

*Anne Gro Hustoft og Grete Olsen*

**Metadata for statistikk om  
personer og husholdninger**  
Forprosjektrapport

## **Forord**

Dette notatet er en forprosjektrapport for prosjektet "Metadata for statistikk om personer og husholdninger" (MOPH). Formålet med hovedprosjektet er å strukturere SSBs metadata ( i første rekke knyttet til person- og husholdningsstatistikkene) slik at relevant dokumentasjon gjøres tilgjengelig internt og eksternt på elektronisk form.

Forprosjektrapporten er ment som et innspill til årets VP-prosess der det må avgjøres om, og med hvilket ressursomfang, prosjektet skal videreføres i et hovedprosjekt i 2001. Målet for forprosjektet har vært å identifisere hvilke statistikker et eventuelt hovedprosjekt skal konsentrere seg om, kartlegge status knyttet til dokumentasjon av disse statistikkene og komme med et forslag til videre arbeid.

## Innholdsfortegnelse

<b>1. Innledning</b> .....	<b>3</b>
<b>2. Statistikk som inngår i første fase av prosjektet</b> .....	<b>3</b>
<b>3. Kartlegging av dokumentasjon</b> .....	<b>3</b>
<b>4. Brukerbehov</b> .....	<b>4</b>
<b>5. Forhold til andre systemer i SSB</b> .....	<b>4</b>
<b>6. Forslag til strukturering av metadata</b> .....	<b>5</b>
<b>7. Forslag til videre arbeid - innspill til VP-prosessen</b> .....	<b>8</b>
<b>VEDLEGG 1</b> .....	<b>11</b>
"Core variables" for Eurostats Social statistics.....	11
<b>VEDLEGG 2</b> .....	<b>12</b>
Liste over person- og husholdningsstatistikker som er tatt med i MOPH.....	12
<b>VEDLEGG 3</b> .....	<b>14</b>
Kartlegging av MOPH-statistikkene.....	14
<b>De sist utgitte publikasjonene i serien Notater</b> .....	<b>34</b>

# 1. Innledning

Formålet med dette prosjektet er å strukturere SSBs metadata (i første rekke metadata knyttet til person- og husholdningsstatistikkene) slik at relevant dokumentasjon gjøres tilgjengelig både internt og eksternt på elektronisk form. All dokumentasjon som gjøres tilgjengelig eksternt, vil også være tilgjengelig internt, men det motsatte vil ikke være tilfellet. Struktureringen av metadata er viktig for bedre å kunne få fram sammenhenger mellom SSBs sosiale indikatorer, samordningen vil lette arbeidet med store tverrgående prosjekter som nytter flere datakilder (f.eks. FoB) og det vil også være et nyttig bidrag i forhold til Eurostats arbeid med sosiale indikatorer. Det er i dag et problem at SSBs dokumentasjon av en del statistikker ikke er lett tilgjengelig, noe som kan forårsake mye ekstraarbeid for medarbeiderne.

Dette notatet presenterer resultatet av en forprosjektfase der målet har vært å identifisere hvilke statistikker vi i første omgang skal konsentrere oss om, kartlegge status for dokumentasjonen knyttet til disse statistikkene og på bakgrunn av dette komme med innspill til VP-prosessen vedrørende første trinn i et hovedprosjekt. Det er lagt vekt på å få til et realistisk ambisjonsnivå, noe som innebærer å bygge på datasystemer som allerede eksisterer, og dele det videre arbeid inn i ulike trinn som hver for seg vil gi en nyttegevinst, uavhengig om en velger å gå videre med et nytt trinn eller ikke. Hvorvidt en velger å bygge ut dette til et "fullstendig" metadatasystem, vil avhenge av tilgjengelige ressurser og erfaringene fra det arbeid som gjøres.

## 2. Statistikker som inngår i første fase av prosjektet

På det første møtet i referanse- og styringsgruppen ble det vedtatt at seksjonssjefene skulle sette opp en liste over hvilke av deres statistikker som skulle inngå i prosjektets første fase. Dette skulle gjøres med utgangspunkt i de 26 kjernevariablene som Eurostat har definert for sin "Social statistics". Disse variablene, samt de av SSBs statistikker som ble ansett å være knyttet til kjernevariablene, er tatt med i hhv. vedlegg 1 og 2.

## 3. Kartlegging av dokumentasjon

For å få en pekepinn om ressursbehovet knyttet til et framtidig metadatasystem - og dermed hvilket ambisjonsnivå som er realistisk ut fra de tilgjengelige ressurser - var det nødvendig å finne ut hvilken dokumentasjon som i dag finnes for de relevante statistikkene.

Prosjektet har derfor kartlagt hva som finnes av dokumentasjon generelt, men med spesiell vekt på datalagring. I SSB skal i prinsippet alle filer som langtidslagres, lagres som flate filer og dokumenteres i Datadok. Kartleggingen viser at dette ikke stemmer helt med de faktiske forhold. Av de undersøkte statistikkene er det bare ca. halvparten av de langtidslagrede filene som er lagret som flate filer (vi har da regnet befolknings- og utdanningsstatistikken som to statistikker fordi alle enkeltstatistikker under disse emnene er lagret på samme måte). Resten er lagret som SAS-filer eller i Oracledatabaser. Av de flate filene er det bare et mindretall som er dokumentert i Datadok. Det er stor variasjon mellom seksjonene når det gjelder hva som er dokumentert, og hvor lett tilgjengelig dette er.

Hvis en skal kunne bruke Datadok som utgangspunkt for et metadatasystem (og i samsvar med SSBs vedtatte prinsipper), er det nødvendig at alle langtidslagrede filer etterhvert finnes som flate filer og dokumenteres i Datadok. Ressursbehovet for å oppnå dette, vil variere fra seksjon til seksjon, alt etter

hvor mye seksjonen allerede har dokumentert. I Oracleutvalget som er nedsatt i SSB, arbeider en nå med å finne ut hvordan en kan få lagt Oraclefiler som skal langtidslagres, ut som flate filer.

Dokumentasjonen av kartleggingen finnes i vedlegg 3, både i en fullstendig utgave og som oppsummering i en tabell.

## 4. Brukerbehov

Som nevnt innledningsvis skal metadatasystemet ivareta både interne og eksterne behov for dokumentasjon. Etter diskusjoner i styrings-, referanse- og prosjektgruppe var det enighet om at de eksterne brukerne i første omgang skulle ha tilgang til informasjonen som finnes i "Om statistikken" og i den tekstlige beskrivelsen i NOS. Informasjonen i disse to publikasjonene vil til en viss grad være overlappende, men NOS-teksten vil på noen punkter være mer detaljert. Det er imidlertid ikke alle statistikker som utgis i NOS, og for disse skal "Om statistikken" (etter den mal som er vedtatt i DM) gi tilstrekkelig informasjon til å nyttiggjøre seg statistikken. Brukere på forskernivå e.l. som trenger ytterligere detaljer og eventuell tilgang på ekstra data, vil som før måtte henvende seg til seksjonene.

For interne brukere skal dokumentasjonen lette arbeidet for den enkelte medarbeider og gjøre kommunikasjonen mellom seksjonene enklere. Det er i dag et problem ved mange statistikker at det kan være vanskelig bl.a. å finne ut hvor data er lagret, hvordan ulike variable er definert og beregnet, hvilke koder som brukes for hva og hvilke endringer som er foretatt fra en statistikkårgang til en annen. Mye av dette er dokumentert, men ligger kanskje i skuffer (eller finnes bare i enkelte medarbeideres hoder), og er derfor ikke generelt tilgjengelig verken for nye medarbeidere på en seksjon, eller for folk på andre seksjoner. Et metadatasystem skal avhjelpe dette ved at den viktigste informasjonen finnes elektronisk tilgjengelig internt i SSB. Hadde et slikt system eksistert i dag, ville det gjort livet atskillig enklere bl.a. for medarbeiderne i FoB som samler inn data fra ulike kilder, og som derfor ofte støter på problemer knyttet f.eks. til manglende variabeldefinisjoner og kodelister der bl.a. koder for ulike typer manglende data ikke er oppgitt.

Etter diskusjoner i prosjektet ble det avgjort at en skulle starte metadatadokumentasjonen med variabeldefinisjoner. Behovet for å få disse på plass ble opplevd som det mest presserende både når det gjaldt intern bruk og for kontakten med eksterne brukere. Arbeidet med å knytte beskrivelsen "Om statistikken" til de ulike frigivelsene (statistikkene) er allerede i gang, og NOS-tekstene finnes selvfølgelig for de statistikker som utgis i NOS.

## 5. Forhold til andre systemer i SSB

I løpet av forprosjektet ble det tatt opp hvorvidt Produktregisteret skulle trekkes nærmere inn i utviklingen av et metadatasystem. Etter en del diskusjon konkluderte en med at det var viktig å trekke inn dette på sikt, og "linke" det til andre deler av metadataprojektet, men at det i prosjektets første trinn ikke var hensiktsmessig å ta det med fordi det var for lite oppdatert. I det videre arbeid må en vurdere hvordan en skal forholde seg både til Produktregisteret, Tidsseriedatabasen (TD) og Regionaldatabasen (RD). Forholdet til STABAS (STAndard dataBASen) må vurderes allerede i neste trinn. Det vil være naturlig, siden vi konsentrerer arbeidet om variablene, å ikke bare kunne få tilgang til deres definisjoner, men også til hvilke standarder de publiseres etter ( om noen). STABAS er under utvikling og de fleste standarder er derfor ennå ikke lagt inn i databasen, men mulighetene for link mellom variabelmetadata og STABAS bør ses nærmere på allerede i neste års arbeid.

## 6. Forslag til strukturering av metadata

Det er laget en figur som gir en grov oversikt over et metadatasystem. Her er kilderegister, fellesregister og sluttfiler skilt ut som adskilte trinn i en prosess, men i noen tilfeller (for noen statistikker) kan disse til en viss grad være overlappende. Det er knyttet metadata til enhver langtidslagring av data.

Vi har prøvd å ivareta

- metadata for registrene
- at ulike trinn i produksjonsprosessen dokumenteres
- at metadata som er beskrevet et sted (f.eks. for et register) skal "arves" av statistikker som bruker disse registrene slik at en ikke trenger definere f.eks. variabler mer enn et sted
- at metadataene er lagt opp slik at NOS og "Om statistikken" skal kunne produseres på grunnlag av dem (de er heller ikke særlig mer omfattende enn det som behøves for å kunne produsere disse to "publikasjonene" )
- at variablene som bruker standarder skal kunne "peke" til de aktuelle standardene i STABAS

Det må også defineres et mer detaljert dokumentasjonsnivå enn det som er beskrevet i nåværende forslag (ulike dokumentasjonsrapporter, flytskjema etc.), men vi vil i denne omgang bare si at det hører med, og så konsentrere oss om det øverste nivået som trengs for å produsere NOS og "Om statistikken".

Vi vil foreslå at vi i størst mulig grad baserer oss på allerede eksisterende dokumentasjonssystemer i SSB. Et av dem er Datadok som hovedsakelig brukes til å lagre teknisk informasjon (format, variabelnavn, kodelister) om flate filer. Mer tekstlige metadata vil det være naturlig å gjøre tilgjengelig i et hypertekst/intranettbasert system. Q:\Dok brukes i dag som lagringssted for tekstlig informasjon (hovedsakelig Wordfiler) knyttet til ulike statistikker (i sin nåværende form mest til utviklingsprosjekter). Hvis en velger å bruke Q:/Dok, blir det nødvendig å opprette en nærmere tilknytning til Datadok slik at en variabel i Datadok f.eks. kan "peke" på et dokument i Q:\Dok der denne variabelen er definert. Slik dokumentene på Q:\Dok er organisert i dag, er de til dels vanskelig tilgjengelige for "uinnvidde". Derfor kan "pekerkontakten" med Datadok også virke strukturerende på dokumentene i Q:\Dok. Det finnes i dag ingen slik sammenheng mellom Datadok og Q:\Dok., den må eventuelt implementeres i hovedprosjektet. En må i det videre arbeid avklare nærmere hvilke data som skal lagres hvor og foreta en grenseoppgang for de ulike delene av metadatasystemet. Men dersom en velger Q:/Dok som lagringssted, vil det også være naturlig, når den tid kommer (for noen statistikkens vedkommende er den allerede kommet), at den mer detaljerte dokumentasjonen også legges der.

Som vist i figuren, skal metadataene som lagres, kunne være utgangspunkt for de to dokumentene som skal være tilgjengelige for "allmennheten"; NOS og "Om statistikken". Disse vil også ligge på statistikkens område på Q:\Dok. (eller annet lagringssted som velges for tekstlige metadata). For å gjøre disse tilgjengelige på internett, må først Wordfilene gjøres om til XML (gjøres av en rutine som S730 har, Excelfiler gjøres om til SGML). Disse filene blir deretter via en annen rutine lagt over på web-serveren. Alt dette ligger fortsatt innenfor brannmuren. For at filene skal gjøres tilgjengelige for andre på internett, må de kopieres ut på en tilsvarende web-server som er bygd opp utenfor brannmuren. Et slikt arbeid må organiseres med ansvarlige på hver seksjon, men det er mulig at rutine som brukes for å legge dokumenter ut på Internett i dag, er tilstrekkelige.

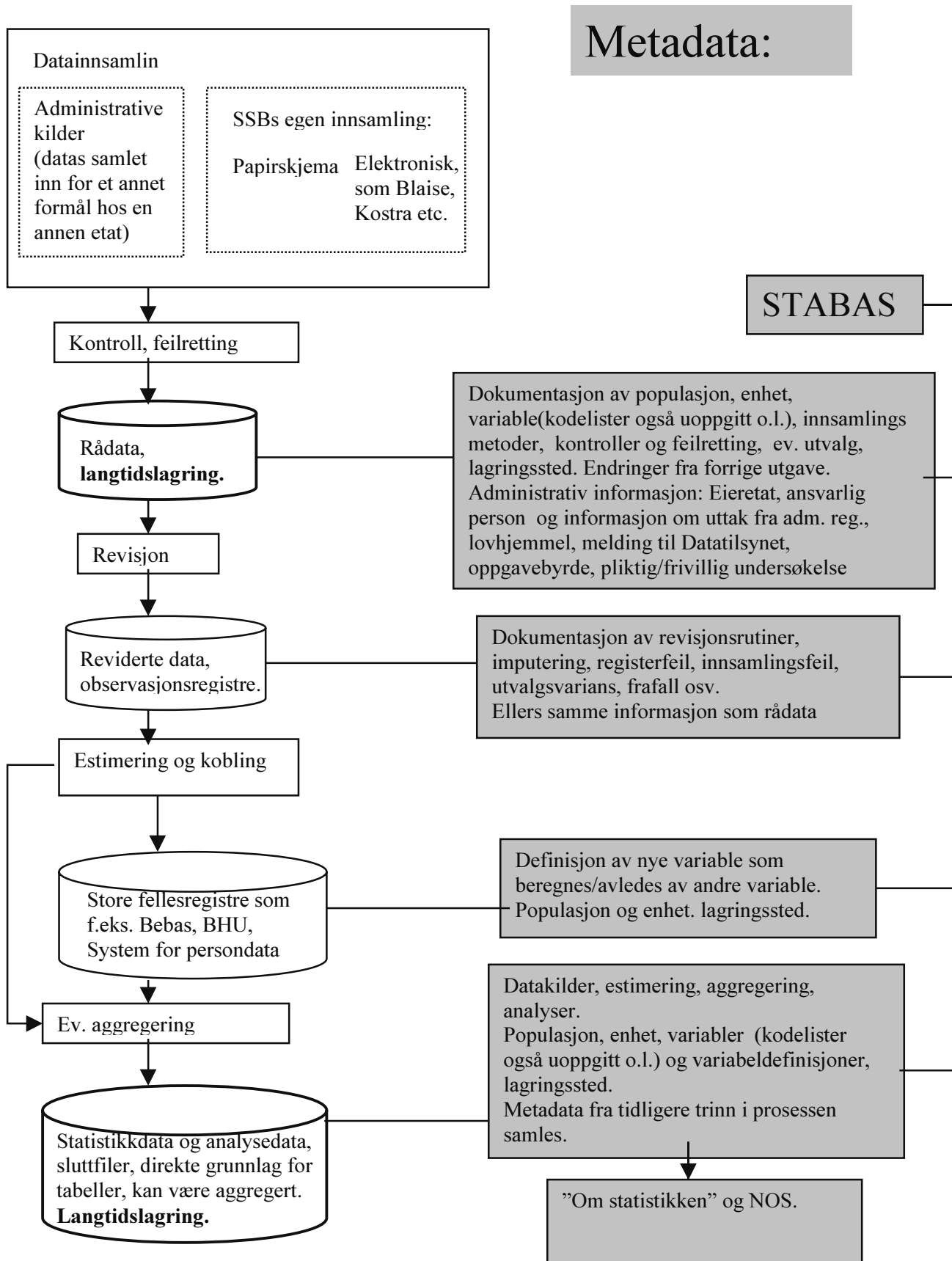
Wordfiler/Excelfiler som skal legges ut på intranettet, må også gjøres om til XML/SGML via de samme rutine. Det har vært diskutert å legge Q:/Dok på intranettserveren, men det er foreløpig ikke gjort mer med denne saken.

Det ble på et møte i styrings- og referansegruppene tatt opp hvorvidt dataadministratorene kunne bidra i forhold til Datadok. Inger Helene Ruud på seksjon S710 opplyser at dataadministratorene nå er i

gang med et arbeid for å fokusere på bruken av navnestandarden i Datadok. De har ingen ressurser avsatt til å dokumentere filer i Datadok, og de har heller ikke nok kunnskap om de enkelte filer til å gjøre dette. Da må i tilfellet de enkelte fagseksjoner foreta selve spesifikasjonen av filene først. Det er ønskelig at en i VP-prosessen diskuterer om dataadministratorene kan trekkes mer inn i selve dokumentasjonsarbeidet, og hvordan arbeidet i såfall kan fordeles mellom dem og fagseksjonene.

Figuren gir som nevnt kun en skjematisk oversikt over hvilke metadata som skal tas med og hvor de skal lagres. Dette må spesifiseres mer nøyaktig, bl.a. via en datamodell, i prosjektets neste trinn. Som tidligere nevnt, skal den informasjonen som blir lagt inn i metadatasystemet, være nok til å kunne produsere "Om statistikken" og teksten i NOS. De som allerede har laget disse publikasjonene, har dermed i utgangspunktet dokumentert hovedtyngden av de metadata som skal inn i systemet. Når det er laget et utkast til datamodell, må seksjonene få dette til gjennomsyn slik at det kan sjekkes om denne også tilfredsstillende de behov seksjonen har til metadatarapportering til internasjonale organisasjoner (IMF, OECD...). En må også koordinere dette med pågående kvalitetsarbeidet i SSB.

# Metadata i produksjonsprosessen





## 7. Forslag til videre arbeid - innspill til VP-prosessen

Aktiviteter i trinn 1 av hovedprosjektet

1. Alle langtidslagrede filer i statistikkene som "tilhører" prosjektet, og som produseres neste år, skal lagres som flate filer og dokumenteres i Datadok. Seksjonene skal også lage en plan over når eldre filer skal gjennomgå samme prosess (vi må bli enige om hvor langt bakover i tid filer skal dokumenteres på denne måten). Det er viktig at navnestandarden følges i dokumenteringsarbeidet, og VP-prosessen må diskutere i hvilken grad dataadministratorene skal trekkes inn i aktivitet 1 i MOPH-prosjektet.

2. Det må spesifiseres hvilke metadata som skal lagres i denne omgang (det minst detaljerte nivået tilpasset "Om statistikken" og NOS-malen), og hvor disse skal lagres. Det må bl.a. avsettes IT-ressurser til å stå for datamodelleringen. Web-folkene må også trekkes inn når "lokalisering" skal avgjøres for å avklare bruken av metadataene i forhold til intranett /internett (f.eks. spørsmålet om Q:/Dok bør legges på intranettserveren, og hvordan wordfilene med tekstlige metadata bør utformes/struktureres for lettest mulig å kunne konverteres til XML-filer og tilpasses intranett/internett.). Arbeidet med denne aktiviteten bør starte så tidlig som mulig i 2001.

De spesifiserte metadatene må også sjekkes i forhold til seksjonenes krav til internasjonal rapportering.

3. Tilknytningen til STABAS må beskrives (hvor langt en kan gå i implementering vil bla. avhenge av tilgjengelige IT-ressurser og framdriften i STABAS). En bør i løpet av 2001, før et eventuelt trinn 2 av hovedprosjektet settes i gang, vurdere forholdet til Produktregisteret, TD og RD.

4. Det må opprettes nødvendige "linker"/"pekere" mellom de ulike dokumentasjonsstedene (f.eks. Datadok og Q:/Dok dersom disse velges som lagringsområder). Disse må også klargjøres slik at de relevante data kan legges inn.

5. Det må lages variabeldefinisjoner til filene som lagres i Datadok. Her bør en starte med noen pilotprosjekter, f.eks. en statistikk og et register. Mulige kandidater her er utdanningsstatistikken og BHU der det allerede er gjort et arbeid via PROMUT.

Alle seksjoner som deltar i prosjektet, må lage en tidsplan for når variabeldefinisjonene for deres person- og husholdningsstatistikker skal være ferdige.

6. Dersom Q:/Dok velges som lagringssted for metadata, må det ryddes opp i den dokumentasjon som allerede finnes der. MOPH-statistikker som bruker Q:/Dok, må finne "forståelige" navn på katalogene (dersom de ikke allerede har det), og legge inn et dokument som forteller hvilken dokumentasjon som finnes i de ulike katalogene.

7. Det er viktig å klarlegge ulike spørsmål knyttet til implementering/bruk av et metadatasystem; f.eks. hvordan håndtere variabler med samme navn, men med ulike definisjoner i ulike statistikker, hvordan håndtere variabler som endrer definisjon, men som (i alle fall ifølge dagens praksis) beholder samme navn, hvordan få til en god håndtering av f.eks. registre som endrer sine variable ofte (mange versjoner) etc.

8. Alle MOPH-statistikkene skal i løpet av 2001 ha dokumentert de metadata som er nødvendige for å lage "Om statistikken" (og ev. NOS-teksten). "Om statistikken" skal foreligge for MOPH-statistikkene.

9. Et eventuelt trinn 2 i prosjektet skal planlegges på bakgrunn av erfaringene fra trinn 1 og ønsket ambisjonsnivå for metadatasystemet.

Anslag over ressursbehov og relevant tidsplan må diskuteres i VP-prosessen. På neste side gis derfor en oversikt i tabells form over oppgaver og ansvarlige (også for å komme med ressursanslag). Det anses som viktig å få startet oppgave 3 så tidlig som mulig neste år da den er forutsetning for mye av det andre som skal gjøres i prosjektet.

I løpet av VP-prosessen må det også avgjøres som prosjektmedarbeiderne skal være de samme som i forprosjektet, eller om andre (eventuelt i tillegg) skal delta i arbeidet.

<b>Oppgaver</b>	<b>Ansvarlig</b>	<b>Hvem anslår ressursbruk</b>
1. Langtidslagrede filer som produseres i 2001, skal lagres som flate filer og dokumenteres i Datadok	Fagseksjonene, ev. sammen med dataadministratorene	Fagseksjonene, ev. også dataadministratorene
2. Plan over når eldre filer skal gjennomgå samme prosess	Fagseksjonene	Fagseksjonene
3. Spesifikasjon av hvilke data som skal inngå i systemet, hvor disse skal lagres og datamodellering.	Prosjektleder, IT-prosjektmedarbeidere (både fra personstatistikk og web-gruppen) Fagseksjonene vil også bidra med innspill om hvilken type data som skal med	Prosjektleder, IT-prosjektmedarbeidere
4. Oppretting av "linker"/"pekere" mellom de ulike dokumentasjonsstedene.	IT-prosjektmedarbeidere	IT-prosjektmedarbeidere
5. "Klargjøring" av dokumentasjonssteder i henhold til datamodell	IT-prosjektmedarbeidere	IT-prosjektmedarbeidere
6. Lage tidsplan for når variabeldefinisjonene skal være ferdige	Fagseksjonene	Fagseksjonene
7. Klargjøre og legge inn variabeldefinisjoner knyttet til utdanningsstatistikken og BHU i metadatasystemet.	Prosjektleder, S320 og IT-prosjektmedarbeidere	S320 og IT-prosjektmedarbeidere
8. Dersom Q:/Dok velges som lagringssted for metadata, må det ryddes opp i det som allerede ligger her.	Fagseksjoner som benytter Q:/Dok	Fagseksjoner som benytter Q:/Dok
9. Evaluering av trinn 1 og eventuell planlegging av trinn 2 for metadatasystemet.	Prosjektleder	Prosjektleder

## VEDLEGG 1

### "Core variables" for Eurostats Social statistics

1. Place of usual residence
2. Region
3. Locality
4. Country of birth
5. Country of citizenship
6. Sex
7. Age
8. Legal marital status
9. De facto marital status
10. Housing unit
11. Private household
12. Household relationship
13. Reference person
14. Labour status
15. Status in employment
16. Hours worked
17. Income
18. Occupation
19. Educational attainment
20. Current education and training activity
21. Economic activity
22. Main social status
23. Health status
24. Socio-economic status
25. Tenure status of household
26. Number of rooms

## VEDLEGG 2

### Liste over person- og husholdningsstatistikker som er tatt med i MOPH

Seksjon	Statistikk
260	2090 AKU 2101 Registerbasert sysselsettingsstatistikk 2106 Registrerte arbeidsløse/tiltak 2111 Sykefravær, registerbasert 2112 Sykefravær, skjemabasert
310	Samboerstatistikk
320	Utdanningsstatistikk. Elever under lov om videregående opplæring Befolkningsstatistikk. 1., 2., 3. og 4. kvartal og endringer i 4., 1., 2. og 3. kvartal Utdanningsstatistikk. Lærlinger og avlagte fagprøver Befolkningsstatistikk. Overgang til norsk statsborgerskap Befolkningsstatistikk. Folkemengde etter alder, kjønn og sivilstand, 1. januar Befolkningsstatistikk. Født Befolkningsstatistikk. Utenlandske statsborgere, 1. januar Utdanningsstatistikk. Lærere, 1. oktober Befolkningsstatistikk. Døde Befolkningsstatistikk. Flyktninger, 1. januar Befolkningsstatistikk. Adopsjoner Befolkningsstatistikk. Barn, 1. januar Befolkningsstatistikk. Flyttinger Befolkningsstatistikk. Inngåtte ekteskap og registrerte partnerskap Utdanningsstatistikk. Studenter ved universiteter og vitenskapelige høyskoler Befolkningsstatistikk. Innvandrerbefolkningen, 1. januar Befolkningsstatistikk. Skilsmisser og separasjoner Utdanningsstatistikk. Elever og studenter, 1. oktober Befolkningsstatistikk. Beregnet folkemengde. Nasjonale og regionale tall, 1. januar BHU
330	Dødsårsaksstatistikken Helseundersøkelsen Pleie- og omsorgsstatistikken Svangerskapsavbrudd Steriliseringer Pasientstatistikk Spesialisthelsetjenesten, psykiatri
350	Utvalgsundersøkelser 0202 Levekårsundersøkelsen. Tverrsnitt 0204 Levekårsundersøkelsen. Panel 0208 Tidsnyttingsundersøkelsen 1201 Forbruk og sparing 5901 Mediebruk 5903 Kultur- og mediebruk

Totaltelling

0671 Barnevernsstatistikk - individ

0731 Sosialhjelpsstatistikk - individ

0771 FD-Trygd (forløpsdatabase med data om trygd, inntekt, utdanning sysselsetting etc)

0810 Familievernstatistikk (ny, data for første årgang i 2001)

0821 Anmeldte lovbrudd

0822 Etterforskede lovbrudd

0823 Reaksjoner for lovbrudd

0825 Fengslinger

370 FoB

420 1610 Inntekts- og formuestatistikk for husholdninger

Omnibus

### Kartlegging av MOPH-statistikkene

#### S260

#### Registrerte helt arbeidsledige og personer på arbeidsmarkedstiltak

Hyppighet:

Kvartalsvis og årlig.

Type statistikk:

Fulltelling, regionalfordelt på kommuner. Kvartals- og årsgjennomsnitt.

Datakilder:

Data kommer som månedsfiler fra aetat Arbeidsdirektoratet én gang per måned, og settes sammen til kvartalsfiler før produksjon. Statistikk for personer registrert som helt arbeidsledige er f.o.m. 1988 basert på opptellinger i SOFA-søker-registeret ved utgangen av hver måned. Tilsvarende data for personer på arbeidsmarkedstiltak finnes f.o.m. 1991. Opplysning om bostedskommune er i hovedsak hentet fra SSBs system for befolkningsstatistikk. Fil med beregnet AKU-arbeidsstyrke på kommunenivå brukes som grunnlag for beregninger av prosentvis arbeidsledighet (denne filen finnes også i SSBs datasystem.)

Koding/revisjon:

Det kontrolleres for nye typer arbeidsmarkedstiltak. Dette gjøres ved å kjøre kontrolltabeller som sammenlignes med aetats egen publisering i "Månedstatistikk om arbeidsmarkedet". Vi tar som regel kontakt for å få oppgitt nye tiltakskoder.

Lagring:

Filer med månedsdata langtidslagres på URSUS som SAS-filer.

Lagring rådata: \$SOFA/DATA/MNDSOEK

Dokumentasjon:

Statistikken dokumenteres i Regionalstatistikk (RS), hvor ledighetstallene publiseres. En enkel dokumentasjon av produksjonssystemet for SOFA-søker-registeret finnes i Datadok (Q:). Ellers er produksjonssystemet dokumentert, men mangelfullt oppdatert.

Standarder:

Kommune/fylke

Alder

Annet:

Arbeidsdirektoratet har det administrative ansvaret for SOFA-søker-registeret, som er hovedkilden for denne statistikken.

#### Arbeidskraftundersøkelsene (AKU)

Hyppighet:

Kvartalsvis og årlig, men månedlig publisering av noen sesongjusterte nøkkeltall (3-måneders glidende gjennomsnitt).

Type statistikk:  
Utvalgsundersøkelser.

Datakilder:

Hovedkilden for AKU er kvartalsvise, representative utvalgsundersøkelser. Datainnsamlingen foretas gjennom intervju per telefon (eventuelt besøk hvis ikke tlf.kontakt). En del opplysninger gitt i tidligere intervju brukes ved etterfølgende intervju. For koding av næring hentes det inn informasjon fra arbeidstakerregisteret og SSBs Bedrifts- og foretaksregister (på individnivå). For øvrig innhentes demografiske data fra personregisteret, og utdanningsdata fra SSBs register over befolkningens høyeste utdanning.

Koding/revisjon:

Under selve intervjuet, som foregår ved bruk av PC-basert spørreskjema (CAI), er det lagt inn flere maskinelle rutiner for å forebygge feil svar og/eller feil registrering av svar, f.eks. når det gjelder antall arbeidstimer i løpet av uka. I noen tilfeller får intervjueren "advarsler" ved registrering av svar, mens det i andre tilfeller er lagt inn maksimumsverdier.

Koding av næring, yrke og utdanning gjøres ved seksjon 450. Til næringskodingen gjøres det bruk av maskinelle oppslag i SSBs Bedrifts- og foretaksregister (koplet til arbeidstakerregisteret på individnivå). Utdanningsdata innhentes maskinelt fra SSBs register over befolkningens høyeste utdanning, men intervjuobjektene spørres om utdanning tatt etter en gitt dato (siste oppdatering av registeret). For personer uten registrert utdanning (mest aktuelt for innvandrere) er slike data basert på innsamling via intervju i AKU.

Lagring:

Dataene er lagret som flate filer og som SAS-dataset, både rådata og statistikkfilene.

Dokumentasjon:

Dokumentasjon av filene foreligger som filbeskrivelser (på papir). Prinsipper og definisjoner er ellers dokumentert i NOS Arbeidsmarkedsstatistikk og på SSBs web-sider.

Filbeskrivelse på q:\dok\aku\notater\tabfil3.doc

Skjema på q:\dok\aku\info\skjema

Et større dokumentasjonsnotat om AKU under arbeid (revisjon av Notater 93/27).

Standard:

Næring er kodet etter Standard for næringsgruppering (NOS C 182).

Yrke er kodet etter Standard for yrkesklassifisering (NOS C521).

Utdanning er kodet etter Standard for utdanningsgruppering (SNS nr. 7 1989).

Fylke

Alder

### Arbeidstakere etter arbeidssted og utvalgte næringer

Hyppighet:

Årlig, pr 4. kvartal

Type statistikk:

Fulltelling. Registerbasert.



Datakilder:

Arbeidsgiver-/arbeidstakerregisteret, Enhetsregisteret/Bedrifts- og foretaksregisteret og Det sentrale personregisteret.

Koding/revisjon:

Omfattende revisjon av næringskoder, arbeidsstedsopplysninger og evt. splitting av bedrifter. Hoveddelen av rettingene gjøres på bedriftsnivå, men også enkelte omkodinger av arbeidssted til bosted på individnivå (gjelder arbeidstakere i forsvaret, sjøfolk og avisbud).

Statistikken omfatter foreløpig ikke primærnæringene, med unntak av fiskeoppdrett.

Lagring:

Revisjonen foregår i Oracle. Det finnes ingen filer som er laget på Unix ennå. Grunnfilene som brukes i statistikken ligger på \$ARBTAKEK/

ASCII-filer: (samme)

Dokumentasjon:

Prinsipper og definisjoner i NOS Arbeidsmarkedsstatistikk.

Standard:

Næring er kodet etter Standard for næringsgruppering (NOS C 182)

Kommune/fylke

Annet:

Rikstrykdeverket har det administrative ansvaret for arbeidsgiver-/arbeidstakerregisteret, som er hovedkilden for denne statistikken.

Arbeidstakere etter bosted, utdanning og avtalt arbeidstid

Hypighet:

Årlig, pr 4. kvartal

Type statistikk:

Fulltelling. Registerbasert

Datakilder:

Arbeidsgiver-/arbeidstakerregisteret, Det sentrale personregister og SSBs register over befolkningens høyeste utdanning.

Koding/revisjon:

Ingen spesielle.

Lagring:

SAS-datasett: \$ARBTAKEK/periode/

ASCII-filer:

Filer som lages under produksjonen ligger på \$ARBMARK/bosted/

Dokumentasjon:

SAS-filer dokumenteres i DataDok. Nye filer (laget på Unix) er ikke ferdigdokumentert ennå.

Prinsipper og definisjoner i NOS Arbeidsmarkedsstatistikk.

Standard:

Utdanning er kodet etter Standard for utdanningsgruppering (SNS nr. 7, 1989).

Kommune/fylke

Alder

Annet:

Rikstrygdeverket har det administrative ansvaret for arbeidsgiver-/arbeidstakerregisteret, som er hovedkilden for denne statistikken.

### **Seksjon 310:**

#### Samboerstatistikken:

Type statistikk: utvalgsundersøkelse, 8000 personer brutto

Hypighet: 4 ganger i året, via SSBs omnibusundersøkelse

Datakilder: intervjuundersøkelse, CAPI, påkobling av utdanningsopplysninger fra register

Koding / revisjon: omnibusundersøkelsen er en oppdragsundersøkelse, det lages datafiler til hver enkelt oppdragsgiver. Til samboerstatistikken lages det en samlet datafil ved årets slutt, der de 4 undersøkelsene kobles sammen. Dette gjøres ved seksjon 340. Seksjon 310 går gjennom datasettet og ved eventuelle feil gis det beskjed til S340 som retter opp dette i Omnibusfila.

Lagring: Rådata fra omnibusundersøkelsene lagres på seksjon 340. S310 har ingen egen lagring av disse dataene.

Dokumentasjon:

### **Seksjon 320**

#### Befolkningsstatistikken:

Denne består av:

Befolkningsstatistikk 1., 2., 3. og 4. kvartal og endringer i 4., 1., 2. og 3. kvartal for folkemengde, fødte, døde og flyttinger.

Overgang til norsk statsborgerskap

Folkemengde etter alder, kjønn og sivilstand, 1. januar

Fødte (totalt, pr. år)

Utenlandske statsborgere, 1. januar

Døde (totalt, pr. år)

Flyktninger, 1. januar

Adopsjoner

Barn, 1. januar

Flyttinger (totalt, pr. år)

Inngåtte ekteskap og registrerte partnerskap

Innvandrerbefolkningen, 1. januar

Skilsmisser og separasjoner

Beregnet folkemengde. Nasjonale og regionale tall, 1. januar

Type undersøkelser: Fulltelling

Datakilder: Befolkningsstatistikken bygger i hovedsak på DSF. S320 har en kopi av dette som oppdateres hver måned, dette skal etter hvert gjøres hver uke.

Adopsjoner: Her låner SSB skjema fra Statens ungdoms- og adopsjonskontor. Disse punches inn i SSB, og leveres så tilbake (strengt konfidensialitetskrav). En håper etter hvert å få skjemaene elektronisk.

Flyktninger: Her kommer det en fil med opplysninger fra UDI.

Opplysninger fra GAB (Statens kartverk) kommer på egen fil.

Koding/revisjon: Det kjøres kontroller på meldingene som kommer inn til BEBAS (befolkningsbasen). Når det tas ut tabeller fra BEBAS brukes SAS-programmer for å sjekke dem, bl.a. ved kjøring av frekvenstabeller. Småting rettes opp manuelt. Det blir gjort få endringer på folkemengde.

Lagring: BEBAS består av ulike tabeller som er lagret i en Oracle database. Basen er sammensatt av av to hovedbaser; "Oppdatert folkemengde" og "TransBef "(innkomne meldinger om endringer i befolkningsvariablene). Det finnes også en tabell Transmot som mottar meldinger fra DSF. Denne blir brukt til å oppdatere andre Oracler tabeller, bl.a. "Oppdatert " (Befolkningstabellen). "Transbef" består av meldingene mottatt til Transmot, samt variabler som er koplet på fra Befolkningstabellen. Oppdateringen av tabellene med de aktuelle variablene, er programmert i Oracle. For noen statistikker hentes dataene rett ut fra meldingsbasen, f.eks. når det gjelder statistikk over flytting. Da kobles andre variable på meldingene, f.eks. alder, økonomiske regioner etc. I dag finnes det 120- 150 ulike variable i BEBAS, men for tiden gjøres det flere endringer her, bl.a. legges det til flere kontrollvariable. Filene over folkemengden blir laget ut fra spesialkjøringer av tabellen Befolkning, mens fødte, døde og de andre filene i årsstatikken lages ut fra tabellen Transbef. Filene hentes ut vha. ulike SAS-programmer, rettes opp og brukes i de ulike statistikkene. De langtidslagres som flate filer. Det er disse langtidslagrede filene som blir publisert. Alle statistikker med prefix Befolkningsstatistikk, har langtidslagrede flate filer (de fleste en, noen flere).

Folkemengde etter alder, kjønn og sivilstand.: Denne er lagret både som "fullstendig fil", dvs. uttrekk fra BEBAS, og som aggregerte data som brukes i ulike publiseringer.

Skilsmisser ligger også lagret på to filer.

Den BEBAS-utgaven som blir brukt som grunnlag for folkemengdefila, lagres i sin helhet som flat fil.

Inngåtte ekteskap og registrerte partnerskap lagres i hver sin fil selv om de stort sett publiseres under ett (som ekteskap).

Det kommer ut endelige tall en gang i året, og endringstall hvert kvartal for folkemengde, fødte, døde og flyttinger. Det gis også ut foreløpige tall for de nevnte variablene, men disse langtidslagres ikke.

Dokumentasjon: De ulike tabellene i BEBAS er dokumentert i Notater 2000/24 "Dokumentasjon av BESYS-befolkningssystemet". De gamle filene fra stormaskinen (som stort sett var dokumentert), vil dokumenteres på UNIX etter hvert som de flyttes over på Oracle.

De statistikkene som lagres som flate filer nå, dokumenteres i Datadok.

SAS-filene som brukes til å hente ut data og til kontroll, er dokumentert med flytdiagram fom. 1998.

Noen av variablene i BEBAS er definert i Notater 2000/24, men det litt mer "tilfeldig" hvilke som er dokumentert.

Datadok passer ennå ikke så godt til lagring av variabler brukt i databaser. Databasevariablene har stort sett for lange navn for Datadok.

#### Utdanningsstatistikken:

Består av:

Utdanningsstatistikk. Elever under lov om videregående opplæring

Utdanningsstatistikk. Lærlinger og avlagte fagprøver

Utdanningsstatistikk. Studenter ved universiteter og vitenskapelige høyskoler

Utdanningsstatistikk. Elever og studenter, 1. oktober

Utdanningsstatistikk. Lærere, 1. oktober

BHU

Type statistikk: Fulltelling

Datakilder: Data kommer inn via LINDA systemet (opptaket til videregående skole, LINDA inneholder også tall over avsluttet grunnskole), M-STAS systemet (brukt stort sett av høyskolene og FS-systemet (brukt stort sett av universitetene). Data oversendes på diskett. Informasjon fra BEBAS og BoF (brukes for å kontrollere organisasjonsnummer på skolene) brukes også. I tillegg nyttes data fra Lånekassens adm. register (for å registrere studenter med fullført utdanning i utlandet).

Koding/revisjon: De innkomne dataene kodes om etter NUS-standarden (bortsett fra M-STAS som selv bruker NUS), og kontrolleres, bl.a. mot BoF. Disse kontrollerte filene utgjør rådataene. Til rådataene legges også variable fra BEBAS (bostedkommune, innvandringskategori og landbakgrunn, fødeland og statsborgerskap) og fra BoF (næringsgruppe, skolekommune og eierforhold). Det gjøres også dublettkontroller når dataene videregående skole samles med dem fra universitet- og høyskoler.

Lagring: Rådataene skal fom. 2000 lagres som arkivfiler (flate filer). Det finnes ei fil for avsluttet og ei fil for igangværende utdanning både for videregående skoler og for universitet/høyskole. Disse er lagret som arkivfiler. Universitets/høyskoledata og data for videregående skoler sammenstilles i ei årgangsfil. Denne lagres også som arkivfil. Innholdet i denne arkivfila koples sammen med informasjon fra Lånekassen (studenter med utdanning fullført i utlandet, land, type utdanning (Kodet om til NUS)) og legges over i BHU. BHU er årgangsfiler som inneholder alle personer over 16 år (eller under 16 dersom de har fullført en utdanning) registrert med sin høyeste utdanning. Den inneholder også et tilleggsfelt for alle som er under utdanning. BHU lagres også som arkivfil. BHU brukes direkte i statistikkproduksjonen.

Dokumentasjon: Alle arkivfilene i utdanningsstatistikken er dokumentert i Datadok.

I forbindelse med den nye forløpsdatabasen for utdanning som nå utarbeides, vil det bli publisert detaljert dokumentasjon av de kontroller som gjøres av dataene. Dessuten skal hele prosessen fra datainnsamling til ferdig statistikk dokumenteres, delvis i NOS, delvis som egne rapporter. I et pågående PROMUT-prosjekt arbeides det med å dokumentere alle variable som publiseres innenfor utdanningsstatistikken.

### **Seksjon 330:**

#### Dødsårsaksstatistikken:

Type undersøkelse: Fulltelling

Datakilder: Kommunelege og sykehus sender et skjema pr. dødsfall inneholdende bl.a. navn og personnr, dødsårsak og -sted, samt ev. underliggende dødsårsak

Koding/revisjon: Skjemaene scannes på Kongsvinger, revisjon og koding gjøres på S330

Lagring: Skjemaene er grunnlaget for Dødsårsaksregisteret som ligger i en Oracle database.

Skjemaene ligger lagret som flate filer ( \*.tif filer ) på NT-serveren.

Dokumentasjon: Revisjonen er godt dokumentert, men finnes sannsynligvis i mer interne notater "i skuffen".

Dødsårsaksstatistikken utgis i NOS, og har derfor dokumentasjon som tilsvarer denne.

Standard: ICD-10

Annet: Helsetilsynet eier dødsårsaksregisteret.

#### Helseundersøkelsen:

Type undersøkelse: Utvalgsundersøkelse, del av den rullerende Levekårsundersøkelsen

Datakilder: Papirskjema utfylt av intervjuer (personlig intervju)

Koding/revisjon: Scannes på Kongsvinger, S340 sjekker dataene

Lagring: Postskjema og utvalg med ICD-10 koder ligger lagret på SAS-filer på Levekårsområdet.

Dokumentasjon: Det finnes en dokumentasjonsrapport knyttet til selve datainnsamlingen (skjema, svarprosent, frafall osv.) Ellers dokumenteres helseundersøkelsen i NOS.

Standard: ICD-10

Annet:

### Pleie- og omsorgsstatistikken

Type undersøkelse: Fulltelling

Datakilder: For kommuner som er med i KOSTRA, kommer skjemaene elektronisk. Fra de andre kommunene (også noen bydeler) kommer det papirskjema. Her leverer hver institusjon summariske data knyttet til egen virksomhet. Det er fire typer skjemaer:

- 1) Hjemmetjenester (typer tjenester, hvor mange mottar, hvor mange ansatte, egenandel osv.)
- 2) Institusjoner, 3) Personell og 4) Boliger (boliger som ikke er def. som institusjoner, f.eks. omsorgsboliger, antall boliger og beboere tas med). Noen kommuner rapporterer også via GERIX (datasystem) som er på individnivå (inneholder bl.a. mottakere av hjemmetjenester og institusjonsbeboere).

En bruker også institusjonsregisteret. I tillegg oppdateres enkelte av variablene ved koblinger mot BOF. RD brukes også for kopling mot folketall i tabellprogrammene.

Koding/revisjon: Ikke-KOSTRA skjema scannes på Kongsvinger. Revisjon gjøres av S330. S330 oppdaterer også institusjonsregisteret. Institusjonsregisteret oppdateres ved at det sendes ut utskrifter fra registeret til den enkelte institusjon. Ved endringer i oppgavene føres dette på og utskriften returneres til SSB. I tillegg sendes det lister over registrerte institusjoner i kommunen/bydelen til den enkelte kommune/bydel (for å fange opp nye institusjoner o.l.). Ved ev. endringer skal listene ajourføres og returneres til SSB.

Lagring: Papirskjemaene lagres på S330, disse inneholder også rettinger. KOSTRA har sin egen lagring av elektroniske skjema i en Oracle database (tallene i skjemaet er lagret for seg, og metadatene er lagret over slik at skjemaet kan gjenskapes). GERIX-dataene lagres som flate filer og SAS-filer på OVIBOS. Foreløpige tall lagres som flate filer (gis ut en gang pr. år). Reviderte/endelige tall lagres i en (annen enn KOSTRAS) Oracle database og som flate filer på UNIX.

Dokumentasjon: Revisjonen for 97 er dokumentert, det tas vare på feilmeldinger. De kommende revisjoner vil sannsynligvis også dokumenteres slik. Variablene i spørreskjemaene er definert på skjemaet. De flate filene er ikke lagret i Datadok, og mangler tilfredsstillende dokumentasjon. Statistikken er ellers dokumentert i NOS (årlig).

Annet:

### Statistikk over utførte svangerskapsavbrudd:

Type undersøkelse: Fulltelling

Datakilder: Papirskjema fra fylkeslegen (sendes inn hvert kvartal). Disse skjemaene blir til sammen et register over svangerskapsavbrudd. SSB får statistikkdelen, den er sannsynligvis ikke mulig å identifisere på person.

Koding/revisjon: Scanning på Kongsvinger, enkel revisjon i S330.

Lagring: Svangerskapsavbruddsregisteret er lagret i en Oracle database. Rådata lagres her, og revideres etter hvert. Statistikken produseres fra denne databasen. Registeret er KANSKJE også lagret som årgangsfiler/ flate filer.

Dokumentasjon: Eventuelle flate filer er sannsynligvis ikke registrert i Datadok. Databaseregisteret er dokumentert i databasen.

Annet: Svangerskapsavbruddsregisteret eies av Helsetilsynet

### Statistikk over steriliseringer

Type undersøkelse: Fulltelling

Datakilder: Som for svangerskapsavbrudd

Koding/revisjon: Som for svangerskapsavbrudd

Lagring: De gamle filene var lagret som flate filer på stormaskin. Nye filer skal lagres i Oracle database.

Dokumentasjon: De flate filene er sannsynligvis ikke dokumentert.

### Pasientstatistikken

Type undersøkelse: Fulltelling

Datakilder: Norsk pasientregister (NPR) mottar disketter fra alle somatiske sykehus, og sender deretter fila over til SSB. Noen få sykehus sender disketter direkte til SSB. Hver pasient har et løpenummer pr. sykehus pr. år, og kan ikke identifiseres.

Koding/revisjon: De fleste kontroller gjøres av SINTEF som har NPR, de skal også gi en kvalitetsvurdering av dataene som leveres (det har vært et problem med kvaliteten på opplysningene fra sykehusene). S330 gjør noen kontroller.

Lagring: Dataene fra NPR (rådataene) kommer på CD-rom, og denne tas vare på som arkivmateriale. De endelige tallene for pasientstatistikken lagres som årgangsfiler (flate filer). De foreløpige tallene langtidslagres ikke.

Dokumentasjon: Statistikken dokumenteres i NOS (det har ikke vært utgitt NOS siden 1991, men det kommer en nå). De flate filene er ikke dokumentert i Datadok.

Standard: ICD-10

Annet: Sosialdep. har opprettet pasientregisteret, SINTEF administrerer det.

### Spesialisttjenesten, psykiatri

Type undersøkelse: Fulltelling

Datakilder: SSB mottar skjema (et pr. år) fra hvert sykehus/klinikk. Disse inneholder summariske tall for pasienter og for antall ansatte (5 KOSTRAFylker, resten rapporterer på papir).

Regnskapstall sendes også inn. Disse legges på egen fil. SSB får også tall fra NPR. Dette er individbaserte tall.

Koding/revisjon: Papirskjemaene som kommer til SSB, revideres i stor grad før de sendes til Kongsvinger til punching. KOSTRAskjemaene lagres i Oracle, og det kjøres kontroller i SAS.

Lagring: Rådataene lagres ikke på annen måte enn at de kanskje kan gjenfinnes som flate filer oversendt på mail etter punching på Kongsvinger. Papirskjemaene lagres, men uvisst hvor lenge. De endelige tallene (fom. 1990) lagres på SAS-filer, filene fra før 1990 ligger som komprimerte filer på UNIX, og disse er svært dårlig dokumentert.

Dokumentasjon: Statistikken dokumenteres i NOS. I KOSTRADelen loggføres feil og endringer (resultat av kontrollene kjørt i SAS)

Standard: ICD-10

Annet: Som for pasientstatistikken

Generelt om dokumentasjon av Oracle databaser: Strukturen (datatype, formater etc.) er godt dokumentert. Kodelister er mer sparsomt dokumentert, og semantikken (variabeldefinisjoner, ev. hvor data kommer fra etc.) er i liten grad dokumentert.

Generelt om dokumentasjon av tabell(SAS)programmer: Disse er generelt dårlig dokumentert, dette gjør det f.eks. vanskelig for nye medarbeidere å finne ut hvilke tabellprogrammer som er brukt i ulike sammenhenger.

## Seksjon 350:

### Levekårsundersøkelsen, tverrsnitt (0202)

Hyppighet: Årlig

Type statistikk: utvalgsundersøkelse

Datakilder: Intervjuopplysninger fra et utvalg av personer 16 år og over som ikke er bosatt på institusjon. Det kobles til en del opplysninger fra inntekts- og formuesundersøkelsen for husholdninger. Datainnsamlingen skjer elektronisk (CAPI).

Koding/revisjon: Næring og yrke kodes de fleste år. Koding skjer ved S450. I år med helseundersøkelse kodes sykdomsdiagnose. Kodingen skjer ved S330.

Lagring: Langtidslagring som SAS-filer.

Dokumentasjon: Innsamling og innsamlingsmaterieell er dokumentert i serien Notater. Koding er ikke dokumentert, men skjer i henhold til standarder for næring, yrke og sykdommer. Dokumentasjonsnotatene er tilgjengelig på web. Noen hovedpunkter i dokumentasjonen vil også inngå i en 3-årlig NOS. Filene er dokumentert som SAS-filer. Programmer for bearbeiding av opprinnelige filer blir oppbevart i UNIX.

Standard: Undersøkelsene følger standardene. De mest vanlig er for næring, yrke, utdanning, sykdommer, husholdning, sosioøkonomisk status, landsdeler.

### Levekårsundersøkelsen, Panel (0204)

Hyppighet: Årlig

Type statistikk: utvalgsundersøkelse

Datakilder: Intervjuopplysninger fra et utvalg av personer 16-79 år som ikke er bosatt på institusjon. Det kobles til en del opplysninger fra inntekts- og formuesundersøkelsen for husholdninger og det er planer om å koble til opplysninger fra FD-trygd. Datainnsamlingen skjer elektronisk (CAPI).

Koding/revisjon: Næring og yrke kodes. Koding skjer ved S450. I enkelte år kodes sykdomsdiagnose. Kodingen skjer ved S330.

Lagring: Langtidslagring som SAS-filer.

Dokumentasjon: Innsamling og innsamlingsmaterieell er dokumentert i serien Notater. Koding er ikke dokumentert, men skjer i henhold til standarder for næring, yrke og sykdommer. Dokumentasjonsnotatene er tilgjengelig på web. Filene er dokumentert som SAS-filer. Programmer for bearbeiding av opprinnelige filer blir oppbevart i UNIX.

Standard: Undersøkelsene følger standardene. De mest vanlig er for næring, yrke, utdanning, sykdommer, husholdning, sosioøkonomisk status, landsdeler.

#### 590(4) Kultur- og mediebruksundersøkelsene

Hyppighet: Mediebruk hvert år, et tillegg med kulturbruksspørsmål hvert tredje år, etter 2000 hvert fjerde år.

Type statistikk: Individutvalg

Datakilder: Elektronisk spørreskjema (Blaise)

Koding/revisjon: Dataene blir kontrollert på 340.

Lagring: SAS-filer

Dokumentasjon: Dataene dokumenteres i prosjektmapper på 340. Fra 1999-undersøkelsen skal det lages egne dokumentasjonsrapporter på 340. På 350 blir også lagret dokumentasjon i prosjektmapper.

Standard: For en del av bakgrunnsspørsmålene benyttes det standard spørreformuleringer.

Annet: Datainnsamlingen til mediebruksundersøkelsen er eksternt finansiert av Norsk medieforvaltning/Kulturdepartementet og Telenor. Kulturdepartementet og Norsk kulturråd finansierer kulturbruksspørsmålene.

#### 0208 Tidsbruksundersøkelsen

Hyppighet: Hvert 10. år, hittil siste gang 2000

Type statistikk: Individutvalg

Datakilder: Elektronisk spørreskjema (Blaise) og føringshefter på papir for tid brukt til ulike aktiviteter.

Koding/revisjon: Dataene blir kontrollert på 340.

Lagring: SAS-filer

Dokumentasjon: Dataene dokumenteres i prosjektmapper på 340. Det skal lages egne dokumentasjonsrapporter på 340. På 350 blir også lagret dokumentasjon i prosjektmapper.

Standard: For en del av bakgrunnsspørsmålene benyttes det standard spørreformuleringer.

Annet: Datainnsamlingen til tidsbruksundersøkelsen er delvis eksternt finansiert av 5 departementer, LO og Telenor.

#### Barnevernstatistikk - individ

Hyppighet: Årlig

Type statistikk: Fulltelling



Datakilder: Statistikken omfatter i dag samtlige barn med tiltak og/eller undersøkelsessak i barnevernet. Barnet er enhet i statistikken og registreres med fullt fødselsnummer (11-siffer). Statistikken inneholder også opplysninger om; *hvem som meldte saken som førte til undersøkelse, meldingens saksinnhold, og resultat av undersøkelsen*. For barn som har mottatt tiltak i statistikkåret registreres det *type tiltak* (hjelpetiltak/omsorgstiltak), *hvilke tiltak* (23 ulike tiltak), og tiltakets varighet målt ved *antall måneder i statistikkåret*. Det registreres også om foreldrene ble fratatt foreldreansvaret i statistikkåret (jfr. BVL § 4-20) og *grunnlag for iverksetting av tiltak* (13 ulike svaralternativ), *barnets statsborgerskap, mors og fars fødeland*. På siste side i skjemaet, som gir oss et situasjonsbilde på telledatoen 31.12., registreres det *hvor barnet bor per 31.12. og hvilke tiltak som er i verk ved utgangen av året* (23 forskjellige tiltak) og om tiltaket er et hjelpetiltak eller omsorgstiltak.

Det gjøres uttrekk fra kommunenes IT-system og sendes inn på diskett. Alternativet er utfylling av ett papirskjema per barn og disse leses optisk. For KOSTRA-kommuner uten fagsystem brukes elektroniske skjema som sendes inn på diskett. All datainnsamling sendes rekommandert i posten.

Koding/revisjon: Ved uttrekk fra fagsystem er dataene ferdig kodet, og skal være kontrollerte ved bruk av kontrollprogrammet for barnevern og rettet opp før innsending. Alle data blir kontrollert i SSB ved hjelp av kontrollprogrammet og kommunene mottar feillister for oppretting.

Lagring: Dataene lagres som flate filer på UNIX.

Dokumentasjon: Både datafiler og SAS-program vil på sikt bli laget i Datadok. Pt. ligger alle SAS-programfiler på UNIX.

Standard: Landkodeliste fra BEBAS

#### Anmeldte lovbrudd (og statistikk over ofre)

Hyppighet: Halvårsstatistikk (Anmeldte lovbrudd)

Type statistikk: Fulltelling.

Datakilder: Politiets sentrale register for straffesaker - STRASAK. Politiets Datatjeneste (PD) leverer utkjøringer av registerdata for siste periode fra STRASAK. Anmeldte lovbrudd registreres elektronisk av politibetjenter ved de enkelte politikammer ved henvendelser fra publikum. Statistikken regnes brukbar fra og med 1991-årgangen, men er av bedre kvalitet fra og med 1995-årgangen.

Koding/revisjon: PD leverer data på bånd - som SSB (S303) lagrer på flate tekstfiler, Unix. Det foretas enkle bearbeidinger m.h.t. geografisk spredning, gjerningstidspunkt og koding av lovbrudd (S350).

Lagring: Rådata lagres permanent som flate tekstfiler (under \$KRIMINAL/anmeldt/inndata). Ferdig bearbeidet datamateriale lagres som SAS filer (under \$KRIMINAL/anmeldt/data).

Dokumentasjon: Prosessen, fra innsamling til ferdige tabeller er dokumentert i Gundersen og Hustad (2000) "Statistikk over anmeldte lovbrudd og registrerte ofre" Notater 2000/13 og på Unix under \$KRIMINAL/anmeldt/notat. Årlig bearbeiding dokumenteres i logger. Statistikken er ellers dokumentert i NOS (årlig).

Standard: Typer lovbrudd, kommuner (1994), politidistrikt, landbakgrunn på ofre

Annet: Eier av register: Justisdepartementet

**NB !**

Rådata fra årgangen 1998 er blitt bearbeidet til **statistikk over ofre** (jf. Gundersen i Samfunnsspeilet 2000/3). Statistikken over ofre krever en mer omfattende bearbeiding bl.a. fordi ofre ikke har plikt til å oppgi personnummer. Bearbeidet statistikk over ofre, blant annet etter tillegging av personnummer ved hjelp av navn og BEBAS, er nærmest en fulltelling av personer som har anmeldt en sak etter å ha blitt offer for lovbrudd. Datamaterialet påføres landbakgrunn. Rådata lagres permanent som flate tekstfiler under \$KRIMINAL/offer/inndata og bearbeidet datasett lagres permanent under \$KRIMINAL/offer/data. Statistikken over ofre vurderes å bli publisert årlig (NOS). Bearbeidingen av statistikken over ofre er dokumentert i Gundersen og Hustad (2000) "Statistikk over anmeldte lovbrudd og registrerte ofre"

Etterforskede lovbrudd

Hyppighet: Årlig statistikk

Type statistikk: Fulltelling.

Datakilder: Politiets sentrale register for straffesaker - STRASAK. Politiets Datatjeneste (PD) leverer utkjøringer av data for siste periode fra STRASAK.

Ferdig etterforskede lovbrudd registreres elektronisk av påtalemyndigheten ved de enkelte politikammer. Ferdig etterforskede lovbrudd innebærer at det har blitt gitt en rettskraftig reaksjon - dvs. at statistikken også omfatter dommer fra domstolssystemet.

Koding/revisjon: PD leverer data på bånd - som SSB (S303) lagrer permanent på flate tekstfiler, Unix.

Det foretas ulike kontroller av ikke aksepterte og ulogiske verdier (S350).

Feillister sendes til politikamrene, rettinger inkluderes i den videre bearbeidingen (S320).

Forskjellige bearbeidinger m.h.t. geografisk spredning, gjerningstidspunkt, koding av lovbrudd og koding av rettskraftige avgjørelser (S350).

Gjerningspersoner påføres landbakgrunn.

Lagring: Rådata lagres permanent som flate tekstfiler. Ferdig bearbeidet datamateriale lagres som SAS -filer. Bearbeidet materiale på Unix fra og med årgang 1987. Før 1992 inneholder datasettet kun forbrytelser.

Dokumentasjon: Prosessen, fra innsamling til ferdige tabeller er dokumentert i Gundersen (1999):

" Statistikk over etterforskede lovbrudd" Notater 1999/52 og på unix under

\$KRIMINAL/eforsk/notat. Årlig bearbeiding dokumenteres i logger. Statistikken er ellers dokumentert i NOS (årlig).

Standard: Bruker blant annet standarder for: Typer lovbrudd, kommuner (1994), politidistrikt, landbakgrunn på siktede

Annet: Eier av register (STRASAK): Justisdepartementet. Har følsomme data om straffeforhold.

Forbruksundersøkelsen

Hyppighet: Database laget årlig siden 1996. Innsamling kontinuerlig siden 1974.

Type statistikk: Utvalgsundersøkelse

Datakilder: Private husholdninger ved intervju og føring av utgiftshefter.  
Koblede opplysninger fra register (inntekt og utdanning).

Koding/revisjon: Automatisert revisjon i Blaise-intervju ved CATI intervjuing. Automatisert krysskontroll ved registrering av hefter på seksjon 450. Manuell revisjon ved seksjon 350 under resultatkontroll.

Koding av yrke, næring og varekoder ved seksjon 450.

Lagring: Data eldre enn 1996 er overført fra stormaskin på flate filer på Unix.

Føringshefter lagres ved seksjon 350. Fra 1996 vil Blaise- intervju (rådata) lagres foreløpig ved seksjon 340, disse bør overføres til seksjon 350. Registrerte, kontrollerte og bearbejdede resultater ligger i Oracle database på OPR3. Eldre årganger på basen bør legges ut på flate filer for lagring, dette arbeidet er ennå ikke påbegynt.

Dokumentasjon: Statistikken dokumenteres i NOS (tre år sammenslått).

Prosessen fra datainnsamling til basetabeller for perioden 1996 - 1999 er dokumentert på Q:\dok\fbureg\

Fra 2000 er hele behandlingen under omlegging. En helhetlig dokumentasjon med kravspesifikasjon, kontroller og dokumentasjon er under arbeid og vil foreligge i løpet av året.

#### Statistikk over straffereaksjoner

Hyppighet: Årlig

Type statistikk: Fulltelling

Datakilder: Det sentrale straffe- og politiopplysningsregister (SSP), Det sentrale bøteregisteret (Bot) og det sentrale innkrevingssystemet Tor. Elektronisk datainnsamling.

Koding/revisjon: Praktisk talt alle opplysninger i det innkomne materialet er kodet. I materialet fra Bot er gjerningssted oppgitt i klartekst og blir omkodet maskinelt i SSB. Det utføres maskinelle kontroller på det innkomne materialet og feillister sendes de ansvarlige for registrene til oppretting. Rettede lister reurneres SSB som foretar oppretting av materialet. Det foretas en del maskinelle beregninger og omkodinger i materialet og det tilføres data fra Det sentrale folkeregister (DSF) før materialet er klart til statistikkproduksjon.

Lagring: Rådata og endelig materiale (revidert og bearbejdet og klart til statistikkproduksjon) lagres som UNIX-filer.

Dokumentasjon: Er under utarbejding og følger i hovedsak retningslinjene gitt i Berby (1999) og Statistisk sentralbyrå (1995, 1997 og 1998c). Vil bli utgitt i serien Notater.

Standarder: Inndeling av kriminalitet i 10 hovedgrupper, Standard for aldersgruppering (med enkelte avvik), Fylker, Statsborgerskap.

Annet: Justisdepartementet er eier av de administrative registrene dataene hentes fra.

#### Fengslingsstatistikk

Hyppighet: Årlig

Type statistikk: Fulltelling

Datakilder: Det sentrale straffe- og politiopplysningsregister (SSP) er hoveddatakilde og datainnhenting skjer elektronisk. I tillegg leverer Kriminalomsorgens utdanningscenter summariske oversikter på papir over enkelte data som ikke ligger i SSP.

Koding/revisjon: Materialet fra SSP er kodet ved innhenting, og statistikken produseres på råmaterialet etter at det er foretatt en enkel maskinell revisjon og bearbeiding.

Lagring: Rådata og endelig materiale (revidert og bearbeidet og klart til statistikkproduksjon) lagres på UNIX-filer.

Dokumentasjon: Et bearbeidingsopplegg utarbeidet i 1991 viste seg i praksis å være for ambisiøs og ressurskrevende. Det ble derfor i perioden 1993 - 1998 utarbeidet statistikk på rådata og etter en meget enkel stikkordpreget beskrivelse. Arbeidet med nytt opplegg og dokumentasjon starter i 2. kvartal 2000.

Standarder: Standard for aldersgruppering (med enkelte avvik).

Annet: Justisdepartementet er eier av SSP som er hoveddatakilde.

### Sosialhjelpsstatistikk - individ

Hyppighet: Årlig

Type statistikk: Fulltelling

Datakilder: Statistikken omfatter alle mottakere av økonomisk sosialhjelp i kommunene. Den enkelte sosialhjelpsmottaker er enhet i statistikken og registreres med fullt fødselsnummer (11 siffer). Skjema som ligger til grunn for innsending inneholder også opplysninger om;

- sivilstand
- barn under 18 år som forsørges av og bor sammen med klient
- viktigste kilde til livsopphold
- forhold til trykkesystemet
- arbeidssituasjon
- hvilke måneder klienten har mottatt stønad
- samlet stønadssum mottatt i løpet av året (fordelt på bidrag og lån).

Det gjøres uttrekk fra sosialkontorenes IT-system som sendes inn på diskett. Kontorer som ikke har elektronisk fagsystem sender inn skjema (ett skjema per mottaker), som leses optisk. For KOSTRA-kommuner som ikke har fagsystem benyttes elektroniske skjema som sendes inn på diskett. Alle forsendelser sendes rekommandert i posten.

Koding/revisjon: Ved uttrekk fra fagsystem er dataene ferdig kodet, og skal være kontrollerte ved bruk av kontrollprogrammet for sosialhjelp og rettet opp før innsending. Alle data blir kontrollert i SSB ved hjelp av kontrollprogram. Kommunene mottar feillister for oppretting.

Lagring: Dataene lagres som flate filer på UNIX.

Dokumentasjon: Både datafiler og SAS-program vil på sikt bli laget i Datadok. Pt. ligger alle SAS-programfiler på UNIX.

Standard: Landkodeliste fra BEBAS.

### Seksjon 420:

#### Inntekts- og formuesundersøkelsen for husholdninger.

Type statistikk: Årlig utvalgundersøkelse. Utvalgsstørrelse og utvalgssammensetning er dokumentert i NOS Inntekts- og formuesstatistikk for husholdninger 1986 – 1996, kap. 2.3.

Husholdningene som inngår i Inntekts- og formuesundersøkelsene er i perioden 1986 til 1996 utvalget har vært et panel av personer 16 – 79 år og deres husholdningsmedlemmer (IF), som har vært med årlig siden 1986. Hvert år blir det til IF trukket et tilleggsutvalg med formål å rette opp skjevheter som oppstår i panelet over tid. I tillegg til IF har Forbuksundersøkelsens (FU) og Levkårsundersøkelsens (LKU) utvalg har vært med flest ganger.

Boforholdsundersøkelsens (BFU) og Helseundersøkelsens (HU) utvalg har også vært med. Fra 1996 innførte Statistisk sentralbyrå ett nytt system for levkårsrelaterte undersøkelser. Fra og med dette året gjennomføres to levkårsundersøkelser i året, - en panelundersøkelse (LKP) i første halvår og en tverrsnittsundersøkelse (LKT) med roterende tema på høsten. I 1991, 1992 og hvert år siden 1995 har også et tilleggsutvalg av personlige næringsdrivende (IFN) vært representert.

Tabell 1 viser utviklingen i antall husholdninger som er med i undersøkelsen fra 1986 til 1998 og hvilke utvalg som inngår de ulike årene.

År	Utvalg	Antall husholdninger
1986	LKU,FU	4 975
1987	IF, FU	3 393
1988	IF, FU	3 423
1989	IF, FU	3 475
1990	LKU, IF, FU	6 046
1991	IF, IFN, FU	8 072
1992	IF, IFN, FU	8 104
1993	IF, FU	3 522
1994	IF, LKU, BFU	12 799
1995	IF, IFN, HU	10 127
1996	IF, IFN, LKP, LKT	14 110
1997	IF, IFN, LKP, LKT	14 679
1998	IF, IFN, LKP, LKT	14 521

Datakilder: Inntekts- og formuesundersøkelsene henter også data fra en rekke andre kilder.

Intervjuopplysninger er grunnlaget for husholdningssammensetningene. Administrative og statistiske registre har vi også brukt i stor utstrekning. Følgende registre har vi hentet data fra:

- Det sentrale folkeregister (ekteskapelig status m.m.)
- Ligningsregisteret (utlignede skatter, nettoinntekt m.m.)
- Lønns- og trekkoppgaverregisteret, LTO (arbeidsledighetstrygd, attføringspenger m.m.)
- Fra Rikstrygdeverket, GR1 (grunn- og hjelpestønad, pensjonsstatus m.m.)

- Fra Sosialdepartementet (sosialhjelp)
- Utdanningsregisteret (høyeste fullførte utdanning)
- Fra Statens lånekasse for utdanning (lån, stipend, gjeld)
- Fra Husbanken (bostøtte)

Koding/ revisjon: Det blir ikke foretatt koding/ revisjon i tradisjonell forstand.

Til og med 1993 ble opplysningene lastet inn i en database i ORACLE og knyttet til ei skjermrutine for kontroll og oppretting (se avsnitt 2.4. om datainnsamling). Materialet ble kontrollert både for registreringsfeil, summariske feil og logiske feil. Det ble også foretatt en konsistenskontroll i forhold til netto inntekt og netto formue i Statistisk sentralbyrås skattestatistikk for forskuddspliktige. Det blir også foretatt kontroller av samme karakter på registeret (fra 1993), men på grunn av stort omfang både i antall observasjoner og beløpsfelter er det ikke mulig å gjennomføre absolutt alle konsistenskontroller internt i materialet. Inntekts- og formuesundersøkelsen benytter dataene fra registeret etter at det er kontrollert som beskrevet over. Personlig næringsdrivende i utvalget gjennomgår ytterligere kontroll. For disse blir det foretatt konsistenskontroller mellom selvangivelsens opplysninger og næringsoppgavene. Husholdnings sammensetningen for de som ikke har svart på husholdningsintervjuet blir kontrollert opp mot tilgjengelige opplysninger på register. I tillegg blir det avledet et antall avledede kjennemerker som f.eks.

- sosioøkonomisk status
- alder
- kommunestørrelse
- inntektstaker/hovedinntektstaker
- hvor mange yrkestilknyttede i husholdningen
- hvor mange personer i husholdningen

mm.  
Dokumentasjon på hvordan de avledede kjennemerkene framkommer finnes i Notater 98/11 Inntekts- og formuesundersøkelsen 1995, som er siste utkomne notat på området. En del kjennemerker er også omtalt i NOS Enkelte kjennemerker/variabler er også beskrevet i NOS Inntekts- og formuesstatistikk for husholdninger 1986 – 1996, kap. 3.

Lagring: Dataene lagres som flate filer på Unix i henhold til retningslinjene gitt i Interne dokumenter 99/1 Datalagring på Unix i Statistisk sentralbyrå. Alle filer er dokumentert i Datadok i henhold til regelverket gitt i Interne dokumenter 99/6 Brukerbeskrivelse for Dokumentasjonssystem i Statistisk sentralbyrå.

For hver årgang lagres :

- en personfil med alle opplysninger
- en husholdningsfil (alle beløp aggregert for alle personer i husholdningen)
- en utvalgsfil (bare personkjennemerker)
- en spesialfil tilpasset behovet for skattemodellen LOTTE.

Dokumentasjon: De enkelte undersøkelsene skal etter planen dokumenteres i serien Notater. Pr. i dag er imidlertid undersøkelsen for inntektsåret 1995 den siste som er publisert i denne serien, 98/11 Inntekts- og formuesundersøkelsen 1995, Dokumentasjon (Vidar Pedersen). De senere årene avviker imidlertid i liten grad fra denne. Undersøkelsen er dessuten dokumentert i NOS Inntekts- og formuesstatistikk for husholdninger 1986 – 1996. Videre er den nye regnskapsoppstillingen i Inntektsregnskapet dokumentert i Notater 98/43 Inntekt etter skatt (Jon Epland). Hvordan utvalgene vektet blir også dokumentert i serien Notater. Den siste utkomne er 1999/78 Beregning av vekter i Inntekts og formuesundersøkelsene 1997 (Per Morten Holt, Sigrun Kristoffersen og Vidar Pedersen).

Standarder: Utdanning, landsdeler, husholdninger, alder, inntektsbegreper mv.

## Oppsummering av kartlegging

Seksjon	Navn på statistikken	Prod nr	Emne	Hypighet	Fulltelling /utvalgsum dersøkelse	Lagring rådata	Lagring statistikkdata	Data dok?	Dokumentasjon	Standard:
310	Samboerstatistikken (utvalget?)	6390	00.02.30	4 ganger i året	Utvalg	S340, Omnibus	Nei	n	?	
330	Dødsårsaksstatistikk en (registeret?)	0520	03.01.10		Fulltelling	Flate filer. Skjemaene lagres som *.tif på NT.		j	Rev. dok. i notater "i skuffen". Stat. dok i NOS.	ICD-10
330	Helseundersøkelsen	0581	03.01		Utvalg	SAS-filer, Levekårsomr.		n?	Dok. rapp. knyttet til datainns. Stat. dok. i NOS.	ICD-10
330	Pleie- og omsorgsstatistikken	0660	03.03		Fulltelling	Papirskj. lagres S330.med rettinger. KOSTRA i Oracle (tall og metadata, skj. kan gjenskapes) GERIX-lagres som flate filer og SAS-filer på OVIBOS.	Foreløpige tall lagres som flate filer (gis ut en gang pr. år). Reviderte/endelige tall i en Oracle og som flate filer på UNIX.	n	Revisjonen 97 er dok, feilmeldinger lagres. Kommende rev. vil sannsynligvis dok. slik. Var. i skj. er def. på skj. Flate filer ikke i Datadok - mangler tilfredsstillende dok. Statistikken er dok. i NOS.	
330	Statistikk over utførte svangerskapsavbrudd	0560	03.01.20		Fulltelling	Oracle database	KANSKJJE også lagret som årgangsfiler/ flate filer.	n?	Flate filer er sannsynligvis ikke reg. i Datadok. Databaseregisteret er dok. i basen.	
330	Statistikk over steriliseringer	0585	03.01		Fulltelling	Gamle filer flate filer på stormaskin. Nye filer lagres i Oracle.		n?	Flate filer sannsynligvis ikke dok.	

Seksjon	Navn på statistikken	Prod nr	Emne	Hypighet	Fulltelling /utvalgsum dersøkelse	Lagring rådata	Lagring statistikkdata	Data dok?	Dokumentasjon	Standard:
330	Pasientstatistikken	0620	03.02.10		Fulltelling	CD-rom fra NPR arkiveres.	Endelige tall i flate årgangsfiler. Foreløpige tall langtidslagres ikke.	n?	Flate filer ikke dok. i Datadok. Stat. dok. i NOS, siste 1991, neste 2000.	ICD-10
330	Spesialisttjenesten, psykiatri	0657	03.02.20		Fulltelling	kan kanskje gjenfinnes som flate filer Papirskjemaene lagres uvisst hvor lenge KOSTRA lagres i Oracle	Endelige tallene (fom. 1990) lagres på SAS-filer Før 1990 komprimerte filer på UNIX, dårlig dokumentert	j?	KOSTRA loggfører feil og endringer fra SAS-kontroller Stat. dok. i NOS	ICD-10
320	Befolkningsstatistikk en	0420	02		Fulltelling	Oracle	flate filer på UNIX	j	Notater 00/24	
320	Utdanningsstatistikk	0330	04		Fulltelling	Flate filer fra 2000		j	NOS og egne rapp	
350	Levekårsundersøkels en, tverrsnitt	0202		Årlig	utvalg		SAS filer	n	Notater, NOS, dokumenterte SAS-filer	Næring, yrke, utdanning, sykdommer, husholdning
350	Levekårsundersøkels en, panel	0204		Årlig	utvalg		SAS filer	n	Notater, NOS, dokumenterte SAS-filer	Næring, yrke, utdanning, sykdommer, husholdning, sosioøkonomisk status, landsdeler



Seksjon	Navn på statistikken	Prod nr	Emne	Hypighet	Fulltelling /utvalgundersøkelse	Lagring rådata	Lagring statistikkdata	Data dok?	Dokumentasjon	Standard:
350	Kultur- og mediebruksundersøkel sen	5904	07.02	medie-årl., kultur - tre/fireårig	Utvalg	SAS-filer	SAS-filer	n	Prosjektmapper påSS340	Standard spm.
350	Tidsbruksundersøkelsen	0208	00.02.30	10 årlig	Utvalg	SAS-filer	SAS-filer	n	Prosjektmapper påSS340	Standard spm.
350	Barnevernsstatistikk - individ	0671	03	Årlig	Fulltelling	flate filer på UNIX	flate filer på UNIX	n	SAS-prog på UNIX Vil bli dok. i Datadok	Landkode fra Bebas
350	Anmeldte lovbrudd - og statistikk over ofte	0821	03.05	Halvårlig	Fulltelling	flate filer	SAS-filer	n	Notater, logg, NOS	Typer lovbrudd, Komm, politidistrikt , landbakgr.
350	Etterforskede lovbrudd	0822	03.05	Årlig	Fulltelling	flate filer	SAS-filer	tja	Notater, logg, NOS	Typer lovbrudd, Komm, politidistrikt , landbakgr.
350	Forbruksundersøkelsen			Årlig	Utvalg	flate filer på UNIX	flate filer på UNIX (før 1996), Oracle	n	NOS, Q.\. Helhetlig dok i løpet av 2000.	
350	Reaksjoner for lovbrudd (Statistikk over straffereaksjoner)	0823	03.05	Årlig	Fulltelling	UNIX-filer	UNIX -filer	n	Under utarbeiding	Inndeling av krim. Aldersgrupp ering (avvik), Fylke, Statsborgers kap.
350	Fengslingsstatistikk	0825	03.05	Årlig	Fulltelling	UNIX-filer	UNIX -filer	n	Under utarbeiding	Aldersgrupp ering (avvik)
350	Sosialhjelpsstatistikk - individ	0731	03.04	Årlig	Fulltelling	flate filer på UNIX	flate filer på UNIX	j	SAS-prog på UNIX Vil bli dok. i Datadok	Landkode fra Bebas

Seksjon	Navn på statistikken	Prod nr	Emne	Hypighet	Fulltelling /utvalgsum dersøkelse	Lagring rådata	Lagring statistikkdata	Data dok?	Dokumentasjon	Standard:
420	Inntekts- og formuestatistikk for husholdninger	1610	05.01	Årlig	Utvalg	flate filer på UNIX	flate filer på UNIX	j	Notater, NOS	utdanning, landsdel, husholdn., alder, inntektsbegreper mv.
260	Registrerte helt arbeidsledige og personer på arbeidsmarkedstilrak	2106	06.03	Kvartal og årlig	Fulltelling	\$SOFA/DATAMN DSOEK	måneddata SAS filer på URSUS	j	RS SOFA i Datadok	Komm/fylke Alder
260	Arbeidskraftundersøkelsene (AKU)	2090	06.01	Kvartal og årlig, månedlig nøkkeltall	Utvalg	flate filer og sas-dataset	flate filer og sas-dataset	n	Q: \dok NOS, web. Notat under arb.	Næring, Yrke, Utdanning Fylke Alder
260	Arbeidstakere etter arbeidssted og utvalgte næringer	2102	06.01	Årlig	Fulltelling. Registerba sert.	\$ARBTAKEK/ASCII-filer (flate)	revisjon i ORACLE	n	NOS	Næring Komm/fylke
260	Arbeidstakere etter bosted, utdanning og avtalt arbeidstid	2103	06.01	Årlig	Fulltelling. Registerba sert.	ASCII-filer (flate) som lages under produksjonen ligger på \$ARBMARK/bosted/	SAS-datasett: \$ARBTAKEK/periode	j/n	SAS-filer dokumenteres i DataDok. Nye filer: dok under arb. NOS	Utdanning Komm/fylke Alder

## De sist utgitte publikasjonene i serien Notater

- 2000/38 R.N. Johnsen: Kommunale gebyrer knyttet til bolig. Januar 2000. 27s.
- 2000/39 J.-A.S. Lie: Revisjon av data til Pleie- og omsorgsstatistikken i 1997 og 1998. 83s.
- 2000/40 Y. Holm, A.H. Tangen og O.M. Tidemann: Forprosjektrapport om etablering av IMF's internasjonale investeringsposisjon (IIP) for Norge. 97s.
- 2000/41 K.O. Olsen: Forsikring i nasjonalregnskapet. 42s.
- 2000/42 J. Johansen og J. Lajord: FD - Trygd: Dokumentasjonsrapport: Arbeidssøkere. 1992-1998. 74s.
- 2000/43 H.V. Sæbø: Til statistikkens pris: Prispolitikk i statistikkbyråene med hovedvekt på elektronisk formidling. 9s.
- 2000/44 E. Rønning: Omnibusundersøkelse - mai/juni 2000. Dokumentasjonsrapport. 32s.
- 2000/45 A. Holmøy og M. Høstmark: Undersøkelse om omfanget av utgifter til helse- og sosialtjenester: Dokumentasjon og tabellrapport. 116s.
- 2000/46 Fagseminar om arealpolitikk og arealstatistikk i opptakten til et nytt årtusen. Seminarrapport 30. mars 2000. 167s.
- 2000/47 Publikasjoner fra forskningsvirksomheten 1991-1999: Revidert versjon. 82s.
- 2000/48 A.-K. H. Grorud: Bedrifts- og foretaksregisteret: Regler og rutiner for ajourhold. 121s.
- 2000/49 T. Hoel, B.R. Joneid og G.E. Wangen: Trekkbas: Brukerdokumentasjon. 35s.
- 2000/50 J.F. Bjørnstad: En innføring i utvalgsundersøkelser. 91s.
- 2000/51 A.G. Pedersen: Oversikt over dødsårsaksregisterets koderegler 1996-1997. 39s.
- 2000/52 O. Klungsøyr: Sammenligning av mikroformler for prisindekser og modelltilpassning. 36s.
- 2000/53 G. Dahl og B.R. Joneid: FD - Trygd: Dokumentasjonsrapport: Inntekt og formue, 1992-1997. 36s.
- 2000/54 K. Flugsrud og G. Haakonsen: Utslipp av klimagasser i norske kommuner: En gjennomgang av datakvaliteten i utslippsregnskapet. 44s.
- 2000/56 T. Hægeland: "Ny" vekstteori: Et nytt forskningsprogram eller naturlig progresjon? Utviklingen innenfor økonomisk vekstteori vurdert i forhold til Imre Lakatos' vitenskapsfilosofi. 19s.
- 2000/57 K.-G. Lindquist: SAS-programmer for korrigerende av data fra industristatistikken og beregning av variable for analyseformål. 53s.
- 2000/58 A. Akselsen: FD - Trygd: Dokumentasjon av uttak til Sandmanutvalget (SHD). 28s.
- 2000/59 J. Johansen og Ø. Sivertstøl: FD - Trygd: Dokumentasjonsrapport: Fødsels- og sykepengene, 1992-1998. 109s.
- 2000/60 K.-G. Lindquist: Database for energiintensive næringer: Tall fra industristatistikken: Oppdatert versjon av notat 97/30. 17s.
- 2000/61 O. Haugen: Utrekning av vekter til inntekts- og formuesundersøkingene 1998. 24s.
- 2000/62 Ø. Kleven: Ferieundersøkelsen 1999 Panel: Dokumentasjonsrapport. 50s.
- 2000/63 P.G. Larssen: Overvakingssystemet for bedrifter i BoF. 29s.
- 2000/64 R. N. Johnsen: Undersøking om foreldrebetaling i barnehagar, august 2000. 36s.
- 2000/65 A. Thomassen: Byggekostnadsindeks for rørleggerarbeid i kontor- og forretningsbygg. 14s.