

Grete Dahl

FD - Trygd

**Type prosjekt, organisering,
forløpsdata og analysemuligheter**

Forord

FD-Trygd (Forløpsdatabasen-Trygd) ble åpnet 1. februar 1999.

I dette notatet er omtalt arbeidet med å utvikle FD-Trygd. Det er gjort rede for organiseringen av og formålet med prosjektet, hvilke data basen inneholder og hvordan de statistiske opplysningene er organisert som forløpsdata. Helt kort er også nevnt de foreliggende planer for utvidelse av prosjektet.

Notatet er skrevet av prosjektleder Grete Dahl.

Statistisk sentralbyrå, april 1999
Johan-Kristian Tønder
fagdirektør

Innhold

1. Type prosjekt	3
2. Historikk og videre planer	4
2.1 Historikk.....	4
2.2 Videre planer.....	6
3. Organiseringen av prosjektet	7
3.1 Styringsgruppa	7
3.2 Prosjektleder og prosjektgruppa.....	7
3.3 Referansegruppa	8
3.4 Brukertjenesten.....	9
4. Formål med prosjektet.....	10
4.1 Tilrettelegge data som er egnet til studier av trygdekarrierer og til å belyse rekruttering til - avgang fra - og overgang mellom velferdsordninger, og mellom velferdsordninger, offentlige tjenestepensjonsordninger og arbeidsmarkedet.....	10
4.2 Gi statistisk merinformasjon til surveydata, input til modellarbeid etc.	12
4.3 Definere og tilrettelegge populasjoner for å trekke utvalg til statistiske undersøkelser.....	12
4.4 Bygge opp en trygdemodul til Fob2000.....	13
5. Data og dataorganiseringen	13
5.1 Administrative data	13
5.2 Typer data	14
5.3 Gjenbruk av data.....	14
5.4 Dataorganiseringen	16
6. Dokumentasjon og informasjon	17
Vedlegg 1. Mandat for styringsgruppa.....	19
Vedlegg 2. Mandat for prosjektleder.....	20
Vedlegg 3. Mandat for referansegruppa.....	21
Vedlegg 4. Sluttrapport fra forprosjekt (inkl. prosjektskriv).....	22

1. Type prosjekt

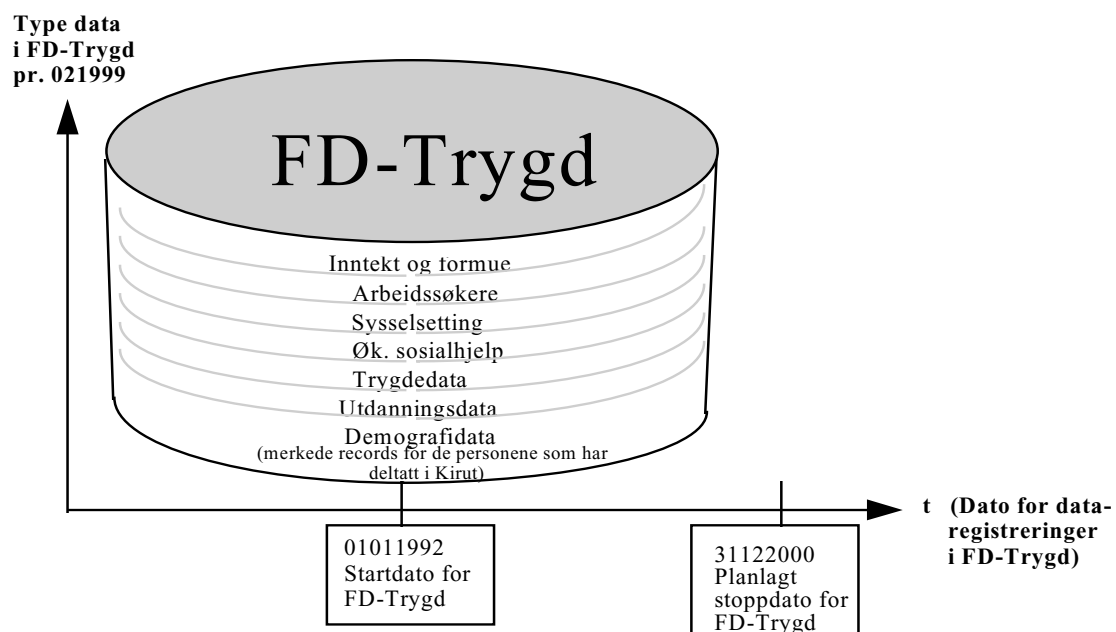
FD-Trygd er en *forløpsdatabase* med opplysninger om *trygdeforhold* og *andre typer statistiske data*. *Individ* er statistisk enhet, og opplysningene i databasen består av registreringer av begivenheter i den enkelte persons livsløp. Disse registreringene kan settes sammen til individuelle hendelses- eller forløpshistorier (*forløpsdata*), og aggregert over en gruppe av individer eller en hel populasjon, gir forløpsdataene oversikt over bevegelsene *mellom suksessive* begivenheter eller *tilstandsendringer* (conf. *forløpsanalyse*). I forløpsdataene i FD-Trygd er eksempler på tilstander at personen "mottar sykepenger", "mottar uførepensjon" etc.

De viktigste typer forløpsdata i FD-Trygd er opplysninger om forskjellige typer pensjoner, som uførepensjon, avtalefestet pensjon etc., og opplysninger om folketrygdytelser som vanligvis er av mer kortvarig karakter, som for eksempel overgangsstønad, stønad til barnetilsyn, fødsels- og sykepenger og attføringspenger. I tillegg inneholder basen forløpsdata som er basert på opplysninger fra en del andre ordninger eller temaområder. Disse områdene er demografi, utdanning, økonomisk sosialhjelp, arbeidssøking og sysselsetting. Det er også planer om å utvide FD-Trygd til å omfatte offentlige pensjonsordninger utenom folketrygden og noen andre, mindre ordninger. En del opplysninger i FD-Trygd er registrert som tverrsnittsdata. Dette gjelder blant annet opplysningene om inntekt og formue.

Dataene i FD-Trygd bygger på opplysninger i *administrative registre* og statistikkregistre som er dannet på grunnlag av registerdata. Basen inneholder opplysninger *fra og med 1992* og skal, ifølge planene, bygges ut til å omfatte data til og med år 2000. Opplysningene gjelder *hele befolkningen*.

Figur 1, nedenfor, viser hvilke hovedtyper av data som inngår i FD-Trygd for årene 1992 og 1993 ved åpningen av basen i februar 1999.

Figur 1. Hovedtyper av data som inngår i FD-Trygd for årene 1992 og 1993 ved åpningen av basen i februar 1999



Utgiftene til oppbyggingen av FD-Trygd har så langt vært delt mellom Statistisk sentralbyrå og ekstern finansiering. Sosial- og helsedepartementet har vært *hovedsponsor* av prosjektet, men noe finansiell støtte er også gitt fra Arbeids- og administrasjonsdepartementet og Finansdepartementet.

FD-Trygd kan ses på som en videreutvikling av prosjektet Kirut (*K*lientstrømmer *i*nn, *r*undt, *u*t av trygdesystemet), som i de senere årene har vært en viktig datakilde for trygde- og sosialpolitisk forskning. Kirut er finansiert av Sosial- og helsedepartementet, men dataene "eies" av Statistisk sentralbyrå. Kirut inneholder data fra og med 1989, og Norsk Samfunnsvitenskapelig Datatjeneste, Bergen, står for utlevering av data fra dette datasystemet.

Til tross for mange likhetspunkter mellom FD-Trygd og Kirut, er det flere forhold som atskiller disse to prosjektene. Kirut omfatter bare 10 prosent av befolkningen mellom 16-67 år, mens FD-Trygd, som nevnt, omfatter hele befolkningen. Opplysningene i FD-Trygd kan derfor i større grad enn Kirut-dataene nyttes til analyser som er avgrenset til *små regionale områder* som for eksempel bydeler/kommuner, eller for *små befolkningsgrupper* som for eksempel innvandrere, vanskeligstilte ungdomsgrupper, marginaliserte grupper etc.

Det er overveiende de samme typer data som inngår i de to prosjektene med unntak for de utvidelser som er planlagt for FD-Trygd, jf. ovenfor. Sammenlignet med Kirut, er imidlertid dataene i FD-Trygd organisert på en måte som sikrer *en mer dynamisk datastruktur*. Dette gjelder spesielt for de opplysningene som er knyttet til demografiske forhold, utdanning og økonomisk sosialhjelp, og sett fra et forskningsmessig synspunkt utgjør dette kanskje det mest interessante skille mellom de to datasystemene.

Det er nedlagt mye arbeid i oppbyggingen av Kirut, og de faglig sett mange likhetspunktene mellom FD-Trygd og Kirut gjør det mulig å legge opp til en eventuell "etterbruk" av Kirutdataene sammen med opplysningene fra FD-Trygd. Det blir lagt til rette for dette ved å merke recordene i FD-Trygd for de personene som har deltatt i Kirut, jf. figur 1.

2. Historikk og videre planer

2.1 Historikk

Det omfattende statistikk-faglige og It-faglige arbeidet med å bygge opp forløpsdatabasen FD-Trygd startet ikke før tidlig i 1996. En del forberedende arbeid på prosjektet ble imidlertid gjennomført i årene 1993-1995.

Prosjektskrivet for FD-Trygd ble utarbeidet høsten 1993 og godkjent i forbindelse med virksomhetsplanleggingen i Statistisk sentralbyrå i november samme år, jf. figur 2. Prosjektskrivet er tatt inn helt til slutt i rapporten fra forprosjektet, vedlegg 4.

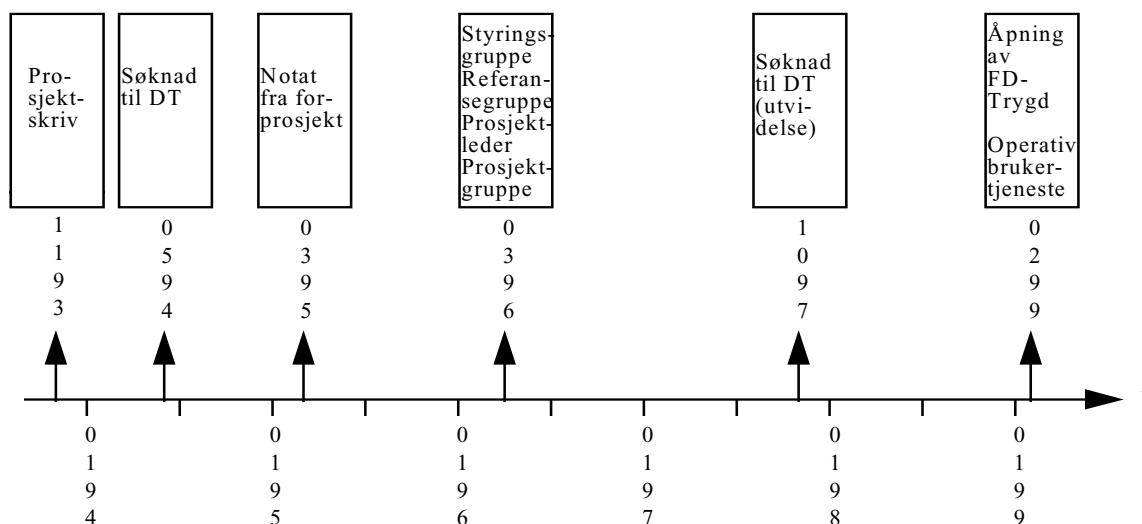
I prosjektskrivet er spesielt framhevet betydningen av å bygge opp hendelsesorienterte statistiske data på individnivå til anvendelse i trygde- og sosialpolitisk forskning. Det er også vektlagt å bruke dataene i FD-Trygd til å gi opplysninger om personer som mottar trygd til Folke- og bolig tellingen år 2000, og til å framskaffe inputdata for anvendelse i modellarbeid. Formålene med FD-Trygd er mer detaljert gjort rede for i kapittel 4.

I prosjektskrivet er fastlagt rammene for oppstart av prosjektet. Prosjektskrivet inneholder også en del detaljert informasjon, blant annet om hvilke hovedtyper av kjennemerker som skal inngå i FD-Trygd. På grunnlag av disse opplysningene ble det våren 1994 utarbeidet en *søknad til Datatilsynet*, og tillatelse til å føre dette registeret/datasystemet ble gitt i mai 1994, jf. figur 2.

I 1994 ble det i fora både i og utenfor Statistisk sentralbyrå orientert om FD-Trygd og de planene som forelå for prosjektet. Blant eksterne institusjoner som var representert i disse orienteringsmøtene var Norges forskningsråd, Norsk samfunnsvitenskapelig datatjeneste og Sosial- og helsedepartementet. Det ble også utført et arbeid med sikte på å planlegge oppbyggingen av en trygdemodul til Folke- og boligtellingsen år 2000. Dette planleggingsarbeidet er dokumentert i "Mot en registerbasert Folke- og boligtellingsen år 2000?", Notater 95/3, Statistisk sentralbyrå. Utover dette nådde arbeidet med FD-Trygd i 1994 ikke opp i prioritet sammenlignet med andre prosjekter innen området trygdestatistikk i Statistisk sentralbyrå. De forberedende arbeidene på FD-Trygd ble derfor ikke tatt opp igjen før i 1995.

Det ble tidlig innsett at FD-Trygd er et prosjekt med høy vanskelighetsgrad. Gjennomføringen av prosjektet krever stor statistikkfaglig kompetanse, så vel som innsikt i og evne til å løse metodiske og It-tekniske problemer. I erkjennelsen av dette ble det ansett som viktig å trekke på kompetanse fra flere fagområder for å sikre en god og vellykket gjennomføring av prosjektet. I Statistisk sentralbyrås virksomhetsplan for 1995 ble ønskene om å nå fram til et tverrfaglig samarbeid på prosjektet FD-Trygd konkretisert, og som en del av disse planene ble det bestemt å nedsette en *arbeidsgruppe* med representanter fra flere fagområder og seksjoner i Statistisk sentralbyrå for å gjennomføre et *forprosjekt* til prosjektet FD-Trygd.

Figur 2. Historikk - FD-Trygd



Arbeidsgruppa holdt sitt første møte i januar 1995 og avsluttet arbeidet i mars samme år med et notat som inneholder resultatene fra forprosjektarbeidet, jf. figur 2. Notatet er tatt inn som vedlegg 4. Her nevnes bare de viktigste punktene som ble diskutert av arbeidsgruppa.

Arbeidsgruppa

- vurderte omfanget av FD-Trygd
- ga en oversikt over typer av data som skal inngå i prosjektet
- skisserte hovedtyper av arbeidsoppgaver som må utføres for å gjennomføre prosjektet
- vurderte dataorganiseringen
- vurderte nødvendig kompetanseoppbygging
- diskuterte type samarbeid og samarbeidsformer
- foreslo en tidsplan for prosjektet.

Etter at forprosjektet var avsluttet, ble bare mindre arbeidsoppgaver utført på FD-Trygd i 1995. Den store satsingen på å bygge opp databasen startet først i 1996.

I mars 1996 ble det valgt en *prosjektleder* for FD-Trygd. Det ble videre oppnevnt en *styringsgruppe*, en *referansegruppe* og en *prosjektgruppe* for prosjektet, jf. figur 2. Sammensetningen av disse gruppene, samt funksjonene for gruppene og for prosjektleder er det gjort nærmere rede for i kapittel 3.

De opprinnelige planene med hensyn til hvilke typer data som skal inngå i FD-Trygd er nedfelt både i prosjektskrivet og i notatet fra forprosjektet. Utover dette bestemte styringsgruppa i 1997 at FD-Trygd skal utvides. En *søknad vedrørende denne utvidelsen* ble sendt til Datatilsynet og tillatelse ble gitt i oktober 1997, jf. figur 2. Utvidelsen er ganske betydelig. Den omfatter blant annet inntak av opplysninger i FD-Trygd om offentlige pensjonsordninger utenom folketrygden, jf. kapittel 2.2.

Til åpningen av basen i 1999 er det opprettet en *brukertjeneste* som skal forestå utlevering av data fra FD-Trygd, jf. figur 2 og nærmere omtale i kapittel 3.

2.2 Videre planer

I tillegg til å levere ut data fra databasen kan arbeidet på FD-Trygd etter åpningen av basen i 1999, deles inn i to virksomhetsområder. For det første vil det skje en videre oppbygging av basen ved at data *for nye årganger* blir lagt inn. For det annet vil den tidligere vedtatte *utvidelsen av prosjektet* bli gjennomført. Dette er tunge arbeidsoppgaver som vil kreve mye ressurser, og gå over flere år. På lengre sikt er det derfor vanskelig å overskue utviklingen i framdriften av prosjektet. I forbindelse med den årlige virksomhetsplanleggingen i Statistisk sentralbyrå blir det imidlertid hver høst laget en plan for virksomheten på FD-Trygd i kommende år. Vi skal her kort referere til planen for arbeidet på FD-Trygd i 1999, selv om det må tas forbehold om at eventuelle endringer i premissene for planen vil kunne påvirke realiseringen av den.

Som vist i figur 1, kapittel 1, er det ved åpningen av FD-Trygd i februar 1999 lagt inn opplysninger i databasen for årgangene 1992-1993 for temaområdene demografi, utdanning, trygd, økonomisk sosialhjelp, sysselsetting, arbeidssøkere og inntekt og formue. For flesteparten av disse områdene skal det i 1999, ifølge virksomhetsplanen, legges inn data i basen for fire nye årganger, det vil si for årgangene 1994-1997. Unntak gjelder for områdene utdanning, inntekt og formue og en liten del av tryggedataene (avslag på uførepensjon). For disse områdene skal det i løpet av 1999 legges inn data for årgangene 1994-1996.

Vedrørende utvidelsen av FD-Trygd skal det i 1999 utføres arbeid med å spesifisere, kontrollere og legge data inn i databasen for

- offentlige pensjonsordninger utenom folketrygden. (Bare arbeidet med spesifisering og kontroll av dataene skal starte i 1999. Dataene skal ifølge planen ikke legges inn i basen før i år 2000)

- vernepliktige og sivilarbeidere
- barnetrygd (spesielt merking av records for enslige forsørgere som mottar barnetrygd)
- personer som tjener opp omsorgspoeng

- statsansatte (merking av records).

For flere av disse nye områdene er det foreløpig uklart om det vil være mulig å legge inn data i FD-Trygd helt tilbake til og med årgangen 1992. Dette vil avhenge både av tilgjengeligheten til grunnlagsdataene og av datakvaliteten.

3. Organiseringen av prosjektet

FD-Trygd er et prosjekt ved Avdeling for personstatistikk, Statistisk sentralbyrå.

De *organene* som styrer og utvikler prosjektet er *styringsgruppa*, *referansegruppa*, *prosjektleder* og *prosjektgruppa*. I tillegg står *brukertjenesten* ansvarlig for utleveringen av data fra databasen.

3.1 Styringsgruppa

Styringsgruppa har den overordnede styringen av prosjektet. Styringsgruppa fastlegger i store trekk prosjektets framdriftsplan og vurderer hvilke tiltak som må settes i verk dersom det oppstår problemer med dataflyten til prosjektet eller it-tekniske kapasitetsproblemer etc. som forårsaker at framdriften av prosjektet blir vanskeliggjort. Styringsgruppa vurderer og gir forslag til løsninger som ivaretar og sikrer en mest mulig effektiv og formålstjenlig utlevering av dataene fra prosjektet, jf. mandat for styringsgruppa, vedlegg 1.

Styringsgruppa består av personer både fra Statistisk sentralbyrå og andre etater. De etatene som i sin tid deltok i oppbyggingen av Kirut er representert i styringsgruppa for FD-Trygd, og dette har sikret at erfaringene fra Kirut er blitt tatt hensyn til i arbeidet med FD-Trygd.

Styringsgruppa består av følgende medlemmer

fra Statistisk sentralbyrå

- fagdirektør Johan-Kristian Tønder (leder)
- rådgiver Eva Birkeland (ny 1997)
- seksjonsleder Berit Otnes
- seksjonsleder Paul Inge Severeide

fra andre etater

- avdelingsdirektør Odd Helge Askevold, Sosial- og helsedepartementet
- direktør Bjørn Henriksen, Norsk samfunnsvitenskapelig datatjeneste
- underdirektør Hans Kure, Arbeidsdirektoratet
- kontorsjef Grete Michaelsen, Rikstrykdeverket
- rådgiver Årstein Risan, Finansdepartementet.

Siden styringsgruppa ble oppnevnt i 1996 er det holdt to møter hvert år i styringsgruppa. Prosjektleder er sekretær for styringsgruppa.

3.2 Prosjektleder og prosjektgruppa

Prosjektleder utarbeider prosjektets gjennomføringsplan og stiller opp årsplaner for virksomheten på prosjektet sammen med andre sentrale aktører, herunder styringsgruppa. Prosjektleder sørger for samarbeid med ulike viktige instanser utenfor prosjektet, og arrangerer møter eller tar initiativ til annen faglig kontakt som kan motivere til innsats for å nå prosjektets mål. Prosjektleder er daglig leder av prosjektgruppa, jf. mandat for prosjektleder, vedlegg 2.

Prosjektleder for FD-Trygd er rådgiver Grete Dahl, Statistisk sentralbyrå.

I prosjektgruppa skjer det daglige arbeidet med prosjektet. Ved prosjektets start i 1996 bestod dette arbeidet i å utvikle den datamodellen som anvendes for organiseringen av dataene i FD-Trygd. For de ulike temaområdene eller hovedtyper av data som inngår i FD-Trygd utfører prosjektgruppa arbeidet med å spesifisere, kontrollere, dokumentere og modellere dataene. Det utføres et systemerings- og programmeringsarbeid for å legge dataene inn i databasen. Arbeidet som prosjektgruppa utfører byr på store faglige utfordringer. Arbeidet med dataene og dataorganiseringen er nærmere beskrevet i kapittel 5.

Prosjektgruppa er sammensatt av personer fra Seksjon for leveårsstatistikk og fra Kontor for IT ved Avdeling for personstatistikk.

Prosjektgruppa består av følgende medlemmer

- rådgiver Grete Dahl (leder)
- førstekonsulent Karsten Bjønnes
- førstekonsulent Torger Eidem (ny 1998, sluttet 1998)
- førstekonsulent Jon Folkedal (ny 1997, sluttet 1998)
- førstekonsulent Johnny Johansen (ny 1997)
- rådgiver Bjørn Roar Joneid
- førstekonsulent Jorunn Lajord
- førstekonsulent Randi Wølner (sluttet 1997).

I tillegg deltar førstekonsulent Anders Akselsen som forestår brukertjenesten på FD-Trygd, jf. kapittel 3.4, i alle møter i prosjektgruppa.

Det holdes jevnlig, som oftest ukentlige møter i prosjektgruppa. På møtene gjennomgås det løpende arbeidet og det føres kontroll med framdriften. Innenfor rammene for de årlige virksomhetsplanene, setter prosjektgruppa fortløpende opp kortvarige tidsplaner for sitt arbeid. Disse planene dekker fra 4-6 uker og er detaljerte med hensyn til de arbeidsoppgavene som skal utføres i perioden.

Prosjektgruppa nytter et elektronisk informasjons- og dokumentasjonssystem for lagring av møtereferat, midlertidige og avsluttede arbeidsdokumenter og forskjellige typer plandokumenter. Systemet er et verktøy som letter den daglige ledelsen av prosjektet. Systemet legger til rette for og bidrar til en gjensidig løpende informasjonsutveksling mellom prosjektgruppas medlemmer, og til at de nærmeste administrative lederne til en hver tid kan ha oversikt over arbeidet på prosjektet.

3.3 Referansegruppa

Referansegruppa bistår prosjektgruppa med faglig veiledning og råd i forbindelse med utviklingen og oppbyggingen av FD-Trygd, jf. mandat for referansegruppa, vedlegg 3.

Ved sammensetningen av referansegruppa er det tatt hensyn til at forskningsmiljøer skal være representert. Forskerrepresentantene er fra Forskningsavdelingen, Statistisk sentralbyrå, og Institutt for økonomi, Universitetet i Bergen. Utover dette er det de samme etatene og avdelingene som er representert i referansegruppa som i styringsgruppa, jf. kapittel 3.1.

Referansegruppa består av følgende medlemmer

fra Statistisk sentralbyrå

- rådgiver Grete Dahl (leder)
- rådgiver Eva Birkeland (ny 1997)
- forsker Tom Kornstad
- seksjonsleder Berit Otnes
- forsker Inger Texmon
- i tillegg er medlemmene av prosjektgruppa og brukertjenesten, jf. kapittel 3.2, faste medlemmer i referansegruppa

fra andre etater

- førstekonsulent Steinar Folkvord, Arbeidsdirektoratet (sluttet 1997)
- rådgiver Carl Gjersem, Finansdepartementet
- førstekonsulent Oddbjørn Haga, Rikstrygdeverket
- rådgiver Frode Hammer, Sosial- og helsedepartementet (sluttet 1998)
- førstekonsulent Frank Jullum, Finansdepartementet (sluttet 1996)
- førstekonsulent Dag Kiberg, Norsk samfunnsvitenskapelig datatjeneste
- rådgiver Michael B. Nielsen, Sosial- og helsedepartementet (ny 1998, sluttet 1999)
- førstekonsulent Anita M. Sivertsen, Sosial- og helsedepartementet (ny 1999)
- førstekonsulent Ola Thune, Rikstrygdeverket
- l. amanuensis Kjell Vaage, Institutt for økonomi, UiB
- førstekonsulent Thomas Øigarden, Arbeidsdirektoratet (ny 1997).

Siden referansegruppa ble oppnevnt i 1996 er det holdt to møter hvert år i referansegruppa.

3.4 Brukertjenesten

Statistisk sentralbyrå samarbeider med Norsk samfunnsvitenskapelig datatjeneste om brukertjenesten på FD-Trygd.

FD-Trygd er fysisk plassert i Statistisk sentralbyrå, og brukertjenesten i Statistisk sentralbyrå skal foreta alle datauttakene fra basen og legge til rette filer for brukerne. Når det gjelder det øvrige arbeidet som brukertjenesten skal utføre, er dette delt mellom de to etatene. Spesielt skal arbeidet med å anonymisere data som leveres ut til forskere og studenter utføres av Norsk samfunnsvitenskapelig datatjeneste, mens Statistisk sentralbyrå gjør dette selv for data som leveres ut til offentlig planlegging og andre brukere.

Retningslinjene som skal følges ved bestilling av data fra FD-Trygd (spesifisering av datauttak, innhenting av konsesjon etc.) er ikke omtalt her. Brukertjenesten gir nærmere informasjon.

Henvendelse for utlevering av data fra FD-Trygd skjer

for forskere og studenter til

Norsk samfunnsvitenskapelig datatjeneste
Hans Homboesgate 22
5007 Bergen

for offentlige planleggere, offentlig forvaltning og andre brukere til
Statistisk sentralbyrå
v/Anders Akselsen S 350
Kongensgate 6
P.b. 8131 Dep
0033 Oslo.

4. Formål med prosjektet

De viktigste formålene med FD-Trygd er å

- tilrettelegge data som er egnet til studier av *trygdekarrierer* og til å belyse *rekruttering til - , avgang fra - og overgang mellom* velferdsordninger, og mellom velferdsordninger, offentlige tjenstepensjonsordninger og arbeidsmarkedet, kapittel 4.1
- gi statistisk *merinformasjon til surveydata, input til modellarbeid* etc., kapittel 4.2
- definere og *tilrettelegge populasjoner* for å trekke *utvalg* til statistiske undersøkelser, kapittel 4.3
- bygge opp en *trygdemodul til Fob2000*, kapittel 4.4.

4.1 Tilrettelegge data som er egnet til studier av trygdekarrierer og til å belyse rekruttering til - , avgang fra - og overgang mellom velferdsordninger, og mellom velferdsordninger, offentlige tjenstepensjonsordninger og arbeidsmarkedet

Hovedformålet med forløpsdatabasen FD-Trygd er å skaffe tilveie et datamateriale som *fortløpende, over tid* viser personers tilknytning til trygd, økonomisk sosialhjelp, offentlige tjenstepensjonsordninger, arbeidssøking og inntektsgivende arbeid. I tillegg inneholder basen opplysninger som både *over tid* og på *gitte tidspunkter* kan nyttes for en nærmere karakteristikk av personene. Dette er for eksempel opplysninger om demografi, utdanning, inntekt og formue. Dataene gir i særlig grad detaljert informasjon om personenes trygdeforhold, og materialet kan brukes til å studere *rekrutteringen til - , avgangen fra - og overgangen mellom* forskjellige trygdeordninger, samt overgangen fra trygdeordninger til andre sosiale ordninger, tjenstepensjonsordninger og arbeidsmarkedet, eller vice versa.

Innen en bred anlagt velferdsforskning er det flere underliggende formål med å anvende forløpsdataene i FD-Trygd slik som beskrevet ovenfor. I det følgende, **pkt. i) - iii)**, er nevnt noen av disse formålene. Det er også vist til noen aktuelle, konkrete anvendelsesområder, men oversikten er ikke på noen måte uttømmende sett i forhold til den forskning som trolig vil bli foretatt på dette datamaterialet.

i) For det første vil forløpsdataene i FD-Trygd kunne nyttes til å framskaffe forskningsbasert kunnskap om hvordan *velferdssystemet fungerer for enkeltindivider*, samt *beskrive de persongruppene* som mottar ytelser fra forskjellige velferdsordninger. I forlengelsen av en slik kartlegging vil det også være interessant å studere hvordan enkeltindividenes situasjon er påvirket ikke bare av velferdssystemets virkemåte, men også av personenes *deltakelse på andre areaner* i samfunnet som for eksempel i arbeidsmarkedet, i forbindelse med å skaffe seg utdanning etc. Dataene kan nyttes til å framskaffe et beslutningsunderlag som kan bidra til å videreutvikle velferdsordningene i samsvar med framtidens behov og forutsetninger. Siden FD-Trygd omfatter hele befolkningen, vil materialet også

kunne anvendes til å fokusere spesielt på *velferdsordningenes virkemåte* overfor *innvandrere* eller andre *små befolkningsgrupper*. En del konkrete anvendelsesområder er som følger

- kartlegge og påvise ulikheter i *trygdekarrierer* for grupper av personer, samt gi en nærmere *karakteristikk av persongruppene* ved opplysninger om demografiske forhold, familietilknytning, bosted, utdanning, tilknytning til arbeid etc. (Stikkord: Trygdekarrierer for uførepensjonister? Hvor blir det av attføringsklientene? Hva kjennetegner langtidssykmeldte som greier å komme tilbake til inntektsgivende arbeid, og hvordan ser deres videre karriere ut - tilbakefall til trygd, deltid i arbeidslivet etc.? Hvem er gjengangere i trygdesystemet, og hva karakteriserer disse personene?)
- identifisere persongrupper som er *sambrukere* av velferdsordninger. (Stikkord: Hva kjennetegner personer som er sambrukere av trygd og økonomisk sosialhjelp? Varierer stønadslengden for økonomisk sosialhjelp med type trygd personene mottar, og er stønadslengden ulik for yngre versus eldre trygdemottakere?)
- beskrive persongrupper som går ut i *tidligpensjonering*. (Stikkord: Hva kjennetegner personer som går ut i avtalefestet pensjon? Synes mønsteret for eventuelle tidligere trygdekarrierer å være forskjellig for henholdsvis statspensjonister og andre alderspensjonister?)
- påvise hvilke persongrupper som *kombinerer arbeid og trygd* eller *utdanning og trygd*. (Stikkord: Hvor vanlig er det at trygdemottakere deltar i yrkeslivet, og hva kjennetegner disse gruppene sammenlignet med trygdemottakere som ikke deltar yrkesaktivt? Hva kjennetegner enslige forsørgere som kombinerer utdanning og trygd? Hvor vanlig er det at enslige forsørgere med overgangsstonad avbryter en påbegynt utdanning?)
- rette søkelys på hvordan velferdssystemet fungerer for *innvandrere* eller andre *små befolkningsgrupper*. (Stikkord: Er det spesielle forhold som karakteriserer innvandreres trygdekarrierer? Hvordan og hvor ofte anvendes velferdsordningene av ungdomsgrupper med ustabil tilknytning både til arbeid og utdanning?).

ii) Et annet viktig formål med FD-Trygd er å nytte forløpsdataene til å *evaluere effektene* av *regelverksendringer* eller *strategitiltak* knyttet til velferdsordningene. Slike analyser kan gå ut på å studere om regelsverksendringene eller tiltakene har hatt tilsiktet virkning eller om det kan påvises konsekvenser som ikke samsvarer med intensjonene.

Kanskje det største og mest omfattende strategitiltak i senere år er satsingen på *arbeidslinja* i velferdspolitikken. Arbeidslinja går ut på sikre at den enkelte i størst mulig grad skal kunne forsørge seg selv gjennom eget arbeid, og velferdspolitikken på 1990-tallet har vært sterkt preget av dette strategivalget. Ved å nytte virkemidler som for eksempel regelverksendringer eller andre tiltak har offentlige myndigheter forsøkt å legge forholdene til rette slik at den enkelte skal kunne stå i inntektsgivende arbeid så lenge det er mulig, samt å hindre utstøting fra arbeidsmarkedet.

Data fra FD-Trygd vil kunne anvendes til å evaluere flere sider ved arbeidslinja. Det kan eksempelvis foretas analyser for å studere

- effektene av tiltak knyttet til *yrkesrettet attføring*
- effektene av *tidsbegrensningen av rehabiliteringspenger*
- effektene på *tilgangen til uførepensjon* av *endringer i regelverk, friinntektsgrenser* etc. for uførepensjonen

- effektene av bruken av ulike *arbeidsmarkedstiltak* overfor registrert *arbeidsledig ungdom* (det vil si ungdom som har opparbeidet rett til dagpenger)
- effektene av *omleggingen av stønadsordningene* for *enslige forsørgere*.

iii) Et tredje formål med forløpsdataene i FD-Trygd er å *kartlegge velferdsutviklingen* for personer og familier som *over tid* har ulik tilknytning til forskjellige typer trygdeordninger eller til kombinasjoner av trygd, andre supplerende stønadsordninger, ledighet og inntektsgivende arbeid. Opplysninger om inntekt kan nyttes som indikator på velferd, og for alle personer kan det på grunnlag av dataene i FD-Trygd dannes en *familieinntekt*, nedbrutt på underposter som trygd, lønn, kapitalinntekt etc. (Inntektsdataene i FD-Trygd er imidlertid *ikke* registrert og organisert i basen som forløpsdata. Fineste tidsenhet for inntektsregistreringene er *år*).

Velferdsordningene skal kompensere for tap av eller mangel på inntekt ved for eksempel sykdom, funksjonshemming, tap av forsørger, alderdom, arbeidsløshet etc., og underliggende formål med analyser som nevnt ovenfor, kan være å

- framskaffe empiri om utilsiktede *sosiale skjevheter*
- studere *velferdsmessige konsekvenser av endringer i regelverk*
- studere hvordan *folketrygden* har virket og virker som *forsørgelsessystem*.
(Stikkord: Likestillingsperspektiv? Betydningen av egen inntekt fra arbeid, både tidligere inntekt og inntekt som trygdemottaker? Bruken av sosiale støtteordninger som for eksempel økonomisk sosialhjelp? Betydningen av framveksten av pensjonsordninger utenom folketrygden?).

4.2 Gi statistisk merinformasjon til surveydata, input til modellarbeid etc.

Det er planlagt å nytte dataene i FD-Trygd til å gi statistisk merinformasjon til *paneldelen* av *Statistisk sentralbyrås levkårsundersøkelse*. Paneldelen til levkårsundersøkelsen består av personer i alderen 16-79 år, og utvalget utgjør (netto) på årsbasis ca. 4 000 personer.

Vi tror at FD-Trygd også vil kunne anvendes for å øke informasjonsmengden i andre surveys eller som *input* til forskjellige typer *modellarbeid*. Det kan også bli aktuelt å koble data fra FD-Trygd til datasett som *brukerne selv eier*, forutsatt at vilkårene i statistikkloven og personregisterloven er oppfylt.

4.3 Definere og tilrettelegge populasjoner for å trekke utvalg til statistiske undersøkelser

Det har tidligere vært en del forespørsler til Statistisk sentralbyrå om å trekke utvalg og gjennomføre utvalgsundersøkelser for nærmere definerte *pensjonistgrupper*. Dataene i FD-Trygd er lagt til rette slik at det skal være mulig å trekke slike utvalg fra opplysningene i databasen.

Forløpsdataene i FD-Trygd bygger overveiende på opplysninger som tidligere er anvendt i offisiell tverrsnittsstatistikk, jf. kapittel 5.3. Av flere grunner er det likevel noen *små* avvik mellom forløpsdataene og tverrsnittstatistikken (pr. de tidspunkt denne statistikken blir produsert). Viktigste årsak til dette er at forløpsdata forutsetter at opplysninger kan stilles sammen på individnivå over tid. Opplysninger som er nyttet i tverrsnittsstatistikk men som ikke oppfyller disse kravene, vil derfor ikke kunne anvendes ved kreering av forløpsdata.

For temaområdene *demografi* og *pensjoner* i FD-Trygd, blir det ved *hvert årsskifte* lagt inn “tilleggstabeller” i databasen. I disse tabellene blir det registrert *eventuelle avvik* mellom personrecords (fødselsnumre) i forløpsdataene i basen og de som er nyttet ved produksjonen av den offisielle tverrsnittstatistikken. Dersom det for disse fødselsnumrene (eller for andre) finnes avvik for sentrale kjennemerkeverdier, blir også disse avvikene registrert i "tilleggstabellene". Ved å nytte opplysningene både fra forløpsdataene i FD-Trygd og fra disse “tilleggstabellene” er det således mulig å trekke *representative utvalg* for forskjellige populasjoner av *pensjonister*. (Opplysningene i “tilleggstabellene” blir imidlertid *ikke* nyttet for å “rette opp” forløpsdataene i FD-Trygd).

4.4 Bygge opp en trygdemodul til Fob2000

Det er planer om å bruke data fra FD-Trygd til Folke- og boligtellingsen år 2000. Det er enda ikke endelig fastlagt hvilke typer data som skal tas inn, men trolig vil det bli nyttet opplysninger for forskjellige typer *folketrygdpensjonister* og pensjonister med *avtalefestet pensjon*.

5. Data og dataorganiseringen

I dette kapitlet er omtalt bare de dataene som inngår i FD-Trygd ved åpningen av basen i 1999, jf. figur 1, ovenfor. Dataene som vil bli lagt inn i FD-Trygd som en følge av den planlagte utvidelsen, jf. kapittel 2.2, er ikke omhandlet.

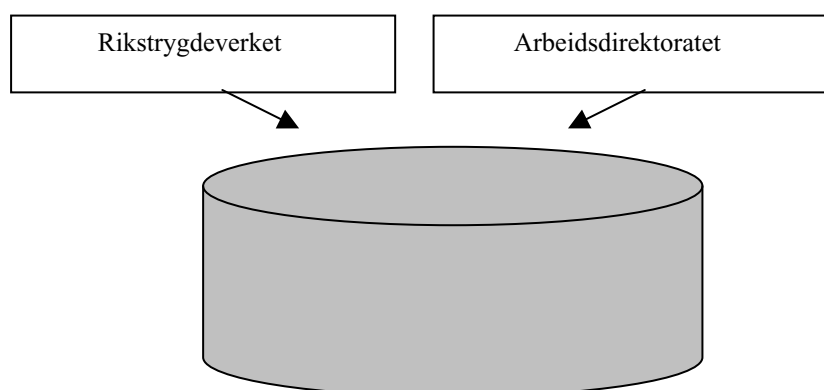
5.1 Administrative data

Det har i senere tid skjedd en tiltagende satsing på bruk av administrative data i produksjon av offisiell statistikk og til forskningsformål. I et slikt bilde har FD-Trygd en naturlig plass. Alle data som anvendes i denne basen er hentet fra administrative registre eller fra store statistikkregistre som er kreert på grunnlag av administrative data.

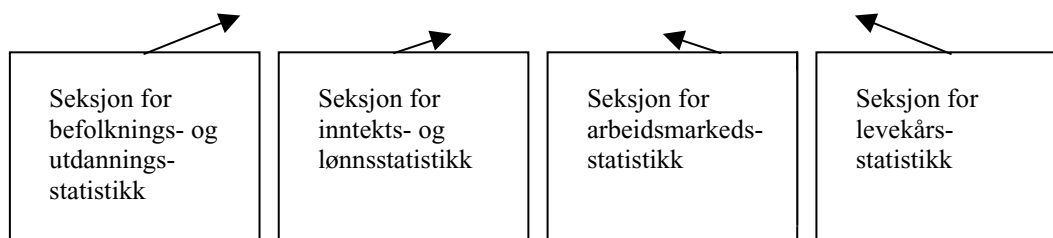
Den største dataleverandøren til FD-Trygd er *Rikstrygdeverket*. Fra Rikstrygdeverkets databaser og store datasystemer overføres til FD-Trygd opplysninger både om langtids- og korttidsytelser fra folketrygden. En annen stor ekstern dataleverandør er *Arbeidsdirektoratet*. Fra Arbeidsdirektoratets registre nyttes opplysninger om arbeidssøkere. *Seksjoner* i *Statistisk sentralbyrå* leverer også data til FD-Trygd. Dette er forskjellige typer administrative data som seksjonene har hentet inn til statistikkformål, og som senere blir nyttet som inputdata til FD-Trygd. Grunnlagsdataene for

demografi- og utdanningsopplysningene i FD-Trygd er hentet fra *Seksjon for befolknings- og utdanningsstatistikk*, opplysningene om sysselsetting fra *Seksjon for arbeidsmarkedsstatistikk* og opplysningene om økonomisk sosialhjelp fra *Seksjon for leveårsstatistikk*. I tillegg nyttes inntekts- og formuesdata fra *Seksjon for inntekts- og lønnsstatistikk*. Figur 3 viser dataflyten til FD-Trygd.

Figur 3. Dataflyt til FD-Trygd fra etater/seksjoner



FD-Trygd



5.2 Typer data

De dataene som er i FD-Trygd ved åpningen av basen i 1999, kan deles inn i *temaområder* eller *ordninger* som vist i figur 4, nedenfor. I figuren er også vist til *hvilken etat* dataene er hentet fra, og for hver ordning er det gitt *eksempler på kjennemerker*. For en fullstendig oversikt over kjennemerkene i FD-Trygd vises til variabelliste og annen tilgjengelig dokumentasjon av dataene, jf. kapittel 6.

5.3 Gjenbruk av data

Flesteparten av de administrative registrene som utgjør datagrunnlaget for FD-Trygd er tidligere nyttet av Statistisk sentralbyrå eller andre etater til å produsere offisiell statistikk. Registrene inneholder opplysninger pr. et tidspunkt, for eksempel pr. utgangen av måned, år etc., eller for et tidsintervall, for eksempel for et år. Opplysningene er nyttet som grunnlag for en omfattende og mangeartet tverrsnittsstatistikk; som trygdestatistikk, sosialhjelpsstatistikk, befolkningsstatistikk etc.

Figur 4. Temaområder eller ordninger som det er opplysninger for i FD-Trygd ved åpningen av databasen i 1999

Etat	Temaområde eller ordning	Eks. på kjennemerker (eller internvariable, jf. kapittel 5.4)
Rikstrygdeverket	Alderspensjon	grunnpensjon-faktor, alderspensjonsdel
	Uførepensjon	grunnpensjon-faktor, primærdiagnose
	Pensjon - etterl. ektefelle	særtilllegg-faktor, fradrag f/faktor
	Pensjon - etterlatt fam.pleier	grunnpensjon-faktor, særtilllegg-faktor
	Pensjon - etterlatt barn	pensjonstype, særtilllegg-faktor
	Avtalefestet pensjon - off. sektor	pensjonstype, sluttpoengtall
	Avtalefeste pensjon - privat sektor	pensjonstype, grunnpensjon-faktor
	Foreløpig uførestønad	uføregrad, primærdiagnose
	Avslag på uførepensjon ¹	primærdiagnose, avslagskategori

	Overgangsstonad	grunnbeløp, forventet inntekt
	Stonad til barnetilsyn	forventet inntekt, stonad til barnetilsyn
	Grunnstonad	grunnstonad kroner, diagnose
	Hjelpestonad	hjelpestonad kroner, diagnose
	Rehabiliterings- og attføringspenger	attføringsgruppe, bortfallsgrunn
	Fødselspenger	uttaksprosent, bortfallsgrunn
	Sykepenger	sykepengegruppe, sykepengegrad
	Fødsels- og sykepenger (aggregert) ¹	sykepengegruppe, konto
Arbeidsdirektoratet	Arbeidssøkere og yrkeshemmede	arbeidssøkerstatus, tiltaksgruppe
	Dagpenger	opnådd prosent, ukesats
Statistisk sentralbyrå	Demografi	sivilstand, type flytting
	Fødeland etc. ¹	fødeland, innvandrerkategori
	Fødsler ¹	type fødsel, registreringsstatus
	Høyeste fullførte utdanning	utdanningens art, fullføringstidspunkt
	Igangværende utdanning	heltid/deltid, AMO-kurs
	Økonomisk sosialhjelp - varighet	
	Mottatt økonomisk sosialhjelp ¹	bidrag kroner, antall måneder
	Sysselsetting - viktigste arbeidsforhold	næring, arbeidsstedskommune
	Antall ikke-viktigste arbeidsforhold ¹	antall forhold, år
	Inntekt ¹	personinntekt, alminnelig inntekt
	Formue ¹	realkapital, bruttofinanskapital

¹ Ikke-forløpsdata.

Det er flere grunner til at disse registeropplysningene kan *gjenbrukes* til å kreere *forløpsdata* i FD-Trygd.

For det første finnes det opplysninger i flere av disse registrene som ikke blir nyttet i tverrsnittsstatistikken, men som er verdifull informasjon ved kreering av hendelsesorienterte data. Dette er variable som angir *når* det skjer endringer i kjennemerkeverdier eller endringer i tilstander i en persons livsløp. Tidsangivelsen for disse variablene er for eksempel år, måned, dag, og disse dateringene anvendes for å danne forløpsdata i FD-Trygd.

For det annet blir samme type register i en del tilfeller *kreert hyppigere* enn pr. de tidspunkt det lages offisiell tverrsnittsstatistikk. Eksempler på dette er filer som blir laget pr. utgangen av hver måned, mens den offisielle statistikken bare blir produsert pr. utgangen av hvert kvartal eller hvert år. Disse månedsfilene anvendes i FD-Trygd. Opplysningene i hver månedsfil sammenlignes med opplysningene i foregående fil, og de månedlige endringsopplysningene som kommer ut av denne prosessen, nyttes til å danne forløpsdata med år, måned som datering for tilstandsendingene.

Omorganiseringen av dataene til forløpsdata er en vesentlig del av prosjektgruppas arbeid. Prosjektgruppa spesifiserer og henter ut fra de administrative registrene de dataene som skal inngå i FD-Trygd. Disse opplysningene lagres i egne filer (miniregistre), og prosjektgruppa foretar en del kontroller og annen type kvalitetssikringsarbeid på opplysningene i miniregistrene. Etter dette skjer selve datamodelleringen, det vil si omorganiseringen av dataene til forløpsdata. Dette arbeidet krever en nøye gjennomgang av opplysningene i miniregistrene og i de opprinnelige administrative registrene, slik at tilstandsendingene blir ivaretatt og registrert best mulig i FD-Trygd. Til hjelp for disse registreringene nyttes egne datovariabler i FD-Trygd, jf. kapittel 5.4.

5.4 Dataorganiseringen

Opplysningene i FD-Trygd er organisert i “tabeller” i databasen (Oracle-database). Det er en tabell for hvert temaområde eller hver ordning; for eksempel en tabell for igangværende utdanning, en tabell for alderspensjon, en tabell for overgangsstønad etc., jf. kapittel 5.2. Opplysningene i hver tabell gir informasjon om *når* en person kommer inn i ordningen eller i deler av ordningen (*tilgang*), og *når* en person går ut av ordningen eller av deler av ordningen (*avgang*). I tillegg er det for hver person registrert *når* det skjer *endringer* i personens forhold til ordningen. For å ta vare på slike endringer er det innen hver ordning/hver tabell innført variable (også kalt internvariable) for alle forhold som det er registrert endringer for i FD-Trygd. Innen ordningen alderspensjon kan for eksempel en endring for en bestemt person bestå i en økning i alderspensjonsdelen fra 50 prosent til 100 prosent. Dette betyr at kjennemerket eller internvariabelen alderspensjonsdel har endret verdi fra 50 til 100 prosent.

Til analyseformål vil det som oftest være ønskelig å hente ut data fra flere tabeller samtidig. I FD-Trygd er dette mulig ved bruk av en felles “nøkkel” (fødselsnummer) som finnes i alle tabellene. Ved hjelp av “nøkkel” kan vi sette sammen opplysninger fra flere ordninger/tabeller både på samme tidspunkt og over tid. Det er for eksempel mulig å ta ut data for personer som er i gang med en utdanning i perioden fra og med oktober 1992 til og med mai 1993, og for disse samtidig få registrert deres eventuelle bruk av overgangsstønad i den angitte perioden.

Forløpsanalyse går ut på å analysere tidsintervall mellom suksessive *tilstandsendinger* (*hendelser*), og med tiden som stokastisk variabel. Både *tilgang til* ordninger, *avgang fra* ordninger og *endringer i* personenes forhold til ordningene er å betrakte som *tilstandsendinger* for personene. I FD-Trygd er disse tilstandsendingene for alle variable registrert ved hjelp av egne *datovariabler*. Datovariablene

har forskjellige verdier. For en del tilstandsendinger inneholder datovariabelen verdien år, måned, dag (YYYYMMDD). For andre tilstandsendinger har datovariabelen verdien år, måned (YYYYMM) eller år (YYYY). Tilstandene er imidlertid ikke bestandig faktisk endret (oppdatert) så hyppig som verdien av datovariabelen (dateringen) gir uttrykk for. For å anskueliggjøre hvordan tilstandsendinger er datert og oppdatert i FD-Trygd har vi nedenfor vist til et par variable fra ordningen Alderspensjon.

Figur 5. Eksempler på dataorganiseringen og dateringen av variable i FD-Trygd

Variabelnavn Kortnavn Datatype (Lengde)	Datering	Oppdaterings kontroll	Koder	Kommentarer
Alderspensjondel apd number (3)	YYYYMM	Fortløpende	Andel av alderspensjon som tas ut: 000=Ingen tas ut eller uoppgitt 001-099=1-99% kan tas ut 100=Hele pensjonen tas ut.	Personer på 67-69 år som ikke har søkt om alderspensjon, er ikke registrert som alderspensjonist.
Grunnpensjon-faktor gpfak number (7)	YYYYMM	Ved tilgang/ avgang/ årsskifte		Det årlige beløp pensjonisten har krav på, regnet som andel av aktuelt grunnbeløp.

Variabelnavn Kortnavn Datatype (Lengde)	Datering	Oppdaterings kontroll	Koder	Kommentarer
				(5 desimaler).

Første kolonne i figur 5 gir opplysninger om variabelnavn og kortnavn for variabelen, samt opplysninger om datatype og lengde. Opplysningene i kolonnene 2 og 3 gir opplysninger om hvordan variablene er henholdsvis datert og kontrollert med sikte på en eventuell oppdatering. I de to siste kolonnene finnes kodeopplysninger og kommentarer til variablene.

Opplysningene i kolonnene 2 og 3 bør ses i sammenheng siden disse to kolonnene tilsammen gir informasjon som viser hvordan tilstandsendringene er behandlet i FD-Trygd. Kolonne 2 angir datering for variablene. Selve datovariablene er utelatt fra dette oppsettet, men datering (i kol. 2) gir informasjon om hvordan verdiene for datovariablene teknisk sett er registrert eller angitt for de enkelte variablene som inngår i FD-Trygd. Kolonne 3 angir hvor ofte den til en hver tid faktiske variabelverdi eller tilstand for en variabel blir kontrollert mot en eventuell "ny tilstandsverdi" med sikte på å foreta en oppdatering, det vil si registrere en tilstandsendring.

For de to variablene (alderspensionsdel og grunnpensjon-faktor) som er tatt inn i eksemplet ovenfor, betyr opplysningene i kolonnene 2 og 3 følgende: Variabelen alderspensionsdel er datert år, måned (YYYYMM), i kol. 2. Denne variabelverdien kontrolleres *fortløpende*, (jf. kol. 3), med sikte på en eventuell oppdatering. Med *fortløpende* menes at kontroll (og mulig oppdatering) skjer så ofte som datering (i kol. 2) gir uttrykk for; i dette tilfelle hver måned i året og alle år (YYYYMM). For variabelen grunnpensjon-faktor er også datering år, måned (YYYYMM). For denne variabelen blir det imidlertid kontrollert for endring bare ved hvert årskifte og ved avgang fra ordningen. I tillegg er variabelverdien tatt inn ved tilgang til ordningen. Dette er angitt i kolonne 3 med teksten *Ved tilgang/avgang/årskifte*.

Flesteparten av dataene i FD-Trygd er *forløpsdata*, og for disse er dataorganiseringen og datering av variablene i prinsippet som vist i eksemplet ovenfor. For de få opplysningene i FD-Trygd som *ikke* er *forløpsdata*, er dataorganiseringen og beskrivelsen av variablene dokumentert tilsvarende de opplysningene som er gitt i kolonnene 1, 4 og 5 i figur 5.

6. Dokumentasjon og informasjon

I prosjektgruppas arbeid blir dokumentasjon tillagt stor vekt. Dette gjelder både dokumentasjon av dataene i FD-Trygd så vel som dokumentasjon av alle aktiviteter som er forbundet med å bygge opp forløpsdatabasen.

For hvert enkelt temaområde blir det laget en *dokumentasjonsrapport* for det arbeidet som utføres i prosjektgruppa fra og med innhentingen av grunnlagsregistrene og til dataene er lagt inn i databasen. Disse dokumentasjonsrapportene inneholder en beskrivelse av arbeidet i forbindelse med å spesifisere og ekstrahere opplysningene fra de opprinnelige administrative registrene. Rapportene dokumenterer videre de kontrollene som er utført på materialet, samt selve datamodelleringen og arbeidet med å omorganisere dataene til forløpsdata. I rapportene er det referert til de programmene som er nyttet for å legge dataene inn i databasen. Dokumentasjonsrapportene trykkes i serien *Notater* i Statistisk sentralbyrå.

Det er utarbeidet en *variabelliste* som detaljert dokumenterer de typer data som er i FD-Trygd ved åpningen av databasen i februar 1999. Det arbeides også med å etablere en *Web-side* for

FD-Trygd på Internett. Etter planen skal det videre utarbeides et notat som omhandler *datakvaliteten* og *usikkerheten* på dataene i FD-Trygd.

På noe sikt vil Statistisk sentralbyrå ytterligere vurdere hvilke muligheter som finnes for å informere brukerne om FD-Trygd. Kontakten med brukerne bør trolig legges opp som en *to-veis kommunikasjon*, slik at også brukernes tilbakemeldinger og erfaringer med dataene i FD-Trygd blir fanget opp. Vi tror at *brukerseminar* eller annen tilsvarende kontakt med brukerne vil være viktig for gjensidig å kunne utveksle informasjon om FD-Trygd i tiden som kommer.

Vedlegg 1

SSB, 1996

MANDAT OG FUNKSJONSTID FOR STYRINGSGRUPPA. TRYGD-FOBHISTORIE¹

1. Styringsgruppa skal overordnet styre arbeidet med utviklingen av prosjektet Trygd-fobhistorie med utgangspunkt i de primære planer som er lagt for gjennomføringen av prosjektet i henholdsvis projektskrivet, «GDA, 10.11.93» og forprosjektrapporten, «Notater 95/22» fra Statistisk sentralbyrå.
2. Styringsgruppa skal i store trekk fastlegge prosjektets framdriftsplan ut fra de ressurser som Statistisk sentralbyrå og andre institusjoner stiller til disposisjon for prosjektet.

3. Styringsgruppa skal vurdere hvilke tiltak som må iverksettes dersom det oppstår problemer med dataflyten til prosjektet, edb-tekniske kapasitetsproblemer eller andre problemer som forårsaker at framdriften av prosjektet blir vanskeliggjort.
4. Styringsgruppa skal vurdere og gi forslag til løsninger som ivaretar og sikrer en mest mulig effektiv og formålstjenlig utlevering av dataene fra prosjektet. Datautleveringen må imidlertid være i overensstemmelse med og skje innenfor SSBs regelverk.
5. Styringsgruppa skal være et kontaktledd mellom prosjektet og ledelsen i de institusjonene som er representert i styringsgruppa.
6. Styringsgruppas arbeid skal i praksis hovedsakelig skje ved møtevirksomhet. Prosjektleder forbereder de sakene som skal behandles i styringsgruppa. Utenom møtene forholder prosjektleder seg til styringsgruppas leder.
7. Fagdirektør ved Avdeling for personstatistikk, Statistisk sentralbyrå, er leder for styringsgruppa.
8. Funksjonstiden for medlemmene av styringsgruppa utløper ved årsskiftet 1998/-99², som er det tidspunkt da det viktigste utviklingsarbeidet på prosjektet ventes å være avsluttet.

¹ Senere skiftet navn til FD-Trygd.

² Mandatperioden er forlenget til utgangen av 2001.

Vedlegg 2

SSB, 1996

MANDAT FOR PROSJEKTLEDER. TRYGD - FOBHISTORIE¹

Prosjektleder skal

- i samarbeid med styringsgruppa bidra til utarbeidelse av et formålstjenlig mandat for prosjektets hovedansvarlige, inklusive seg selv
- utarbeide prosjektets gjennomføringsplan i samarbeid med øvrige sentrale aktører, herunder styringsgruppa
- velge ut prosjektets operative medarbeidere sammen med linjeledelsen og andre, slik at prosjektleder i rimelig grad får lov til å tilkjenne personlige preferanser

- holde kontakt med de organer som stiller ressurser til prosjektet for å sikre at ressursene er tilgjengelige når prosjektet trenger dem
- lede prosjektgruppas daglige arbeid, det vil si orientere, holde kontakt med prosjektdeltakere, arrangere viktige møter og motivere til innsats for å nå prosjektets mål og formål
- sørge for samarbeid med og informasjon til ulike viktige instanser utenfor prosjektet
- holde seg underrettet om framdriften. Kontrollere innsats mot planer og resultater mot mål for til enhver tid å være klar over prosjektets status
- foreslå og iverksette tiltak ved forsinkelser og overskridelser
- kunne gi medlemmene av prosjektgruppa egne ansvarsområder, selv om prosjektleder beholder det overordnede faglige ansvar
- sørge for tilfredsstillende dokumentasjon, rapportering og arkivering.

¹ Se note 1, side 19.

Vedlegg 3

SSB, 1996

MANDAT OG FUNKSJONSTID FOR REFERANSEGRUPPA. TRYGD-FOBHISTORIE¹

1. Referansegruppa skal bistå Statistisk sentralbyrå (prosjektgruppa for prosjektet Trygd-fobhistorie) med faglig veiledning og råd i forbindelse med utviklingen og oppbyggingen av forløpsdatabasen Trygd-fobhistorie.
2. Referansegruppas råd og veiledning skal omfatte alle de faglige områdene som er berørt av prosjektet (statistikkfaglige vurderinger, edb-spørsmål, metodespørsmål og spørsmål knyttet til bruk av dataene/output fra prosjektet).

3. Referansegruppas arbeid vil i praksis bli utført ved at SSB (prosjektgruppa for prosjektet) avholder møter for gjennomdrøfting av ulike tema. Utover disse møtene kan medlemmene av referansegruppa bli bedt om å gi skriftlige merknader eller annen tilbakemelding på tilsendte notater eller andre faglige henvendelser fra prosjektgruppa.
4. Funksjonstiden for medlemmene av referansegruppa utløper ved årsskiftet 1998/-99², som er det tidspunktet da det viktigste utviklingsarbeidet på prosjektet ventes å være avsluttet.
5. Prosjektleder er leder av referansegruppa.

¹ Se note 1, side 19.

² Mandatperioden er forlenget til utgangen av 1999.

Vedlegg 4

NOTATER 95/22

SLUTTRAPPORT

fra **forprosjektet**

til prosjektet "Trygd-fobhistorie" ¹

I n n h o l d

1. Forprosjektet - og hovedprosjektet "Trygd-fobhistorie"
2. Forprosjektet. Bakgrunn og formål
3. Arbeidsgruppas sammensetning og arbeid
4. Resultater fra forprosjektarbeidet
5. Plan for gjennomføring av hovedprosjektet "Trygd-fobhistorie"

Vedlegg: Prosjektskriv

¹ Senere skiftet navn til FD-Trygd.

1. Forprosjektet - og hovedprosjektet "Trygd-fobhistorie"

Dette notatet gir orientering om erfaringene og resultatene fra et forprosjekt til prosjektet "Trygd-fobhistorie". Arbeidet med forprosjektet er foretatt av en Arbeidsgruppe med representanter fra Seksjon for helse- og sosialstatistikk (330), Gruppe for EDB (303) og Seksjon for EDB utvikling (710).

Hovedprosjektet "Trygd-fobhistorie" ble godkjent som prosjekt i Statistisk sentralbyrå ved årsskiftet 1993/1994, (produkt nr. 077 1). Prosjektet er gjort rede for i pkt. 5 i prosjektskrivet "Trygdedata og trygdehistorikk for anvendelse i modellbefolkningsprosjektet og Fob2000", GDA 10.11.93, se vedlegg.

Det viktigste siktemålet med prosjektet "Trygd-fobhistorie" er å legge opp og utvikle et datasystem som skal inneholde "historiske" trygdeopplysninger og annen type statistisk informasjon som viser personers tilknytning til forskjellige trygdeordninger, arbeid/ledighet og andre sosiale situasjoner (utdanningssituasjon/sykdom/attføring/"utestenging" fra trygd/mottak av sosialhjelp etc.) i løpet av 1990-årene. Dette datagrunnlaget skal anvendes til forløpsanalyser og til annen type forskning som er relatert til trygde- og sosialpolitiske forhold i samfunnet, til produksjon av statistikk i tverrsnitt, samt som underlag for utvikling av statistikk ved Folke- og boligtellingsen år 2000.

Som vist til i prosjektskrivet, er det også flere underliggende formål med prosjektet "Trygd-fobhistorie". Blant disse er kanskje det viktigste formålet å gi grunnlag for en forbedret input av trygdedata til modellbefolkningsprosjektet (LOTTE).

Prosjektet "Trygd-fobhistorie" vil kunne bli en "arvtaker" til forløpsdatabasen KIRUT. På en del områder vil imidlertid prosjektet "Trygd-fobhistorie" inneholde flere opplysninger enn de som nå ligger inne i KIRUT. Prosjektet "Trygd-fobhistorie" skal også omfatte hele befolkningen, mens KIRUT bare inneholder opplysninger for et 10-prosent-utvalg av befolkningen i alderen 16-66 år.

Våren 1994 ble det sendt melding til Datatilsynet om opprettelse av datasystemet/registeret "Trygd-fobhistorie". Tillatelse til å føre registeret ble gitt av Datatilsynet, i brev datert 24.05.94.

I 1994 er det ved en rekke anledninger orientert om prosjektet "Trygd-fobhistorie". Det er redegjort for prosjektet både i Statistisk sentralbyrå og i fora utenfor SSB. Blant eksterne institusjoner som har vært representert i disse orienteringsmøtene, er for eksempel Norges forskningsråd, Norsk samfunnsvitenskapelig datateneste, Sosial- og helsedepartementet etc.

Med utspring i prosjektet "Trygd-fobhistorie" er det også i 1994 utført et arbeid med sikte på å planlegge og opprette en trygdemodul for anvendelse i Folke- og boligtellingsen år 2000. Dette planleggingsarbeidet er dokumentert i "Mot en registerbasert Folke- og boligteiling år 2000?", Notater 95/3, Statistisk sentralbyrå.

2. Forprosjektet. Bakgrunn og formål

Bakgrunn

Prosjektet "Trygd-fobhistorie" er et stort og omfattende prosjekt. Prosjektet har ganske høy vanskelighetsgrad, og gjennomføringen av prosjektet forutsetter kompetanseoppbygging på flere områder. Dette gjelder statistikkfaglig kompetanse, så vel som evne til å løse metodiske og Edb-tekniske problemer som vil oppstå både i startfasen av prosjektet og senere under selve arbeidet med prosjektet. Erkjennelsen av disse forhold ble gjort allerede ved utformingen av projektskrivet, og i

dette er det vist til som ønskelig å trekke på kompetanse fra og samarbeide med flere forskjellige fagområder, for på den måten å sikre en god og vellykket gjennomføring av prosjektet, se vedlegg.

I forbindelse med virksomhetsplanleggingen høsten 1994 ble ønskene og ambisjonene om å nå fram til et tverrfaglig samarbeid på prosjektet "Trygd-fobhistorie" konkretisert, i hvert fall til en viss grad. I virksomhetsplanen for året 1995 planla Trygdegruppa ved Seksjon 330, det vil si de statistikkansvarlige for prosjektet, å opprette kontakt og samarbeid med Gruppe for EDB (303) for å planlegge og starte gjennomføringen av prosjektet "Trygd-fobhistorie". Som en del av disse planene ble det også bestemt å nedsette en arbeidsgruppe med representanter fra Trygdegruppa og fra Gruppe for EDB, med det formål å gjennomføre et forprosjekt til prosjektet "Trygd-fobhistorie" tidlig i 1995.

Formål

Hovedformålet med forprosjektet har vært å ta opp og gjennomdrøfte viktige forhold som må avklares før hovedprosjektet "Trygd-fobhistorie" kan starte opp og bli gjennomført. Disse forhold er mangeartede. Blant de viktigste som Arbeidsgruppa skulle diskutere, var følgende:

- Klarlegge omfanget av hovedprosjektet "Trygd-fobhistorie"
- gi en oversikt over type data som skal inngå
- skissere hovedtyper av arbeidsoppgaver som må utføres for å gjennomføre hovedprosjektet
- vurdere dataorganiseringen
- vurdere nødvendig kompetanseoppbygging
- diskutere type samarbeid og samarbeidsformer som er nødvendig
- sette opp tidsplan for gjennomføringen av hovedprosjektet.

Det ble satt som mål for forprosjektarbeidet at dette skulle være avsluttet med en rapport fra Arbeidsgruppa innen 15. mars 1995.

3. Arbeidsgruppas sammensetning og arbeid

Arbeidsgruppa har bestått av personer som har ivaretatt det statistikkfaglige og trygdefaglige arbeidet i tilknytning til forprosjektet. I tillegg har Arbeidsgruppa hatt en sterk representasjon av personer med Edb-kompetanse. Arbeidsgruppa har hatt følgende sammensetning:

Grete Dahl (Seksjon 330, leder av forprosjektet)
Jorunn Lajord (Seksjon 330)
Else Flittig (Seksjon 330)
Knut Strøm (Seksjon 330)
Truls Thirud (Gruppe 303)
Hans Kristian Torvbråten (Seksjon 710)

I tillegg til disse faste representantene i Arbeidsgruppa, har Leif Korbøl, Seksjon 330, deltatt noe i dette arbeidet. Det er også i løpet av forprosjektarbeidet trukket på ekstra Edb-kompetanse fra personer ved Gruppe for EDB (303). Disse personene er Randi Wølner (tilrettelegging av "minidatabase" + opplæring) og Kristian Lønø (opplæring i SAS-Unix).

Arbeidet med forprosjektet har vært sammensatt. For en stor del har dette arbeidet likevel vært knyttet direkte til de drøftinger og beslutninger som er foretatt på møter i Arbeidsgruppa.

Arbeidsgruppa holdt sitt første møte 26. januar 1995. Det er holdt i alt 4 møter i Arbeidsgruppa, og siste møte ble holdt 8. mars 1995.

Denne sluttrapporten gir en dokumentasjon av arbeidet med forprosjektet. Rapporten er skrevet av Grete Dahl.

4. Resultater fra forprosjektarbeidet

Omfang av prosjektet "Trygd-fobhistorie"

Det datasystemet som skal kreeres i prosjektet "Trygd-fobhistorie", skal inneholde opplysninger for hele befolkningen og dekke alle årene fra og med 1992 til og med år 2000. Dersom prosjektet blir vellykket, ser vi for oss en videreføring på mer permanent basis.

Opplysningene som skal inngå i datasystemet skal hentes fra 13 forskjellige hoveddatafiler. Antallet inputfiler kan imidlertid vise seg å bli en god del større. Dataene skal organiseres som forløpsdata, og for å ta vare på endringene/"historikken" over tid, kan det bli aktuelt å nytte flere statistikkfiler (endringsfiler) som input i tillegg til disse 13 hoveddatafilene. Antallet inputfiler vil være avhengig av type systemløsning som blir valgt, av organiseringen av datasystemet, og av det forarbeid som blir gjort med hensyn til organisering av dataene før de legges inn i det endelige datasystemet. Vi viser for øvrig til punktet **Dataorganisering**, jf. nedenfor.

Til tross for at det kan bli aktuelt med flere inputfiler enn de 13 hoveddatafilene som er nevnt ovenfor, skal vi i det følgende gi en del opplysninger om disse hoveddatafilene. Disse filene er av ulik størrelse. Dette gjelder både med hensyn til antallet records som finnes på filene, og antallet kjennemerker som skal nyttes fra de forskjellige filene til prosjektet "Trygd-fobhistorie". Vi har nedenfor gitt omtrentlige anslag over antallet records på disse inputfilene. For en orientering om kjennemerker eller type data som skal inngå i datasystemet vises til punktet **Type data som skal inngå**, jf. nedenfor.

Type hoveddatafiler (inputfiler) og antallet records på disse filene:

	Antall records
- Situasjonsfil/fødelandsfil - befolkningsstatistikk	ca. 4.5 mill
- Utdanningsregisteret	3.8 "
- Inntektsregistre (både i SSB og RTV)	3.8 "
- Sosialstatistikkregisteret	0.2 "
- Sysselsettingsfil (A-/A-registeret)	3.1 "
- Sofasøkerregisteret	0.2 "
- GR1	0.9 "
- GR3	0.9 "
- Barnetrygdregisteret	0.3 "
- Fødsels- og sykepengeregisteret	0.3 "
- Attføringspengeregisteret	ca. 70 000 rec.
- Registeret for enslige forsørgere	" 60 000 "
- Registeret over avslag på uførepensjon	" 6 000 "

Av inputfilene danner filene fra befolkningsstatistikken basis for de opplysningene som skal inngå i datasystemet. Alle andre filer skal kobles til befolkningsstatistikkfilene. Blant de øvrige 12 hoveddatafilene vil det kunne være "overlapp" i den forstand at en og samme person på en og samme tid er registrert i to eller flere av disse filene. (En del personer vil også være registrert i to eller flere filer på en og samme tid på grunn av feil i registrene).

Type data som skal inngå

I oversikten som følger har vi gitt en grov oversikt over type data eller kjennemerker som skal nyttes fra de forskjellige hoveddatafilene til prosjektet "Trygd-fobhistorie". Vi gjør oppmerksom på at dette bare er en foreløpig oversikt. I arbeidet under fase I av hovedprosjektet, jf. kapittel 5, nedenfor, skal denne oversikten gjøres fullstendig.

Fra:

- Situasjonsfil/fødelandsfil- befolkningsstatistikk

- fødselsnr
- kjønn
- alder
- ekteskapelig status
- bostedskommune
- fødeland
- første oppholdsdato
- landbakgrunn
- innvandrerkategori
- kategorier for enslige forsørgere
- kategorier for samboere
- antall barn
- barnas alder

- Utdanningsregisteret

- utdanningsnivå (fullført utdanning/under utdanning)

- Inntektsregistre

- samlet inntekt (og grove inndelinger i inntektsposter)
- skatt

- formue
- pensjongivende inntekt (pensjonspoeng, omsorgspoeng etc.)
- disponibel inntekt

- Sosialstatistikkregisteret
- mottatt sosialhjelp/beløp
- perioder i året når sosialhjelp er mottatt

- Syssettingsfil (A-/A-registeret)
- arbeidsforhold i året, med start- og stoppdato
- forventet arbeidstid

- Sofasøkerregisteret
- registrert ledighet
- dagpenger
- arbeidssøkerstatus og tiltakskode

- GR1 og GR3
- alderspensjon
- uførepensjon
- etterlattepensjon
- AFP-pensjon
- uføregrad
- diagnose

- sluttpoengtall
- tilleggspensjon

- Barnetrygdregisteret
- status for enslig forsørger/samboer (Tilsv. kategorier hentes også fra andre registre)

- Fødsels- og sykepengeregisteret
- mottak av sykepenger, med start- og stoppdato
- sykepengegrunnlag
- kontonr
- utbetalt beløp
- bortfallsgrunn

- Attføringspengeregisteret
- mottak av attføringspengestønad, med start- og stoppdato
- type attføring
- utbetalt beløp
- bortfallsgrunn

- Registeret for enslige forsørgere
- overgangsstønad
- stønad til barnetilsyn

- Registeret for avslag på uførepensjon
- registrert avslag og når i året avslag er gitt
- diagnose
- avslagsårsak.

Dataorganisering

Opplysningene i datasystemet "Trygd-fobhistorie" skal organiseres som forløpsdata. Dataorganiseringen må likevel være slik at det på grunnlag av opplysningene i dette datasystemet er mulig å ekstrahere tverrsnittsdata som "stemmer med" offisiell statistikk på de forskjellige statistikkområdene. Dette gjelder både statistikk utarbeidet av Statistisk sentralbyrå, så vel som offisiell statistikk fra andre institusjoner, for eksempel Rikstrygdeverket.

Arbeidsgruppa har brukt en del tid til å drøftet hvilken dataorganisering som vil være mest hensiktsmessig å nytte i prosjektet "Trygd-fobhistorie". Det synes å være et valg mellom henholdsvis et svært system av flate filer og en databaseløsning. KIRUT, som er et lignende datasystem, er lagt opp som en forløpsdatabase, og litt mer praktisk anlagte lærebøker i forløpsanalyse synes også å anbefale en slik løsning for lignende type data.

De EDB-messige utfordringene i prosjektet utgjøres dels i håndtering av den store datamengden som skal vedlikeholdes, og dels i å etablere et system som er hensiktsmessig både for kontroll og analyseformål. Rundt regnet er datamengen på 7 GB, noe som tilsvarer ca. 5 ganger størrelsen på det nye befolkningsregistersystemet som er bygget opp i seksjon 320. Når det gjelder analysesiden, så er det nødvendig å spesifisere nærmere hva som menes med forløpsanalyse, bl.a. hvor store utvalg man skal analysere på, hvilke Edb-verktøy man ønsker å bruke, og hvilke krav til ytelse man krever eller forventer.

De Edb-faglige utfordringene i å etablere et fungerende system er store, og det er helt klart at systemet ikke vil likne noe annen statistikk rutine ved avdelingen. Dette betyr at det hersker stor usikkerhet om

utforming av den ferdige løsningen, noe som igjen peker i retning av at vi tar i bruk en strukturert arbeidsmetode i hovedprosjektet. Prosjektet bør ta stilling til verdien av å bruke systemutviklingsmetoden som SSB er i ferd med å etablere. Hvis metoden skal benyttes, bør den benyttes fra starten av hovedprosjektet.

Arbeidsgruppa kom ikke fram til noen endelig konklusjon på spørsmålet om dataorganisering, men Gruppa vurderte det som hensiktsmessig at de statistikkansvarlige (Trygdegruppa) skaffet seg noe kjennskap til prinsippene for databaser, dataprogram/verktøy brukt mot databaser etc., slik at de senere kan være bedre i stand til å delta i vurderingen av hvilken type dataorganisering som er mest hensiktsmessig. Det ble opprettet en "minidatabase" med data fra 4 aktuelle filer for et 2-prosent-utvalg av befolkningen, slik at Trygdegruppa kan nytte denne basen framover til å skaffe seg denne nødvendige innsikten. I denne forbindelse vil Trygdegruppa også lære seg en del om QBE-vision og eventuelt SAS-Access.

Kompetanseoppbygging

Prosjektet "Trygd-fobhistorie" skal legges på Unix-plattform, og det er stort sett programmer i Windows som vil bli anvendt til dette prosjektet. Dette betyr at Trygdegruppa vil trenge mye Edb-opplæring for å gjennomføre dette prosjektet, og denne kompetanseoppbyggingen kommer i tillegg til den statistikkfaglige "opplæringen" som nødvendigvis må følge som en del av prosjektet.

I den perioden forprosjektet har vart, har Edb-opplæringen kommet godt i gang. Gruppe for Edb (303) har gitt korte innføringer både i QBE-vision (på database) og SAS-Unix. Trygdegruppa har også deltatt i Excel-kurs.

Under den tiden prøveprosjektet har vart, har Trygdegruppa også blitt rustet opp med ny maskinvare. Trygdegruppa har fått to nye PC-er (med Windows), og en tredje PC er satt i stand for bruk av Windows.

Type samarbeid og samarbeidsformer

Som nevnt under kapittel 2, foran, er "Trygd-fobhistorie" et stort prosjekt som inviterer til samarbeid på tvers av faggrupper og kompetanseområder. Det er derfor sett på som vesentlig å etablere arbeids- og referansegrupper for dette prosjektet, samt å opprette og ha best mulig kontakt med andre seksjoner i SSB som skal være leverandører av data til prosjektet. Det er en fordel at det velges fleksible samarbeidsformer slik at disse kan tilpasses de krav til kompetanse som vil melde seg, og kanskje endre seg, etter hvert som prosjektet gjennomføres.

Når det gjelder kontakt med andre seksjoner i SSB, har Trygdegruppa allerede gjennomført et møte med Seksjon for befolkning og utdanning (320) for å drøfte og fastlegge hvilke data som skal hentes fra befolkningsstatistikkfilene og utdanningsfilene til prosjektet "Trygd-fobhistorie". Prosjektleder av forprosjektet har også tatt kontakt med Utne ("Personregnskapet"), og noen forberedende drøftinger er foretatt med tanke på å samordne datatuttakene fra befolkningsstatistikkfilene til henholdsvis "Trygd-fobhistorie" og "Personregnskapet".

I møte i Sosial-og helsedepartementet, 20. januar i år, ga Rikstrygdeverket et tilbud om å delta i en arbeidsgruppe sammen med Trygdegruppa for å være behjelpelig med overføring og forståelse av trygdefiler, samt å overføre erfaringene fra KIRUT til Trygdegruppa og til prosjektet "Trygd-fobhistorie". Denne arbeidsgruppa (med RTV) vil bli opprettet straks. Trygdegruppa vil også i nærmeste tid ta kontakt med NSD, Bergen, for å innhente informasjon om opplegget av KIRUT.

I løpet av sommeren 1995 vil det også bli opprettet en referansegruppe for prosjektet "Trygd-fobhistorie". Referansegruppa skal ha en sterk ekstern representasjon.

5. Plan for gjennomføring av hovedprosjektet "Trygd-fobhistorie"

På nåværende tidspunkt synes det rimelig å dele arbeidet med gjennomføringen av hovedprosjektet "Trygd-fobhistorie" inn i 3 faser med følgende tidsplan for gjennomføringen:

Fase I - Utviklingsfasen. Gjennomføring i perioden 15.03.95 - 15.09.95.

Fase II - Dataorganiseringsfasen. Gjennomføring i perioden 15.09.95 - 30.06.96.

Fase III - Produksjonsfasen. Gjennomføringen starter i 2. halvår 1996.

Fase I - Utviklingsfasen. Denne fasen strekker seg fra 15. mars til 15. september 1995. I denne fasen vil arbeidet med prosjektet bestå av følgende hovedoppgaver

- gjennomgå, kontrollere og dokumentere alle kjennemerker som skal inngå i prosjektet "Trygd-fobhistorie" for årene/årgangene 1992 og 1993. (Når det gjelder datakontroll vil denne bare omfatte kontroll internt innen hvert register. Kontroll av opplysninger på tvers av registre vil ikke være mulig å gjennomføre før i fase II av prosjektet)
- opprette de arbeidsgrupper som er nødvendig for å utføre arbeidet under fase I
- opprette referansegruppen. Et møte i denne gruppen skal arrangeres under fase I. Formålet med dette møtet er å få synspunkter på datainnholdet i prosjektet (type kjennemerker etc.)
- fortsette kompetanseoppbyggingen, via kurs, seminarer etc.
- ha ukentlige møter eller ta kontakt etter behov med Gruppe for EDB (303) for å drøfte/få hjelp til å løse Edb-problem etc.
- ta kontakt med NSD, Bergen, vedrørende opplegget av KIRUT
- innhente bistand (til prosjektet) i den hensikt å utøke kompetansen vedrørende dataorganisering og metodikk i forbindelse med forløpsanalyse. Denne bistanden/kompetansen bør det også kunne trekkes på under fase II av prosjektet
- lage en sluttrapport fra fase I. I tillegg til å beskrive oppnådde resultateter under fase I, bør denne rapporten også inneholde en tidsplan og et detaljert opplegg av arbeidet som skal utføres under fase II.

Grete Dahl er prosjektleder av fase I. Hun vil etter behov delegere arbeidsoppgaver til Trygdegruppa, og ellers ha kontakt med "berørte" etater, seksjoner/avdelinger i SSB og samarbeids- og referansegrupper.

De viktigste arbeidsoppgavene som skal utføres under fase II vil være å velge endelig dataorganisering/systemløsning, samt å sette sammen opplysninger på tvers av filer. Som en følge av denne databearbeidingen vil det oppstå behov for ytterligere kontrollarbeid. Parallelt med arbeidet under fase II vil det også bli laget foreløpige opplegg og planer for utarbeiding av statistikk og eventuelt analysearbeid med basis i det datagrunnlaget som vil inngå i "Trygd-fobhistorie". Det konkrete arbeidet med statistikkene/analysene skal etter planen starte i 2. halvår 1996, jf. omtalen av fase III.

I fase III av prosjektet regner vi med å ha kommet så langt at prosjektet er inne i en produksjonsfase, det vil si at dataene i datasystemet kan anvendes til forskjellig statistikk- og analyseformål. I denne fasen vil det likevel bli utført ganske mye arbeid med tillegg av nye data, kontroll av data etc., siden prosjektet "Trygd-fobhistorie" skal strekke seg over mange år. Som nevnt ovenfor, vil det bli lagt nærmere planer for fase II (og eventuelt for fase III) i forbindelse med utarbeiding av sluttrapporten fra fase I.

GDA, 10.11.93

Vedlegg

PROSJEKTSKRIV

Trygdedata og trygdehistorikk for anvendelse i modellbefolkningsprosjektet og i Fob 2000.

1. Dette er to nye, løpende prosjekt med start i 1994.
2. Prosjektene har som formål å bygge opp og utvikle datasystem som inneholder relevante trygdedata og annen statistisk informasjon for anvendelse i
 - modellbefolkningsprosjektet, og i
 - Fob 2000. (Som nevnt nedenfor, pkt. 5, er det imidlertid mange formål med dette prosjektet, utover den direkte anvendelsen av dataene i Fob 2000).
3. I det som følger er prosjektene omtalt som henholdsvis
 - Trygdedata til modellbefolkningsprosjektet, (forkortet: **Trygd-modellbefolkning**), pkt. 4, og
 - Trygdedata og trygdehistorikk til Fob 2000, (forkortet: **Trygd-fobhistorie**), pkt. 5.

Arbeidet som skal utføres under disse to prosjektene har faglig sett klare berøringspunkter, spesielt fordi kunnskap om og behandling av trygdedata vil stå meget sentralt i begge prosjektene.

4. Trygd-modellbefolkning.

Det årlige arbeidet med å koble trygdedata til modellbefolkningsprosjektet vil bestå i å

- i) innhente trygdefiler som skal anvendes i prosjektet, og foreta nødvendige kontroller av disse registeropplysningene.
- Kontakt med seksjon 510 og RTV.

ii) koble de opplysningene som ønskes fra trygdefilene til fødselsnumre i modellbefolkningsprosjektet (inntektsundersøkelsen). Spesifikasjon av hvilke trygdeopplysninger som skal nyttes, utarbeides av seksjon 510.
Kontakt med seksjon 510 og seksjon 420.

iii) gi innspill/forslag til eventuelle endringer i uttaket av trygdedata til modellbefolkningsprosjektet. (Innspill/forslag til endringer som følge av for eksempel endringer i registre i RTV, eller for å komplettere de trygdeopplysningene som til nå er tatt inn i prosjektet, jf. bl. a. pkt. 5, nedenfor).
Kontakt med seksjon 510.

iv) sette opp (i 1994) **en tidsplan for det årlige arbeidet** som skal utføres under pkt. ii).
Kontakt med seksjon 510 og seksjon 420.

(Vi regner med at arbeidet under pkt. i) vil følge den tidsplan som er nedfelt i SSBs samarbeidsavtale med RTV, og at eventuelt arbeid under pkt. iii) naturlig vil falle på omtrent samme tidspunkt av året som arbeidet under pkt. i)).

Ressursanslag for 1994 for arbeid utført ved seksjon 330: 200 tv.

5. Trygd-fobhistorie.

i) **Hovedformålet** med dette prosjektet er som nevnt å legge opp og utvikle et datasystem som skal anvendes i Fob 2000. Datasystemet skal gi "historiske" trygdeopplysninger og annen type statistisk informasjon som viser personenes tilknytning til trygd, arbeid/ledighet og andre sosiale situasjoner (utdanningssituasjon/sykdom/attføring/"utestenging" fra trygd/mottak av sosialhjelp etc.) i løpet av 1990-årene. Opplysningene skal foreligge (kryptert) på personnivå, og de vil i vesentlig grad tilføre en **sosial dimensjon** til øvrige registerdata og andre typer data som skal anvendes i Fob 2000. Den sosiale dimensjon er kjennetegnet ved at datasystemet vil inneholde "forløpsdata", som anvendt i Fob 2000, viser personenes tilknytning i løpet av 1990-årene, til ulike trygdeordninger/"utestenging" fra trygd/tilknytning til andre offentlige stønadsordninger/arbeid etc. I bred forstand vil datasystemet således gi vesentlig merinformasjon om personenes sosiale situasjon og forankring i år 2000, sammenlignet med for eksempel tilsvarende kjennemerker knyttet til bare i tverrsnitt (for året 2000).

Definisjoner og standarder i offisiell statistikk vil bli anvendt i dette prosjektet. Men i en del tilfeller nyttes ulike definisjoner for ett og samme begrep (eller variabel, kjennemerke etc.) i SSBs statistikk. Begreps- og definisjonsavklaringer som kan bidra til en mer ensartet praksis i SSB, er derfor forhold som vil bli tillagt vekt under arbeidet med prosjektet. **Spesielt for trygdedataenes del** vil trygdegruppa ha ansvar for å **utvikle og foreslå faste og ensartede definisjoner** for anvendelse i SSBs statistikk.

I tillegg til dette er det flere **underliggende formål** med prosjektet: - Prosjektarbeidet vil bidra til å avdekke inkonsistenser i ulike dataregistre (i SSB og i RTV/trygdeetaten), og således bedre grunnlaget for den registerbaserte statistikken i SSB.

- Dataene vil på sikt gi grunnlag for en "forbedret" input av trygdeopplysninger til modellbefolkningsprosjektet (pkt. 4, foran).
- Utover den direkte anvendelsen av datamaterialet i Fob 2000, vil datamaterialet også før år 2000 kunne nyttes som basis for statistikk i SSB. Materialet vil for eksempel være en rik datakilde for neste (og senere) Samlepublikasjon(er) om trygd.
- Datamaterialet vil kunne anvendes til spesielle analyseformål, både i og utenfor SSB (inntektsgivende oppdrag). Det vil føre for langt i dette notatet å skissere forskningsanvendelser av dette materialet. Men sett fra mitt ståsted, ser jeg spesielt betydningen av å bruke materialet til forskningsarbeid som er relatert til trygde- og sosialpolitiske forhold i samfunnet.

ii) Opplysningene i datasystemet skal knyttes til alle personer registrert bosatt **pr. 1. januar hvert år. Starttidspunktet er 1. januar 1992.** Når vi har valgt å starte i 1992 skyldes dette at datakvaliteten på

en del av de registrene som skal anvendes, ikke synes "god nok" for tidligere år. Perioden på 8 år (fra 1992 til 2000) skulle likevel være lang nok til å gi et brukbart bilde (i år 2000) av personenes tidligere "sosiale forhold".

iii) For befolkningen pr. 1. januar hvert år skal knyttes til opplysninger både for **samme år** og **året forut**, og hver gang knyttes til **samme type** opplysninger, jf. pkt. iv) - vi), nedenfor. På den måten innføres en link i systemet som i grove trekk gir opplysningene et preg av forløpsdata.

iv) Opplysningene i datasystemet vil bestå dels av **data pr. tidspunkt** og dels av opplysninger **for en gitt periode (ett år)**. Eksempler på data pr. tidspunkt er alder, kjønn, mottak av pensjon eller overgangsstønad pr. utgangen av året etc. Eksempler på opplysninger for året er attføringsperiode(r) i året (angitt med f.eks start- og stoppdato/type attføring), avslag på uførepensjon i året (med opplysning om f.eks avslagsårsak), antall arbeidsforhold i året (angitt med f.eks start-og stoppdato/arbeidstid) etc.

Avsnittene v) og vi) skisserer hvilke opplysninger datasystemet skal inneholde og hvilke registre som skal anvendes. Som vist til i pkt. vii), nedenfor, vil imidlertid en del av arbeidet med prosjektet i 1994 bestå i å spesifisere opplegget av datasystemet i detalj. Vi tar derfor forbehold om komplettering av de opplysningene som er gitt i avsnittene v) og vi).

v) Opplysninger i datasystemet.

Pr. tidspunkt:

- kjønn, alder, ekteskapelig status. Pr. 1.1.
- statsborgerskap, fødeland, botid, oppholdsgrunn og andre relevanteomkodede kjennemerker (for innvandrere). Pr. 1.1.
- utdanningsnivå(fullført utdanning), utdanningsnivå(under utdanning). Pr. 1.10.
- mottak av alders-, uføre-, etterlattepensjon (etterl. ektefeller, etterlatte fam. pleier, etterl.barn). Pr. 31.12.
(Vi vil her også ta inn mer trygdespesifikke kjennemerker. Kjennemerker som vil bli vurdert tatt inn er bl.a. minstepensjonist/ikke minstepensjonist, sluttpoengtall, størrelse på tilleggspensjonen, AFP-pensjonist/ikke AFP-pensjonist, uføregrad, diagnose etc.)
- mottak av overgangsstønad og stønad til barnetilsyn (evt. også utdanningsstønad). Pr. 32.12.
- kode for enslig forsørger. Pr. 31.12.
- arbeidssøkerstatus og tiltakskode. Pr. 31.12.

For året:

- arbeidsforhold i året, med start- og stoppdato, forventet arbeidstid.
- sykepengetilfeller i året, med start- og stoppdato, sykepengegrunnlag, bortfallsgrunn.
- attføringspengetilfeller i året, med start- og stoppdato, type attføring, bortfallsgrunn.
- avslag på uførepensjon, avslagsårsak.
- mottak av sosialhjelp i året, når i året sosialhjelp er mottatt.

vi) Registre som skal anvendes:

- personregisteret/fødelandsfil
- GR1 og GR3
- barnetrygdregisteret
- miniregister for enslige forsørgere
- register for avslag på uførepensjon
- fødsels- og sykepengeregisteret
- attføringspengeregisteret
- utdanningsregisteret

- sosialhjelpsdata
- sofasøker-registeret
- årsfil fra A-/A-registeret.

vii) Dette er som tidligere nevnt et løpende prosjekt. Arbeidet i 1994 vil bestå i å lage et detaljert opplegg for datasystemet. I den forbindelse må vi gjøre en del faglige og metodiske avklaringer, samt legge opp et system som edb-teknisk kan fungere tilfredsstillende. I tillegg vil de første koblingene bli foretatt. Før utgangen av 1994 skal det lages en rapport som skisserer det arbeid som er utført i løpet av året. I rapporten vil det bli lagt vekt på å dokumentere de dataene som (så langt) er lagt inn i systemet, vise til inkonsistenser etc.

Under arbeidet med dette prosjektet forutsetter vi å ha nær kontakt med flere seksjoner og personer i SSB. Vi nevner spesielt seksjonene 260, 510 og 320 + prosjektleder for Fob 2000 og registerprogrammet.

Søknad til Datatilsynet vil bli sendt umiddelbart etter godkjenning av prosjektet i SSB.

viii) Etter hvert bør det vurderes om dette datasystemet bør utvides, eller om det bør kobles til opplysninger fra andre data/datasystem som utvikles i SSB bl.a. med sikte på anvendelse i Fob 2000. Vi tenker her spesielt på tilknytning til et datasystem for private pensjonsordninger, hvis og når et slikt system blir opprettet, tilknytning til inntektsdata, familieopplysninger, og eventuelt til et system for mer detaljerte opplysninger om personers arbeidsmarkedstilknytning.

Ressursanslag for 1994:

- | | |
|--|-----------|
| - for arbeid utført ved seksjon 330: | 1 000 tv. |
| - for assistanse i metodiske spørsmål, fra metodegruppa: | 100 tv. |
| - for assistanse i edb-spørsmål, fra edb-gruppa: | 100 tv. |