

Forord til 4. reviderte utgave

4. reviderte utgave inneholder mye av det samme som versjon 3 (ID99/1), men med noen justeringer. Endringene er et resultat av siste års satsing på arbeidet med dataadministrasjon i SSB, og er drøftet i møter med lokale dataadministratører.

Dataadministrasjon i SSB ble satt i fokus i forbindelse med VP for 2000. I retningslinjene stod følgende:

"I forbindelse med at alle data i SSB lagres på ny plattform fra årsskiftet, etableres det rutiner for å følge opp at data lagres i henhold til vedtatte retningslinjer, at de er dokumentert og at de på andre måter er sikret mot utilsiktet tap eller misbruk. Dette krever blant annet konkrete ansvarshavende i alle seksjoner. Det forutsettes at det etableres lokale dataadministratører i seksjonene som blir kontaktpunkt og ansvarlige for oppfølging av seksjonens lagring av data. Den sentrale oppfølgingen skal foretas av Seksjon for IT."

Gruppe for sentral dataadministrasjon (SDA) ble opprettet i Seksjon for IT, og har et sentralt ansvar for å følge opp retningslinjene for lagring og dokumentasjon av data. Samtidig har alle seksjoner opprettet lokale dataadministratører (LDA). Seksjonssjef er formell eier av seksjonens data, og LDAs oppgave er å bistå seksjonsleder i arbeidet med å kontrollere og følge opp seksjonens bruk og lagring av data. LDA er samtidig et viktig knutepunkt både internt i seksjonen og i forhold til andre seksjoner når det gjelder spørsmål om seksjonens data.

Retningslinjene er opprinnelig skrevet av Håkon Berby. Hans innsats med reglene for navnestandarden i SSB er en viktig pilar for det arbeidet som SDA skal føre videre.

Endringene er i hovedsak i kapittel 4, Retningslinjer for lagring av data og navnestandarden for lagringsområder og filer.

Det er tre nyheter i versjon fire av navnestandarden:

1. Bruk av store bokstaver:
 - Store bokstaver godkjennes i substamme, filklasse, versjon og filhale, men vi oppfordrer til å bruke små bokstaver da dette faktisk er det mest "fornuftige" fordi unix er "casesensitive".
2. Bruk av substammer:
 - Før var det kun tillatt med en substamme i hver filstreng. Nå er det imidlertid tillatt å bruke flere substammer i forbindelse med alle faste underkataloger unntatt i "arkiv" filstrenger.
3. Nye faste underkataloger:
 - "famedb"
 - "trolldb"
 - "sasdb"
 - "incl"

Denne reviderte utgaven er godkjent i IT-utvalget i 30.mai 2001.

Kongsvinger/Oslo, 13. august 2001
Svein Longva

Forord, 1. utgave

Dette notatet definerer de grunnleggende retningslinjene for datalagring på SSBs UNIX-maskiner. Retningslinjene er utarbeidet over et lengre tidsrom i samarbeid med et antall medarbeidere i EDB-gruppene på avdelingene for økonomisk statistikk og personstatistikk og Driftskontoret såvel som i Seksjon for EDB utvikling i Avdeling for samordning og utvikling. Noen har bidratt med mer grunnleggende tekniske innspill, andre har bidratt med kritikk av, og forslag til forbedringer i, framstillingen. Av disse har Stein Turtumøygard og særlig Hanne Modahl og Liv Daasvatn bidratt med meget verdifulle, kritiske gjennomganger av både form og innhold. Det tekniske "gjennombrudd" for differensiert adkomstregulering til stammer og filer skyldes Kjell Stræte; beskrivelsen av reguleringsmekanismen, og eventuelle problemer med å forstå den, skyldes forfatteren.

Mange behov og hensyn er søkt ivaretatt, noe som på en del områder har resultert i kompromisser mellom ulike ønsker og behov på den ene siden og tekniske begrensninger i UNIX på den andre siden. Kompromissene gjelder både prinsippene for navning av lagringsområder og filer, samt reguleringen av adkomsten til dem.

Retningslinjene vil bli ajourført på grunnlag av erfaring i bruk, teknisk utvikling samt endrede behov.

Oslo, 24/11-1995

Håkon Berby

Forord til 2. reviderte utgave

2. reviderte utgave følger de samme hovedprinsippene som ble lagt til grunn for 1. utgave. Det som er nytt eller endret er en konsekvens av ønsker fra brukersiden om forbedringer etter at retningslinjene har vært anvendt på flere statistikkområder. Noen av endringene forutsetter utbygging og styrking av systemet for datadokumentasjon. Dette gjelder bl.a. for substammer, hvor det ikke lenger er krav til prefiks i navnene på substammene.

Den første utgaven ble behandlet i EDB-utvalget i november 1995 og der vedtatt som gjeldende navnestandard. Den ble publisert i serien Interne dokumenter (ID 95/13 "Datalagring på UNIX i Statistisk sentralbyrå"). Den foreliggende revisjonen har vært behandlet i EDB-utvalgets møte den 19. september d.å. og utvalget vedtok at denne versjonen (i) definerer gjeldende navnestandard, (ii) danner spesifikasjonsgrunnlaget for automatiske rutiner for sletting av "gamle" arbeidsfiler og filer som ikke følger navnestandarden og (iii) vil også stille krav til utbygging av systemet for datadokumentasjon.

Oslo, 11/10-1996
Håkon Berby

Forord til 3. reviderte utgave

3. reviderte utgave følger de samme hovedprinsippene som ble lagt til grunn for de to tidligere utgavene (2. reviderte utgave, ID 96/14). Det er tre nyheter når det gjelder navnestandarden:

1. Vedrørende den faste underkatalogen for lagring av arkivdata
 - Navnet på den faste underkatalogen for lagring av arkivdata er endret fra **/data** til **/arkiv** for en mer intuitiv forståelse av hva slags data den skal romme. I en ubestemt overgangsperiode, vil katalognavnet **/data** fremdeles være tillatt. Uansett vil overgangen ikke skje før år 2000 p.g.a. at nødvendige ressurser må prioriteres til teknologiskiftet og eventuelle "år 2000 problemer".
 - I de to tidligere utgavene av disse retningslinjene har en "åpning" for å merke navnet på en enkeltfil med prefiks for arbeidsdata dessverre "hengt igjen" fra tidlige utkast til navnestandard hvor den faste underkatalogen **/data** ikke var reservert for langtidslagring av arkivdata. Denne "feilen" er rettet opp i denne utgaven slik at det heretter ikke kan legges inn arbeidsdata i underkatalogen for arkivdata. I en overgangsperiode, i det minste til etter år 2000, kan etablerte filer med arbeidsdata allikevel beholdes.
2. Vedrørende navnsettingen av kataloger og filer:
 - Begrensningen på 8 posisjoner for navnene på underkataloger og filer under stammekatalogene er fjernet. Begrensningen på 7 posisjoner for stammekataloger er opprettholdt av administrative grunner.
 - Navnestandarden har hatt som implisitt, og uskrevet, forutsetning at "standard UNIX konvensjoner" blir fulgt. Siden det ikke finnes noe *obligatorisk* kurs som inneholder alle slike konvensjoner på UNIX har denne forutsetningen ikke slått til for alle UNIX-brukere i SSB. Et vesentlig punkt i disse konvensjonene er at katalog- og filnavn skal skrives med små bokstaver. Dette fordi det er betydningsforskjell mellom store og små bokstaver på UNIX, og brukerne skal være sikre på skrivemåten for katalog- og filnavn til enhver tid. Derfor er disse tidligere implisitte forutsetningene skrevet inn (se punkt 4.2).

Denne reviderte utgaven er godkjent i IT-utvalget i mars 1999. Det vil bli tatt initiativ til å opprette en bredere sammensatt arbeidsgruppe som skal behandle og revidere:

- Inndelingen i datatyper.
- Hvilke data som skal defineres som arkivdata.
- Regler og rutiner for sletting av arkivdata.

Etter at denne revisjonen er godkjent, er det planlagt å gi ut retningslinjene i serien Håndbøker.

Oslo, 11/03-1999

Håkon Berby

Innhold

1.	INNLEDNING	6
2.	OVERORDNET DATASTRUKTUR I SSB.....	6
2.1.	DATATYPER	8
2.2.	BEARBEIDING, ORGANISERING OG LAGRING AV DATA	10
2.3.	NOEN ANDRE BEGREPER	12
3.	DATAADMINISTRASJON, LANGTIDSLAGRING OG DATAARKIV	12
3.1.	DATAADMINISTRATIVE FORHOLD	12
3.2.	ARKIVDATA.....	13
3.3.	DATAARKIV.....	14
3.4.	OVERFØRING TIL RIKSARKIVET (RA)	14
4.	RETNINGSLINJER FOR LAGRING AV DATA OG NAVNESTANDARD FOR LAGRINGSOMRÅDER OG FILER	14
4.1.	RETNINGSLINJER FOR LAGRING AV DATA.....	14
4.1.1.	<i>Langtidslagring av arkivdata.....</i>	<i>14</i>
4.1.2.	<i>Lagringsformat for arkivdata (i flate filer).....</i>	<i>14</i>
4.2.	NAVNESTANDARD FOR UNIX-FILER	17
4.2.1.	<i>Formål.....</i>	<i>17</i>
4.2.2.	<i>Omfang.....</i>	<i>17</i>
4.2.3.	<i>Skrivemåte for katalog- og filnavn.....</i>	<i>18</i>
4.2.4.	<i>Brukernes hjemmekataloger.....</i>	<i>18</i>
4.3.	BRUKSOMRÅDER, LAGRINGSOMRÅDER OG DATAFILER.....	18
4.3.1.	<i>Hovedbegreper for identifisering av SSBs datafiler.....</i>	<i>18</i>
4.3.2.	<i>Hvordan hovedbegrepene kan representeres i katalog-/filstrukturen.....</i>	<i>19</i>
4.3.3.	<i>Stamme.....</i>	<i>21</i>
4.3.4.	<i>Substamme.....</i>	<i>22</i>
4.3.5.	<i>Fast underkatalog ./arkiv/.....</i>	<i>25</i>
4.3.6.	<i>Øvrige faste underkataloger.....</i>	<i>30</i>
4.3.7.	<i>Filhale.....</i>	<i>31</i>
4.4.	KODEKATALOGER FOR STANDARD KLASSEKJEMPE- OG ANDRE GRUPPERINGSKJENNEMERKER	31
4.5.	REGLER FOR BRUK AV DATA FRA ANDRE BRUKSOMRÅDER.....	32
5.	HVORDAN STAMMEKATALOGENE ER KNYTTET TIL FILSYSTEMET PÅ UNIX.....	32
6.	PROGRAM FOR KONTROLL AV NAVNESTANDARDEN.....	34
6.1.	EKSEMPEL PÅ NAVNESTANDARDKONTROLLPROGRAM.....	35
DE SIST UTGITTE PUBLIKASJONENE I SERIEN STATISTISK SENTRALBYRÅS HÅNDBØKER.		37

1. Innledning

Dataforvaltning er en viktig og sentral oppgave for Statistisk sentralbyrå. Forvaltningen har som mål at dataene lagres på en sikker og velorganisert måte for de ulike anvendelsesområder. Dette notatet beskriver en overordnet datastruktur som skal ivareta:

- Sentrale elementer av dagens produksjonsmønster.
- Elementer som ennå ikke er ferdig etablerte. Eksempelvis tverrsektorielle mikrodatabaser og referansedata (et begrep hentet fra Eurostats DSIS prosjekt).

SSBs data må lagres slik at de kan føres inn under den overordnede datastrukturen. Notatet beskriver regler og retningslinjer for navnsetting av lagringsområder og filer på UNIX.

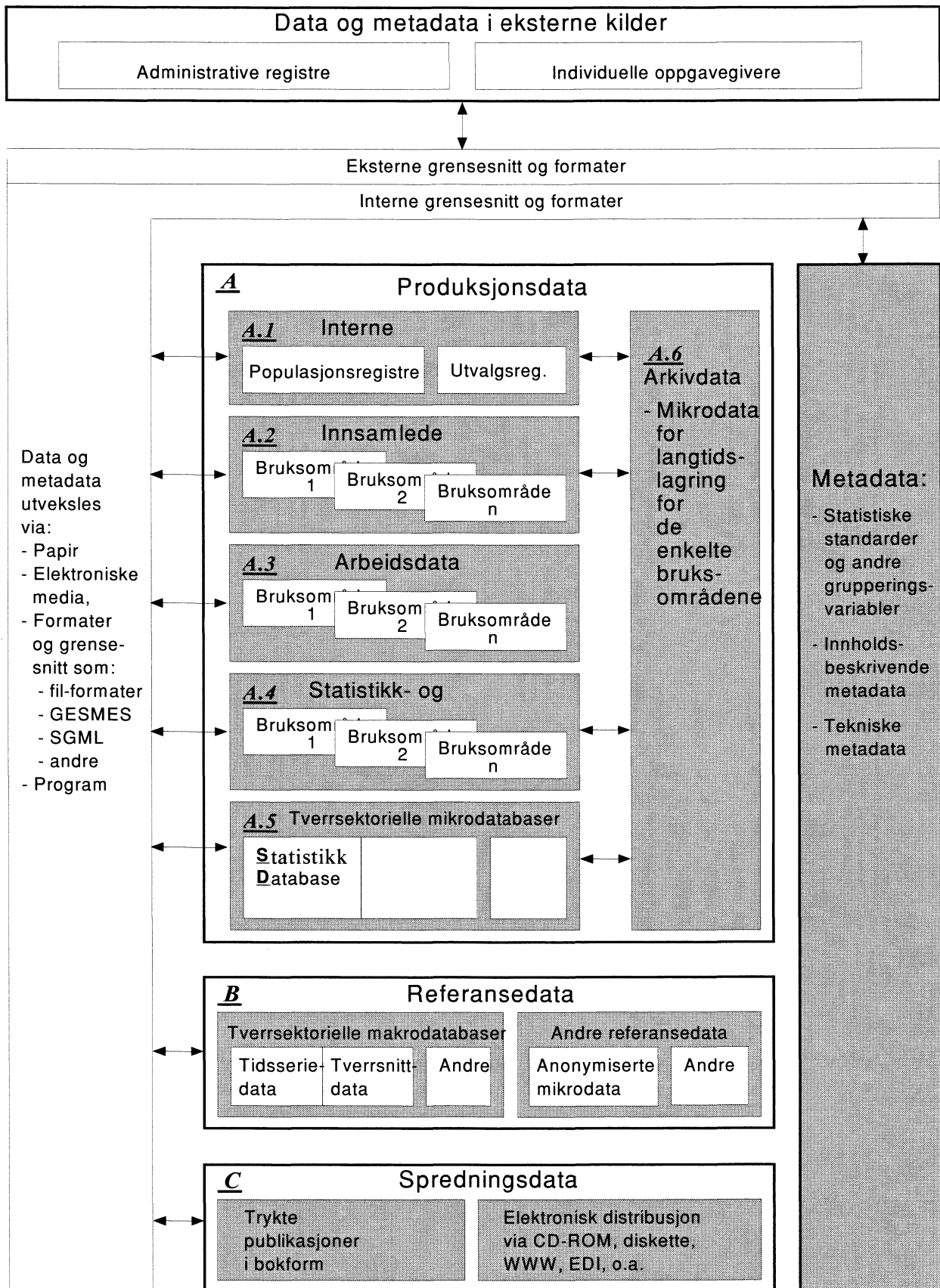
Retningslinjene tar også utgangspunkt i lover, regler og retningslinjer for forvaltning av SSBs data bl.a. i statistikkloven, Lov om personregister og konsesjonsbetingelsene til Datatilsynet, regler om overføring av data til Riksarkivet. I tillegg er tilgjengelig teknologi og kompetanse viktige rammevilkår. Retningslinjene forutsettes revidert i takt med endringer i lover og forskrifter, tekniske rammebetingelser samt strategier for produksjon og dataforvaltning.

Deler av avsnittene om navnstandard og adkomstregulering forutsetter at en del grunnleggende begrep innen UNIX filsystem er kjent.

2. Overordnet datastruktur i SSB

Den overordnede datastrukturen er illustrert i figur 2.1 etter et statusorientert perspektiv. Det vil si at fokus er lagt på datatyper og ikke på de prosessene som bruker og produserer data. Tallene i boksene i figuren refererer til datatypene som er beskrevet nedenfor.

Notatet "*Beskrivelse av ulike datatyper i SSB*" (GOI/HBe/JoL, 5. desember 1994) utviklet et begrepsapparat for å beskrive en overordnet datastruktur for SSB. Det tas her utgangspunkt i disse begrepene.



Metadata:

- Statistiske standarder og andre grupperingsvariabler
- Innholdsbeskrivende metadata
- Tekniske metadata

Data og metadata utveksles via:

- Papir
- Elektroniske media,
- Formater og grensesnitt som:
- fil-formater
- GESMES
- SGML
- andre
- Program

Figur 2.1 Overordnet datastruktur - datatyper i et statusorientert perspektiv
 Dataene innen de ulike datatypene, i figuren, tilhører ulike bruksområder.

Med **bruksområde** menes her et bestemt *statistikk-, forsknings- eller analyseområde*. Et arbeidsområde som omfatter organisering og forvaltning/behandling av tverrsektorielle mikro- eller makrodata, er også definert som et bruksområde.

Den overordnede lagringsstrukturen vil komme til uttrykk både i datastrømmer og produksjonsprosesser innen og mellom de enkelte bruksområdene, i overordnede metadatastrukturer i en dokumentasjonsbase, samt i brukernes innganger til data i ulike meny- og ikonstrukturer.

2.1. Datatyper

Datastrukturen over omfatter både data og metadata.

Vi nevner her bare kort at metadata er data om data. Metadata beskriver og dokumenterer data. Eksempler på metadata er:

- *Statistiske standarder inkl. kodelister for standard klassifikasjons- og andre grupperingskjennemerker*
- *Dokumentasjon og innholdsbeskrivelser av datafiler og -baser*
- *Dokumentasjon av kjennemerker*
- *Dokumentasjon av produksjonsprosessen som behandler data*

Bortsett fra UNIX-filer som inneholder dokumentasjon og kataloger over kodelister for standard klassifikasjons- og andre grupperingskjennemerker, er dette notatet stort sett viet datadelen i strukturen.

Vi skiller i hovedsak mellom tre hovedtyper av data etter et mønster fra Eurostats DSIS-prosjekt (Distributed Statistical Information System):

- A Produksjonsdata
- B Referansedata
- C Spredningsdata

A. Produksjonsdata

Produksjonsdata omfatter ulike typer av data som inngår i produksjonen av primærstatistikk og av analyseoppdrag. Produksjonsdata vil normalt inneholde identifiserbare opplysninger. Inndelingen i basisregistre, rådata, statistikk- og analysedata samt tverrsektorielle mikrodata er basert på den rolle de har i produksjonsprosessen. Arbeidsdata og arkivdata er betegnelser på data ut fra hvordan de skal lagres. Arkivdata er data som skal lagres for alltid i åpent format på særskilte områder og være dokumentert i DataDok.¹ Arbeidsdata er alle andre data lagret i formater tilpasset aktuelle programverktøy i databaser eller arbeidskataloger. *Den følgende kategoriseringen av datatyper er noe overlappende.*

1. Basisregistre

Med basisregistre menes registre som definerer populasjonene for statistikkrutinene. Basisregistrene danner trekkegrunnlaget for mange statistikkrutiner, og utdrag fra registrene er direkte grunnlag for andre statistikkrutiner. De viktigste basisregistrene er:

- BoF, Det sentrale bedrifts- og foretaksregisteret
- BEBAS, et "skyggeregister" av DSF (Det sentrale folkeregisteret)

¹ Se beskrivelse på intranettet.

2. Rådata

Rådata omfatter data som er samlet inn direkte fra oppgavegivere eller indirekte fra administrative registre før de eventuelt er stilt sammen for revisjon og bearbeiding i produksjonsrutinene for de enkelte bruksområdene.

3. Arbeidsdata

Arbeidsdata omfatter datafiler som opprettes i løpet av revisjons- og bearbeidingsprosessen innen et bruksområde, og som kan slettes når prosessen er ferdig.

4. Statistikk- og analysedata

Statistikk- og analysedata er endeholdeplassen for bearbeidingsrutinene for de enkelte bruksområdene. De utgjør grunnlaget for publisering av statistikk og av analyse- og forskningsresultater samt for videre analyse.

5. Tverrsektorielle mikrodata

Med tverrsektorielle mikrodata menes organiserte samlinger av statistikkdata, ofte harmonisert etter internasjonale standarder. Eksempler på slike data er:

- Folke- og boligtellinger
- Barnefilen / mødrefilen
-

6. Arkivdata

Arkivdata er data som skal danne basis for gjenbruk i SSB og følgelig lagres for evig tid (langtidslagres). De kan også være gjenstand for overføring til Riksarkivet.

Arkivdata rommer hovedsakelig:

- statistikkdata/analysedata, som i utgangspunktet alltid skal langtidslagres,
- Grunnregelen er at alle rådata skal langtidslagres (de skal kunne danne basis for kommende kvalitetskontroll og -rapportering)
- situasjonsfiler fra basisregistrene.

Fagseksjonene har ansvaret for å vurdere hvilke datafiler som skal gis status som arkivdata.

B. Referansedata

Med referansedata forstår vi data som er offisielt godkjente og publiserbare. Referansedata bør være grunnlaget for elektronisk publisering og internasjonal rapportering.

I første omgang vil referansedata bare omfatte makrodata, dels i form av tverrsnittorientert statistikk (Regionalstatistisk database), dels i form av tidsserieorientert statistikk (Referansedatabase for tidsserier). Utdrag fra det nasjonale referansemiljøet vil være tilgjengelig som en del av det europeiske referansemiljøet når dette blir etablert.

Man vil også kunne tenke seg anonymiserte utdrag av mikrodata som en del av referansemiljøet, formidlet via standardiserte grensesnitt. Dette materialet vil f.eks. kunne være tilgjengelig for Eurostat, OECD og for spesielt interesserte forskere.

C. Spredningsdata

Med spredningsdata forstår vi *i prinsippet* utdrag av data fra referansemiljøet beregnet på spredning mot ulike brukergrupper. I dagens situasjon vil statistikkdata *i praksis* ofte være kilden for spredningsdata.

2.2. Bearbeiding, organisering og lagring av data

I et bearbeidingsperspektiv kan følgende nevnes i tilknytning til datatypene:

A.1 De grunnleggende basisregistrene foreligger i SSB som strukturerte databaser som oppdateres kontinuerlig og/eller som situasjonsuttak. *Databasene* som oppdateres kontinuerlig oppfattes her som en form for *arbeidsdata*. *Situasjonsuttakene* kan være *rådata* til statistikkproduksjonen eller *statistikk/analysedata*. De skal uansett (i) være definert som arkivdata, (ii) organiseres i flate filer og (iii) lagres på UNIX i følge navnestandarden for arkivdata. Omfatter eksempelvis:

- BoF (Bedrifts og foretaksregisteret)
 - BoF oppdateres bl.a. regelmessig med meldingsfiler fra Enhetsregisteret. Meldingene er å betrakte som rådata som skal langtidslagres
 - Den strukturerte databasen som oppdateres kontinuerlig, er arbeidsdata.
 - Aktuelle situasjonsuttak skal lagres som arkivdata.
- BEBAS (skyggeregister av DSF (Det sentrale folkeregisteret) som er et eksternt administrativt register som forvaltes av Sentralkontoret for folkeregistrering som igjen hører inn under Skattedirektoratet).
 - SSB ajourfører BEBAS regelmessig ved hjelp av transaksjonsmeldinger fra DSF; i intervaller på ca 14 dager (februar 1999).
 - BEBAS er ikke en kopi av DSF, da meldingene til en viss grad videreføres av SSB.
 - Transaksjonsmeldingene fra DSF er i SSB å betrakte som rådata som skal defineres som arkivdata.
 - Aktuelle situasjonsuttak fra BEBAS skal defineres som arkivdata.

A.2 Innsamlede rådata, omfatter:

- Data som er valgt ut og hentet fra administrative registre i form av situasjonsuttak
- Direkte innsamlet på skjema
- Direkte innsamlet på elektronisk form
 - Meldinger i EDI-format
 - "filer" overført via magnetiske lagringsmedia (bånd/diskett) eller overført via linje/nettverk

Rådataene er oftest samlet inn og organisert etter et bestemt registreringstidspunkt eller -periode. De fleste rådata skal bli gitt status som arkivdata. Disse må i tilfelle (i) organiseres i flate filer (ii) navnes, (iii) lagres på område for arkivdata og (iv) dokumenteres etter retningslinjene.

Andre rådata organiseres og lagres som arbeidsdata tilpasset den øvrige produksjonsprosessen.

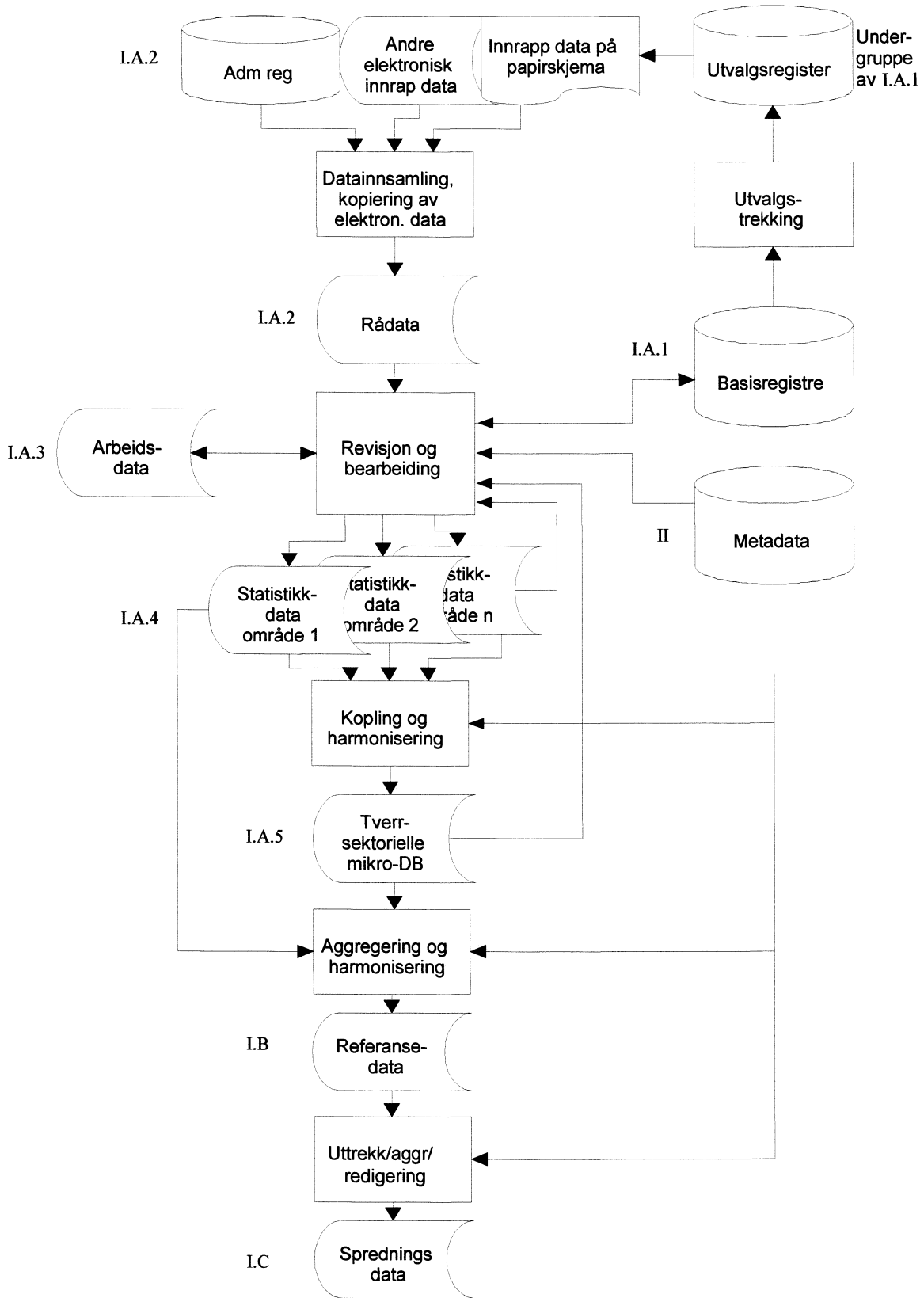
A.3 Arbeidsdata, omfatter:

- Data under revisjon/koding, lagres i dag oftest i databaser for aktuelt registreringstidspunkt (evt. -periode), i mange tilfelle også med data for tidligere registreringsperiode for kontroll og revisjon.
- Data under annen bearbeiding, lagres ofte i filer i proprietære formater som avhenger av det verktøyet som anvendes i bearbeidingen.

A.4 Statistikk- og analysedata

- Den såkalte endeholdeplassen for dataene fra de enkelte statistikkrutinene. Dataene skal (i) *organiseres* i flate filer, (ii) *navngis/lagres* i områder for *arkivdata* på UNIX og (iii) *dokumenteres* med henblikk på gjenbruk og evt. overføring til Riksarkivet.

Se også figur 2.2 om datatyper i produksjonsorientert perspektiv på neste side.



Figur 2.2 Datatyper i et produksjonsorientert perspektiv

2.3. Noen andre begreper

Lagringsområder

Et lagringsområde er et avgrenset område i filsystemet til en eller flere datamaskiner i et nettverk. I et operativsystem som UNIX vil et slikt lagringsområde være definert som en katalog ("directory"), eventuelt med underkataloger, som rommer en samling filer. Slike lagringsområder skal bidra til å:

- regulere adkomsten til dataene
- skape god oversikt over datafilene
- administrere det samlede datalageret

Dataeier

Seksjonsleder er dataeier og ansvarlig for seksjonens data.

Fileier

Fileier (og katalogeier) er et fast begrep innen UNIX operativsystem. Det er betegnelsen på den brukeridenten som har opprettet en fil (eller en katalog). Brukere som har adkomst via denne brukeridenten har bestemte rettigheter i forhold til bruk og administrasjon av de filene/katalogene som er opprettet via brukeridenten.

3. Dataadministrasjon, langtidslagring og dataarkiv

Dette avsnittet behandler forbindelsen(e) mellom avsnitt 2 "Overordnet datastruktur i SSB" (heretter betegnet "datastrukturen") og avsnitt 4 "Retningslinjer for lagring av data og navnestandard for lagringsområder og filer" (heretter betegnet "navnestandarden").

3.1. Dataadministrative forhold

Navnestandarden er et virkemiddel innen SSBs dataadministrasjon. Innenfor denne rammen, har navnestandarden bl.a. som formål (jf. avsnitt 4.2.1) å støtte lagring og gjenfinning av datafilene for de ulike bruksområdene. For dette formålet bidrar navnestandarden til

1. å definere sikre lagringsområder (via stammekataloger)
2. å identifisere datafilene for bruk og gjenbruk (via stamme, filklasse og generasjon).

Navnestandarden søker også å håndtere et grunnleggende skille mellom datatypene for å lette (evt. automatisere) den mer fysiske dataadministrasjonen. Et hovedskille går mellom:

- i. filer med de sentrale data som skal tas vare på over tid for arkiv- og beredskapsformål, dvs.:
 - rådata som skal langtidslagres
 - meldingsfiler for oppdatering av basisregistrene er å betrakte som rådata som skal langtidslagres
 - statistikkdata og analysedata som skal langtidslagres
 - situasjonsuttak fra basisregistrene betraktes å være på linje med statistikkdata som derfor skal langtidslagres
 - referanse- og spredningsdata
- ii. de øvrige filene, inkludert analysedata som ikke skal langtidslagres, kan oppfattes som arbeidsdata og skal slettes av automatiske rutiner.

De filene som skal *langtidslagres* for gjenbruk og eventuell overføring til Riksarkivet skal være identifisert med begrepene stamme, filklasse og generasjon og lagret i den faste underkatalogen .../arkiv/... , som skal være plassert under en stammekatalog eller en substamme når stammen rommer substammer. De identifiserende begrepene, inkludert substamme, skal være registrert i DataDok (det sentrale datadokumentasjonssystemet).

De filene som rommer *arbeidsdata*, skal være lagret i egne underkataloger for arbeidsdata og trenger ikke å være registrert i det sentrale datadokumentasjonssystemet.

Rådata som skal langtidslagres (inkludert meldingsfiler til basisregistrene), vil av praktiske grunner ofte bli etablert i én eller flere filer (i ymse formater) i underkataloger for arbeidsdata. For langtidslagring må disse dataene konverteres (og eventuelt reorganiseres) til åpent format for overføring til underkatalogen .../arkiv/... .

Rådata som ikke skal langtidslagres eller som er lagret i *produktspesifikke formater*, bør rent administrativt bli håndtert på linje med arbeidsdata, og *skal være lagret i underkataloger for arbeidsdata*. Filer med slike data trenger heller ikke å være registrert i det sentrale datadokumentasjonssystemet.

Statistikk- og analysedata vil ofte bli dannet som resultat av en revisjons-/bearbeidingsprosess, eventuelt være et ekstrakt av en eller flere "revisjonsdatabaser", andre statistikk-/analysedata, etc. Dataene vil deretter bli brukt i tabellproduksjon og analyseformål som normalt vil kreve at de er lagret i filer i produktspesifikke formater i underkataloger for arbeidsdata. Situasjonstakkene fra basisregistrene kan i seg selv være analysedata eller være del av rådataene til enkelte statistikkrutiner; de skal derfor langtidslagres som arkivdata.

Når bearbeidingsrutinene er gjennomført (for den aktuelle generasjonen eller evt. engangsundersøkelsen), konverteres statistikk-/analysedataene til en flat fil i åpent format (se pkt. 4.1) og overføres til ../arkiv/.. katalogen. Normalt skal datafilen i produktspesifikt format slettes etter dette, men ved behov (f.eks. beredskapsformål) kan dataene også beholdes i produktspesifikke formater i underkataloger for arbeidsdata.

Referanse- og spredningsdata vil som regel bli håndtert, tildels også bli kontinuerlig oppdatert, via særskilte systemer og funksjoner i produktspesifikke formater så lenge de er aktuelle. De må derfor være langtidslagret i de produktspesifikke formatene (databaser), og trenger normalt ikke langtidslagring i åpent format.

3.2. Arkivdata

For alle typer produksjonsdata, unntatt arbeidsdata, vil det finnes datafiler som langtidslagres og som derfor skal defineres som arkivdata. Se også definisjon av arkivdata i avsnitt 2.1 samt avsnitt 4.1.

De langtidslagrede arkivfilene skal danne basis for gjenbruk i SSB over et langt tidsrom. En del av de langtidslagrede dataene vil også være gjenstand for overføring til Riksarkivet. For å dekke begge disse formålene skal arkivdataene lagres i et så åpent format som mulig:

De langtidslagrede arkivdataene skal være lagret i form av ("flate") sekvensielle datafiler i tegn- eller tekstformat og med fast postlengde.

Tegnstandard skal være: ISO 8-bits (8859-1) standard.

Det vil si at dataene skal være pakket ut fra de produktspesifikke (proprietære) filformatene til de ulike programleverandørene før overføring til langtidslageret. De enkelte kjennemerkene må heller ikke være lagret i felt med pakkede formater som for eks. binære heltalls- eller flyttallsformater.

Data som inngår i tverrsektorielle mikrodatabaser, referansedata og data i basisregistrene skal være tilgjengelige på kort varsel via etablerte systemer. De må derfor normalt være langtidslagret (også) i sine produktspesifikke formater i arbeidskataloger for dette formålet og på en slik måte at de ikke blir slettet (husk - automatisk sletting skjer kun etter overskredet tidsramme *etter siste gangs bruk*).

3.3. Dataarkiv

Dataarkivet skal romme alle langtidslagrede data. Det vil omfatte de datafilene som ligger på plattelager under de ulike stammenes faste underkatalog /arkiv/.

3.4. Overføring til Riksarkivet (RA)

Det er foreløpig ikke utarbeidet rutiner for overføring av filer til RA, og heller ikke avklart rutiner for loggføring av slike forsendelser.

4. Retningslinjer for lagring av data og navnestandard for lagringsområder og filer

4.1. Retningslinjer for lagring av data

4.1.1. Langtidslagring av arkivdata

Retningslinjene for lagring av data inneholder en detaljert beskrivelse av navnestandarden for lagring av arkivdata i punkt 4.2 (navnestandard for UNIX-filer). Gyldige lagringsformat for arkivdata er "åpne format i flate filer", se punkt 4.1.2.

Når det gjelder arbeidsdata har navnestandarden kun et grunnleggende krav om at arbeidsdata som er lagret under filsystemet på UNIX skal være lagt til særskilte underkataloger som indikerer hvor lang tid de kan være lagret før automatisk sletting. For arbeidsdata lagret i databaser, gjelder de regler som måtte finnes for det aktuelle databasesystemet. Se også punkt 2 foran for skillet mellom arkiv- og arbeidsdata. Lagringsformatene vil være avhengig av det formatet som er mest hensiktsmessig for det aktuelle verktøyet.

4.1.2. Lagringsformat for arkivdata (i flate filer)

Vi har et grunnleggende skille mellom **i) nominelle** kjennemerker og **ii) metriske** kjennemerker. Betegnelse nominelle og metriske er hentet fra "Publiseringshåndboken"; i) og ii) er også blitt betegnet henholdsvis *kvalitative* og *kvantitative*. I tillegg vil datafiler også kunne inneholde **iii) ren tekstlig informasjon**. Andre formater som f.eks. grafiske formater og bilder vil ikke bli behandlet her.

i) Nominelle kjennemerker (variabler)

Dette rommer alle former for kjennemerker som er representert ved koder, uansett om kodene er representert med *tallkoder*, andre koder eller en blanding. Denne typen informasjon skal alltid være representert i kjennemerker som har formater av typer som:

- Alfamerisk
- Alfabetisk
- "Character"

Hvordan formatet skal spesifiseres når arkivfilen produseres, er avhengig av det verktøyet som anvendes.

Eksempler på slike kjennemerker er:

- Kommunekode
- Næring (etter Standard for næringsgruppering)
- Kjønn
- Aldersgrupperinger

Spesielt om dato

Dato er egentlig et nominelt (kvalitativt) kjennemerke, men har spesielle krav som gjør at det kan bli gjenstand for beregninger. Dette gjør at når dato lagres i verktøyspesifikke formater kan de lagres i kjennemerker med numerisk format eller om nødvendig med "datoformat".

Ved overføring av data med datoformat fra produktspesifikke SAS-datasett eller Oracle-databaser, skal datoformatet pakkes ut til datofelt med rekkefølgen år, måned, dag (åååå-mm-dd, dvs. antall siffer = $\langle 4+2+2 \rangle^2$). Dato vil da bli lagret "riktig" i filer for arkivdata og vil dekke formål som sortering, presentasjon og datokalkulasjon. Fra en arkivfil med dato lagret som åååå-mm-dd, kan dato leses med rent tegnbasert (alfabetisk, alfanumerisk, tekst, "character") format. Tidspunkt skal lagres med time, minutt, sekund (tt:mm:ss hvor: tt = 00-23). Ønsker man å lagre kombinasjon av dato og tidspunkt, bruker man oppsett: åååå-mm-ddTt:mm:ss. Dette følger internasjonal standard for datering (ISO8601).

Man oppgir så mye dateringsinformasjon som er nødvendig:

År:	åååå
Måned:	åååå-mm
Dag:	åååå-mm-dd
Time:	åååå-mm-ddTt
Minutt:	åååå-mm-ddTt:mm
Sekund:	åååå-mm-ddTt:mm:ss

ii) Metriske kjennemerker (variabler)

Metriske (kvantitative) kjennemerker rommer numerisk informasjon som kan inngå i beregninger. Denne typen informasjon skal alltid være representert i kjennemerker som har "åpne" numeriske formater (uavhengig av verktøy):

- med *ledende blanke*
- med *ledende fortegn ("flytende")* hvis negative tall (Husk! ekstra posisjon til fortegn for ikke å få "avkuttete" tall)
- hvis et numerisk kjennemerke har uoppsett for en enhet skal den lagres som *blank*

² For eldre filer som f.eks overføres fra andre maskinplattformer som MVS (IBM, Comparex) til UNIX, kan de gamle datoformatene beholdes forutsatt at det aktuelle datoformatet dokumenteres i DataDok.

- hvis et kjennemerke har oppgitt 0 (null) for en enhet, skal den lagres med *ledende blanke* foran 0 (se eksemplene nedenfor)

Eksempler på tall for et kjennemerke som skal kunne holde heltall på 12 siffer inkl. negative tall:

- $\frac{123456789012}{}$
- $\frac{-123456789}{}$
- $\frac{}{}$ (= uoppgitt)
- $\frac{0}{}$ (= oppgitt null)
- $\frac{1}{}$

Kvantitative kjennemerker med desimalpunkt skal i utgangspunktet lagres med *implisitt desimalpunkt*.

Eksempler på tall for et kjennemerke som skal kunne holde tall på 12 siffer pluss to desimaler inkl. negative tall:

- $\frac{12345678901200}{}$ (= 123456789012,00)
- $\frac{-12345678936}{}$ (= -123456789,36)
- $\frac{100}{}$ (= 1,00)
- $\frac{0}{}$ (= 0,00)

Denne lagringsmåten med ledende blanke vil medføre at kodene til kvalitative kjennemerker som "feilaktig" har vært holdt i numeriske kjennemerker, kan bli lagret feil dersom ikke de aktuelle kjennemerkene lagres med eksplisitt formatering.

- F.eks. vil kommunenummer (dvs. *kommunekode*) 0301 kunne bli lagret som 301 istedenfor 0301.

Alternativt kan kvalitative kjennemerker lagres *med desimalpunkt*. Det må da settes av en karakter for lagring av desimalpunktet, som kan være:

- ',' (komma) eller
- '.' (punktum)

Valg av desimalpunkt må dokumenteres. Eksempler på utfylling (2 desimaler, komma som desimalpunkt):

- $\frac{123456789012,00}{}$
- $\frac{-123456789,36}{}$
- $\frac{1,00}{}$
- $\frac{,00}{}$ (= 0,00)

Andre formater enn de som er nevnt over er **ikke tillatt**.

HUSK! Når et kontinuerlig kvantitativt kjennemerke (eks. alder, inntekt, osv.) "omkodes" til et grupperingskjennemerke (eks. aldersgruppe, inntektsgruppe, osv.) vil det omkodede kjennemerket være et nominelt kjennemerke og skal derfor lagres med format som andre nominelle kjennemerker (alfanumerisk, alfabetisk eller character avhengig av det verktøyet som er brukt).

iii) Tekstlig informasjon

Tekstlig informasjon (inkluderer også spesialtegn) skal alltid være lagret i kjennemerker med formater som:

- Alfamerisk
- Alfabetisk
- "Character"

Hvilket format som skal spesifiseres når arkivfilen produseres, er avhengig av verktøyet.

4.2. Navnestandard for UNIX-filer

4.2.1. Formål

Navnestandarden skal bidra til:

- **at brukerne kan lagre data på en *sikker og velorganisert* måte uten unødig dobbeltlagring**
- **å støtte *gjenfinning, bruk***
- **å støtte *administrasjon* av data, inkludert automatisk sletting av filer**
- **at dataenes eierseksjon skal kunne ha *full oversikt* over de data som den har ansvaret for**
- **at *adkomsten* til data kan reguleres effektivt.**

Følgende forhold kan også nevnes i tilknytning til disse retningslinjene:

- Navnestandarden er i stor grad uavhengig av endringer i SSBs organisasjonsstruktur, produktinndeling og -numre og skal bare berøre plassering og adkomstregulering til datafilene i filsystemet på UNIX samt kopling til dokumentasjon av dataene.
- Navnestandarden skal støttes av et dokumentasjonssystem med funksjoner for gjenfinning av datafiler basert på dataeier (jf. pkt 2.3), register meldt Datatilsynet, produktnummer, o.a.
- Navnestandarden har et fast rammeverk, men kan til en viss grad tilpasses ulike produksjonsbehov. Den kan bli revidert etter erfaring i praktisk bruk.
- Navnestandarden støtter etablering av automatiske rutiner for å slette arbeidsfiler så vel som udokumenterte arkivfiler.

4.2.2. Omfang

Disse retningslinjene dekker en navnestandard for *alle filer med produksjons-, referanse-, sprednings- og metadata samt program-, parameter- og dokumentasjonsfiler* som:

- er lagret i *sekvensielle filer eller andre filer lagret direkte under UNIX' filsystem*
- tilhører et *bruksområde i SSB*.

De dekker, med andre ord, alle rådata (inkludert kopier av datafiler mottatt fra eksterne oppgavegivere), arbeidsdata og statistikkdata som oppstår i forbindelse med bearbeidingen av data innen et bruksområde, foruten arkivdata. Retningslinjene gjelder derfor også for SAS-datasett og for filer som håndteres av FAME, *men ikke* for databasetabeller (-filer) lagret under et

databasehåndteringssystem. Retningslinjene gjelder ikke "private" filer som den enkelte bruker kan lagre på sin hjemmekatalog og som ikke tilhører et av SSBs bruksområder.

4.2.3. Skrivemåte for katalog- og filnavn

- Gyldige tegn i katalog- og filnavn er :
 - a - z, æøå er ikke tillatt
 - 0 - 9, sifrene fra null til ni.
 - A - Z, tillatt brukt i enkelte underkataloger (substamme, filklasse, versjon og filhale).
 - ÆØÅ, er ikke tillat brukt
- NB! Katalog/filnavn skal være mest mulig beskrivende !

- Gyldig spesialtegn i katalog- og filnavn er:
 - _ ("underline")
 - (bindestrek)
 - . (punktum)
- Gyldig spesialtegn for Emacs katalog- og filnavn er:
 - ~ (tilde) er tillat
 - # (grind), NB! Gyldig bare for "autosave"
- Gyldig spesialtegn for Java kompilerte programmer er:
 - \$ (dollar) er tillat i fast underkatalog ../prog/... med ".class" som filhale.

Merk også!: Disse spesialtegnene (unntatt # for autosave fra Emacs) må ikke stå i første posisjon i et katalog-/filnavn.

- Andre spesialtegn enn de som er nevnt over er ikke tillatt.
- Tegnet "**blank**" er ikke tillatt som skilletegn av administrative grunner.
- Stammenavn skal bestå av maks 7 tegn. I øvrige katalog- og filnavn er antall posisjoner valgfritt.
- Både første og siste tegn i alle katalog- og filnavn må være a -z, A-Z,(se over) eller 0-9.

4.2.4. Brukernes hjemmekataloger

Det er ikke tillatt å lagre produksjons-, referanse- eller spredningsdata på brukernes hjemmekataloger uansett plattform.

4.3. Bruksområder, lagringsområder og datafiler

4.3.1. Hovedbegreper for identifisering av SSBs datafiler

Notatet definerer et sett med hovedbegrep for å organisere datalageret og identifisere filene på UNIX:

- stamme
- filklasse
- generasjon

Her følger først en kort presentasjon av begrepene og prinsippene for organisering av det hierarkiske filsystemet. En mer detaljert beskrivelse av begrepene følger i egne avsnitt.

Alle datafiler, samt andre typer filer³ på UNIX, som tilhører et bestemt bruksområde og dets bearbeidingsrutiner, skal være organisert og tilgjengelig under et eget avgrenset lagringsområde i SSBs filsystem på UNIX. Et slikt lagringsområde kalles en *stamme* (punkt 4.3.3.) og identifiseres med stammenavn. For en del statistikkområder hvor det er hensiktsmessig vil en stamme også kunne deles inn i *substammer* (punkt 4.3.4.).

Alle arkivfiler (punkt 4.3.5.) innen et bruksområde har et *gitt formål og status*. F.eks. *statistikkdataene* for en undersøkelse (engangs eller periodisk) har som formål å danne input til en rutine som etablerer det endelige statistikkproduktet. En slik samling filer med statistikkdata, vil ha lik, eller tilnærmet lik, recordstruktur over de ulike generasjonene (statistikkperiodene) av filene med statistikkdata. Alle statistikkdatafilene for et slikt formål, over alle generasjoner, defineres som en *filklasse* som skal dokumenteres i *DataDok*. Innen et bruksområde kan det være én eller flere filklasser. Alle arkivfiler må tilordnes en filklasse. Se også definisjon av filklasse i punkt 4.3.5.1.

Vi må også kunne identifisere én eller flere filer ut fra *generasjon*; hvilke(n) periode(r) eller tidspunkt som dataene dekker (punkt 4.3.5.2.).

Mer om disse hovedbegrepene følger i avsnittene nedenfor, men først noen ord om sammenhengen mellom dem og katalog-/filstrukturer på UNIX.

4.3.2. Hvordan hovedbegrepene kan representeres i katalog-/filstrukturen

Stammene definerer et overordnet katalognivå for bruksområdenes filer på tvers av SSBs nett av UNIX-maskiner. Hver stamme definerer en såkalt *stammekatalog*. Et grunnleggende og viktig prinsipp for å forstå bruken av begrepene i forhold til katalog-/filstrukturene er følgende:

Under stammekatalogene vil filklasse og generasjon definere hvert sitt nivå i den hierarkiske katalog-/filstrukturen på UNIX.

Hver **filklasse** og hver **generasjonsbetegnelse** kan derfor gi navn til *underkatalog(er)* eller *fil(er)*.

For lesbarhet, er **filklasse** skrevet med **halvfete** typer, generasjon med normale typer. Brukte filklasser i eks. 1-4: **prisdata** og **varedata**.

I alle eksemplene er *navnene til filer med arkivdata gitt filhalen* (extension):

.fil (eksempel: \$KONSUM/arkiv/**prisdata**/g1991.fil)

Dette er gjort for å klargjøre de tilfellene hvor begrepene filklasse og generasjon er lagt til filnivå.

Det er også viktig å merke seg:

- at filhalene kan velges fritt av brukeren og
- at filhalene ikke har noe med de begrepene som skal dokumenteres i *DataDok*; dvs.: **filklasse** og *generasjon*.

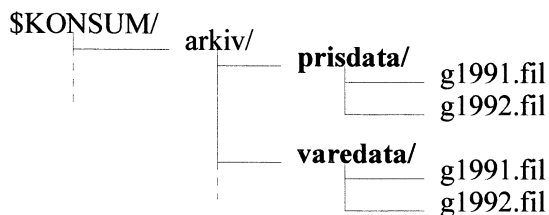
Rekkefølgen til begrepene filklasse og generasjon for arkivdata er valgfritt.

³ Bl.a. program- og dokumentasjonsfiler samt filer med kataloger over statistiske standard klassifiseringer og andre grupperingskjennermerker eller andre metadata.

Nedenfor er det vist fire eksempler på plassering av begrepene i en katalog-/filstruktur. De tre første eksemplene viser et utsnitt av katalogstrukturen med *data som skal langtidslagres*. Det fjerde eksemplet vier en mer fullstendig struktur som også omfatter underkataloger for lagring av arbeidsdata, program, kodekataloger og filer for dokumentasjon.

Eks. 1: \$STAMME/arkiv/filklasse/generasjon.fil

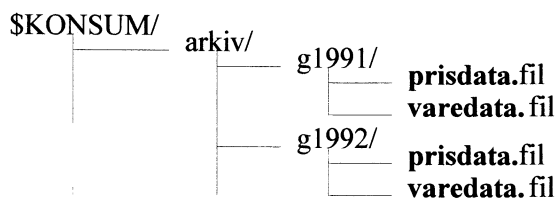
(**filklassenavnene** danner underkatalog, generasjon danner filnavn)



Eksempel 1 viser en katalog-/filstruktur for arkivdata.

Eks. 2: \$STAMME/arkiv/generasjon/filklasse.fil

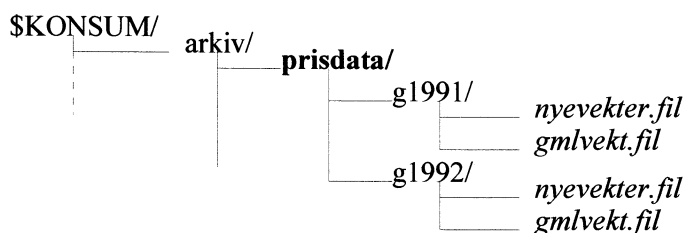
(generasjon danner underkataloger, **filklassenavn** danner filnavn)



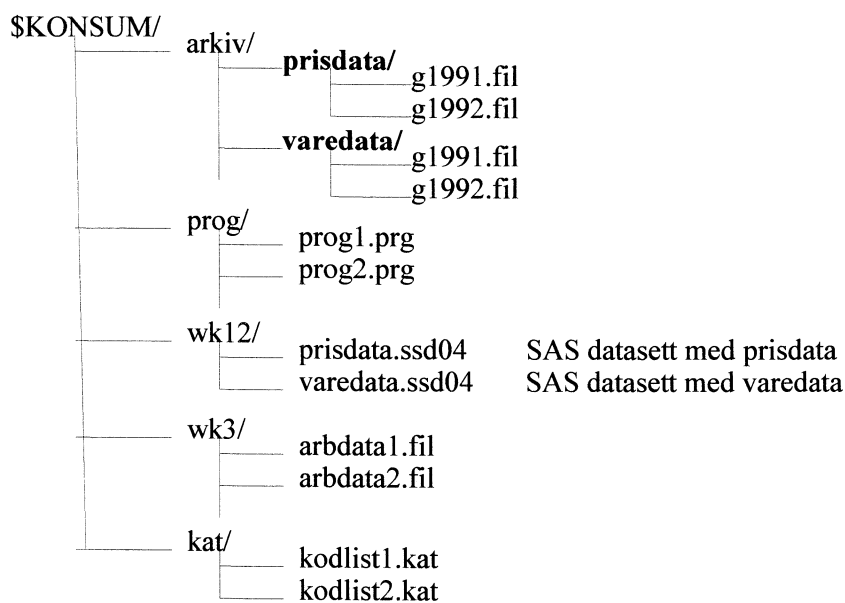
Eksempel 2 viser en annen katalog-/filstruktur for arkivdata.

Eks.3: \$STAMME/arkiv/filklasse/generasjon/versjon.fil

(**filklassenavn** og generasjon danner underkataloger, dataenes *versjon* danner filnavn)



I eksempel 3 tenker vi oss at vi skal *langtidslagre* filer med *gamle* og *nye vekter* i hver sin versjon. Vi har derfor anvendt både filklasse og generasjon for å danne underkataloger under *.../arkiv/...*. Dataenes *versjon* er brukt for å gi navn til filene.

Eks. 4: Utvidet og mer fullstendig katalogstruktur

Eksempel 4 viser en mer fullstendig katalogstruktur med underkataloger som rommer filer med programmer, arbeidsdata, kodekataloger, etc.

4.3.3. Stamme

En **stamme** definerer lagringsområdet for filene til et bruksområde.

- Alle stammer skal være definert i datadokumentasjonssystemet.
- Stammene defineres i form av akronymer som er stabile over tid (eks. "**konsum**" eller "**lotte**").
 - For stammer med produksjonsdata kan stammenavnene velges fritt (*unntak, neste pkt*).
 - Stammer med referansedata skal ha prefiks "**ref**" (pga. andre skript enn produksjonsdata).
 - Spredningsdata vil normalt ha verktøyspesifikke navn (WWW).
- Alle stammebetegnelser må være unike innen SSBs nett av UNIX-maskiner.
- En stamme realiseres i form av en katalog (directory) i UNIX' filsystem. Denne katalogen, "**stammekatalogen**", skal alltid være **overordnet** i forhold til andre kataloger som organiserer filene innen et bruksområde.
 - *Samtidig med opprettelsen av stammekatalogen, blir det opprettet to "UNIX-grupper" (en les og en skrivegruppe som navnesettes tilsvarende stammen) som regulerer adkomstrettighetene til stammekatalogen og dens underkataloger og filer.*
 - *Den lokale dataadministrator (LDA) har ansvaret for å få oppretta nye stammer på seksjonen. Stammer og gruppedefinisjoner defineres i filsystemet av Kundestøtte.*

- Vi skal kunne referere til stammene, og filene under stammene, ved hjelp av miljøvariabler som har navn lik stammenavnet med store bokstaver, f.eks. \$KONSUM. For mer detaljert forklaring av stammekatalogenes plassering i filsystemet og bruken av miljøvariabler, se kapittel 5.
- Alle filer som tilhører et bruksområde skal samles under bruksområdets stammekatalog.

4.3.4. Substamme

En substamme er en underkatalog som kan romme en særskilt bearbeidingsrutine med tilhørende data, program, arbeidsdata, kodekataloger, osv, i sine respektive underkataloger innen et bruksområde.

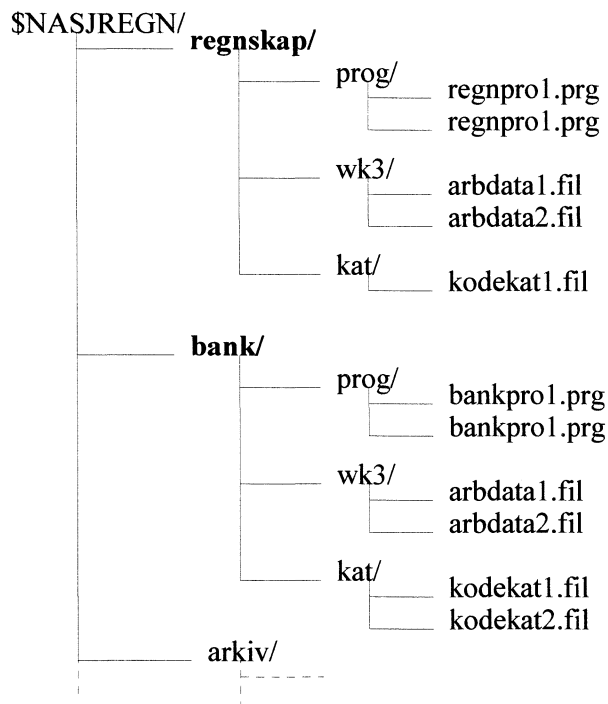
Substammer kan brukes:

- 1) For bruksområder med spesielle behov for å organisere *særskilte bearbeidingsrutiner* atskilt fra den sentrale bearbeidingsrutinen for et bruksområde
Eksempel:
Til nasjonalregnskapet hentes det inn data fra et meget stort antall andre bruksområder i SSB. Dataene fra disse bruksområdene trenger harmonisering og omredigering før de tas inn i de sentrale filene og rutinene for nasjonalregnskapet. Innen stammen for nasjonalregnskapet kan det derfor opprettes særskilte underkataloger for slike bearbeidingsrutiner.
- 2) Hvor det er hensiktsmessig at flere nært beslektede bruksområder kan samles under en overordnet stamme med felles adkomstregulering. Dette gjelder f.eks. hvor man ved en "normal" organisering av stammer etter bruksområder får et stort antall stammer som resulterer i at vi får et for stort antall "UNIX-grupper" som brukerne må være medlemmer av.

Underkatalogene for substammer betegnes med *fritt valgte brukerdefinerte og beskrivende unike navn innen en bestemt stamme.*

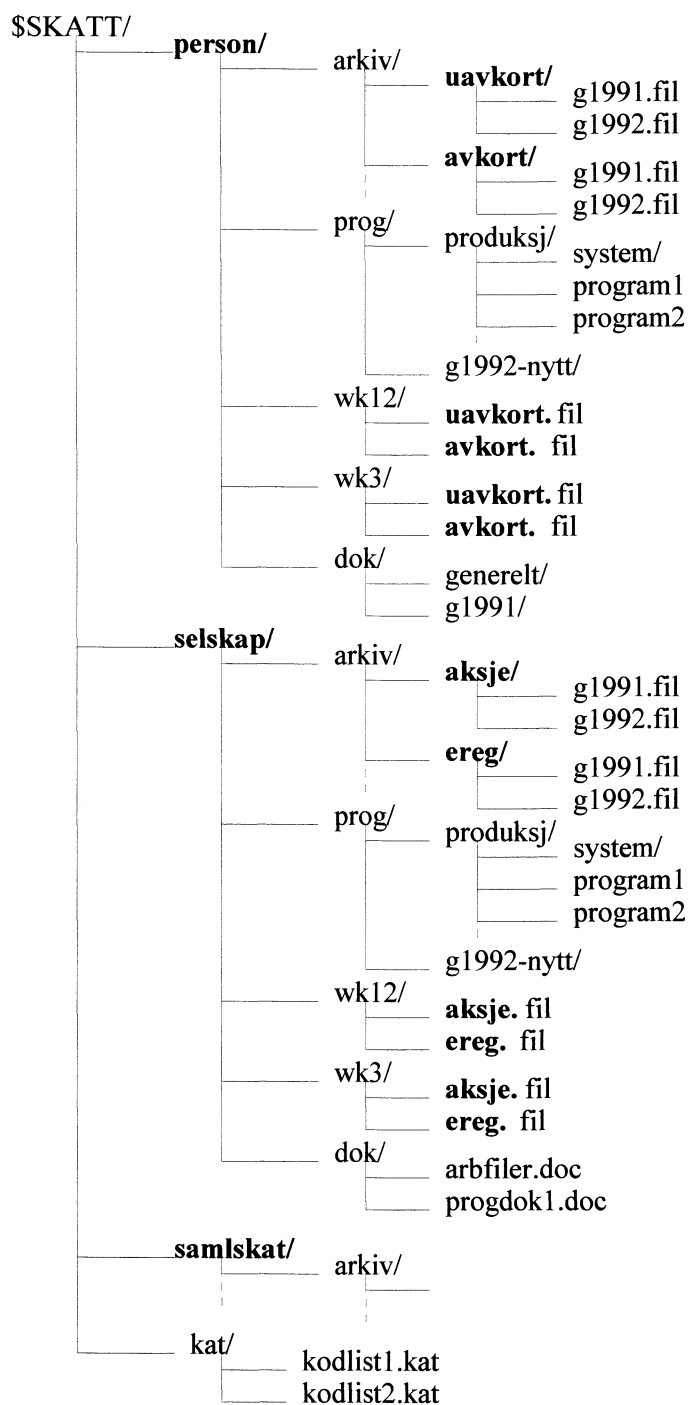
Alle faste underkataloger kan knyttes til en substamme. I forbindelse med fast underkatalog "arkiv" kan substamme kun forekomme en gang, mens sammen med de andre faste underkatalogene kan det benyttes flere substammer i samme filstreng. Vi anmoder dog om en gjennomtenkt og restriktiv bruk.

Eksempel på bruk av substammer basert på behov som nevnt i pkt. 1) over:



Her er **.../regnskap/...** og **.../bank/...** substammer for behandling (*uten langtidslagring*) av data hentet fra regnskaps- og bankstatistikken før harmoniserte data fra disse og andre kilder tas inn i de sentrale datafilene for nasjonalregnskapet (*med langtidslagring* under \$NASJREGN/arkiv/...) før videre bearbeiding (navnene **regnskap** og **bank** er konstruerte eksempler).

Eksempel på substammer basert på behov som nevnt i pkt. 2) over:



Her er **person**, **selskap** og **samlskat** substammer som rommer bearbeidingsrutiner for bruksområder som er samlet under en felles stammekatalog, \$SKATT (katalog- og filnavnene er konstruerte). Underkatalogen .../kat/... under stammen rommer kodekataloger som er felles for alle bruksområdene under stammekatalogen. De er lagt lokalt under stammekatalogen (istedenfor under stammen \$FELLES) siden de er spesifikke for disse bruksområdene.

4.3.5. Fast underkatalog ../arkiv/...

Uansett rekkefølge i katalogstruktur, gjelder følgende grunnleggende prinsipper :

Alle filer med produksjonsdata (unntatt arbeidsdata og rådata som ikke skal langtidslagres) tilhører en filklasse og skal samles:

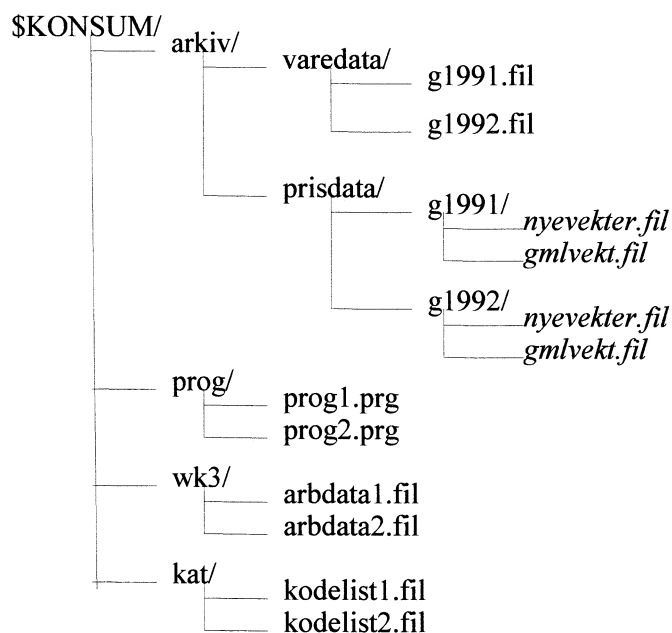
i den *faste* underkatalogen ../arkiv/... .

Dette gjelder også for bruksområder med *referansedata* og *spredningsdata*.

Filer med denne typen data skal være dokumentert i DataDok.

NB! Ikke dokumenterte filer vil bli slettet etter resultatløs purring på dokumentasjon.
 Dette gjelder også filer med navn som ikke følger standarden.

Eksempel på en katalog-/filstruktur med bruk av underkatalogen ../arkiv/... for å samle produksjonsdata som skal langtidslagres:



I eksemplet viser vi en struktur hvor vi ønsker å langtidslagre *prisdata* med både nye og gamle vekter.

4.3.5.1. Filklasse

Filklasse er navnet på fil eller en gruppe filer innen et bruksområde hvor dataene har lik **struktur**, eller over **ulike generasjoner**, har **tilnærmet lik struktur**:

En utdyping:

- *Lik, eller tilnærmet lik, struktur betyr her at filene har felles **enhetstype(r)** og **grovt sett de samme kjennmerkene** beskrevet i den samme post- eller filbeskrivelsen⁴ eller generasjonsbetingede **revisjoner** av filbeskrivelsen.*
- *For tradisjonelle periodiske undersøkelser vil én filklasse dermed omfatte filer hvor **filbeskrivelsen** for enkelte **generasjoner** kan være **revidert** med:*
 - *nye kjennemerker,*
 - *endring av formatet på kjennemerker (eks. lengden økt fra 5 til 7 posisjoner),*
 - *bortfall av kjennemerker eller*
 - *endret rekkefølge på kjennmerkene.*

Filklassenavnet skal være beskrivende for sitt bruksområde.

*Filer som slik sett har lik eller tilnærmet lik struktur, kan dermed **dokumenteres** under ett som én **filklasse** med én **felles filbeskrivelse**, eventuelt med tillegg av **generasjonsbetingede revisjoner** av filbeskrivelsen. Dokumentasjonen skal skje i DataDok.*

Et eksempel på etablering av filklasser: I skattestatistikken er det for "selskapsligningen" for hvert år lagret, én fil som rommer en kopi av skattedirektoratets administrative "aksjeregister" og en annen fil som rommer skatteregisteret for alle "etterskuddspliktige skattytere" for det aktuelle året. Hver fil har for hvert år (i) samme enhetstype og (ii) tilnærmet lik struktur over tid; dvs *hvilke kjennemerker* filene rommer, kan variere noe fra generasjon til generasjon. I dette tilfellet kan vi definere en filklasse med navnet "**aksje**" og en annen med navnet "**ereg**" (se avsnitt 4.3.4, figuren over substammene person, **selskap** og samlskat).

- Hvis det fra en generasjon til en annen, lagres et nytt kjennemerke i generasjonsfilene for en av disse filklassene (f.eks. **aksje**), skal det registreres en ny revisjon av "Filbeskrivelsen" (alt ellers likt). Hver revisjon av filbeskrivelsen for en filklasse knyttes til sine respektive generasjoner. Filene for filklassen **aksje** vil da ha *tilnærmet lik struktur* på tvers av dette generasjonsskiftet.

Et annet eksempel: La oss tenke oss at vi i befolkningsstatistikken for hver generasjon (år) har en fil med statistikkdata for "folkemengden pr. 1/1 pr år" under filklassenavnet **folkmeng**.

- I forbindelse med 6-årsreformen i skoleverket vil vi anta at det fra og med en bestemt generasjon av filklassen ble lagt inn et nytt kjennemerke for å romme en ny aldersgruppering (alt ellers likt). Det måtte da defineres en ny revisjon av filbeskrivelsen for filklassen **folkmeng** for å dekke generasjonsfilene med den nye aldersgrupperingen.

⁴ Én filklasse kan i noen tilfelle også romme én eller flere filer med **ulike recordtyper**, hvor hver posttype kan dekke:

- *hver sin enhetstype med hvert sitt sett av kjennemerker eller*
- *alternative redigeringer av format og/eller rekkefølge på kjennmerkene.*

I filer med flere recordtyper dokumenteres hver recordtype som en egen **subklasse** av den grunnleggende filklassen.

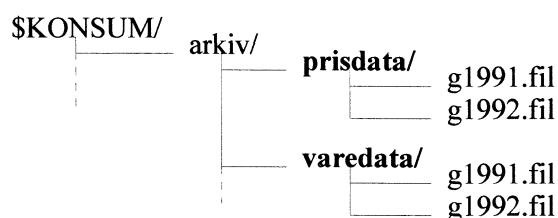
- Det er under etablering en ny regional standard "REGIN4" som skal erstatte den gamle standarden for handelsdistrikt. Fra og med den generasjonen hvor det skal publiseres data etter REGIN4 istedenfor handelsdistrikt, vil "handelsdistrikt" bli erstattet med "REGIN4" (alt ellers likt). En ny revisjon av filbeskrivelsen for **folkmeng** må lages for de generasjonene hvor REGIN4 erstatter handelsdistrikt.

Filene for filklassen **folkmeng** vil da ha *tilnærmet lik struktur* på tvers av begge generasjonsskifter; det første hvor en ny aldersgruppe ble etablert, det andre hvor handelsdistrikt ble erstattet av REGIN4. Med begge endringene over ett generasjonsskifte ville vi gjort endringen med kun en revisjon av filbeskrivelsen.

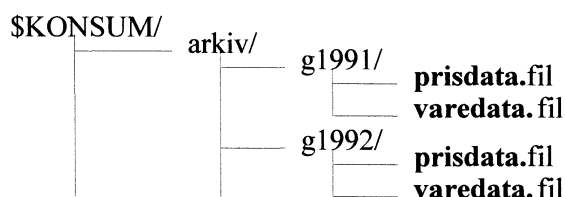
I de fleste eksemplene som er brukt som filklassenavn i disse retningslinjene, er det brukt konstruerte navn i et tenkt bruksområde. Filklassenavnene i eksemplene er **prisdata** og **varedata** (se også avsnitt 4.3.2).

I filsystemet på UNIX kan en filklasse:

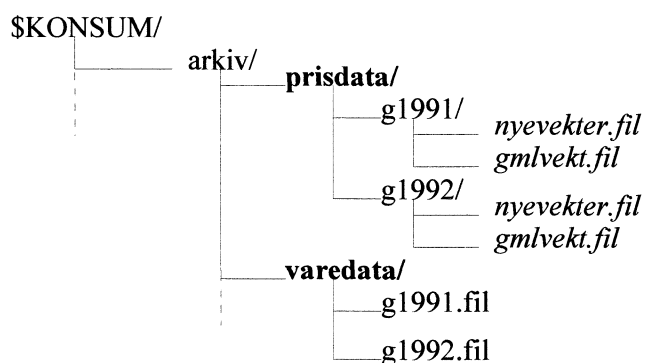
- danne en underkatalog som ligger under den faste underkatalogen .../arkiv/...,



- eller være representert i form av et filnavn ,



NB! Det følgende alternativet er aktuelt når vi trenger å langtidslagre mer enn én enkelt versjon for hver filklasse for hver generasjon.



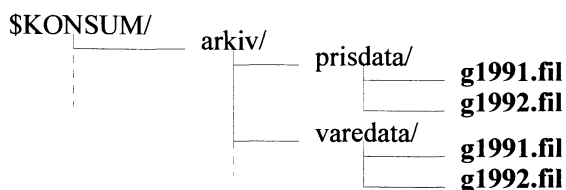
For alle filer med *produksjonsdata*⁵ som skal langtidslagres, skal filklassene være registrert og dokumentert i *DataDok*. For filer med *arbeidsdata* er det ikke nødvendig å registrere filklassene i dokumentasjonssystemet; de skal lagres under særskilte underkataloger for arbeidsdata (se avsnitt 4.3.6). Det antas å være tilstrekkelig at slike filer er dokumentert i bearbeidingsrutinene for bruksområdet.

4.3.5.2. Generasjon

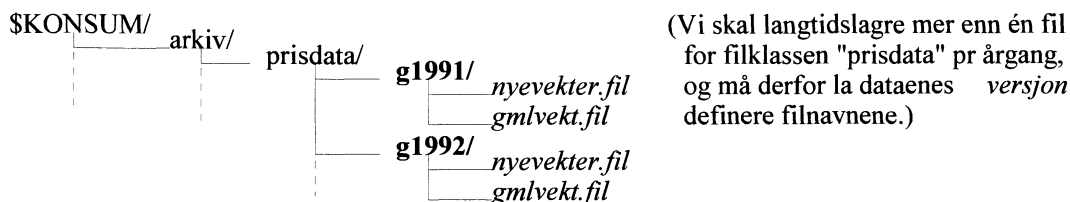
De fleste statistikkrutinene er periodiske. Tids- eller periodeangivelser kan spesifiseres ved å bruke standardbetegnelser for generasjon. Generasjonsbetegnelsene kan definere *underkataloger* eller *filnavn* etter behovene for bearbeiding av dataene på det aktuelle bruksområdet.

NB! Generasjonsbetegnelsen skal vise dataenes *registreringstidspunkt* eller *registreringsperiode* (dette gjelder uansett når dataene behandles og lagres eller når et eventuelt situasjonsuttak er utført!).

Generasjonsbetegnelsen kan for arkivdata være representert i filnavnet:



Et annet eksempel er at generasjon er underkatalog mellom filklasse og versjon(=filnavn):



Generasjonsbetegnelsen skal *som standard* kodes med en innledende **g** (unntak **f** ved framskrivninger og **gs** for å angi startpunkt) og alfanumeriske koder som identifiserer hvilken periode/tidsrom/tidspunkt som dataene gjelder for. Årstall skal alltid skrives med fire tall uansett sammenheng. Eks. på koding av generasjon:

- | | |
|---------------------|---|
| • Årlig fil | g2001 = årgang 2001 |
| • Halvårlig | g2000h2 = år 2000, 2. halvår (h[1-2]) |
| • År/kvartal | g2001k1 = år 2001, kvartal 1 (k[1-4]) |
| • År/tertiel | g2000t2 = år 2000, tertial 2 (t[1-3]) |
| • År/bimånedlig | g1999b2 = år 1999, 2. to-månedts termin (b[1-6]) |
| • År/md | g1999m09 = år 1999, måned 09 (m[01-12]) |
| • År/uke | g1999u04 = år 1999, uke 04 (u[01-53]) |
| • Dagnummer i et år | g2000d161 = år 2000, dag 161 (d[01-366]) |
| • År/md/dag | g1999m09d25 = år 1999, måned 09, dag 25 (m[01-12]d[01-31]) |

⁵ Omfatter rådata, arbeidsdata, statistikkdata, tverrsektorielle mikrodata og arkivdata (se def. i pkt. 2)

- Flere årganger i én fil **g1998g2000** = årgang 1998 - 2000
- Fra år n, månedlige utvidelser **gs1999m04** = **gs** : filens datainnhold g starter april 1999
- Fra år n, kvartalsv. utvidelser **gs1999k1** = **gs** : filens datainnhold g starter 1 kv. 1999
- Fra år n med framskriving **g2000f2005** = fil med basisdata for år **2000** framskrevet til år **2005**
- Kombinasjon av spesifikke årsoppdelinger over flere år kan også brukes:
F.eks. **g1999b4g2000b2** = fom 4.termin 1999 tom 2.termin 2000
g2001m01f2003m12 = fom jan 2001, framskrivninger tom desember 2003.

Prefiksene **g**åååå, **gs**åååå og **f**åååå (hvor åååå står for år) er reserverte betegnelser i navnestandarden og må ikke brukes på annen måte i en katalog-/filstruktur.

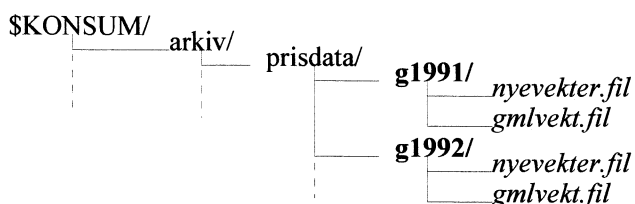
Kontakt SDA for evt. utvidelse av listen med standardkoder for generasjon.

4.3.5.3. Versjon

Versjon er et begrep som kan brukes for å skille mellom f.eks. (i) flere filer i samme filklasse har ulike beregningsgrunnlag (befolkningsframskrivninger med ulike vekter), (ii) ulike "foredlingsnivå" på filer (rådatafiler kontra ferdig reviderte filer) eller (iii) ulike seleksjoner av poster fra en "samlefil" (eks. bestemte filer med utvalgte kommuner selektert fra en "kommunefil"). Versjon kan bare benyttes som filnavn, og må være dokumentert i DataDok.

Ytterligere begreper som kunne vært benyttet i en navnestandard for identifisering av datafiler i SSB er utelatt. Det kan gjelde begreper som **puljer** av data ved f.eks. innsamling av data til undersøkelser med svært store datamengder:

Slike begreper er ikke en del av navnestandarden for identifikasjon av data. Et skille mellom denne type begreper må spesifiseres i filnavnet. Innenfor navnestandarden (og i DataDok) vil det derfor oppfattes som et skille mellom versjoner og må dokumenteres som sådan. Forutsetning er at de aktuelle filene tilhører samme filklasse; dvs. at de har samme filbeskrivelse innen samme generasjon. Filnavnene kan velges fritt av brukeren, men må ikke kollidere med andre katalog- eller filnavn som blir brukt i den aktuelle delen av katalogstrukturen eller andre faste betegnelser (dvs. katalognavn) som er omtalt ellers i disse retningslinjene. Eksempel:



(Vi skal langtidslagre mer enn én fil for filklassen "prisdata" pr årgang, og må derfor la dataenes *versjon* definere filnavnene.)

4.3.6. Øvrige faste underkataloger

4.3.6.1. Fast underkatalog ../work/...

Filer med *arbeidsdata* og andre midlertidige datafiler skal legges i særskilte underkataloger med faste navn for arbeidsdata. Disse underkatalogene skal betegnes med en egen kode:

../wknn/...

hvor **wknn** inneholder informasjon om standard lagringstid før automatisk sletting av de underliggende filene i den aktuelle underkatalogen. Eksempler på koder som kan benyttes:

../wk00/..	arbeidsfiler som slettes automatisk i løpet av det første døgnet
../wk01/..	arbeidsfiler som slettes automatisk etter 1 måned
..	
../wk24/..	arbeidsfiler som slettes automatisk etter 24 måneder <i>NB! 24 måneder er maksimum tillatt lagringstid!</i>

Unntak:

../wk99/..	arbeidsfiler som ikke kan tilpasses innenfor den normale maksimumsgrense på 24 måneder. Behovet for lagring i 99 måneder skal framgå av systemdokumentasjon som ligger på q:\dok\prosjekt\ITanalyse\..., og skal godkjennes av SDA i hvert tilfelle. "wk99" blir i kontrollprogrammet knytta opp mot de stammene som har fått tillatelse til å bruke denne faste underkatalogen. I praksis betyr dette at den som har behov for "wk99", må sende en mail til S710-SDA mailboks. Mailen skal inneholde stamme (ev. også substamme) og vise til hvor behovet er dokumentert.
------------	---

NB! *Slettekriterium* skal være **n** måneder etter *siste gangs bruk av filen*.

Eksempler:.

\$STAMME/wk00/.../filnavn.fil	<i>ingen lagringstid</i> , filen kan slettes umiddelbart
\$STAMME/wk01/.../filnavn.fil	filen skal slettes 1 mnd etter siste gangs bruk
\$STAMME/substamme/..wk12/.../filnavn.fil	filen skal slettes 12 mnd etter siste gangs bruk

4.3.6.2. Flere faste underkataloger

Vi skal også kunne lagre både program, kataloger osv. Disse faste underkataloger har ingen krav til den øvrige navnsettingen annet enn at filstrengen alltid må inneholde kun en fast underkatalog.

- ../kat/.. underkatalog for kodekataloger som bare brukes innen den aktuelle stammen

- **../prog/..** underkatalog for programmer til verktøy som ikke har sine egne "proprietære" systembibliotek og parameterfiler for program o.l. Underkatalogen kan bl.a. struktureres med henblikk på applikasjoner som hver kan skilles i en testversjon og en produksjonsversjon.
- **../incl/..** underkatalog for programmer som er ferdig kompilert og linket, f.eks. i SAS-Frame.
- **../famedb/..** underkatalog for famedatabaser
- **../trollldb/..** underkatalog for trolldatabaser
- **../sasdb/..** underkatalog for sasdatabaser. Den skal kun benyttes til sasfiler som er under kontinuerlig oppdatering dvs. at de kan defineres som en database (f.eks. istedenfor Oracle). Den skal ikke benyttes til langtidslagring for permanente sasfiler. Behovet for "sasdb" skal framgå av systemdokumentasjon som ligger på q:\dok\prosjekt\ITanalyse\..., og skal godkjennes av SDA i hvert tilfelle. "sasdb" blir i kontrollprogrammet knytta opp mot de stammene som har fått tillatelse til å bruke denne faste underkatalogen. I praksis betyr dette at den som har behov for "sasdb", må sende mail til S710-SDA mailboks. Mailen skal inneholde stamme (ev. substamme) og vise til hvor behovet er dokumentert.
- **../dok/..** underkatalog for faktadokumentasjon eller henvisning til systemdokumentasjon. F.eks. for opptellinger av ulike slag, kvitteringer for hva som er lagt ut på web eller filbeskrivelser for filer som ikke skal lagres i DataDok (ikke "arkiv"-filer). Men det er viktig å merke seg at all systemdokumentasjon slik som spesifikasjoner, datamodeller, projektskriv osv., skal lagres på NT under q:\dok<stamme>\ . Ref. "Håndbok i utvikling av statistikkssystemer".

4.3.7. Filhale

Skal starte med "." (punktum) og etterfølges av fritt valgt antall og type standardkarakterer. Noen programmer tildeler filhale automatisk, slik at det enkelt å se hvilken type fil det er. Det er ikke pålagt å bruke filhale, men vi vil likevel oppfordre til å benytte det.

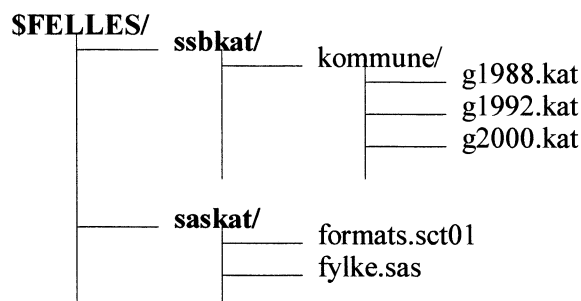
4.4. Kodekataloger for standard klassifikasjons- og andre grupperingskjennemerker

De kodekatalogene som er felles for flere bruksområder skal være lagret i et sentralt lagringsområde som er felles for hele SSB. Dette lagres på en egen katalog; **\$FELLES**. Kodekatalogene skal der legges under faste underkataloger som:

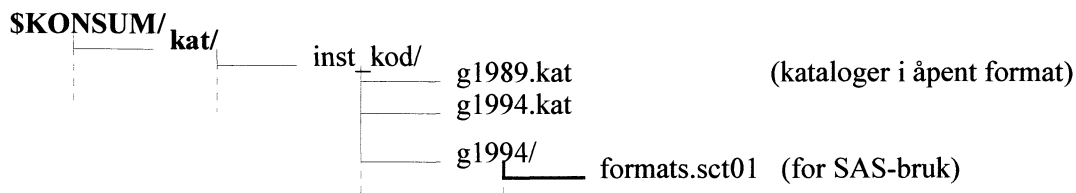
../ssbkat/... for kodekataloger med "*åpent*" format
../saskat/... for kodekataloger i *SAS-format*

Filer med kodekataloger *for SAS-bruk* vil, som for andre filer i SAS interne format, være kjennetegnet med spesielle filhale håndtert av SAS.

Eksempel på katalogtre for kodekataloger i åpent format (under ../ssbkat/...) og i SAS-format med et formatbibliotek og en fylkeskatalog for generell bruk:



De kodekatalogene som bare brukes innen ett enkelt bruksområde, skal legges under den spesielle underkatalogen `.../kat/...` under det aktuelle bruksområdets stamme. I tillegg kommer navn som viser hvilken kodekatalog det er tale om og eventuelt betegnelser for generasjon og revisjon. Strukturen blir med andre ord som for filer med data. Eksempel:



Under katalogen \$FELLES er det også satt av plass for generelle subprogram for felles bruk. Underkatalogen `.../sasprog/...` skal dekke generelle subprogram/rutiner for bruk i SAS.

Eksempel: `$FELLES/sasprog/fnr sjekk.sas`

"fnr sjekk" er en generell rutine for kontroll av fødselsnummer.

4.5. Regler for bruk av data fra andre bruksområder

I mange tilfelle vil data fra et statistikkområde bli brukt innenfor andre bruksområder enn det området hvor de oppstår og bearbeides.

For at dataeieren (ansvarlig seksjon) skal kunne ha full kontroll over de data som tilhører et bruksområde, skal slike data alltid være lagret under den opprinnelige stammen. I de tilfelle hvor dataene skal brukes i et annet bruksområde (enten det gjelder et statistikk-, analyse- eller et forskningsprosjekt) skal **leseadkomst gis til dataene der de ligger under den opprinnelige stammen** via tidsbegrenset adkomst til filen(e).

Dataene skal ikke kopieres uendret over til andre bruksområders stammekataloger.

Bearbeidede utdrag av dataene, evt. etter kopling med andre data, lagres normalt under stammekatalogen til det andre bruksområdet. Ansvarlig eier av disse dataene blir da seksjonen som er ansvarlig for dette bruksområdet

5. Hvordan stammekatalogene er knyttet til filsystemet på UNIX

Stammen definerer en hovedkatalog i filsystemet på UNIX. Denne hovedkatalogen vil alltid være relativ til overliggende kataloger i filsystemene på de enkelte maskinene i nettet. Navnene til disse overliggende katalogene er bestemt av maskin- og diskplassering. De fullstendige filnavnene vil derfor bli svært lange og sammensatte. For at man skal få så enkel adkomst som mulig til filene under stammene, blir det opprettet "*miljøvariabler*" som peker på stammene.

Dersom det viser seg å være behov for omorganisering av filsystemet, kan Seksjon for IT-drift når som helst, og uten forvarsel, flytte et lagringsområde (stammekatalogen med alle underliggende kataloger og filer) til en annen plass i filsystemet. Dette vil endre de fullstendige navnene til stammekatalogen og dens underkataloger og filer.

*Brukersystemene må ikke bli påvirket av slike endringer, derfor skal **all bruk av filene i stammene være via stammenes respektive miljøvariabler, ikke via de overliggende katalogene.***

Eksempel på miljøvariabler og sammenhengen med filsystemet:

KONSUM /ssb/castor/a1/konsum

Ved definisjon som miljøvariabel gis KONSUM innholdet: /ssb/castor/a1/konsum.

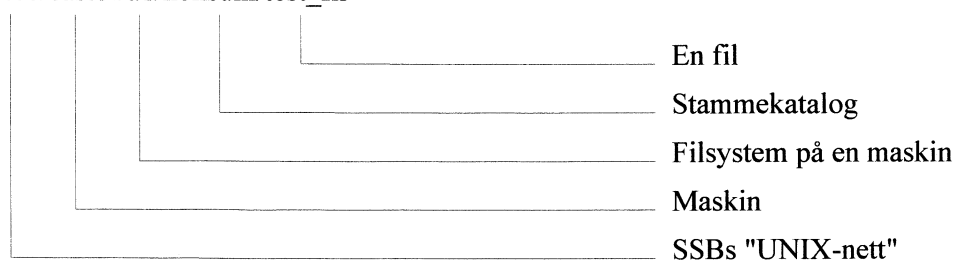
Vi refererer til filer under stammen via miljøvariabelen ved å bruke prefikset \$ foran miljøvariabelen.

Eks: \$KONSUM/test_fil



Dette er maskinens tolking:

/ssb/castor/a1/konsum/test_fil



Alle stammekataloger som etableres på UNIX, vil bli lagt direkte under et "UNIX-filsystem" som i det "ordinære platelageret" vil ha navn som:

an eller ann.. (hvor **a** er *konstant* og **n** er siffer fra 0 - 9)

Alle "hjemmekataloger" blir lagt til et "UNIX-filsystem" som regel vil ha navn som:

hn eller hnn.. (hvor **h** er *konstant* og **n** er siffer fra 0 - 9)

Det er etablert konvensjon innen "UNIX-miljøene" at **miljøvariabler skal ha STORE bokstaver**. Denne konvensjonen er derfor etablert som **en del av SSBs navnestandard**.

Definisjon av stammekataloger, miljøvariabler og brukergrupper utføres av Kundestøtte som også sørger for at aktuelle miljøvariabler er definert for autoriserte brukere.

6. Program for kontroll av navnestandarden.

Programmet "navnstd" kontrollerer fil- og katalognavn mot Unix navnestandard. Ulike kontroller brukes avhengig av hvor mye av filstrengen som skal sjekkes.

Alle kan kjøre dette programmet på stammer de har lesetillatelse til. Programmet startes med kommandoen **navnstd** og deretter stammenavn og eventuelt deler av filstreng som skal kontrolleres.

Eks. på bruk:

navnstd (ved å taste kun navnstd får man opp en kortfattet brukerveiledning av programmet).

navnstd \$STAMME (sjekker alle kataloger/filer på stammen, og lister ut alle feil på skjermen)

navnstd \$STAMME -ok (her får man listet ut alle filer/kataloger, også de som er riktige.)

navnstd \$STAMME/arkiv (her blir kun arkiv-katalogen sjekket)

6.1. Eksempel på navnestandardkontrollprogram.

Eksempel på testutkjøring etter kontroll ifølge versjon 4. Det er laget en egen rutine på Unix for å sjekke at filtrestrukturen stemmer overens med navnestandarden. Den startes med kommando "navnstd".

```
> navnstd $DATADOK/ -ok
```

```
-----
Rapport fra program navnstd, 2001-03-02T11:07:36, kjørt av 'root':
Kontrollerer navn på startpunkt og underliggende filtreobjekter mot Unix navnestandard.
NB! Kan ta noen sekunder avh. av antall som skal sjekkes. Kan avbrytes m. 'ctrlC'
-----
```

```
Resultat av navnestandardkontroll v. 3.1 under stamme: $DATADOK ('/ssb/ursus/a1') [710].
```

Eier	Startpunkt	Status	Std.	DataDok type
root	'/datadok/'	OK	Arkiv 1	. Stamme
hkt	'/datadok/navnstd_test/#TESTFeilDir/'	Feil		
	- Ugyldig karakter brukt i navn. ('#')			
hkt	'/datadok/navnstd_test/#TESTFeilDir/MereFeil.Fil###'	Feil		
	- Ugyldig/mangler fast underkatalog. (Gyldige valg: /arkiv /dok /famedb /incl /kat /prog /trolldb /wk[0-24,99])			
	- Ugyldig karakter brukt i navn. ('#')			
hkt	'/datadok/navnstd_test/'	OK	Arkiv 2a	. Substamme arkiv
hkt	'/datadok/navnstd_test/dok/'	OK	Fri 3b1.	Dokumentasjon (../dok/)
hkt	'/datadok/navnstd_test/dok/testfil.txt'	OK	Fri 3b3.	Dokumentfil (../dok/../dokfil)
hkt	'/datadok/navnstd_test/famedb/'	OK	Fri 3c1.	FAME db. (../famedb/)
hkt	'/datadok/navnstd_test/FeilTestFil.Faenls'	Feil		
	- Ugyldig/mangler fast underkatalog. (Gyldige valg: /arkiv /dok /famedb /incl /kat /prog /trolldb /wk[0-24,99])			
hkt	'/datadok/navnstd_test/sasdb/'	OK	Fri 2b	. Substamme
hkt	'/datadok/navnstd_test/trolldb/'	OK	Fri 3g1.	TROLL db. (../trolldb/)
hkt	'/datadok/navnstd_test/wk99/'	OK	Fri 3h1.	Arbeidsdata (../wkNN/)
hkt	'/datadok/sbskatt7'	OK	Arkiv 2a	. Substamme arkiv
hkt	'/datadok/sbskatt/arkiv/'	OK	Arkiv 3a1.	Arkiv (../arkiv/)
hkt	'/datadok/subskatt/'	OK	Arkiv 2a	. Substamme arkiv
hkt	'/datadok/subskatt/arkiv/'	OK	Arkiv 3a1.	Arkiv (../arkiv/)

2 Datalagring på UNIX i Statistisk sentralbyrå

Melding:

```
-----  
|  
| NB! Stammen følges opp automatisert hver natt i forhold til navnestandarden.  
|  
-----
```

Sluttrapport

```
Resultat av navnestandardkontroll under '/ssb/ursus/al/datadok/'  
Antall katalognavn kontrollert .....: 12  
  Antall OK .....: 11  
  Antall feil .....: 1  
Antall filnavn kontrollert .....: 3  
  Antall OK .....: 1  
  Antall feil .....: 2  
Antall symbolske linker kontrollert ..: 0  
  Antall OK .....:  
  Antall feil .....:  
Lapse tid ..: 2.00 sek.  
Brukertid ..: 0.10 sek.  
Systemtid ..: 0.03 sek.  
-----
```

De sist utgitte publikasjonene i serien Statistisk sentralbyrås håndbøker

- 45 Håndbok i datasikkerhet og fysisk sikring. Revidert utgave, november 1998. 1998. 83s.
- 46 Telefonkatalog. 1998. 89s.
- 47 EØS-avtalen. Det statistiske samarbeid og konsekvenser for Statistisk sentralbyrås statistikkproduksjon. 1994. 55s.
- 48 Håndbok i tilsetningssaker. 1994. 32s.
- 49 Oppgaveplikt og tvangsmulkt. 1995. 55s.
- 50 Emneinndeling 1995. 1995. 43s.
- 51 Intervju: EDB-arbeidsbok. 1995.
- 52 Intervju: EDB-oppslagsbok. 1995.
- 53 Intervju: Opplæring og administrasjon. 1995.
- 54 Internkontroll: Revidert utgave 1997. 25s.
- 55 Nordisk statistikk på CD-ROM: Veiledning. 20s.
- 56 PC-Axis versjon 2.2: Brukerhåndbok. 69s.
- 57 Produktregister versjon 4.0: Brukerveiledning. 49s.
- 58 Håndbok i prosjektstyring. 20s.
- 59 Personalreglement for Statistisk sentralbyrå. 22s.
- 60 Produktnummerkatalog pr. 28.02.1996. 55s.
- 61 Innkjøpshåndbok. 1996.
- 62 Timeplan versjon 3.0: Brukerveiledning. 16s.
- 63 Håndbok i EDB-metode. 52s.
- 64 Publiseringshåndbok: Regler og retningslinjer for publisering i Statistisk sentralbyrå. 93s.
- 65 Håndbok i utvikling av statistikksystemer: Med vekt på IT-metode. 52s.
- 66 Håndbok i datarevisjon. 48s.
- 67 Arkivnøkkel for Statistisk sentralbyrå. 76s.
- 68 Rapporteringshåndbok for KOSTRA-regnskap 1999: Oppslagshefte til hjelp ved filuttrekk for KOSTRA-rapportering. 52s.
- 69 Yrkeskatalog for innrapportering av yrke til arbeidstakerregisteret. 86s.
- 70 Håndbok for KOSTRA-rapportering 2000. Oppslagshefte til hjelp ved filuttrekk for KOSTRA-rapportering. 74s.
- 71 Håndbok i SAS. Del 2: Oppslag. 243s.
- 72 Yrkeskatalog pr. november 2000. Korrigert utgave. 172s.
- 73 Håndbok i SAS. Del 1: Innføring. 65s.

B Returadresse:
Statistisk sentralbyrå
N-2225 Kongsvinger

Statistisk sentralbyrå

Oslo:
Postboks 8131 Dep.
0033 Oslo

Telefon: 21 09 00 00
Telefaks: 21 09 49 73

Kongsvinger:
2225 Kongsvinger

Telefon: 62 88 50 00
Telefaks: 62 88 50 30



Statistisk sentralbyrå
Statistics Norway