

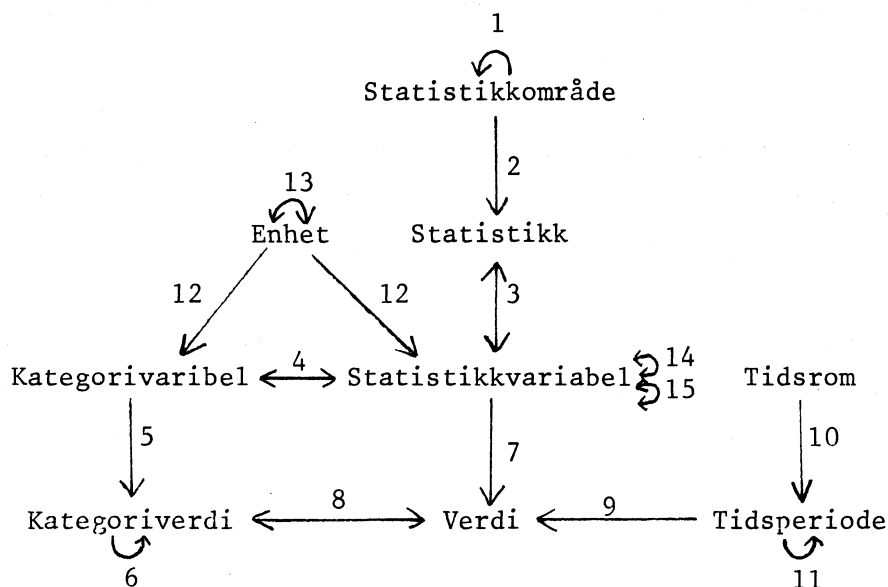
Statistikkområde:	Eks.: Regionalstatistikk
Statistikk:	De enkelte statistikker innenfor et statistikkområde
Statistisk variabel:	(Kvantitative kjennetegn, dataelementer) Eks.: Antall innbyggere, brutto nasjonalprodukt
Kategorivariabel:	(Klassifikasjonsenhet, kvalitativt kjennetegn) Eks.: Geografisk inndeling, kjønn, aldersgruppering
Tidsperiode:	(Tidsrom, tidsskala) Eks.: Måned, år (1970-1981)
Enhet:	(Objekt, element) Eks.: Person, bedrift
<u>(Relasjon)</u>	(Eks.: Person <u>er ansatt</u> i bedrift)

En statistikk vil inneholde verdier (tall) for en (eller flere) statistiske variable fordelt på 0 eller flere kategorivariabeler over et visst tidsrom. (Eks.: Inntekt pr. kommune, formål, art, tidsrom i kommunal regnskapsstatistikk.)

En variabel (kjennetegn) vil som regel være knyttet til en bestemt type enhet, beskrive en egenskap ved en enhet. Det er spesielt på mikrodatanivå (individualdata) at det er naturlig å definere enhetstyper (klasser) (person, bygning, skole, kommune etc.), og knytte dataene til disse. På makrodatanivå er

det som regel mest hensiktsmessig å strukturere dataene bare ved hjelp av kategorivariabler (klassifikasjonssystemer), da det her dreier seg om svært avledede (aggregerte) data som stammer fra et stort antall forskjellige enhets-typer på mikronivå.

Sammenhengene mellom begrepene foran kan illustreres ved figuren:



Forklaring på relasjonene:

1. Et statistikkområde kan være delt opp i flere underområder som igjen kan være videre oppdelt osv.
2. Et statistikkområde (laveste nivå) inneholder en eller flere statistikker.
3. En statistikk gjelder en eller flere variable og omvendt en statistisk variabel kan opptre på en eller flere statistikker.
4. En statistisk variabel kan være fordelt på et visst antall (0 eller flere) kategorivariable, og en slik variabel kan anvendes på flere forskjellige statistiske variable.
5. Kategorivariable antar verdier (k-verdi) og er som regel hierarkisk oppbygd (ofte i 2-5 nivåer). Aggregeringsnivåene er gitt ut fra disse.
- 6.
7. Statistisk variable antar verdier som gjelder en viss kombinasjon av
8. kategorivariable (nivå, verdi) og en bestemt tidsperiode.
- 9.
10. Tidsrom består av tidsperioder (hierarkisk).
- 11.
12. Enhet har tilknyttet variable.
13. En enhetstype kan referere til (være relatert til) andre enhetstyper.
14. En variabel kan bestå av en gruppe av andre variable.
15. En variabel kan være avledet av andre variable.

Figuren foran er en modell for å beskrive selve datagrunnlaget i praktisk statistikk (generell statistikk-database). De viktigste funksjonene (mot en slik database) vil være inngivelse (registrering og kontroll) av data, statistikkproduksjon og statistisk analyse (forskning).

Inngivelse av data skjer ved hjelp av såkalte skjemaer som inneholder de aktuelle opplysningene fra statistikkiverne. Dataene i skjemaene er gjerne disaggrerte individualdata fra givere som er representert som enheter i modellen foran.

Statistikkproduksjon dreier seg først og fremst om å lage faste publikasjoner som framstiller statistikkene på forskjellige måter (tabeller, grafiske presentasjoner etc.). De samme uttaksformene er også aktuelle ved mer tilfeldige, ad hoc-pregede aktiviteter (spørring mot databasen, videre bearbeiding av data ved analyser o.l.).

For å beskrive brukere og bruken av statistikkdataene bør altså dokumentasjonen inneholde opplysninger om statistikkiverne, produsenter, konsumenter, (input-)skjema, publikasjoner, tabeller, grafiske framstillinger etc.

Dokumentasjonsopplysninger

Dokumentasjonen (på konseptuelt nivå) skal altså være knyttet til "objekter" ("ting, elementer") som faller inn under de forskjellige begreper i praktisk statistikk (variable, tabeller etc.).

En del opplysninger vil være generelle. Det er slikt som navn, definisjoner, kommentarer etc. I tillegg vil det for hver dokumentasjonstype være spesielle opplysninger og referanser til andre elementer.

Ved å bestemme hva (dokumentasjonstyper) som skal dokumenteres, hvilke opplysninger som skal inngå, samt beskrivelsesregler, kommer en fram til en formalisert måte (standard) og dokumentere (spesifisere) på. I praksis kan dokumentasjonen utføres på spesielt utformede blanketter (skjemaer). Ved maskinell behandling av dokumentasjonen vil en benytte tilsvarende skjermbilder.

På de etterfølgende sider er noen slike "blanketter" skissert.

Dato: _____

VARIABEL Kort navn: _____ Var.-nr. _____

Beskrivende navn: _____

Synonyme navn: _____

Definisjon: _____

Datatype: _____ Måleenhet: _____

[Lovlige verdier: _____]

[Spes.kontroller: _____]

[Avledning: _____]

[Enhet: _____]

[Statistikkområde: _____]

Sign.: _____

Dato: _____

ENHET Navn: _____

Definisjon: _____

[Kommentarer: _____

_____]

Antall forekomster: _____ (gj.sn., min., max.)

Tilgang/avgang: _____

Variabel-liste: _____

<var. navnliste>

[Kontroller ved lagring/endring/sletting: _____

_____]

Sign.: _____

Dato: _____

TIDSSKALA Navn, nr.: _____

[Definisjon: _____

_____]

[<Kommentarer>: _____

_____]

[Starttidspunkt (år evt. måned, dato): _____

Perioditet (år, måned): _____

[Brudd, tidspunkt: _____]]*

[Merknader: _____

_____]

(Antall perioder med foreløpige tall.)

Sign.: _____

Dato: _____

KATEGORIVARIABEL Navn: _____

Definisjon: _____

Antall nivåer: _____

Nummersystem: _____
(tekst)

[Gjelder fra dato: _____

Gjelder til dato: _____]

Sign: _____

Dato: _____

TABELL Navn: _____

 Nummer: _____

Definisjon: _____

Produsent: _____

[Publikasjon: _____]

Periodisitet: Årlig/kvartalsvis/månedlig/ad hoc

[Statistikk: _____]

Variabel (-liste): _____

Tabellhode: Tekst, kategorivariabelliste (tidsskala)

Forspalte: Tekst, kategorivariabelliste (tidsskala)

Sign.: _____

Bruk av dokumentasjon/spesifikasjon

En fullstendig dokumentasjon representerer systeminformasjon på mange nivåer (problemorientert datateknisk). Den problemorienterte (konseptuelle) delen vil som regel i store trekk være ferdig før selve utviklingen av datasystemet starter, og vil derfor fungere som kravspesifikasjon. Denne nyttes altså under design og implementering av systemet, og senere også under vedlikehold og drift. Den datatekniske dokumentasjonen er av interesse spesielt for programmerere og databaseadministrator, og da særlig i vedlikeholdsøyemed.

Ved maskinell lagring og behandling av dokumentasjonen åpner det seg størst muligheter for effektiv utnyttelse av dokumentasjon. Den vil for det første være lettere å ajourholde, og en kan søke på og ta ut de delene en til hver tid er interessert i. Konsistenssjekking og analysering av dokumentasjonen vil også lettere kunne foretas. En slik dokumentasjonsdatabase kan dessuten brukes av forskjellige implementeringsverktøy (kompilatorer, databaseschemageneratorer, programgeneratorer) slik at systemutviklingen kan foregå mer eller mindre automatisk.

Dokumentasjonen vil altså være en aktiv del i datasystemet som den skal dokumentere. Den vil ofte være så integrert i den virkelige databasen at skillet mellom metadata (data om data) og data blir meget flytende. I såkalte kunnskapsbaser (databaser som "vet" hva de inneholder) er det en slik integrasjon av databasesystem og dokumentasjonssystem (dictionary). I store statistiske databaser som inneholder mange dataelementer er det også aktuelt med en slik struktur for å oppnå interaktive, konverserende (veiledende) brukerdialoger og høy grad av datauavhengighet i programvaren.

85

BESKRIVELSE AV DATAELEMENT

DA

9. Dato for utarbeidelse

1. Versjon VERS	2. Status STAT
	F

GENERELLE OPPLYSNINGER

2. Kort navn (max 8 kar.) KORT KOMMNR		4. Beskrivende navn (max 36 kar.) BESK KOMMUNE-NUMMER																							
5. Beskr. av SIGN TSV	6. Beskr. dato DATO 280981	7. Endr. dato ENDR	8. Klasse KLAS		9. Elementtype ELTY																				
10. Avledn. regel AVLP		12. Definisjon DEFI (se memo)		9. Elementtype ELTY																					
11. Måleenhet MALE		13. Definisjon DEFI *		9. Elementtype ELTY																					
		<p>Elementet er et nummer som identifiserer fylkeskommune eller kommune. Fylke bestemmes av elementets to første posisjoner.</p>		<table border="1"> <tr> <td>Navn (nøkkel) 1</td> <td><input checked="" type="checkbox"/></td> <td>Klokkeslett 6</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Kode 2</td> <td><input type="checkbox"/></td> <td>Tekst 7</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Teller 3</td> <td><input type="checkbox"/></td> <td>Kontroll 8</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Mengde 4</td> <td><input type="checkbox"/></td> <td>Konstant 9</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Dato 5</td> <td><input type="checkbox"/></td> <td>Annet 10</td> <td><input type="checkbox"/></td> </tr> </table>		Navn (nøkkel) 1	<input checked="" type="checkbox"/>	Klokkeslett 6	<input type="checkbox"/>	Kode 2	<input type="checkbox"/>	Tekst 7	<input type="checkbox"/>	Teller 3	<input type="checkbox"/>	Kontroll 8	<input type="checkbox"/>	Mengde 4	<input type="checkbox"/>	Konstant 9	<input type="checkbox"/>	Dato 5	<input type="checkbox"/>	Annet 10	<input type="checkbox"/>
Navn (nøkkel) 1	<input checked="" type="checkbox"/>	Klokkeslett 6	<input type="checkbox"/>																						
Kode 2	<input type="checkbox"/>	Tekst 7	<input type="checkbox"/>																						
Teller 3	<input type="checkbox"/>	Kontroll 8	<input type="checkbox"/>																						
Mengde 4	<input type="checkbox"/>	Konstant 9	<input type="checkbox"/>																						
Dato 5	<input type="checkbox"/>	Annet 10	<input type="checkbox"/>																						
14. Avledningssetning AVLS		15. Referanse REFE KOMMUNE-TABELL		16. Eksempel på forekomst EKSE 0414, 2. Vang kommune i Hedmark fylke																					

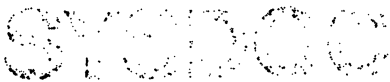
INPUT-OPPLYSNINGER

17. Skalr. faktor (10N) SKAL	18. Min. lengde MINI 4	19. Max. lengde MAXI 4	20. Blanke i streng BLAN	21. Justert mot JUSJ
			Ja 1 <input type="checkbox"/> Nei 2 <input checked="" type="checkbox"/>	Venstre 1 <input type="checkbox"/> Høyre 2 <input type="checkbox"/>
22. Kontroller som alltid skal utføres ved innlesing KONT *				23. Kontroll (se pros) KONP
Kontroll mot kommune-tabell				KOMMTAB
				24. Koding (se memo) KODM
				KODM

OUTPUT-OPPLYSNINGER

25. Max. lengde MAXU 4	26. Skalr. faktor SKAU	27. Justert mot JUSJ	28. Dekoding (se memo) DEKM
		Venstre 1 <input type="checkbox"/> Høyre 2 <input type="checkbox"/>	
29. 1. linje READ 'KOMMUNE'	30. 2. linje HEAL	31. Utskriftsregel UTSK PIC '9999'	

Copyright © 1983 GTR



DA

C. Data enh. i
DOKU

1. Versjon VERS

2. Status STAT

F

GENERELLE OPPLYSNINGER

3. Kort navn (max. 8 kar.) KONT LØNN		4. Beskrivende navn (max. 35 kar.) BESK LØNN-PR-UKAP-PR-POST						
5. Beskr. av SIGN TSV	6. Beskr. dato DATO 280981	7. Endr. dato ENDD	8. Klasse KLAS				9. Elementtype ELTY	
10. Avleid. regel AVLP (se pros.)		11. Måleenhet MALE		12. Definisjon DEFN (se memo)		Navn (nøkkel) 1 <input type="checkbox"/> Klokkeslett 6 <input type="checkbox"/> Kode 2 <input type="checkbox"/> Tekst 7 <input type="checkbox"/> Numerisk 1 <input checked="" type="checkbox"/> Teller 3 <input type="checkbox"/> Kontroll 8 <input type="checkbox"/> Alfabetisk 2 <input type="checkbox"/> Mengde 4 <input checked="" type="checkbox"/> Konstant 9 <input type="checkbox"/> Alfamerisk 3 <input type="checkbox"/> Dato 5 <input type="checkbox"/> Annet 10 <input type="checkbox"/> Streng 4 <input type="checkbox"/>		
13. Definisjon DEFI *								
<p>Elementet inneholder sum lønn for et bestemt underkapittel (pos. 1-3 av kontoen) og post (pos 5-7 av kontoen). Verdien er i hele tusen kroner.</p>								
14. Avledningsseier AVLS *								
15. Referanse REFE								
16. Eksempel på forekomst EKSE 790 7: kr. 790.000,00								

INPUT-OPPLYSNINGER

17. Skalerfaktor (10N) SKAL 10 ³	18. Min. lengde MINI 9	19. Max. lengde MAXI 9	20. Etende i streng BLAN Ja 1 <input type="checkbox"/> Nei 2 <input checked="" type="checkbox"/>		21. Justert mot JUST Venstre 1 <input type="checkbox"/> Høyre 2 <input checked="" type="checkbox"/>
22. Kontroller som <u>edit</u> skal utføres ved innlesing KONT *				23. Kontroll (se pros) KONT	
				24. Koding (se memo) KODM	

OUTPUT-OPPLYSNINGER

25. Max. lengde MAXL	26. Skalerfaktor SKAU	27. Justert mot JUST Venstre 1 <input type="checkbox"/> Høyre 2 <input checked="" type="checkbox"/>		28. Dekoding (se memo) DEKM
29. 1. Headline HEAD ' LØNN '	30. 2. Headline HEAL		31. Utskriftsregel UTSK PIC 'Z(9)'	

KRITERIER FOR VALG AV DATABASESYSTEM (DBMS)

Type kriterier

- økonomiske
- strategiske
- operative
- funksjonelle
- tekniske

Økonomiske

- kjøp/leie/leasing (pris)
- årlig vedlikeholdskostnader
- investering i maskiner
- investering i spesialsoftware (f.eks. data-dictionary)
- kjøreutgifter
- investering i opplæring/kompetanseoppbygging
- implementering (brukersystemer)
- tilpasninger, utvidelser
- overgang til neste generasjonssystem etter n-tid

Strategiske

- ansvarlig leverandørs soliditet
- nåværende og senere antatt utbredelse av systemet
- systemets antatte levetid

Operative

- tilgjengelighet
- krav til maksimal reetableringstid ved systemutfall
- krav til operatør/betjening
- sikkerhet mot tap, misbruk og ødeleggelse av data (datasikkerhet)
- krav til maskinressurser
- krav til behandlings-/svarstider (batch/online)
- minimum antall samtidige brukere

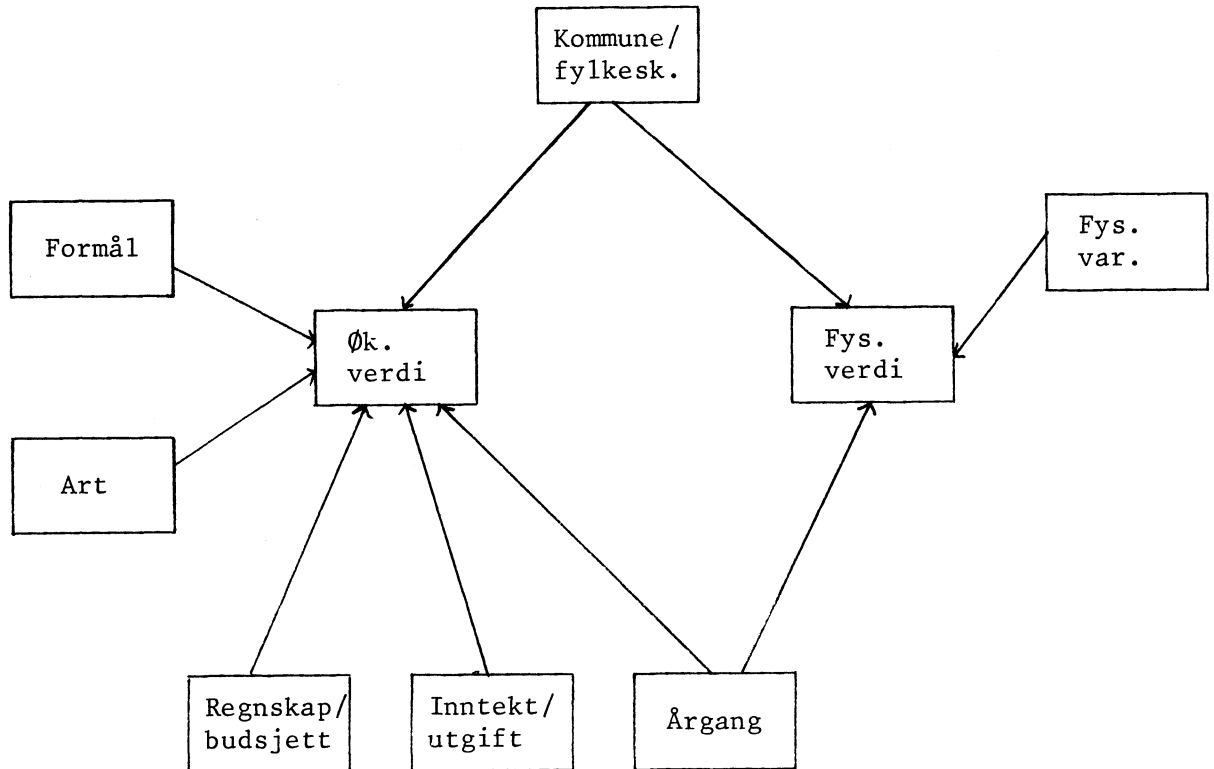
Funksjonelle

- maskinuavhengige ("frie" kontra maskinleverandørens systemer)
- type vertsspråk/brukerspråk
- høyere nivå brukerspråk (rapportgen., spørreprogram osv.)
- feilfinningsprogrammer (monitor, trace, utilities)
- dokumentasjonshjelpemidler (data katalog/dictionary)
- binding brukerprogram - database (logisk/fysisk datauavhengighet)
- sikkerhet (identifikasjon, autorisasjon, kontroll, tilgjengelighet, tillit til data)

Tekniske

- evne til å oppta større belastninger
- flere samtidig brukere
- aksessmetoder
- datastruktur (logisk/fysisk, relasjoner)
- datauavhengighet (funksjonelt og teknisk)
- sikkerhetskopiering
- reetableringsprogrammer
- reorganiseringsprogrammer

DATAMODELL FOR KOMMUNALØKONOMI



En økonomisk variabelverdi vil gjelde regnskap eller budsjett, inntekt eller utgift, et bestemt for mål (kap.), en bestemt art (post) og et bestemt år for en kommune eller fylkeskommune.

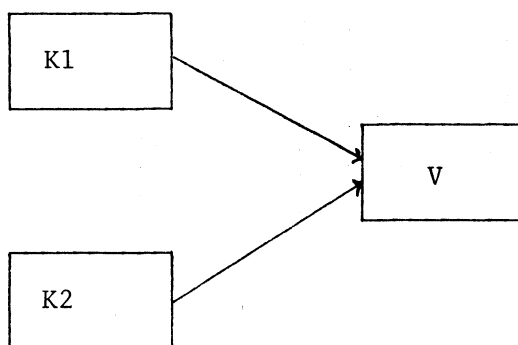
En fysisk variabel-verdi gjelder en bestemt kommune/fylkeskommune og årgang.

Beskrivelser av kommuner, formål, art, fysiske variable m.m. representerer små datamengder og kan lagres i tabeller og hjelperegistre som tar liten plass.

Det er lagringen av selve verdiene som vil utgjøre datavolumet i basen. Verdiene må lagres slik at de kan identifiseres ved hjelp av kommune-, formål- art- En økonomisk verdi vil derfor logisk sett ha tilknyttet en identifikasjonsdel som består av id-begrepene for henholdsvis kommune, art, formål etc. Jeg skal i det følgende beskrive og vurdere to metoder for fysisk implementasjon av identifikasjonsdelen til et verdi-entry.

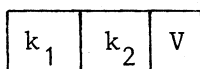
IMPLEMENTERING AV STATISTIKK-DATABASE

For oversiktens skyld betraktes variabler som er gruppert etter 2 kategorietyper, slik at problemet blir to-dimensjonalt. Resonnementet kan uten videre generaliseres til flere dimensjoner. En datamodell for dette tilfellet vil ha følgende struktur:



METODE I

k_1 og k_2 lagres sammen med v slik at en får recordstrukturen:



k_1 , k_2 identifiserer en record av denne typen. For å kunne aksessere en spesifikk rekord av denne typen kan en opprette en primærnøkkel bestående av k_1 og k_2 . En kan også opprette sekundærnøkler for k_1 og/eller k_2 slik at en kan hente ut alle rekorder for en gitt verdi av k_1 (eller k_2). På denne måten kan en oppnå stor fleksibilitet med hensyn på søking i databasen.

Selv om en logisk rekord her kun inneholder et aktuelt tall (i verdidelen), kan en effektivisere på fysiske aksesser ved å gruppere rekorder som brukes ofte sammen (f.eks. rekorder med samme k_1 -verdi) innenfor en felles fysisk blokk.

Denne metoden kan bli svært plasskrevende spesielt hvis det er mange dimensjoner (kategorietyper). Rekorden vil inneholde en liten datadel (v) og en stor identifikasjonsdel ($k_1, k_2 \dots$). For å kunne søke på de ulike kjennetegn (kategorier) må det opprettes indekser eller inverterte lister for disse. Plassforbruket til disse aksessmekanismene vil være proporsjonal med nøkkellengden og antall rekorder i filen. En primærnøkkel (bestående av k_1 og k_2) vil for de fleste indeksemekanismene kreve mer plass enn selve datafilen i dette tilfellet.

METODE II

Datamodellen foran kan ses på som en 2-dimensjonal matrise med K1 og K2 som rekker og rader (eller omvendt):

		K1		1 2 3 4						kolonnenr.
		K2		k ₁						
rad- nummer	1	k ₂								
	2									
	3									
	4									

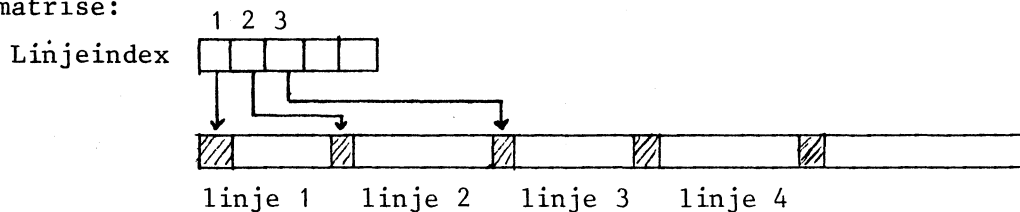
En verdi i matrisen identifiseres ved hjelp av kolonne- og radnummer. For et gitt k_1, k_2 -par må en derfor finne tilsvarende matriseindekser (i_1, i_2) før en aksesserer matrisen: (k_1, k_2, i_1, i_2 adresse til matriseelement).

Matriseindeksene må enten kunne beregnes direkte fra kategoriverdiene eller må ligge lagret sammen med disse slik at de finnes ved å slå opp på kategoriverdi (i K1 og K2 tabellene). Hvis slike kategoritabeller er veldig små kan eventuelt posisjonsnummer innenfor tabellen benyttes som matriseindeks.

Matrisen lagres kolonne- eller radvis, og bør komprimeres slik at ruter som ikke er i bruk eller mangler verdi ikke beslaglegger plass. Linjene i den lagrede matrisen blir således av variabel lengde. En må derfor sammen med matrisen lagre informasjon om lengde og startpunkt for de ulike linjene. Det er derfor kun matriser over en viss størrelse som vil lønne seg å komprimere.

Den fysiske lagringsstrukturen for en matrise vil dermed kunne se ut som følger:

Lagret matrise:



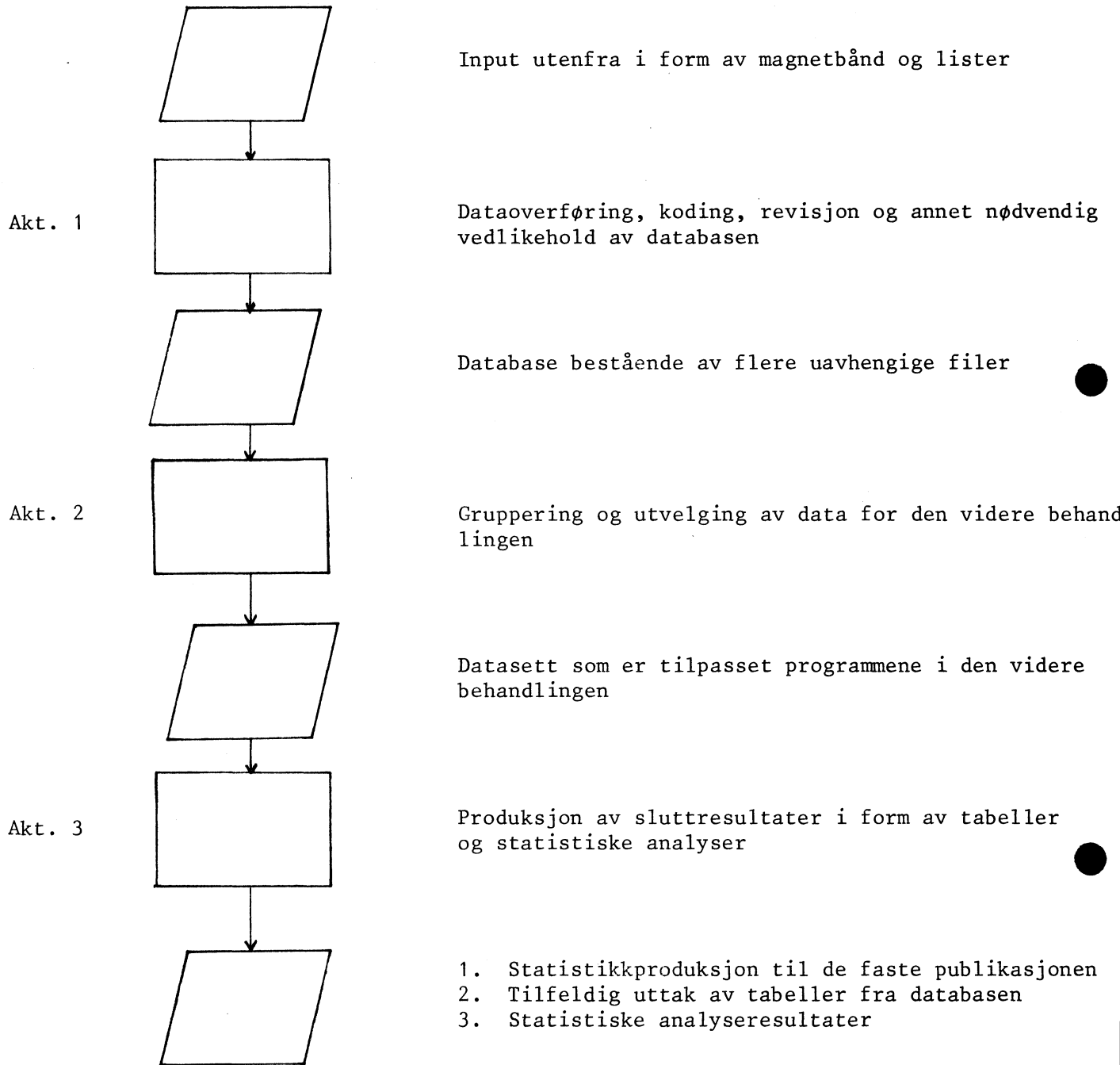
(En linje er enten en rad eller kolonne.)

I linjene må det markeres hvilke elementer som ikke har verdi i matrisen. Dette kan f.eks. gjøres ved en bit-map.

Merk at linjeindexen kun trengs hvis matrisen pakkes.

Denne metoden vil altså kreve liten ekstra lagerplass for å realisere alle søkemulighetene mot variabelverdiene i databasen. En må ha rask random-aksess mot kategoriverdien i K1 og K2, da en alltid må slå opp i disse for å finne matriseindekser. Disse tabellene er oftest så små at de kan ligge resident i kurtiglager.

VURDERING AV PROGRAMPAKKER

1. Skjematisk framstilling av informasjonssystemet

2. Brukere av systemet

Brukere vil i første omgang være 3. kontor, senere også brukere utenfor Byrået.

3. Oppbygging av systemet

Systemet skal være brukerorientert. Assistansen fra Systemkontor og Driftskontor skal være liten når systemet er etablert. Det betinger at alt arbeid med å få data inn og ut av systemet må skje interaktivt fra terminal.

I aktivitet 1 er det lagt vekt på at databasen skal ha en hensiktsmessig form med tanke på bearbeidningen av dataene og med tanke på det maskinelle ressursbehovet.

For at aktivitet 3 skal bli så fleksibel som mulig med hensyn på programvalg og effektiv utnyttelse av maskinen, er det nødvendig med aktivitet 2 for å tilrettelegge data for de ulike programmene.

Det finnes slagkraftige filebehandlingsprogrammer på markedet. Men antakelig vil behovet kunne dekkes enklere ved egenproduserte programmer. Eventuelt bruke SFB og SOP dersom disse ble mer tilpasset interaktiv bruk. I de fleste programpakkene det kan bli aktuelt å bruke vil det dessuten finnes muligheter for å gruppere og selektere dataposter og kjennemerker.

Sluttproduktene i aktivitet 3 kan deles i 3. Statistikkproduksjonen til de faste publikasjonene krever få endringer i programmer og parametersett fra gang til gang. Opplegget bør likevel være slik at endringene kan gjøres ved fagkontoret.

Programmer for dette formålet kan spesialprogrammeres. Det vil i så fall gjøre det mer eller mindre vanskelig å gjøre endringer i tabellene.

Standardprogrammet TEKSTTAB kan bare brukes til å produsere tekstdelen til tabellene. Talldelen må produseres ved hjelp av STP eller spesialprogram.

TAB-68 produserer tekst og tall. Det har sine begrensninger. Men kombinert med et tekstbehandlingsprogram vil det kunne brukes. Intervjukontorets tabellprogram bør også vurderes i denne sammenheng. På litt sikt bør kravet være at programmet kan produsere input til et fotosettingsprogram.

Analysedelen vil for en stor del være regresjonsanalyser. For dette formålet finnes flere programpakker som SAS, SPSS/SCSS, DDPP, BMDP, APL og flere frittstående programmer. Systemet bør være slik at det tillater bruk av flere programmer SPSS/SCSS synes å ha visse fordeler. Det er lett å tilpasse IBM-anlegget, det er enkelt å bruke og det er kjent på fagkontoret. Ulempen kan være at det tillater begrensede datamengder. Derfor er det viktig at filbehandlingen i aktivitet 2 kan tilrettelegge velegnete filer.

FOTOSATS VED TRYKKING AV TABELLER MV.

Vedlagte tabell 16 "Personer som var på ferietur høsten 1978" er en prøve som Dataprint A/S har laget ved hjelp av fotosats. Et magnetbånd, produsert i Byrået, med enkle koder som styrer redigeringen (skifte av skrifttype, kolonnebredde mv.), er brukt som grunnlag. Det er mulig å bruke mange forskjellige skrifttyper, store og små bokstaver, "trinnløs" variasjon av avstanden mellom linjene/ordene/bokstavene m.m.

Dataprint A/S er villig til å samarbeide med Byrået (vederlagsfritt) ved å bygge inn spesialrutiner i sine systemer som tar hensyn til Byråets krav til tabellhoder, forspalter mv. I tillegg må det lages en programmodul som tar hånd om nødvendig koding av skriftlinjene, og som må kunne brukes i alle tabellprogrammer i Byrået som lager grunnlag for trykking. Dette gjelder også tekster og tabellhoder som Tekstbehandlingskontoret lager på sitt EDB-utstyr.

Tekstbehandlingskontoret og Trykningskontoret må eventuelt trekkes inn i planleggingen av dette.

Dataprint A/S leverer korrektur dagen etter at de har fått datagrunnlaget (magnetbånd, diskett o.l.), og leverer eventuelt ferdig trykte hefter et par dager (maks. 14 dager) etter at ferdig korrektur foreligger.

Byrået står ellers fritt mht. å velge trykkeri, eventuelt å trykke selv, når satsen er ferdig fra Dataprint A/S.

TABELL 16. PERSONER SOM VAR PÅ FERIE TUR HØSTEN 1978, I GRUPPER FOR KJØNN/ALDER/FERIESTEDET/KOMMUNETYPE FOR BOSTED, ETTER TYPE NATUR PÅ FERIESTEDET. PROSENT PERSONS ON HOLIDAY AUTUMN 1978, IN GROUPS FOR SEX/AGE/HOLIDAY AREA/TYPE OF MUNICIPALITY OF RESIDENCE, BY TYPE OF NATURE IN HOLIDAY AREA. PER CENT

		TYPE NATUR PÅ FERIESTEDET TYPE OF NATURE IN HOLIDAY AREA						TALLET PÅ PERSONER SOM SVARTE NUMBER OF RESPONDS
		I ALT TOTAL	STØRRE BY LARGER CITY	KYST- OMRÅDE COASTAL AREA	FJELL- OMRÅDE MOUNTAIN AREA	SKOGS- TRAKTER FOREST- AREA	INNLANDET ELLERS REST OF INLAND COUNTRY	
ALLE PERSONER	ALL PERSONS	100	28	36	16	6	13	286
KJØNN SEX								
MENN	MALES	100	23	34	24	8	10	124
KVINNER	FEMALES	100	31	38	10	4	15	162
ALDER AGE								
15-24 ÅR	YEARS	100	30	40	13	5	13	40
25-34	"	100	38	32	13	11	6	47
35-54	"	100	29	29	23	5	13	84
55-74	"	100	22	43	14	5	15	115
FERIESTEDET HOLIDAY AREA								
NORGE	NORWAY	100	22	23	29	9	16	157
UTLANDET	ABROAD	100	35	53	1	2	9	129
KOMMUNETYPE FOR BOSTED TYPE OF MUNICIPALITY OF RESIDENCE								
LANDBRUKS- OG FISKERI- KOMMUNER, ANDRE KOMMUNER AGRICULTURE- AND FISHING MUNICIPALITIES, OTHER MUNICIPALITIES		100	26	33	14	5	17	42 1)
SENTRALE OG MINDRE SENT- RALE BLANDEDE LANDBRUKS- OG INDUSTRIKOMMUNER CENTRAL AND LESS CENTRAL MIXED AGRICULTURE AND MANUFACTURING MUNICI- PALITIES		100	18	34	20	6	20	50
SÆRLIG SENTRALE, BLANDEDE TJENESTEYTINGS- OG INDU- STRIKOMMUNER. HIGHLY CENTRAL MIXED SER- VICE AND MANUFACTURING MUNICIPALITIES		100	28	40	15	9	9	128
ØVRIGE BLANDEDE TJENESTE- YTINGS- OG INDUSTRI- KOMMUNER OTHER MIXED SERVICE AND MANUFACTURING MUNICI- PALITIES		100	35	33	17	2	12	66

1) AV DISSE: UOPPGITT 5 PST.

1) OF THESE: UNKNOWN 5 PCT.

Datatekst-entreprenører

Dataprint A/S representerer noe helt spesielt i norsk data- og trykkeribransje. De som står bak kaller seg helst *entreprenører* - som tar seg av hele produksjonsprosessen for en trykksak inklusive distribusjonen. Til dette formål har de utviklet sine egne datasystemer. Kjernen i det hele er et fotosatsanlegg.

Fagbladet Norsk Grafisk Tidsskrift vet å fortelle at Dataprint er ganske alene i landet på sitt spesialfelt: Kompliserte trykksaker - både store og små - som *hurtigtrykk*. Kataloger, prislister, konsernregnskaper er typen. Norges Lover satt og ferdig ombrukset (ferdig monterte sider) på 14 dager. Arbeidsdirektorets lister over ledige stillinger, inntil 68 A4-sider, oppdateres kl. 10 på formiddagen og ligger ferdig distribuert over hele landet kl. 8 neste morgen.

SYSTEMHUS

Fagbladet Datatid vil anse Dataprint for å være et «software-hus», en programvare-spesialist, i den grafiske sektor. Den informasjon de skal arbeide med må ligge klar på data. Om den så ikke gjør, tas i disse dager i bruk en optisk leser som kan «læres» til å lese *alle* skrifter (Kurzweil Data Entry Machine). Deres spesialsystemer klarer så konverteringene videre.

Det har tatt tre års intens programutvikling for å få til det man gjør på Dataprint i dag. Tid omregnet i kroner gir en utviklingskostnad på nær en halv million. Markedsprisen er selvsagt noe ganske annet. De er ikke til salgs.

Bedriftens egentlige styrke ligger i organisasjonsformen og det faglige miljø som er skapt. Datafolk og typografer arbeider side om side. De er medeiere i bedriften.

KONVERTERINGS-VERKTØY

Tyngdepunktet i Dataprints programutvikling ligger på et program de kaller «FESS», som er et konverteringsverktøy for record-oppdelte data til grafisk tekstbehandling. Det dreier seg om ADB-data som dermed kan typograferes og lages lay-out på helt fritt.

DataGuidens IBM-konverterte CP/M-disketter er egentlig ikke noe typisk eksempel på Dataprints spesielle evner. Agentur- og produktregisterne i DataGuiden var en snartur innom FESS. Fri-tekstkonvertering for de vanlige tekstinnslagene skjer med enklere verktøy. Poenget er likevel at det ikke er så mange som har lært seg kunsten.

Siden de fleste store institusjonene - offentlige og private - nå har sine «data», teller Dataprints kunderegister svært mange i denne gruppen; departementer, direktorater, etater og institutter - storindustri, rederier og vare-distributører. Eksempelvis Arbeidsdirektoratet, Lovsamlingsfondet (Norges Lover), Utenriksdepartementet, Postdirektoratet, NSB, Elektrolux, Hoegh's rederi, Saga Solreiser m.fl.

Og de trykksakene disse i stor grad har behov for er nettopp all verdens kolonner og registre slik de nå har dem inne i sine regnskaps- og ADB-systemer: prislister, messe kataloger, resultatlister, rutetabeller, leksika etc. (Dataprint er med andre ord en minimaskin«konkurrent» til DataGuidens eget mikromaskinsystem når det gjelder databehandlingen for fotosettingen.)

INFORMASJONS-BEDRIFT

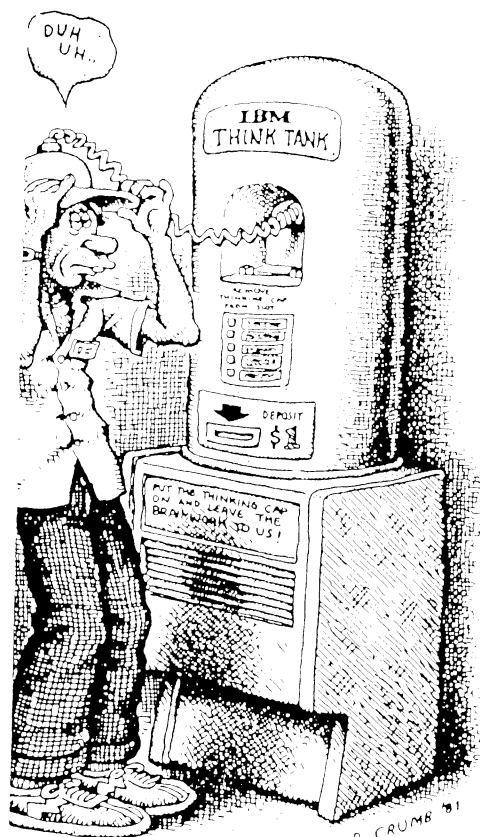
At Dataprint dukket opp i DataGuidens programvare-jungel illustrerer denne artikkels hovedpoeng: «Det er ikke så lett som man tror.» (Ei heller for datafolket.)

Alle oppgaver kan løses, men veien frem går gjennom en jungel som det finnes få og uføllstendige kart over. De få kjentfolk som har gått opp en rute frem til målet må man ut i jungelen for å finne.

Og her har altså Dataprint besatt et høydedrag som det er mulig å få øye på.

Dataprints kunnskaper dekker egentlig hele informasjons-området. De holder databaser for sine kunder på eget anlegg (2 stk. Nord 10/S med 900MB masselager). På anlegget kjøres databaseverktøy som SIBAS, egne distribusjonssystemer, NORTEXT m.m. i tillegg til FESS.

Foreløpig er Dataprint «bare» en trykksakprodusent og -distributør. Men med sin viten om informasjonssystemer og den programjungel som omgir disse, vil vi nok få se dem opptre også i andre sammenhenger etter hvert. På eget forlag har de allerede flere bransjekataloger ute. Dette er store databaser som nok vil dukke opp i det «elektroniske informasjonssamfunn» - så snart det begynner å ta form.



© Robert Crumb, Winters, Calif
Gjengitt i Datatid etter avtale

GRAFISK DATABEHANDLING FOR FOB-80.

Under bestilling fra Tektronix :

- Tektronix 4027 Color Graphics Terminal inkl.
16K display memory og
96K grafisk memory.

(4.mai-pris : 103125,- + moms)

- GPIB Peripheral Output

(4.mai-pris : 4085,- + moms)

- Video Output

(4.mai-pris : 660,- + moms)

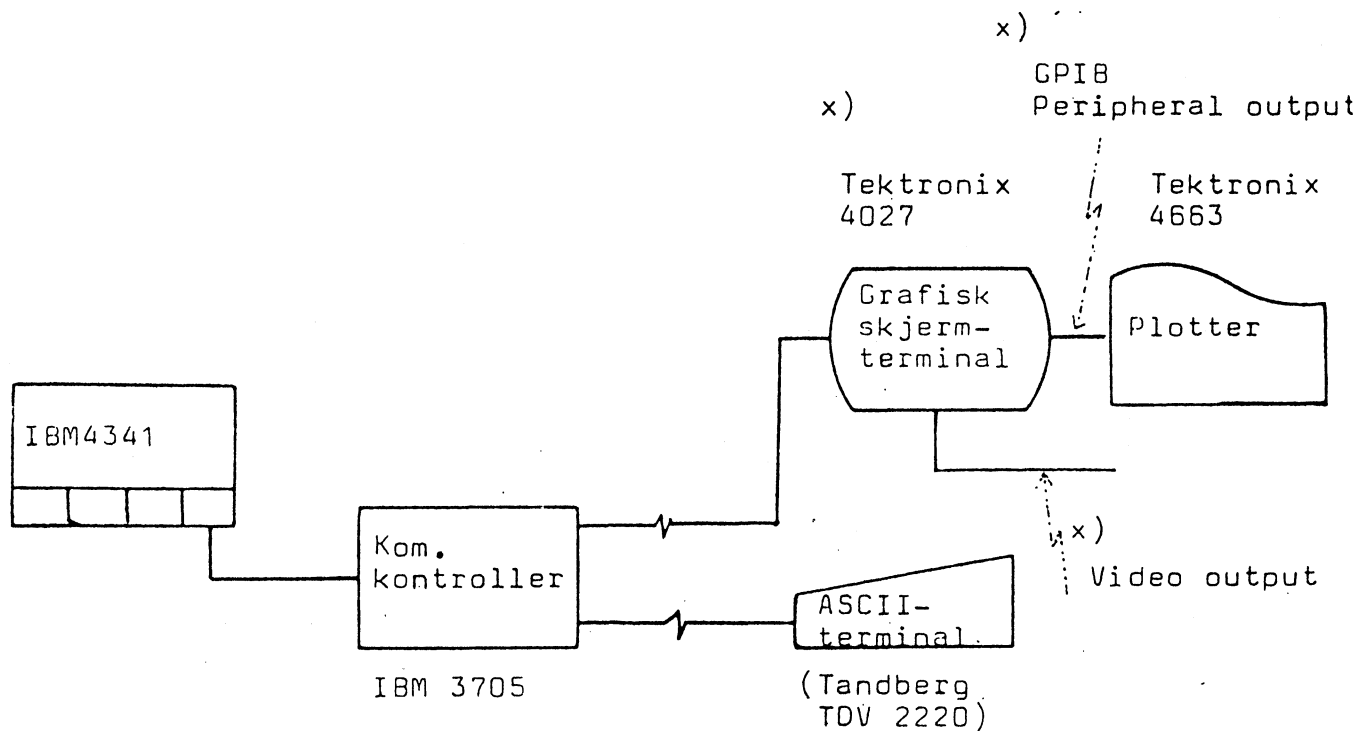
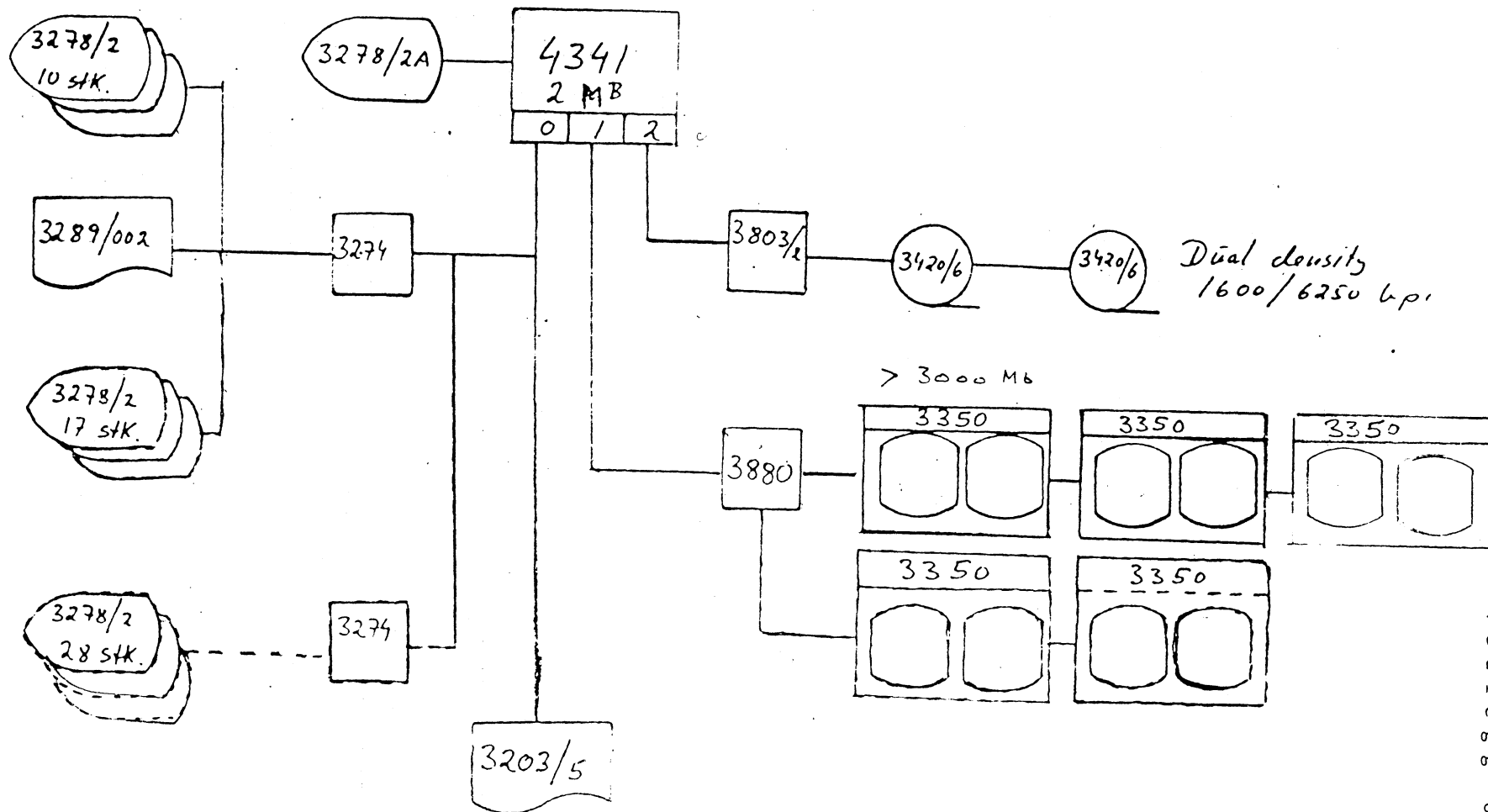


Fig.1 Endelig konfigurasjon for grafisk db. for FOB-80 ?
Utstyr merket x) er under bestilling.

MASKINKONFIGURASJON IBM 4341

OS/VS1 under VM/370



Enheter som er stiplet vil bli installert
i tiden 1/10-80 til 31/1-81

BESKRIVELSER AV DATABASER I DRIFT OG UNDER PLANLEGGING I BYRÅET

Bedrifts- og foretaksregisteret

Inneholder ca. 400 000 enheter, 38 kjennemerker pr. enhet. Størrelse 60 MB.

Databasen ligger på Statens Driftssentrals Honeywell-Bull-maskin (L66/60P), og er implementert ved hjelp av IPS (indekssekvensielt filsystem) og IDS (databasehåndteringssystem).

Ingen eksterne brukere har direkte tilgang til databasen, som er i daglig online bruk fra 17 tilknyttede skjermterminaler ved SSBs Registerkontor.

Antall transaksjoner mot basen er ca. 4 000 pr. uke. Svartidene ved terminalene varierer fra 1-8 sekunder, avhengig av belastningen på databasen og Honeywell-Bull-anlegget.

Det er opprettet 13 "read only" del-databaser for de forskjellige fagkontorer ved SSB, til direkte oppslag på enkeltopplysninger, til løpende statistikkproduksjon, og til ad-hoc-oppdrag.

Personregister

I forbindelse med det sentrale personregister som er tapebasert, er det opprettet et delregister (1981) med mulighet for direkte oppslag på personnummer. Dette inneholder samtlige enheter fra personregisteret (ca. 5 millioner), men har kun noen kjennemerker som er av ikke-konfidensiell natur. Denne databasen skal derfor i prinsippet kunne benyttes direkte av eksterne brukere. Vegdirektoratet vurderer en slik løsning, men foreløpig utføres alle oppdrag mot registeret av Byråets folk.

GAB-registeret (Grunneiendom-, Adresse-, Bygning)

GAB-systemet er under politisk styring av Miljøverndepartementet, mens Statistisk Sentralbyrå har det faglige hovedansvar. Registeret bygges opp kommunervis ved de kommunale datasentralene (7 stk.) og vil bestå av en database for grunneiendommer, en for adresser og en for bygninger. På landsbasis vil det bli ca. 2 mill. enheter av hvert slag (totalt 2 200 MB). De kommunale etater vil utgjøre den største brukermassen. Systemet er kommet i drift i 1981 med G-delen og etter planen skal denne være komplett for alle kommuner i 1983.

Folketellingsdatabase 1980

Databasen for folketellingen i 1980 består av personregister, boligregister, bedrifts- og foretaksregister, yrkeskatalog, kommunekatalog. Dataene ligger på en IBM 4341 med en diskkapasitet på over 3 000 MB. Systemet benytter VSAM-filer, CICS som transaksjonsmonitor og er programmert i COBOL. Ca. 60 skjermterminaler er tilkopleet.

Industristatistikk

Et online databaseopplegg på minimaskin (NORD) benyttes for registrering og kontroll av industristatistikk. Databasen omfatter 18 000 foretak/bedrifter og 160-170 000 varer (ca. 40 MB). Det er tilknyttet 10 skjermterminaler.

Systemet er implementert vha. COBOL med relativadresserbare filer og skjermhåndteringssystemet NSHS.

Tabellproduksjon fra databasen skjer i dag ved at det ekstraheres filer som overføres til den sentrale stormaskinen (Honeywell-Bull) hvor standardprogrammer for tabellgenerering benyttes.

Abonnementdatabase

Databasen inneholder ca. 1 000 abonnenter med bestillinger av tabeller fra utenrikshandelstatistikken. Basen er implementert i IDS på Honeywell-Bull-anlegget. Bruken av databasen har hittil kun skjedd via satsvise transaksjoner.

Database (kartotek). Inngivere av korttidsstatistikk (4. kontor) (1980)

For de ulike statistikkene (industri) er det opprettet datafiler med tilhørende indeksfiler slik at en kan søke direkte på identifikasjonsnummer. For hver bedrift er det lagt inn kvalitative kjennemerker som navn, adresse, kontaktperson osv. ("slippopplysninger"). Systemene brukes for produksjon av slipper, purringer og forskjellige lister. Hvert kartotek (8-10 stk.) opererer på minimaskin (NORD) som et enbrukersystem. COBOL og NSHS er benyttet for implementasjon.

Database for utførselsdeklarasjoner (utenrikshandelstatistikk)

Databasen brukes for registrering og kontroll av utførselsdeklarasjoner (fra Tolldirektoratet). Disse behandles månedvis og basen vil inneholde ca. 60 000 deklarasjonsrekorder med tilhørende varerekorder for en måned. Kontrollen tar gjerne mer enn en måned slik at data for 2 måneder vil ligge inne samtidig

(separeres på 2 fysiske baser). Fra basen ekstraheres 4 delregistre (sekvensielle filer) for henholdsvis Tolldirektoratet, Norges Bank, Skattedirektøren og Statistisk Sentralbyrå som hver for seg bruker og bearbeider dataene videre til sine formål. Systemet kjører under TDS/IDS på Honeywell-Bull. Det benyttes indekssekvensielle filer.

Database for innsjekking av intervjumateriell

I forbindelse med intervjuundersøkelser benyttes et enkelt filopplegg hvor identifikasjonene til intervjuobjektene i en undersøkelse legges. Disse filene ligger på en NORD-maskin, som er koplet til Honeywell-Bull-anlegget, og systemet henter også en del data (navn, adresse etc. til intervjuobjektene) fra dette hovedanlegget for bl.a. utskrift av purringslister.

Systemer for statistisk analyse

For statistisk analyse bruker Byrået i dag DATSY, TROLL, TSP, SPSS som alle er nokså generelle programpakker (språk) som har sin styrke på litt forskjellige områder. De har integrert i seg en del funksjoner for fil/databasehåndtering. Disse er derimot ikke så generelle og omfattende at de kan brukes på store sentrale statistiske databaser.

Database for personaladministrasjon

Databasen benyttes av Personalkontoret og de viktigste enhetene i basen er personer og stillinger. Systemet er implementert på NORD og databasehåndteringsdelen er selvlagd. Systemet er under utvidelse/omarbeiding.

DATABASER UNDER PLANLEGGING

Tidsseriedatabase

Denne skal inneholde 3-5 000 makroøkonomiske tidsserier (arbeidsmarked, nasjonalregnskap, utenrikshandel etc.). Perodesiteten er månedlig, kvartalvis, årlig. Databasen skal benyttes for oppslag interaktivt, tabellproduksjon og særlig for analyseformål (TROLL). Databasen skal etableres i løpet av 1982. Det er ennå ikke helt avgjort om et eksisterende system kan benyttes (AXIS, Sverige, Datenbanken, Østerrike, SUBJECT, University of California, vurderes bl.a.) eller om en satser på egen utvikling.

Database for tidsskriftsirkulasjon

Denne er under planlegging, og det er mulig at det kjøpes et system utenfra som kan tilpasses Bibliotekets behov. Systemet skal implementeres på NORD-maskin.

Database for energistatistikk

Databasen skal inneholde aggregerte tall på forskjellige nivåer for forskjellige energivarer (elektrisitet, fast brensel, olje etc.) fordelt på ulike klassifikasjonsenheter (geografisk inndeling, næring etc.), tidsperioder og statistikkområder. Databasen skal etter planen implementeres på IBM 4341 (1982). I størrelsesorden vil den omfatte 1-10 MB. Mesteparten av inputdataene skal kopieres inn maskinelt, mens mindre mengder skal kunne registreres via skjerm. Ved registrering skal det foretas enkle kontroller av datafeltene. Uthenting og oppretting av data skal kunne utføres online via skjermterminaler. Søking mot basen skal kunne gjøres interaktivt av brukere som har liten kjennskap til innholdet i basen. Den viktigste output-funksjonen forøvrig vil være tabellproduksjon (balanser).

TIDS- OG RESSURSPLAN FOR HOVEDPROSJEKTET. PROSJ.NR. 3137 DATABASE FOR KOMMUNALØKONOMI FOR 1982

Aktivitet	Time- verk	Utføres av	Jan.	Febr.	Mars	Apr.	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Des.
Problemorientert systemering, prosjektledelse mv. ved 3. kontor (nedenfor er det angitt hvor 3. kontor må medvirke)	900		-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Prosjektledelse (maskinorientert del), syst.analyse og design	- 150 150 250	3.ktr. Eks.kons. 1 " " 2 Syst.ktr.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Programutvikling og testing														
- Progr. for oppdatering av dataene	600	Eks.kons. 1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Progr. for ajourhold av filer ..	550	" " 1 " " 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Progr. for estimering ^{ekstrahering} av data fra basen	700	" " 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Sortering, backup, restore	200	" " 3	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Dokumentasjonssystem	150 150	" " 3 syst.ktr	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Tilpasse standard tabellprogram	520	" " "	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
- Opplegg for faste tabeller	150	" " "	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Systemtest, dokumentasjon	- 75 170 165	3.ktr. Driftsktr. Eks.kons.1 og 2 Syst.ktr.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Implementering	- 75 170 65	3.ktr. Driftsktr. Eks.kons.1 og 2 Syst.ktr.	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



STATISTISK SENTRALBYRA

Kommunal- og arbeidsdepartementet
Postboks 8112 Dep.
OSLO 1

DERES REF.:

VÅR REF.:
BB1/HaR

STED OG DATO:
Oslo, 9. oktober 1981

SØKNAD OM FINANSIERING TIL ETABLERING AV EN DATABASE FOR KOMMUNALØKONOMISK
STATISTIKK

Statistisk Sentralbyrå satte i 1981 i gang planleggingen med sikte på å opprette en database for kommunaløkonomisk statistikk. I Byråets brev av 8. juli d.å. til Kommunaldepartementet ble det søkt om støtte til arbeidet med et forprosjekt for denne databasen. Departementet ga i sitt svarbrev av 20. juli tilsagn om inntil 50 000 kroner til konsulentbistand til dette forprosjektet, som skulle klarlegge omfanget, egnede systemer og kostnader ved en kommunaløkonomisk database.

Byrådet etablerte en forprosjektgruppe i august, og tilskøttet fra departementet ble nyttet til å engasjere en konsulent fra Kommunedata Østlandet til denne gruppa. Arbeidet med dette forprosjektet er nå under avslutning, og Byrådet har derfor det nødvendige grunnlaget for å søke departementet om finansieringsbistand til etablering av selve databasen.

Formål

Formålet med databasen er først og fremst å få til en rasjonalisering av statistikkproduksjonen. Dette vil komme brukerne til gode gjennom en raskere tabellproduksjon og dermed en mer aktuell statistikk. Samtidig vil også brukerne kunne få kjørt ut ulike spesialtabeller fra Byrådet hurtigere enn i dag. Viktige eksterne brukere vil senere, mot betaling, kunne kople seg til databasen med terminal.

I vedlegg 2, punkt 1-4, er det gjort rede for omfang, oppbygging og brukeregenskaper ved den databasen som arbeidet i forprosjektet har ledet fram til. I forprosjektet er kostnadene forbundet med utvikling og implementering av databasen blitt beregnet.

På dette grunnlaget søker Byrået med dette om at Kommunaldepartementet stiller til rådighet inntil 610 000 kroner i 1982 for å dekke deler av disse kostnadene. Nedenfor er det gitt en samlet oversikt over utviklings- og implementeringskostnadene:

Kostnader ved utvikling og implementering av database:

1982	Kostnader i alt	Kroner		
		Dekkes over Byråets budsjett	Dekkes ved søknad til RFSP	Dekkes ved denne søknaden
Prosjektledelse, planlegging, systemering, programmering ...	968 000	130 000	288 000	550 000
Teknisk utstyr, datamaskin- kapasitet mv.	-	60 000
Kostnader i alt	288 000	610 000

En tilhørende fordeling av timeverk til utvikling av databasen, fordelt på de enkelte arbeidsoppgaver, er gitt i vedlegg 1.

Ved kalkuleringen til denne søknaden er eksterne priser lagt til grunn, slik at Byrået (fra ekstern datasentral) regner med å få 2,2 årsverk for de 550 000 kr til systemering og programmering som denne søknaden omfatter. Eksterne priser er lagt til grunn fordi Byrået i dag ikke selv har den nødvendige systemkapasitet, slik at det vesentlige av systemerings- og programmeringstjenestene vil måtte kjøpes utenfor Byrået. Det forutsettes at Kommunaldepartementet i tilfelle betaler kjøpesummen direkte til vedkommende datasentral på grunnlag av faktura attestert av Byrået.

I tillegg til de kostnader som er tatt med i oversikten ovenfor, vil det over Byråets budsjett fra slutten av 1982 bli dekket utgifter til datamaskinkapasitet.

Til dette samme prosjektet har Byrået allerede søkt RFSP om inntil 288 000 kroner. Kopi av denne søknaden er oversendt RFSP til orientering.

Kostnader med innlegging av data og ved drift av databasen, som ikke er tatt med ovenfor, vil bestå av utgifter i forbindelse med bruk av program- og maskinvare, samband mv. som er i bruk også for andre databehandlingsoppgaver i Byrået. Det er i dag ikke mulig å si hvor mye dette vil utgjøre. Disse kostnader må søkes dekket over Byråets ordinære budsjett, og dessuten i form av inntekter fra betalte oppdrag.

Hvis Byrået får tilsagn om det beløpet som er forutsatt, vil det bli utviklet et operativt EDB-system for en database med det innhold og de brukeregenskaper som er angitt i vedlegg 2, punkt 1-4. Vi forutsetter at dette systemet vil måtte videreutvikles i de kommende år.

Med hilsen

Odd Aukrust

Liv Bjørnland
Liv Bjørnland

Vedlegg

RESSURSBRUK (TIMEVERK) VED UTVIKLING OG IMPLEMENTERING AV DATABASEN

	Timeverk i alt	Av dette: Dekkes ved denne søknaden
1982:		
- Prosjektledelse, systemdesign	920	350
- Programmer for å opprette databasen med ferdig reviderte data	660	660
- Ekstraheringsprogram(mer) for del- og sumdatabaser til bruk i tabeller (og senere i generelle programpakker)	850	850
- Tilpasse standard tabellprogram til IBM 4341 (muliggjøre interaktiv behandling, generere koder for foto-setting m.v.)	520	-
- Opplegg for faste tabeller	150	-
- Sortering, backup, restore	250	250
- Data dictionary (datadokum.). Nødvendig del for å starte databasen ..	300	-
- Ajourhold av filer. Nødvendig del av rutine for kontroll/koding/revisjon som må gjøres i 1982	980	630
- Systemtest, implementering, brukerdokumentasjon m.v.	900	340
Sum timeverk i 1982:	5 530 (4 årsverk)	3 080 (2,2 årsverk)

1. OVERFØRING AV DATA TIL DATABASEN

1.1. Omfang

Databasen skal inneholde regnskapsdata og i tillegg noen variable av ikke-regnskapsmessig art fra hver enkelt kommune og fylkeskommune. For tiden har vi 454 kommuner og 18 fylkeskommuner, dvs. i alt 472 enheter.

I databasen vil det bli lagt inn de regnskaps- og budsjettdata som 3. kontor innhenter for kommunene. Dette omfatter data fra skjema 2 "Nasjonaløkonomisk gruppering av utgiftene og inntektene på kommuneregnskapet", skjema 3 "Kommunens balansekonto" og skjema 1 "Nasjonaløkonomisk gruppering av utgiftene og inntektene på kommunebudsjettet". Tallene er i 1 000 kr.

1.1.1. Foreløpige og endelige tall

Brukerne av kontorets statistikk har sterkt behov for å få denne på et tidlig tidspunkt. Det vil derfor bli mulig å legge inn foreløpige tall som senere skiftes ut med endelige tall. Brukerne vil på denne måten få muligheten til å få en grov oversikt over den kommunaløkonomiske situasjonen på et meget tidlig tidspunkt.

1.1.2. Ikke-regnskapsmessige data

For hver enkelt fylkeskommune og kommune vil det bli lagt inn i basen noen utvalgte fysiske variable som totalt antall innbyggere og antall innbyggere i ulike aldersgrupper, antall elever og antall klasser i grunnskolen osv. Slike data ligger hovedsaklig lagret på forskjellige magnetbånd i Byrået. Det vil senere være aktuelt å supplere denne variabelliste med flere variable som kan knyttes direkte til kommunens utgifter til ulike formål. Foreløpige data for antall innbyggere osv. skal enkelt kunne skiftes ut med endelige tall.

For å komme fram til et bedre mål for utgifts- og inntektsendringene i kommunesektoren, vil det bli innarbeidet lønns- og prisindekser i systemet. Lønnsindeksene vil bli utarbeidet innenfor de enkelte kapitler. Felles for alle kapitler vil det bli lagt inn prisindekser for utstyr, for nybygg og nyanlegg, for vedlikehold av bygg og anlegg og for andre driftsutgifter.

1.1.3. Fleksibilitet med hensyn til antall variable og enheter

For at systemet skal kunne bli tilpasningsdyktig til utviklingen over tid, vil det bli lagt opp slik at det er mulig å innpasse endringer i antall enheter, kapitler og poster. Kommuner kan bli delt eller slått sammen, derfor vil det kunne være noe forskjell i antall enheter på to ulike tidspunkt. (Det har skjedd svært få slike endringer i perioden 1974-80.) I kapittel- og postgrupperingen vil det også bli endringer. Disse vil i første omgang være av lite omfang, men på lang sikt vil vi kunne få en hovedrevisjon av denne inndelingen.

Også for variablene som er av ikke-regnskapsmessig karakter, vil det kunne bli behov for utvidelser. Spesiell interesse vil det være for å få inn flere variable som er direkte knyttet til de enkelte kapitlene, slik at en etterhvert kan beregne bl.a. enhetskostnader.

Det kan også være behov for å legge midlertidig inn spesielle variable i basen, som skal analyseres i sammenheng med andre variable i basen.

1.1.4. Datainnleggelsen gjennomføres i 3 trinn

1. Trinn. Data for årene 1974(-72)-1981 legges inn fra skjema 2. For disse årene vil regnskapsdataene ligge ferdig revidert og kontrollert på magnetbånd i Byrået. I tilknytning til innleggelsen av regnskapsdataene vil også ikke-regnskapsmessige data bli lagt inn for de samme årene, disse ligger også hovedsaklig på magnetbånd i Byrået.
2. Trinn. Årlig ajourhold av databasen fra og med årgang 1982. Inntil det er etablert et opplegg for integrert databearbeiding via skjermterminaler ved 3. kontor, vil regnskapsdataene overføres ferdig revidert og kontrollert fra magnetbånd i Byrået til databasen.
3. Trinn. Innleggelse av budsjettall, første årgang 1983 og balansetall fra og med 1980 (1983).

2. UTTAK AV DATA

Databasen vil bli organisert på en slik måte at dataene til enhver tid er eller kan bli tilgjengelige. For de fleste formål vil det bli interaktiv adgang til databasen, men det kan også være aktuelt med satsvise kjøringar ved større arbeidsoperasjoner.

2.1. Tabeller

2.1.1. Ferdig redigerte tabeller

Basen vil kunne lagre tabeller som Byrået gir ut, og som kan være av interesse for viktige brukergrupper. I tilknytning til disse tabellene lages det et tabellregister som fortløpende, blir oppdatert automatisk.

2.1.2. Tabellprogrammer som skal benyttes årlig

Det vil bli mulig å sette opp tabellprogrammer som benyttes f.eks. i Strukturtallpublikasjonen, en gang, slik at uttaksprogrammet ikke må skrives hvert år. I tillegg til uttaksprogram for hver deltabell vil det også kunne lages et uttaksprogram som styrer utkjøringen av alle tabeller til en publikasjon, slik at tabellene kommer ut ferdig sortert og nummerert.

2.1.3. Tabellprogrammering

Databasen vil etterhvert i økende grad kunne nyttes til tabelluttak ut fra interne og eksterne brukeres egne spesifikasjoner. Det vil derfor bli lagt opp til et brukervennlig og konverserende system for å få dette til uten for store problemer. Ved større tabellutkjøringer som det ikke er rasjonelt å kjøre ut direkte, bør brukeren få første del av tabellen ut, for å se om tabellen er i overensstemmelse med behovet. Det bør for slik bruk være lagt opp til et system som gjør det enkelt å rette opp, og-endre de spesifikasjoner som alt er gjort.

Etter at brukerne har spesifisert de regnskapsdata de ønsker, vil de få spørsmål om det er behov for å kombinere disse med ikke-regnskapsmessige variable, f.eks. folketall. Dersom brukeren ønsker å benytte alle variable i tall pr. innbygger, vil programmet automatisk utføre denne operasjonen. Ønsker brukerne å kombinere regnskapsdataene med ulike fysiske størrelser, skal det også gis anledning til dette på en enkel måte.

2.1.4. Redigering og utlisting av tabeller

Databasen vil gjøre det mulig å redigere tabellutkjøringene på en slik måte at de kan benyttes som offsetoriginal direkte.

Dette forutsetter:

1. At utlistingen av databasen skal kunne skje på en printer
2. At tabellteksten kan settes inn og eventuelt endres etter behov
3. At basen inneholder navnene på kommunene og fylkeskommunene. Disse skal skrives ut i tabellenes forspalte i tillegg til det 4-sifrede kommunenr.
4. At en skal kunne omredigere tabeller på en enkel måte. F.eks. bytte om på tabellkolonnenes rekkefølge

2.1.5. Omregning til nasjonalregnskapets standard

Det skal til kapittel og post knyttes koder som tilsvarer kodene i nasjonalregnskapet for produksjonssektor, art og formål, slik at tallene i kommuneregnskapene skal kunne omgrupperes til nasjonalregnskapets standard. Konkret vil det si at de fleste av variablene skal tilordnes koder som gjør det mulig å omkode kommuneregnskapet til nasjonalregnskapets standard. Dette vil muliggjøre en utvikling der kommunenes og fylkeskommunenes økonomi på et mer detaljert nivå kan bli innarbeidet i planleggingsmodeller, f.eks. i en delmodell til Modis.

2.1.6. Tidsserietabeller

Det er behov for at basen gir muligheter for å studere utviklingen i kommunaløkonomien over tid. Systemet skal bygges opp så brukervennlig at når brukeren har spesifisert de årgangene som skal analyseres samt variabelspesifikasjonen for ett av årene, finner programmet automatisk fram samme variabel for de øvrige år.

I forbindelse med tidsserieanalyse er omregning til verdibeløp i faste kr og omregninger for endringer i kapittel- og postgrupperingen sentral. Omregninger til faste priser vil bli foretatt ut fra lønns- og prisindekser i basen, og brukerne skal gis tilbud om automatiske omregninger av denne typen, og skal også kunne velge basisår. Seleksjonsrutiner for kommuner som ikke finnes i hele perioden, vil også bli utviklet for tidsserieanalyser.

Databasen vil videre bli lagt opp slik at det er mulig å lage tidsserier for enkeltkommuner eller grupper av kommuner. Ved hjelp av et tidsseriediagram, vil en også grafisk kunne se utviklingen over tid for ulike kommuner.

2.1.7. Sortering, seleksjon og utlisting

Sortering

I systemet vil det være en sorteringsrutine som gjør det mulig å sortere enhetene (kommune) i tabellutkjøringer etter andre variable enn kommunenr. Men i de overveiende fleste tabeller vil det være sortering etter kommunenr. som er det ønskelige.

Seleksjon

Det vil bli mulig å trekke ut kommuner som skal analyseres for seg. F.eks. kan en ønske å lage analyser for kun en kommune, en gruppe av kommuner, f.eks. kommuner med færre enn 1 000 innbyggere eller for alle kommuner unntatt Oslo. Det vil ofte bli behov for å kjøre den samme analysen med ulike seleksjoner.

Horisontal og vertikal utlistering av datamatriksen

I databasen (matriksen) vil det være enkelt å få laget utlisteringer både vertikalt og horisontalt. Ved dette vil en kunne få utlistet verdiene for alle kommuner på en variabel eller alle variabelverdiene for en kommune.

Beregninger av statistiske mål for den enkelte variabel bør lagres for at slike utkjøringer som her er beskrevet, ikke skal bli unødig kostbare.

2.1.8. Flexibilitet; endringer i tabell- og analysebehov

Over tid utvikles nye teknikker for tabell og analyse. Systemet vil derfor bli lagt opp på en slik måte at det blir relativt enkelt å gjøre bruk av nye programpakker eller deler av disse etterhvert som de blir utviklet. Dersom vi ikke legger stor vekt på dette, vil systemet bli forbedret på få år. Systemet vil derfor kunne utvikles etter behov i form av nye tilleggsprogrammer.

2.1.9. Tilknytning for eksterne brukere

Et siktemål vil være at større brukere utenfor Byrået etterhvert skal kunne betjene seg selv via terminaltilknytning til databasen. Da det ikke skal legges inn beskyttede eller sikkerhetsmessig graderte data i basen, er det ikke nødvendig å innføre noen sperringer i databasen for eksterne brukere.

3. INTEGRERT DATABEARBEIDING

Som en del av arbeidet med å få til en rasjonalisering av statistikkproduksjonen, vil det med utgangspunkt i databasen bli lagt opp et system for integrert dataregistrering, revisjon og kontroll av regnskapsoppgavene. Ved denne integrerte databearbeidningen vil disse oppgavene bli kontrollert mot de reviderte oppgavene fra foregående år som allerede ligger i basen. Brukt på denne måten vil databasen dermed føre til en bedre aktualitet på statistikken også ved at bearbeidingstiden i Byrået reduseres.

4. DATADOKUMENTASJON

De variable som legges inn i databasen, vil bli godt dokumentert. Spesielt vil det bli lagt vekt på å dokumentere endringer i lover og forskrifter som fører til brudd i tidsserier.

4.1. Dokumentasjonsinformasjon

Dokumentasjonen av databasen vil bli gitt ut som egen publikasjon, som vil inneholde:

1. Variabelliste med variabelnr. og variabelbetegnelse
2. Variabeldefinisjoner og endringer i disse over tid
3. Enkel statistikk for hver variabel, 2 typer gjennomsnitt, max-min., gjennomsnittsavvik, totalsum etc. Denne listen vil også kunne brukes som en første tilnærming til materialet

5. FRAMTIDSPERSPEKTIVER

5.1. Statistiske analyseteknikker

Systemet bør etterhvert kunne gi beregninger som er fullt på høyde med de mest benyttede samfunnsvitenskapelige datapakker.

I systemet vil det ikke være noen begrensninger med hensyn til hvilke variable som kan kjøres mot hverandre. Slike begrensninger er heller ikke nødvendige, da basen ikke vil inneholde sensitive data.

Omkodinger og konstruksjoner av nye variable ut fra datamateriale vil det bli lagt vekt på å få til så brukervennlig som mulig.

5.2. Grafisk presentasjon

For visuelt å få fram sammenhengen mellom variable, er det behov for grafiske løsninger. I denne forbindelse vil det bli innarbeidet plott-diagrammer. I flere av de nyere samfunnsvitenskapelige programpakkene er det inkludert grafisk presentasjon i tillegg til statistiske analyseteknikker.

5.3. Kommunikasjon med andre databasesystemer

3. kontor ønsker at basen etterhvert også skal kunne tilknyttes eventuelle andre databaser på kommunenivå, slik at det blir mulig, på en relativt enkel måte, å kjøre variable i en base mot variable i andre baser.

5.4. Kartproduksjon

I et samarbeid mellom Byrået og NSD er kommunegrensene i Norge blitt koordinatfestet. Program er anskaffet av Byrået som kan gjøre bruk av dette grunnlagsarbeidet. For å bedre presentasjonen av kommunaløkonomiske data, er det ønskelig å nytte dette kartskraveringsprogrammet i tilknytning til databasen.

5.5. Modeller

I tilknytning til databasen vil det være mulig å bygge opp delmodeller for å beregne konsekvenser av f.eks. endringer i overføringsregler, lønns- og prisnivå, lovfestede pålegg fra staten etc.

6. VARIABELLISTE

6.1. Regnskapsvariable

Variabel nr.	Kommuneregnskap inntekter og utgifter (skjema 2)			
1	Kapittel:	1.07	Maskinanskaffelse	Post: 000-599 Bruttoutgifter
2	"	"	"	" 000-099 Lønn
3	"	"	"	" 100-149 Utstyr
⋮	⋮	⋮	⋮	⋮ ⋮ ⋮
⋮	⋮	⋮	⋮	⋮ ⋮ ⋮
3600	Kapittel:	1.99	Årets regnskapsover-/underskott	Post: 900-999 Finansieringsinntekter

Variabel nr.	Kommunenes balansekonto (skjema 3)			
3601	Kapittel:	2.70	Kasse, postgiro, bankinnskott	1. Kommunekassen
3602	"	"	" " "	12. Total sum
3603	"	"	" " "	13. Havnekasse
3604	"	"	" " "	14. Pensjonskasse
⋮	⋮	⋮	⋮ ⋮ ⋮	⋮
⋮	⋮	⋮	⋮ ⋮ ⋮	⋮
3964	Kapittel:	2.84	Annen gjeld	14. Pensjonskasse

Variabel nr.	Kommunebudsjett (skjema 1)			
1	Kapittel:	1.07	Maskinanskaffelser	Post: 000-599 Bruttoutgifter
2	"	"	"	" 000-099 Lønn
3	"	"	"	" 100-149 Utstyr
⋮	⋮	⋮	⋮	⋮ ⋮ ⋮
⋮	⋮	⋮	⋮	⋮ ⋮ ⋮
3600	Kapittel:	1.99	Årets regnskapsover-/underskott	Post: 900-999 Finansieringsinntekter

6.2. Ikke-regnskapsmessige variable

6.2.1. Fysiske variable 3. kontor benytter i dag

Folketall

Elever pr. klasse :

Antall senger og kurdøgn ved helseinstitusjoner

Antall elever ved grunnskoler og gymnas

Antall klasser ved grunnskoler og gymnas

Elever pr. klasse, gjennomsnitt

6.2.2. Variable av ikke-regnskapsmessig art som det er behov for (1)

6.2.2.1. Produktenheter som kan knyttes til fylkenes/kommunenes produksjon

Antall klasser i grunnskolen K

Antall elever i grunnskolen K

Antall klasser i videregående skoler F

Antall elever i videregående skoler F

Antall aldershjem

Antall plasser ved aldershjem

Antall pasienter ved aldershjem

Km kommunal vei

Km fylkesvei

Km riksvei

Antall kombinerte alders- og sykehjem

Antall plasser ved alders- og sykehjem

Antall pasienter ved alders- og sykehjem

Antall somatiske sykehus F

Antall plasser ved somatiske sykehus F

Antall pasienter gjennomsnittlig ved somatiske sykehus F

Antall barnehager K

Antall barnehageplasser K

Antall barn som benytter barnehageplassene K

(1) Denne listen er foreløpig, og må kunne utvides ettersom behovet for å kombinere regnskapsdata med andre variable øker.

6.2.2.2 Mengdedata som kan knyttes til produksjonsfaktorene (disse data kan legges inn fra 1982)

Kommunalt ansatte i sentraladministrasjon (antall)

"	"	" undervisning	"
"	"	" helsevern	"
"	"	" sosialomsorg	"
"	"	" kirke, kultur	"
"	"	" utb. og boligformål	"
"	"	" forretningsdrift	"
"	"	" ymse	"

6.2.2.3 Lønns- og prisindekser (disse indekstallene vil være felles for alle kommuner. Muligens vil de kunne beregnes særskilt for kommuner og fylkeskommuner)

Prisindeks andre driftsutgifter

Prisindeks nybygg og nyanlegg (1 stk.)

Prisindeks vedlikehold av bygg og anlegg (1 stk.)

Prisindeks utstyr (1 stk.)

Lønnsindekser (18 stk.) forskjellig for hvert kapittel og for endel underkap.

6.2.2.4 Sosioøkonomiske variable

Landsdel

Befolkningstetthet

Folketall

Kommuner gruppert i 13 grupper etter folketall

Kommunetypologier (9-delt)

Aldersfordeling antall personer i ulike aldersgrupper

0-5	år
5-7	"
7-14	"
14-16	"
16-20	"
20-30	"
30-40	"
40-50	"
50-60	"
60-67	"
67-70	"
70-75	"
75-80	"
80-85	"
85-90	"
over 90	år

Antall personer i de 16 aldersgruppene, endring fra året før

Kjønnfordeling: Antall menn

Antall kvinner

Andel av innbyggere som er skattytere

Inntekt gjennomsnitt pr. skattyter

Inntekt gjennomsnitt pr. innbygger

Antall arbeidsløse

Antall uføretrygdede

Antall med sosial stønad

Skilsmissefrekvens

Antall enslige forsørgere

Yrkesfrekvens kvinner

Yrkesfrekvens menn

Antall enslige over 67 år

Antall enslige 30-67 år

Dødelighet

Sykepengedager pr. innb. gjennomsnitt

Innpendling

Utpendling

Antall husholdninger

Andel av sysselsatte ansatt i primærnæringer

Andel av sysselsatte ansatt i sekundærnæringer

Andel av sysselsatte ansatt i tertiærnæringer

Boligmassens aldersstruktur

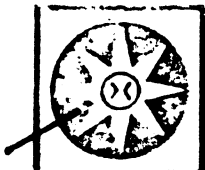
Andel av boliger bygd før 1920

" " " " 1921-1945

" " " " 1945-1970

" " " " etter 1970

Kommunestyrets politiske sammensetning (ca. 10 variable)



Søknad til
Norges almenvitenskapelige forskningsråd

Munthesgt. 29, Oslo 2. Telefon (02) 56 52 90

121

K. nr.:

Saksbehandler:

NAVF

Følg vedlagte veiledning.

Alle rubrikker skal fylles ut nøyaktig med maskin.

1	Råd RFSP	Fagfelt/Emneområde/Disiplin Sosialøkonomi Statsvitenskap Sosiologi	Kommunaløkonomi Database	Tidsrom for søknaden F.o.m ... 1/1-82 ... t.o.m. 31/12-82 ...
2	Søkerens navn eller ansvarlig institusjon: Statistisk Sentralbyrå (1)			Fødselsnummer:
	Nåværende arbeidssted: Statistisk Sentralbyrå, Dronningsgt. 16 Oslo 1	Telefon: (02) 41 38 20		Stilling:
	Privatadresse: Henv. Liv Bjørnland eller Jan Tore Pedersen	Telefon:		Akad. tittel:
3	Kort prosjektittel (maks 100 anslag) Database for kommunaløkonomi			

4 Søknadsbeløp fordelt på hovedposter. Spesifiseres på neste side.
Sosiale utgifter fylles ut av NAVF.

Hovedposter	1982	1983	19	19	19	Adm. notat Skriv ikke her
4.1 forskerlønn	63 077	100 000				
4.2 vikarutgifter						
4.3 vit. ass. lønn						
4.4 tekn./ktr.-ass. lønn	225 000	40 000				
4.5 driftsutgifter						
4.6 prosjektreiser						
4.7 utstyr, instrumenter						
4.8 kongresser						
Søkes dekket av NAVF ialt	288 077	140 000				
4.9 Dekkes av egen institusjon (2)						
4.10 Søkes dekket av andre	300 000	200 000				
Totalsum						

Adm. notat. (Skriv ikke her.)

- (1) Søknaden sendes under forutsetning av at prosjektet blir endelig godkjent av Statistisk Sentralbyrå høsten 1981
- (2) Totalbeløp for dette prosjekt er enda ikke kalkulert og det er derfor vanskelig å gi det nøyaktige beløp som skal dekkes av Statistisk Sentralbyrå

For 1972

4.1	Konsulent l.tr.20 i 7 mdr.	63 077
4.4	Konsulenttjenester i databasekonstruksjon	225 000

5 **Prosjektbeskrivelse. Presiser problemstilling, mål og metode samt framdriftsplan.**Database for kommunaløkonomi

På det kommunaløkonomiske felt er det behov for en database (jfr. st.meld.nr. 79 Langtidsprogrammet 1982-85). Vi har inneværende år fått 50 000 kr fra Kommunaldepartementet for å sette ned en forprosjektgruppe for det innledende arbeid med denne databasen. Departementet har lovet å stille seg imøtekommende til å støtte prosjektet i 1982.

Databasens omfang

1. Databasen kan etableres for årene fra 1974 da vi har god dokumentasjon for dette året (en evt. utvidelse bakover til 1972 kan overveies da vi har regnskapstall på magnetbånd fra dette år). Databasen skal senere holdes løpende ajour. Det må utvikles gode rutiner for oppdatering slik at foreløpige tall lett kan skiftes ut med endelige tall.
2. Følgende data skal overføres til databasen:
 - A) Samtlige data Byrådet samler inn over kommunenes og fylkenes regnskaper. Byrådet kan tenke seg en rutine hvor budsjettall først legges inn og løpende skiftes ut med foreløpige og endelige regnskapstall.
 - B) Fysiske data i relasjon til regnskapstallene slik at det etterhvert blir mulig å beregne enhetskostnader.
 - C) Pris- og lønnsindeksen for å gjøre det mulig å foreta beregninger i faste priser.

prosjektleders faglige bakgrunn for prosjektet.

K. nr.

Prosjektleder Jan-Tore Pedersen er mag.art. i sosiologi og har arbeidet i mange år med kommunaløkonomiske problemstillinger. Han har igjennom dette arbeidet fått praktisk erfaring i bruk av EDB på det samfunnsvitenskapelige felt. I en styrings-prosjekt-gruppe vil dessuten byråsjef Liv Bjørnland og førstekonsulent Bjørn Bleskestad delta fra Kontoret for finansstatistikk.

7 Prosjektets betydning for samfunnsplanlegging? Angi i tilfelle hvilken.

En viktig forutsetning for planlegging og forskning på området kommunaløkonomi er en god og lett tilgjengelig statistikk. Det er blant annet behov for et bedre informasjonsgrunnlag for å kunne analysere utviklingen i kommunale utgifter og inntekter. På denne bakgrunn foreslås det å opprette en database for kommunaløkonomisk statistikk i Byrådet. I tilknytning til denne databasen tenker en seg å bygge opp delmodeller for å beregne konsekvensene av f.eks. endringer i overføringsanalyse, lønns- og prisnivå etc. for kommunene.

8 Offentliggjøring av forskningsresultatene og formidling til almenheten og spesielle brukergrupper.

Et siktemål med databasen er at brukere utenfor Byrådet etterhvert skal få direkte kontakt med databasen. Da det ikke skal legges inn beskyttede data i basen, er det ikke nødvendig å innføre sperringer ved eksterne brukeres kontakt med basen. I tillegg vil Byrådet bli i stand til, på en rask måte, å foreta kommunaløkonomiske utkjøringer på bestilling. I dag kan slike bestillinger ta svært lang tid, da de ofte er svært arbeidskrevende uten noe databasesystem.

9 Referanser — høyst tre navn på forskere som kjenner søkeren og forskningsplanen og som NAVF eventuelt kan henvende seg til.

2) .Ekspedisjonssjef. Per. Øygard.....
Kommunaldepartementet

1) Avdelingsdirektør Juul Bjerke
Statistisk Sentralbyrå

3)

10 Liste over vedlegg.

Database for kommunaløkonomisk statistikk

.....
Sted og dato

.....
Underskrift

SYSTEM- OG DATABEGREPER

Et informasjonssystem er en samling prosedyrer, manuelle eller automatiske som samler inn og behandler data og formidler dataene som informasjon til brukere eller oppbevarer dataene for senere bruk.

Informasjonssystemet må utformes slik at det velger ut, behandler og arrangerer data slik at de danner meningsfylt informasjon til bedriftens behandlingsprosesser.

En bruker uttrykket system såvel om selskapets totale informasjonssystem som om mer spesialisert rutinesamlinger for bestemte formål, f.eks. et faktureringsystem.

Program

Et program er en entydig og detaljert beskrivelse av hvorledes en gitt oppgave skal utføres, på en måte som er akseptabel for datamaskinen.

Utilities

Med utilities forstås programmer og systemer som leveres som en del av datamaskinen for å lette bruken av den, f.eks. sortering eller kopiering av filen.

Styreprogram

Styreprogrammene er en del av operasjonssystemet som foreligger i datamaskinen for

- å tillate kommunikasjon mellom operatør og maskin
- å styre eksekvering av en kjede av brukerprogrammer
- å sørge for korrekt forbindelse med input/output og datamaskinen
- å kontrollere den fortløpende inn- og utmating av maskinen.

Operasjonssystemet

Er den samlingen av styreprogrammer, oversettelsesprogrammer, hjelpeprogrammer osv. som er nødvendige for på en effektiv måte å styre den produktive del av databehandlingen (brukernes jobber).

File eller register

En file er et maskinlesbart kartotek. Det består av en samling logiske records som minst har en felles karakteristikk (identifikasjonsbegrep) og som er organisert slik at gjenfinning er mulig.

IDENTIFIKASJONSBEGREP

Er et dataelement som identifiserer den enkelte record (kundenummer, varenummer).

Record

Er en samling dataelementer (felt) som refererer seg til samme identifikasjonsbegrep (kunderecord, varerecord).

Dataelement

Er den minste opplysningsenhet i en record som behandles. (Kvantum, pris, navn, kundenummer.)

Masterfile

Masterfile er et register eller tabellarisk oversikt over hva som er faste opplysninger eller hva som må betraktes som faste opplysninger på kort sikt. Eksempler på slike opplysninger er:

1. Navn, adresser og våre nummer for kunder
2. Navn, adresse og våre nummer for lagre
3. Navn, detaljer og våre nummer for varer

Transaksjoner

Transaksjoner, data f.eks. varelinjer på en faktura leses inn i et eller flere program for bearbeidelse. Opplysningene på en transaksjonfile vil variere fra kjøring til kjøring i motsetning til opplysninger på en masterfile og gir ofte grunnlag for data i rapporter, fakturaer, kontokuranter eller lign.

Datahistorikk

Data for en periode både for salg i volum, verdi, dekningsbidrag, omkostninger, utestående fordringer etc. lagres på datamaskinens media. Data oppbevares over flere perioder både for bruk i neste periode, men også over lengre tidsrom for spesifisert bruk som historikk vesentlig senere. Denne historikk er av stor betydning ved strategiske vurderinger, analyser og diverse andre formål. De begreper som disse opplysningene kan registreres på bør være permanente og uberørte av flest mulig typer omlegninger en kan tenke seg.

Jobb

En jobb består av et eller flere programmer samt filer som styres av operasjonssystemet via et sett styrekort (JCL).

Step

Et step er en del av en jobb som omfatter eksekvering av et program som kan være et systemprogram (utility) eller et brukerprogram. I JCL for steppet angis hvilket program som skal eksekveres og hvilke filer og maskinressurser som kan benyttes.

Abend

En unormal avslutning av et program. Programmet stoppes vanligvis og gir tilbake en kode som forteller om feilens art. Feilen kan ligge i styrekortene til programmet eller i selve programmet. Hvis det siste er tilfelle, finner man årsaken til ovennevnte "abend" ved å bruke en

dump

som automatisk blir generert hvis en "abend" oppstår.

Dette er en hexadesimal (et tallsystem med 16 som grunntall) utskrift av hva som befinner seg i den del av maskinhukommelsen som programmet bruker.

Med hjelp av visse regler for ovennevnte "abends", kan en finne årsaken til den unormale avslutningen av programmet.

Return-code

Dette er et spesielt register som det operative system eller det opererende program kan gjøre seg bruk av. Med andre ord dette er en melding som enten er lagt av systemet eller av programmet og sier noe om utfallet av programkjøringen, f.eks. for automatisk å dirigere den videre kjøring av jobben.

Backup

Med backup forstås tiltak for å kunne rekonstruere vitale data dersom de skulle bli ødelagt eller gå tapt på annen måte. Et typisk backuptiltak er kopiering av diskfiler til tape. En har også backup av selve datamaskin gjennom avtale med annen innehaver av samme type maskin om gjensidig assistanse ved maskinhavari.

Programbibliotek

Den mest alminnelige måte å oppbevare og å jourholde programmer på under maskinsystemet OS er å legge dem ut på bestemt område avsatt på disk. Dette diskområdet opprettes som et vanlig EDB datasett, men organiseres på en måte som kalles "partitioned organisation". Et "partitioned" datasett brukt som bibliotek opprettes i to former, source og load form. Biblioteket i source form inneholder programmer i den form det er gitt av programmererne, biblioteket i load form inneholder aksekverbare programmer. Med dette menes et program som er oversatt til maskinkoder og hvor de forskjellige deler av programmet er kjedet sammen slik at det kan utføre det som er kodet av programmereren.

Før et man skal finne det program man ønsker, slår systemet opp i en innholdsfortegnelse over datasettet. Dette kalles "directory" og forteller hvor i datasett programmene befinner seg.

Programmene i source-versjon må inneholde beskrivelse av recordene for alle filene programmet benytter seg av. Ved standardisering av feltnavn vil alle beskrivelser av samme record være like uansett program. Derfor har en opprettet et eget bibliotek, COPY-biblioteket som inneholder recordbeskrivelsene og annen kode som er felles for mange programmer. Disse kodesekvensene blir da automatisk forent med koden på sourcebiblioteket når programmet skal kompileres, det vil si at det skal lages et program i load form ut fra source programmet. Ved bruk av COPY-biblioteket oppnår man plassbesparelse på sourcebiblioteket, lettere programvedlikehold og raskere utvikling av nye programmer (programmereren slipper å skrive alle koder selv).

Tape og JCL bibliotek

Tapebibliotek er opprettet for lettere å kunne føre kontroll med taper som er i bruk og innholdet på disse. Fra tapen hentes opplysninger som lagres i et bibliotek. Det kan f.eks. være informasjon om hvilket system tapen tilhører, hvilken jobb i systemet tapen brukes, dato den er tatt i bruk og måten dataene er lagret på.

JCL biblioteket inneholde styrekort som brukes fast i produksjoner. Disse er registrert pr. system, job og step og kan på en enkel måte endres og tilpasses endringer i produksjonsopplegget. Styrekort som er gjort produksjonsklare på biblioteket hentes ut og legges feilfrie på jobbkøen klar til kjøring.

Databaser

Stersom programmeringsteknikken er blitt bedre har også kravet til selve måten å lagre data på økt. Det er flere ting å ta hensyn til ved lagring av data-plasutnyttelse, accesstider osv.

Et annet ønske var å gjøre programmene uavhengige av lagringsmåten, dvs. å gjøre dem datauavhengige. Flere leverandører av maskiner og software har laget systemer som mer eller mindre fyller disse kravene. Generelt kalles denne type systemer for Data Base Management Systemer eller Data-base-systemer.

Noen definisjoner:

PERSONINTEGRITET (PRIVACY)

Individets rett til å kontrollere data om sin egen person.
Andre individers rett til å få tilgang til disse data.
Dette er ikke et datateknisk problem alene, men snarere
juridisk, politisk og datateknisk.

DATA SIKKERHET (DATA SECURITY)

Beskyttelse for et datasystem mot:

- feilaktig modifisering
- ødeleggelse

Tiltak for begrensning/eliminering av skadevirkninger
etter ovenstående.

BEGRENSET TILGANG

Fastsettelse av regler for hvem som gjør hva i systemet.

IDENTIFIKASJON

Systemets mulighet for klassifisering og særbehandling
av ulike resurser som f. eks. brukere, terminaler,
programmoduler.

AUTORISASJON

Systemets måte å avgjøre om en bestemt (identifisert)
resurs har tilgang til en annen resurs (f. eks. dataelement/
modul)

KONTROLL

Funksjonene for kontroll og oppfølging av datasikkerhet.
(Inneholder funksjoner for umiddelbart å oppdage og
forhindre forsettelige eller tilfeldige hendelser som kan
ødelegge/bryte datasikkerheten)

FYSISK BESKYTTELSE

Beskyttelse av utrustning og personale mot påvirkning
og ulykker, samt beskyttelse mot spionasje.

SYSTEMINTEGRITET

Et mål for beskyttelse mot og følgende av ulike feil,
ulykker etc.

TILGJENGELIGHET

Den tid som systemet kan brukes i % av den tid som
systemet er bestemt å kunne brukes.

TILLIT TIL DATA

Sansynligheten for at de data som systemet produserer er
korrekte.

