

*Tonje M. Köber, Hossein Moafi,
Elisabeth Rønning og Øyvind Sivertstøl*

Bruk av forløpsdatabaser i Statistisk sentralbyrå

Forord

Dette notatet presenterer resultatene av prosjektet "bruk av forløpsdatabaser i Statistisk sentralbyrå (SSB)". Prosjektet ble levert inn fra Seksjon for leveårsstatistikk som forslag til utviklingsprogrammet Prosjekt- og samarbeid i SSB. SSB's forløpsdatabaser inneholder en stor mengde av opplysninger som er viktig for samfunnsforskning og offentlige utredninger. Formålet er å vise nytten av å ta i bruk data fra SSB's forløpsdatabaser med hovedvekt på FD-Trygd, i tillegg Nasjonal utdanningsdatabase (NUDB).

Prosjektgruppen bestod av Tonje M. Köber (prosjektleder, S260), Hossein Moafi (S230), Elisabeth Rønning og Øyvind Sivertstøl (S350). Grete Olsen har vært LOS, mens Berit Otnes og Elisabetta Vassenden har vært med i styringsgruppe for prosjektet.

Vi takker for nyttige innspill og tilbakemeldinger i arbeidet med forslag til Webside fra Thorleiv Valen, Egil Løksa, Halvor Strømme og Jan Bruusgaard.

Sammendrag

"Forløpsdatabaser i SSB" er en av de fem forslagene til prosjektarbeid som ble levert inn til utviklingsprogrammet Prosjekt- og Teamarbeid som varte fra 31. oktober 2001 - 5. april 2002. Hovedformålet med prosjektet er å synliggjøre SSB's forløpsdatabaser, og å vise nytten av å ta i bruk disse basene med vekt på FD-Trygd. Prosjektet omfatter også en kartlegging av i hvilken grad databasen FD-Trygd er tatt/tas i bruk, og hva som bør gjøres for å kunne øke bruken av denne forløpsdatabasen både internt og eksternt. En studie av dagens saksbehandlingsprosesser og eksisterende dokumenter inkludert veiledninger inngår også i prosjektets rammer. Målet er å drøfte hvordan ulike dokumenter eventuelt kan forbedres og saksbehandlingsprosessen forenkles.

For å oppnå målet om synliggjøring av SSB's forløpsdatabaser, holdt vi i alt 6 foredrag for 8 seksjoner i SSB som vi mente kunne være potensielle brukere av forløpsdatabaser. Foredragene ble innledet med en presentasjon av forløpsdatabasene FD-Trygd og NUDB. Denne presentasjonen ble etterfulgt med eksempler på ulike data fra FD-Trygd med det formål å illustrere nytten av å ta i bruk forløpsdatabaser. Aktiv dialog med møtedeltakerne vedrørende deres kjennskap og eventuelle fremtidige planer om å søke data fra forløpsdatabaser stod sentralt i foredragene.

Vi fikk positive tilbakemeldinger fra de som deltok på foredragene på hvor interessante og nyttige forløpsdatabaser virker/er. Samtidig ga noen uttrykk for en del hinder og problemer i den sammenheng. At de som jobber med utvalgsundersøkelser ikke får lov til å bruke registerinformasjon fra FD-Trygd ble nevnt blant hindrene de møter når de ønsker data fra FD-Trygd. Ellers vanskelighetsgraden med tanke på tekniske (IT) spørsmål når en bruker data fra basen, og manglende tid og ressurser for gjennomføring av omfattende analyser ble trukket frem som noen av problemene / hindringene interne potensielle brukere har.

Prosjektgruppen gjennomførte en spørreundersøkelse blant de som har brukt/er i ferd med å bruke data fra databasen FD-Trygd. Formålet med undersøkelsen var å kartlegge bruken av basen siden 1999, og ikke minst brukernes behov. Vi var ute etter å finne ut hvem som søker om data, hva slags data de vil ha, om basen er kjent og tilgjengelig, om basen mangler viktige kjennemerker osv. I undersøkelsen spurte vi respondentene også om deres møte med brukertjenesten FD-Trygd. Undersøkelsen omfattet både interne og eksterne brukere. Med interne brukere mener vi brukere som jobber i SSB, med eksterne mener vi brukere utenfor SSB.

Ut i fra svarene å dømme syntes samtlige respondentene at FD-Trygd er en interessant base med mange nyttige opplysninger som kan brukes i velferdsforskning og offentlige utredninger. De aller fleste oppga at de hadde fått data i samsvar med sine ønsker, men noen pekte på manglende kjennemerker ved basen, slik som sivilstand, barnetrygdregistret og opplysninger om statsansatte.

Ut i fra antall førstegangssøkere siden basens oppstart kom det frem at mens antall eksterne førstegangssøkere øker fra år til år, holder antall interne førstegangssøkere seg nokså stabilt. De fleste søkere blant både interne og eksterne søker om individdata. Vårt datamaterial tyder på at de som søker om data fra FD-Trygd jobber med ulike problemstillinger. Problemstillinger rundt enslige forsørgere, bruk av trygd og deltakelse på arbeidsmarkedet, førtidspensjonering og uførepensjonering er noen av de problemstillingene en har ønsket å belyse med data fra FD-Trygd.

En vurdering og studie av dagens saksbehandlingsprosess viser at denne prosessen er komplisert og inneholder flere uklare ledd. Det mest problematiske er prosessen frem til spesifisering av oppdraget, og klarering av forhold som berører juridiske spørsmål. I tillegg kan en peke på uklare retningslinjer om hvem i SSB skal ta seg av hva/hvilket oppdrag. Prosjektgruppen kommer med forslag om hvordan en kan forenkle saksbehandlingsprosessen. Det ene forslaget er en generell veiledning ved bestilling av mikrodata av personer til forskning og planlegging. Det andre forslaget er retningslinjer for deling av oppdrag mellom Brukertjenesten FD-Trygd og fagseksjoner.

Prosjektgruppens hovedkonklusjoner på det punktet er at for det første bør veiledninger til forløpsdatabaser berikes med mer informasjon om respektive databaser, og utfyllende informasjon om hvem som kan søke, hva en tilstrekkelig søknad bør inneholde, hvilke dokumenter som må legges ved søknaden osv. Elektroniske søknader lagt ut på nettet kan lette denne prosessen, og ikke minst forkorte saksbehandlingsprosessen.

For det andre bør en jobbe mot mer brukervennlige dokumenter om forløpsdatabaser. De eksisterende dokumenter er preget av teknisk faglige informasjon som noen ganger kan virke forvirrende for en "vanlig" bruker. En bør redigere dokumentene med å ta hensyn til hva en sluttbruker faktisk trenger av informasjon og opplysninger.

Til slutt må vi understreke viktigheten av å markedsføre SSB's forløpsdatabaser både internt og eksternt. Vi føler behovet for markedsføring av slike databaser i en mye større omfang enn det er tilfelle i dag. Det er opplagt at ikke bare forskere, men også politikere og byråkrater kan gjøres kjent med hvilke analysemuligheter som ligger i SSB's forløpsdatabaser. Strategier knyttet til synliggjøring av forløpsdatabaser bør settes opp i SSB's dagsorden. En ny utfordring i denne sammenhengen er å orientere seg mot det internasjonale markedet.

Innhold

1. Innledning.....	7
2. Forløpsdatabaser i SSB	9
2.1 Forløpsdatabaser	9
2.1.1 Datadesign	9
2.1.2 Dataorganisering og datering av tilstandsendringer.....	9
2.1.3 Beskrivelse av datamodellen for forløp	10
2.2 Forløpsdatabasen-Trygd (FD-Trygd)	12
2.2.1 Bestilling/forespørsler av data fra FD-Trygd.....	13
2.3 Nasjonal utdanningsdatabase (NUDB).....	17
3. Metode	19
3.1 Foredrag.....	19
3.1.1 Planlegging, gjennomføring og bearbeiding av foredragene	19
3.2 Dokumentanalyse	20
3.3 Spørreskjemaet	20
3.3.1 Datainnsamling	20
3.3.2 Behandling av data fra spørreskjemaene	20
4. Vise nytten - øke bruken	22
4.1 Informasjon - foredraget.....	22
4.2 Kommentarer og dialog	22
4.3 Oppsummering	23
5. Kjennskap, tilgjengelighet, bruk og nytte - brukernes erfaringer	24
5.1 Analyse av spørreskjemaene	24
5.1.1 Kjennskap til forløpsdatabasen FD-Trygd og det interessante med den.....	24
5.1.2 Bruk og markedsføring	24
5.1.3 Tilgjengelighet.....	25
5.1.4 Bruk og nytte	28
5.1.5 Mangler, ønsker og eventuelle problemer.....	30
5.2 Oppsummering	31
6. Forslag til forbedringer - fremtidige utfordringer	33
6.1 Dagens situasjon, og hva som bør forbedres	33
6.2 Forslag til forbedringer.....	34
6.2.1 Veiledning på Web	34
6.2.4 Retningslinjer for deling av oppdrag	40
6.2.2 Forenklet saksbehandlingsprosess	42
6.3. Utfordringer.....	44
6.3.1 Generelt dokumentasjonsnotat.....	44
6.3.2 Opplysninger om uttaksvalg	44
6.3.3 Veiledning på Web og deling av oppdrag.....	44
Vedlegg:	
1. Prosjektskriv	46
2. Fra miniregister/grunnlagsstatistikken til forløpstabell - et eksempel.....	48
3. Veiledning til bestilling av data som ligger på Web-siden til FD-Trygd	50
4. Foredrag.....	51
5. Spørreskjema	53
6. Grunnlagstabeller - resultater fra spørreskjema.....	57
De sist utgitte publikasjonene i serien Notater.....	61

1. Innledning

Statistisk sentralbyrå (SSB) innehar flere databaser på både individ og aggregert nivå som danner grunnlaget for produksjon og publisering av forskjellige typer statistikk, samt analyse av og forskning på ulike samfunnsfelt. En type database i SSB som ofte brukes i forskningssammenhenger og offentlige utredninger er forløpsdatabaser. Statistisk enhet i forløpsdatabaser er *individ*, "og opplysninger i databasen består av registreringer av begivenheter i den enkelte persons livsløp. Disse registreringene kan settes sammen til individuelle hendelser- eller forløpshistorier (*forløpsdata*), og aggregert over en gruppe av individer eller en hel populasjon, gir forløpsdataene oversikt over bevegelsene, *mellom suksessive begivenheter eller tilstandsendringer*" (Grete Dahl, "Notater 99/24", SSB1999). Forløpsdatabaser er en logisk sammenstilling av data hentet fra allerede eksisterende grunnlagsfiler for offisiell statistikk.

Data i forløpsdatabaser er strukturert og tilrettelagt for analyse av sammenhenger mellom fortiden, nåtiden og fremtiden. Det vil si at forløpsdatabaser gjør det mulig å ta ut individdata med ønskede kjennemerker til både tverrsnitts- og forløpanalyser når som helst. I hovedsak registreres tre typer hendelser i forløpsdatabaser: Tilganger, avganger og endringer. På den måten er det mulig å gjenskape alle tidligere situasjoner.

SSB's forløpsdatabaser inneholder stor mengde av data og opplysninger som er gull verdt for samfunnsforskning, og ellers i arbeidet med å belyse ulike problemstillinger knyttet til offentlige utredninger. Spørsmålet er hvor godt SSB's forløpsdatabaser er kjent, med andre ord er disse databasene blitt synliggjort overfor potensielle brukere både innen og utenfor SSB, og i hvilken grad blir disse databasene tatt i bruk? Det er også interessant å stille spørsmålet om tilgjengelighet, og spørsmålet om saksbehandlingsprosess fra innlevering av søknad om data frem til dataleveranse. Er det lett å finne frem til brukertjenester i SSB? Finner man nødvendige veiledninger både om datainnholdet (variabellister) og fremgangsmåten for å bestille data på SSB's Websider? Hvor lang tid det tar fra behandling av søknader og utlevering av data? Hvilke formelle lover og regler følges i saksbehandlingsprosessen? Brukervennligheten av utleverte data i forhold til riktig format og koblinger med andre datasett er også et annet spørsmål. Dette er forhold som ønskes belyst gjennom prosjektet.

Vi har forsøkt å kartlegge dagens situasjon med tanke på i hvilken grad disse databasene er "synlige" og tas i bruk, hvilke problemstillinger som en forsøker å belyse med hjelp av data fra forløpsdatabaser, samt å vise hvordan saksbehandlingsprosessen er. Dessuten har vi forsøkt å foreslå alternative løsninger og forenklinger der vi har funnet dagens situasjon problematisk og eller komplisert. En vurdering av eksisterende dokumentasjon med tanke på brukernes behov har vært en annen problemstilling i dette prosjektet.

Med hensyn til prosjektets omfang og ressurser, og av ulike metodiske og praktiske grunner har vi valgt å begrense vår studie til å omfatte forløpsdatabasene FD-Trygd og den nasjonale utdanningsdatabasen (NUDB). Mens den første basen har eksistert siden 1999, skal den andre basen være klar til bruk høsten 2002. Derfor har fokuset i prosjektet vært mot FD-Trygd, men erfaringene kan brukes i forhold til andre forløpsdatabaser i SSB.

I det kommende kapitlet gir vi først en presentasjon av FD-Trygd og NUDB ut i fra et historisk perspektiv og hva disse forløpsdatabasene inneholder av data og informasjon. En beskrivelse av prosessen fra forespørsel til levering av data inngår også i dette kapitlet.

I kapittel 3 diskuterer vi prosjektets avgrensninger noe nærmere, og beskriver datainnsamlings-metodene vi har brukt. Vi gjør også rede for systematisering og analyse av datamaterialet.

Vi foretar ulike analyser av datamaterialet og presenterer resultatene i kapitlene 4 og 5.

I kapittel 6 diskuterer vi mulige løsninger på svake og problemfylte punkter i dagens situasjon på grunnlag av hovedresultatene presentert i de to foregående kapitlene. Forslag til hvordan en kan bidra til mer synliggjøring av forløpsdatabaser, forbedringer av dokumentasjonsmaterialet og forenklinger av kompliserte saksbehandlingsprosesser er også blant de tema som diskuteres i kapittel 6. Kapitlet inneholder også fremtidige utfordringer. Med fremtidige utfordringer mener vi aktuelle problemstillingene som har falt utenfor dette prosjektets rammer, og videre arbeid med forslag til forbedringer.

2. Forløpsdatabaser i SSB

2.1 Forløpsdatabaser

Formålet med forløpsdatabaser på individnivå er først og fremst å tilby statistiske data til samfunnsforskning, offentlige utredninger og for intern produksjon av offisiell statistikk og publisering.

Forløpsdatabaser er organisert på en måte som gjør det mulig å anvende forskjellige forskningsdesign ved analyser som gjennomføres på grunnlag av data fra basen. Det er tatt spesielt hensyn til at de statistiske opplysningene i basen skal kunne nyttes til forløpsanalyser, det vil si til analyser av tilstander og hvor lenge en person befinner seg i en bestemt status før overgang til en ny og endret status, og hvilke faktorer som påvirker at personen går fra en status til en annen. Slike studier forutsetter data på individnivå, samt opplysninger over tidsperioder for hver person (forløpsdata eller hendelsesorienterte data).

De generelle formålene med forløpsdatabaser er:

- Tilrettelegge data som er egnet til studier av tilstander og til å belyse rekruttering til -, avgang fra - og overgang mellom ulike tilstander.
- Gi statistisk merinformasjon til surveydata
- Definere og tilrettelegge populasjoner for å trekke utvalg til statistiske undersøkelser
- Danne grunnlaget for offisiell statistikk
- Ta i bruk avansert verktøy til kontroll av løpende statistikk
- Base til oppslag

Forløpsdata kan f.eks. *fortløpende, over tid* vise personers tilknytning til trygd, økonomisk sosialhjelp, offentlige tjenstepensjonsordninger, arbeidssøking, inntektsgivende arbeid og varighet innen utdanning. I tillegg inneholder basene opplysninger som både *over tid* og på *gitte tidspunkter* kan nyttes for en nærmere karakteristikk av personene. Dette er for eksempel bakgrunnsopplysninger som demografi, inntekt og formue.

2.1.1 Datadesign

Tverrsnittsdata gir generelt opplysninger om tilstanden til et sett med individer på et gitt tidspunkt. Paneldata impliserer at det finnes opplysninger for de samme individene på to eller flere forskjellige tidspunkter. Bare forløpsdata eller hendelsesorienterte data gir opplysninger om tilstander mer eller mindre kontinuerlig over tid.

De aller fleste opplysningene i forløpsdatabaser er registrert som forløpsdata eller hendelsesorienterte data. Opplysningene består av registreringer av tilstandsendringer eller hendelser i den enkelte persons livsløp og av målinger av varigheten mellom suksessive hendelser. En hendelse for en person kan for eksempel være at en persons trygdestatus endres. Både tidligere trygdestatus, det vil si den status som personen skifter fra, og personens nye trygdestatus er da registrert, samt den tiden personen var i sin tidligere trygdestatus før overgang til ny status.

2.1.2 Dataorganisering og datering av tilstandsendringer

Opplysningene er organisert som “tabeller” i en Oracle-database. Det er en tabell for hvert temaområde eller hver ordning; for eksempel en tabell for uførepensjon, en tabell for igangværende utdanning, en tabell for viktigste kurs, en tabell for overgangsstønad til enslig forsørger etc. Opplysningene i hver tabell gir informasjon om når en person kommer inn i ordningen (tilgang), og når en person går ut av ordningen (avgang). I tillegg er det i hver tabell innført en del variable som måler om det skjer endringer i personens forhold til ordningen og eventuelt når slike endringer skjer.

Forløpsanalyse går ut på å analysere tidsintervall mellom suksessive tilstandsendringer eller hendelser med tiden som stokastisk variabel. Både tilgang til ordninger, avgang fra ordninger og endringer i personenes forhold til ordningene kan ses på som tilstandsendringer for personene. Disse tilstandsendringer registrerer vi ved hjelp av datovariabel. Datovariablene reflekterer hvor hyppig tilstandsendringene er målt.

Registrering og datering av hendelser

Avhengig av hvilke egenskaper ved tidsdimensjonen en ønsker å ivareta, har løsningene derfor en tendens til bli ganske uensartede. I FD-Trygd og NUDB har vi valgt en modell hvor vi i så stor grad som mulig prøver å registrere data i “kontinuerlig” tid, hvor kontinuerlig må forstås med utgangspunkt i den detaljeringsgraden til dateringene som faktisk forekommer på grunnlagsdataene. Det sentrale poenget er at en ved registrering i databasen ikke har valgt spesielle tellingstidspunkter, som f.eks. utgangen av hver måned. I stedet er det konstruert ulike typer dato- og hendelsesvariable som skal gjøre det mulig å fortløpende registrere de konkrete hendelse, med deres tilhørende tidspunkt.

Datoformater i databasen

Alle datoer i databasen er registrert med 4 siffer for århundre. På endel temaområder registreres data kun med årstall (YYYY), f.eks. inntekt og formue. På andre temaområder registreres data med år og mnd (YYYYMM), f.eks. utdanningsopplysninger, pensjonsytelser i folketrygden, stønader til enslige forsørgere, sosialhjelp etc., eller kalenderdag (YYYYMMDD), f.eks. fødsels- og sykepenger, attføringspenger, sysselsetting etc.

Metoder for registrering av daterte hendelser

Oppdatering av forløpstabellene er basert på at det kun foretas registrering hvis det skjer noe “nytt”, dvs. at en ny hendelse inntreffer. En person som kommer inn i en ordning blir registrert med en tilgangsrecord hvor startdato for tilfellet, som er første dato vedkommende er aktiv i ordningen, er angitt både i tilgangsdato (tilgdato) og en registreringsdato (regdato) som skal alltid være lik sist daterte opplysning. Tilgangsrecorden med de tilhørende kjennemerker representerer deretter personens tilstedeværelse i ordningen inntil noen av variablene (kjennemerkene) endrer verdi, eller vedkommende går ut av ordningen. Hvis en variabel endrer verdi blir det registrert en endringsrecord, hvor regdato angir hendelsestidspunktet for endringen. Hvilken variabel som har endret verdi angis via variabelens tilhørende datovariabel, som skal ha samme verdi som regdato. Denne nye endringsrecorden vil fra hendelsestidspunktet (regdato) representere status for personen innenfor ordningen fram til en eventuelt ny endring oppstår, hvorpå det vil registreres en ny endringsrecord med en ny hendelsesdato osv. Dette er eksemplifisert i vedlegg 2.

Når en person går ut av en ordning registreres det en avgangsrecord, hvor siste aktive dato for tilfellet angis i avgdato (og regdato). Hvis samme person kommer tilbake på et senere tidspunkt blir et nytt tilfelle startet opp med en ny tilgangsrecord og tilgdato.

2.1.3 Beskrivelse av datamodellen for forløp

Det datadesign som her er beskrevet kan sammenfattes i følgende datamodell for oppstilling av forløpsdata:

Generell tabellstruktur for forløpstabeller:

<i>Kortnavn</i>	<i>Beskrivelse</i>	<i>Kommentar</i>
fnr	Fødselsnummer	Identifiserer person
tilgdato	Tilgangsdato (tilfelle)	Identifiserer startdato for et tilfelle, og er alltid utfyllt. Alle records hvor fnr + tilgdato er like refererer til samme tilfellet
avgdato	Avgangsdato (tilfelle)	Identifiserer stoppdato for et tilfelle. Kun utfyllt på avgangsrecords. Tilfeller som ikke har en slik avgangsrecord er fortsatt aktive.
temakode	Gruppe og hendelse	Identifiserer gruppe innenfor aktuell ordning, samt hvilken hendelse den aktuelle recorden beskriver. Betydningen av gruppekodene varierer mellom de ulike temaområdene, mens hendelseskodene er standardisert på tvers av temaområdene. Hendelseskodene er: 0 = Avgang 1 = Tilgang 2 = Endring 3 = Venstresensur (tilgang)
regdato	Registreringsdato	Første gyldighetsdato (referansedato) for den registrerte hendelsen
tematilg	Tilgangsdato (gruppetilfelle)	Identifiserer startdato for et gruppetilfelle. Alle records hvor fnr + tematilg er like hører til samme gruppetilfellet. Records som hører til samme gruppetilfellet skal samtidig høre til samme tilfelle.
temaavg	Avgangsdato (gruppetilfelle)	Identifiserer stoppdato for et gruppetilfelle. Kun utfyllt på gruppeavgangsrecords. Gruppetilfeller som ikke har en slik avgangsrecord er fortsatt aktive.
iv1	Variabel 1	Kjennemerke 1 knyttet til (gruppe) tilfellet
iv1dato	Variabeldato 1	Datering av første gyldighetsdato for denne kjennemerkeverdien
iv2	Variabel 2	Kjennemerke 2 knyttet til (gruppe) tilfellet. Konstant, dvs. ingen datering
...		
ivn	Variabel n	Kjennemerke n knyttet til (gruppe) tilfellet.
ivndato	Variabeldato n	Datering av første gyldighetsdato for denne kjennemerkeverdien

Utfyllende kommentarer til tabellen

- Alle opplysninger om forløp er knyttet til en person via fødselsnummeret. Videre er alle opplysninger knyttet til et tilfelle via kombinasjonen fnr + tilgdato. Tilgdato representerer startdato for en periode hvor en person er “aktiv” innenfor et temaområde. Alle records som er tilknyttet dette tilfellet, dvs. tilgangs- og avgangsrecord samt evt. endringsrecords, har samme verdi for tilgdato.
- I alle forløpstabellene skal de periodene en person har vært “aktiv” innenfor et tema (ordning), være definert ved datoverdiene i tilgdato og avgdato som angir start- og stoppdato hhv. for det aktuelle tilfellet. Hvis det ikke forekommer noen avgangsrecord for et tilfelle, med referansedato innenfor observasjonsperioden, er tilfellet fortsatt aktivt ved utgangen av denne perioden.
- Det er knyttet en egen datovariabel til hver variabel som kan endre verdi i løpet av den perioden et tilfelle strekker seg over (illustrert ved iv1 og iv1dato), slik at det skal være mulig å se når denne eventuelt skifter verdi. Variable som ikke kan/vil endre verdi i løpet tilfellets periode (illustrert ved iv2) behøver ikke noen egen datovariabel, siden dateringen av variabelens verdi er bestemt av tilfellets (samlede) periode. Alle endringsdateringer er knyttet opp mot gruppenivå (hvis det er mer enn et nivå).

- Venstresensur er en hendelseskode for tilgang som angir at dateringen av tilgangen skyldes tidspunkt for oppstart av databasen, og ikke at vedkommende faktisk kom inn i ordningen på dette tidspunktet. På de temaområdene hvor det finnes egne datovariabler for tilgangsdato på grunnlagsfilene kan den registrerte tilgangsdatoen ligge lenger tilbake i tid enn tidspunktet for oppstart av databasen. I disse tilfellene registreres denne dato som tilgangsdato i databasen, mens venstresensur-koden angir at vi har en spesiell tilgang.
- I de faktiske tabellene erstattes prefikset *tema* som regel med en kortform som indikerer hvilket tema det her er snakk om.

Avvik fra den generelle tabellen

På enkelte temaområder forekommer kun én gruppe (dette gjelder f.eks. for sosialhjelp og foreløpig uførestønad), og datovariablene tilgdato/avgdato blir da sammenfallende med tematilg/temaavg. Forløpstabellene for disse temaområdene inneholder kun ett par datovariabler for tilgang og avgang, siden det andre paret er overflødig.

Recordtypenes informasjonsinnhold

Generelt vil man ikke finne ferdig definerte perioder på den enkelte record, med unntak for avgangsrecord som angir start- og stoppdato for hele tilfellet. Periodene må i stedet bygges opp ved å se recordene for samme tilfellet i sammenheng. For å forstå logikken knyttet til ulike typer uttak er det viktig å være klar over hva slags informasjon som er registrert på de ulike recordtypene. Nedenfor følger en oversikt som beskriver dette i detalj.

Recordtyper og informasjonsinnhold:

- Venstresensurerte tilgangsrecords inneholder opplysninger om;
 - Tilgangsdato for det aktuelle tilfellet (tilgdato), venstresensurert
 - Tilgangsdato for det aktuelle gruppetilfellet (tematilg), venstresensurert
 - Status for alle intervareblar fom. referansedato (regdato = iv-dato = første dag/mnd i 1992)
- Ordinære tilgangsrecords inneholder opplysninger om;
 - Tilgangsdato for det aktuelle tilfellet (tilgdato)
 - Tilgangsdato for det aktuelle/første gruppetilfellet (tematilg)
 - Status for alle variable fom. referansedato (tilgdato/regdato = iv-dato)
- Endringsrecords inneholder opplysninger om;
 - Tilgangsdato for det aktuelle tilfellet (tilgdato)
 - Tilgangsdato for det aktuelle gruppetilfellet (tematilg)
 - Status for alle variable på referansedato (regdato)
 - Status for hver variabel fom. sist registrerte iv-dato tom. referansedato (regdato)
 - Hvilke(n) variable som har endret verdi og “forårsaket” denne endringsrecorden
- Avgangsrecords inneholder opplysninger om;
 - Tilgangsdato for det aktuelle tilfellet (tilgdato)
 - Tilgangsdato for det aktuelle/siste gruppetilfellet (tematilg)
 - Status for alle variable på referansedato (regdato)
 - Status for hver variabel fom. sist registrerte iv-dato tom. referansedato (regdato)

Tilgangs- og endringsrecords inneholder ikke opplysninger om “stoppdato”, hverken for det aktuelle tilfellet eller variabelverdiene som er registrert på recorden. Siste gyldighetsdato for verdiene på disse recordene må hentes fra en etterfølgende endrings- eller avgangsrecord.

2.2 Forløpsdatabasen-Trygd (FD-Trygd)

Som pilotprosjekt ble forløpsdatabasen FD-Trygd konstruert for å dekke behovet innen velferdsforskningen og for å levere tryggedata til folketellinga FOB2000. Trygdesystemet har en sentral plass i velferdssystemet og opplysninger om personers trygdestatus og trygdeforhold utgjør en stor del av dataene i basen.

FD-Trygd inneholder opplysninger primært om trygdeforhold, hvor det er opplysninger om forskjellige typer pensjoner, som uførepensjon, avtafestet 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 temaområder. Disse områdene er demografi, økonomisk sosialhjelp, arbeidssøking og sysselsetting. FD-Trygd ble åpnet for offentlig bruk i februar 1999.

Temaområdene i FD-Trygd

- Bakgrunnsopplysninger
 - demografi (1992-1997)¹
 - inntekt og formue (1992-1999)
 - utdanning (1992-1997)²
- Trygdeopplysninger
 - folketrygdpensjoner (1992-2000)
 - grunn- og hjelpestønader (1992-2000)
 - foreløpig uførestønad (1992-1998)
 - enslige forsørgere (1992-1999)
 - økonomisk sosialhjelp (1992-1997)
 - fødsels- og sykepenger (1992-1999)
 - attføring (1992-1999)
- Arbeidsmarkedsdata
 - sysselsetting (1992-1998)
 - arbeidssøkere (1992-1999)
 - statsansatte (1992-1999)

Ved hjelp av forløpsdataene i basen er det for eksempel mulig å studere hvordan personer over tid "beveger seg" mellom forskjellige typer trygdestatuser eller for eksempel fra trygd til arbeid.

Hovedformålet med FD-Trygd er å bygge opp en database med statistiske data på individnivå til velferdsforskning. Et underliggende mål er å gi opplysninger om hvordan velferdssystemet fungerer for enkelt individer, samt beskrive de persongruppene som mottar ytelser fra forskjellige velferdsordninger. I tillegg evaluere effektene av regelverksendringer eller strategiltak knyttet til velferdsordningene.

Dataene i FD-Trygd bygger på opplysninger i administrative registre som er hentet inn fra Rikstrygdeverket (RTV), Arbeidsmarkedsstaten (Aetat) og Statistisk sentralbyrå (SSB). Basen inneholder opplysninger fra og med 1992 og skal bygges ut til å omfatte data til og med år 2000. Prosjektperioden er til og med årgang 2000, men det er planlagt at basen skal videreføres til og med 2010. Opplysningene gjelder hele befolkningen.

FD-Trygd kan ses på som en videreutvikling av prosjektet Kirut (Klientstrømmer inn, rundt, ut av trygdesystemet), som i de senere årene har vært en viktig datakilde for trygde- og sosialpolitisk forskning. Kirut inneholder data fra 1989-1996, og Norsk Samfunnsvitenskapelig Datatjeneste (NSD), står for utlevering av data fra dette datasystemet. Kirut inneholder 10 prosent av befolkningen mellom 16-74 år, men som nevnt tidligere FD-Trygd inneholder hele befolkningen.

2.2.1 Bestilling/forespørsler av data fra FD-Trygd

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 for filer for brukerne. Når det gjelder det øvrige arbeidet som brukertjenesten skal utføre, er dette delt mellom de to etatene. Spesielt

¹ Hvor tallene i parentes viser hvilke årganger som per 1.januar 2002 er i basen.

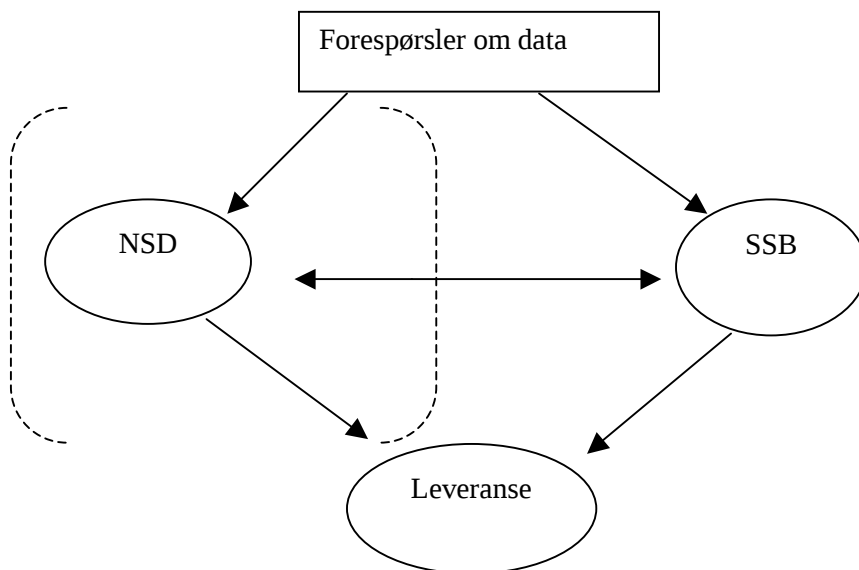
² Avsluttet som temaområde i FD-Trygd, vil etter hvert erstattes av Nasjonal utdanningsdatabase (NUDB)

skal arbeidet med å anonymisere data som leveres ut til forskere og studenter utføres av NSD, mens SSB gjør dette selv for data som leveres ut til offentlig planlegging og andre brukere. SSB har inngått en avtale med NSD om videreformidling av data fra SSB til forskere. Avtalen går i korthet ut på at NSD mottar aidentifiserte data fra SSB som NSD anonymiserer før utlevering. På internettssidene til FD-Trygd ligger det informasjon/veiledning om hvordan bestille data fra FD-Trygd (vedlegg 3).

Prosessdiagram 1 og 2 viser prosessen fra forespørsel til datautlevering av FD-Trygd data. Vi har valgt å vise prosessen ved datautlevering fra FD-Trygd, men denne prosessen vil være tilnærmet lik for de fleste utleveringer av mikrodata fra SSB.

Prosessdiagram 1 nedenfor viser hvordan forespørsler fra forskere, studenter, offentlige etater og andre går til NSD eller SSB og til endelig leveranse av data. Henvendelser til NSD og samarbeidet mellom NSD og SSB er utenfor prosjektets område. Men som nevnt tidligere er det retningslinjer for hvilken instans som skal behandle de ulike søknadene.

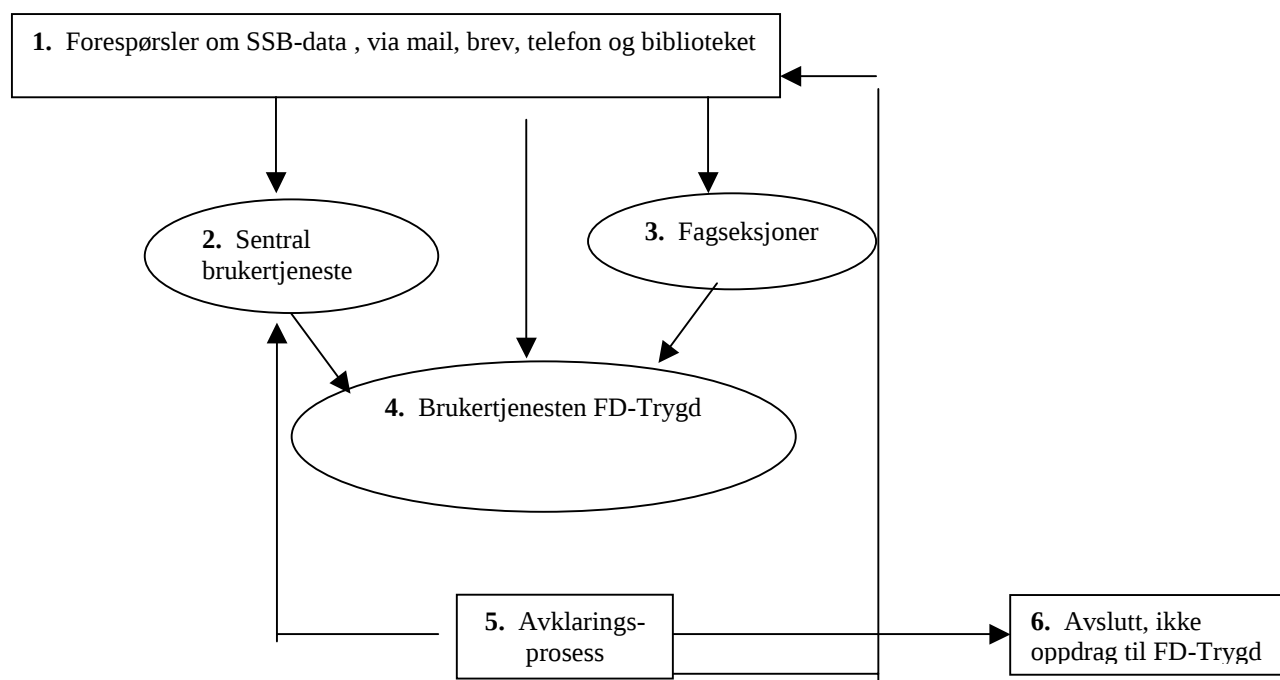
Prosessdiagram 1. Oversikt over mottak av forespørsler av data fra FD-Trygd.



Prosessdiagram 2a og 2b nedenfor viser prosessen ved forespørsler innad i Statistisk sentralbyrå.

Hvor **Prosessdiagram 2a** viser den **innledende prosessen**. Det vil si oversikt over prosessen ved forespørsler av FD-Trygd data til SSB og til avklaringsprosessen. Henvendelser om FD-Trygd data kommer inn via biblioteket, mail, telefon eller brev til brukertjenesten FD-Trygd (BT), Sentral Brukertjeneste (SB) eller fagseksjoner (boks 1-4). Alle henvendelser må igjennom en avklaringsprosess (boks 5), hvor det blir oppnevnt en saksbehandler ofte utført av brukertjenesten FD-Trygd, hvor bestillingen blir konkretisert og bekreftet. Noen henvendelser vil bli avsluttet allerede etter saksbehandling/avklaringsprosessen (boks 6), fordi FD-Trygd ikke var det etterspørteren var ute etter.

Prosessdiagram 2a. Innledende prosess - fra forespørsler av FD-Trygd data til SSB og til avklaringsprosessen



Prosessdiagram 2b viser den **formelle prosessen**. Det vil si oversikt over prosessen ved fordeling av oppdrag, samarbeid i SSB og vurdering av søknader og til utlevering av data. Det enkleste tilfelle av oppdrag vil være der hvor kunden kun ønsker data fra FD-Trygd. Disse dataleveransene vil som oftest utføres av brukertjenesten FD-Trygd, men ved noen oppdrag vil Sentral brukertjeneste utføre leveransen (boks 10a). Det vil også være tilfeller hvor andre fagseksjoner bidrar med data. Dette utføres ofte ved at kunden selv henvender seg til den aktuelle fagseksjonen (boks 11) som et eget oppdrag. En annen mulighet er at Sentral brukertjeneste utfører hele oppdraget ved å bestille data fra FD-Trygd og den aktuelle fagseksjonen og leverer samlet ut data til kunden (boks 10b). Når flere seksjoner bidrar med data eller Sentral brukertjeneste koordinerer oppdraget vil det være noe saksbehandlingstid mellom de ulike instanser innad i SSB, for å avklare/bekreft bestillingen. Det kan også være tilfeller av avslutt av oppdrag (boks 8) i denne delen av saksprosessen.

Bestillingen skal inneholde beskrivelse av formålet med datautleveringen, hva det skal brukes til og også om det skal kobles til egne utvalg eller spørreundersøkelser. En foreløpig henvendelse (muntlig eller skriftlig) kan avklare hvilke opplysninger som er nødvendig å oppgis for videre saksbehandling. Databehandling og leveranse kan imidlertid ikke starte før formell og skriftlig bestilling er endelig godkjent. Dette impliserer også at tabellvalg og variabeltilfanget må være spesifisert, grunnet personvern hensynet i forhold til sensitive opplysninger.

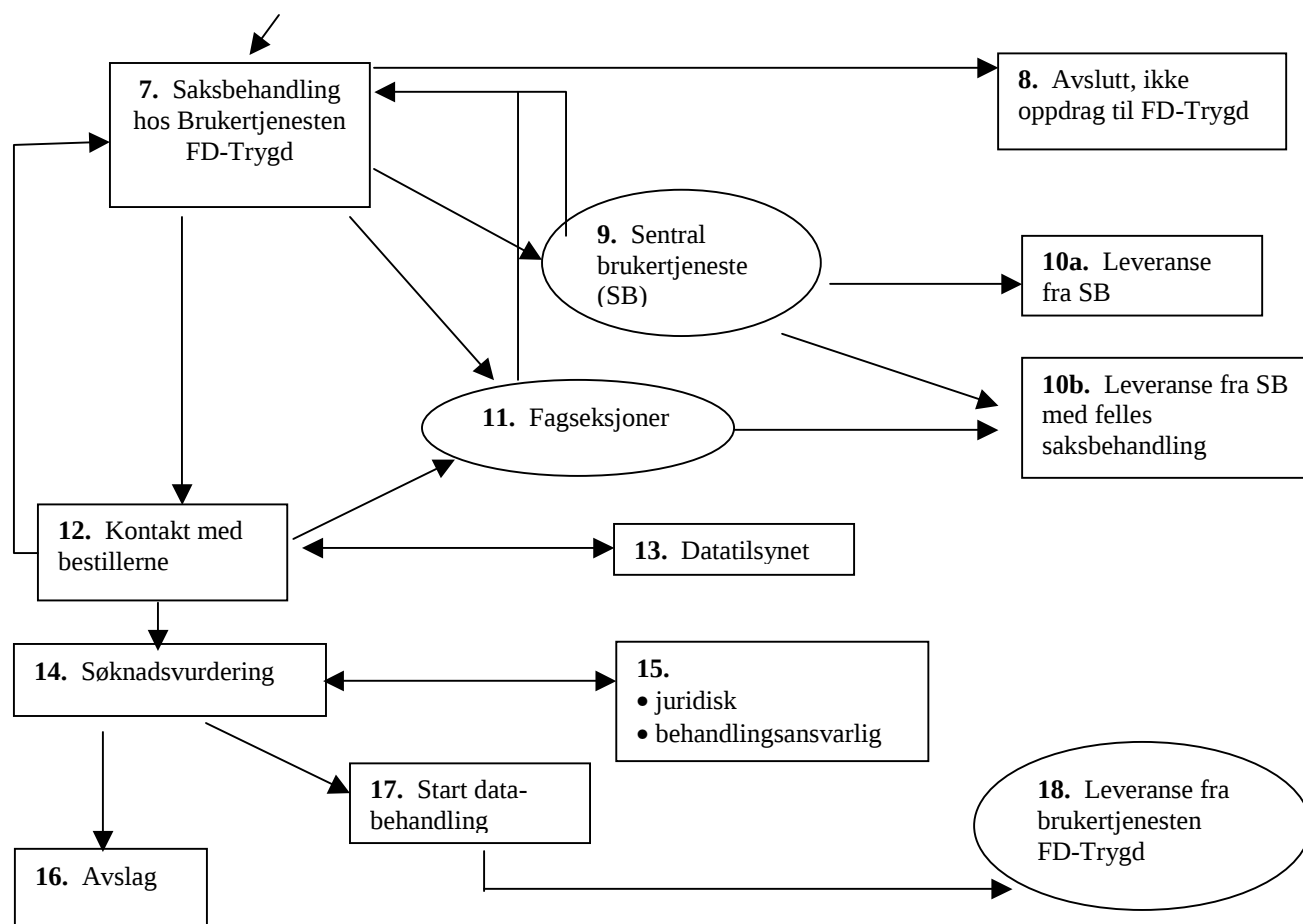
Det er to hovedtyper data som leveres ut, aidentifisert eller anonymisert. Hvis bestillingen inneholder sensitive data må brukeren få konsesjon fra Datatilsynet. Hvis det er ikke-sensitive data, men personidentifiserbare krever SSB melding fra Datatilsynet. Anonymiserte data krever ingen konsesjon eller melding. Dette fører til at brukeren selv må ta kontakt med Datatilsynet (boks 13), men det er ønskelig at kopi av melding eller konsesjon er vedlagt søknaden. I tillegg må brukertjenesten FD-Trygd i kompliserte tilfeller samarbeide med andre fagpersoner (boks 15) internt, for å avklare det juridiske spørsmålet eller om bestillingen er i samsvar med konsesjon/ melding til Datatilsynet. I tillegg om beskrevne prosedyrer for kobling og etterbehandling av data er i samsvar med SSBs krav og rutiner (sikkerhetshåndboken) og om dataleveransen krever dispensasjon fra taushetsplikten i folketryktdloven/ forvaltningsloven.

Det er vesentlig for saksbehandlingstiden at kunden beskriver formålet med dataleveransen ut ifra disse kriterier på et så tidlig stadium som mulig.

Annen type spesifisering som bestemmelse av utvalget, periodevalg - tverrsnittsuttak, type opplysninger (start- og stopp-, eventuelt endringsopplysninger) kommer i tillegg til den avklaringsprosess som så langt er beskrevet og er gjerne en del av den videre saksbehandling. Når oppdraget er bekreftet og eventuelt spørsmålet om konsesjon er avgjort vil selve databehandlingen og leveringstiden avhenge av oppdragets størrelse/ kompleksitet og mengde oppdrag som ligger i kø.

Det vil da være tre mulige dataleveranser av FD-Trygd data fra Statistisk sentralbyrå; leveranse fra Sentral brukertjeneste med data kun fra basen (boks 10a) eller leveranse fra Sentral brukertjeneste med felles saksbehandling der hvor flere seksjoner bidrar med data (boks 10b) eller leveranse fra brukertjenesten FD-Trygd (boks 18).

Proses diagram 2b. Formelle prosessen - fra fordeling og samarbeid av oppdrag i SSB til utlevering av data.



2.3 Nasjonal utdanningsdatabase (NUDB)

Det er utviklet en ny og mer detaljert forløpsdatabase for utdanningsområdet. Nasjonal Utdanningsdatabase, NUDB, som inneholder muligheten til å beregne gjennomstrømningshastighet og normert studietid etter bestemte regler for de som har avsluttet en utdanning. Og det er ytterligere spesifisert skillet mellom kursnivå og personnivå enn i FD-Trygd. Nasjonal utdanningsdatabase er klar til bruk ved utgangen av år 2002, og er utviklet på modell etter FD-Trygd.

Hovedformålet med utdanningsdatabasen er å sammenstille utdanningsdata for å få et statistikkssystem som i større grad enn tidligere, enkelt gir muligheter til å gjennomføre bredere og større analyser av all den informasjonen utdanningsstatistikkfilene inneholder, spesielt gjennomstrømningsanalyser. I tillegg skal denne databasen brukes til statistikkproduksjon og som verktøy i kvalitetssikring av statistikken.

NUDB omfatter elever og studenter fra 1974, avsluttet utdanning fra 1974 og befolkningens høyeste fullførte utdanning fra 1970. De viktigste opplysningene er utdanningsvariable som utdanningens art, klassetrinn, studieretning, vektall og fullført/ikke fullført. I tillegg inneholder basen skoleopplysninger og personopplysninger som for eksempel statsborgerskap, innvandringskategori og foreldrenes utdanningsnivå.

Nasjonal utdanningsdatabase kan ta utgangspunkt i en hvilken som helst gruppe utdanninger/personer/kull/skoler osv., på et hvilket som helst tidspunkt, og følge disse enten fremover eller bakover gjennom utdanningssystemet.

3. Metode

Problemstillingen i prosjektet "Bruk av forløpsdatabaser i SSB" er for det første å vise nytten og øke bruken av forløpsdatabaser, og for det andre å vurdere eksisterende dokumentasjon med tanke på brukernes behov.

Hvordan går vi fram for å finne ut noe om dette? Når det gjelder den første delproblemstillingen, å vise nytten av og øke bruken, fant vi at foredrag var egnet datainnsamlingsmetode.

For den andre delproblemstillingen, å vurdere eksisterende dokumentasjon med tanke på brukernes behov, måtte vi for det første selv gå gjennom eksisterende dokumentasjon for forløpsdatabasen FD-Trygd. Vi mente også at det var viktig å ta kontakt med brukerne av en forløpsdatabase og vi valgte derfor å lage et spørreskjema, som ble sendt ut til brukere av forløpsdatabasen FD-Trygd. Personlig intervju med ulike brukere ble også vurdert, men på grunn av prosjektets begrensede ressurser, ble spørreskjema med en del åpne spørsmål og svar valgt.

Utvalget vårt kan deles opp på to måter. For det første interne og eksterne brukere. Med interne mener vi brukere som jobber i SSB, med eksterne mener vi brukere utenfor SSB. For det andre kan vi dele brukerne i aktuelle og potensielle brukere. Med aktuelle mener vi de som bruker forløpsdata per i dag og med potensielle mener vi personer eller institusjoner som kan ha nytte av å bruke forløpsdata, men som ikke gjør dette ennå. Både interne og eksterne brukere kan deles opp i aktuelle og potensielle. Vårt prosjekt avgrenser seg til aktuelle og potensielle interne og aktuelle eksterne brukere.

3.1 Foredrag

3.1.1 Planlegging, gjennomføring og bearbeiding av foredragene

Utvalget til foredragene ble begrenset til potensielle interne brukere av forløpsdatabaser i SSB. Vi tok for oss organisasjonskartet i SSB og valgte ut seksjoner som jobber med personstatistikk, og som vi dermed mente kunne ha nytte av å bli informert om forløpsdatabaser. På avdeling for personstatistikk holdt vi foredrag for Seksjon for levekårsstatistikk, Seksjon for helsestatistikk, Seksjon for befolknings- og utdanningsstatistikk og Seksjon for demografi- og levekårsforskning. På avdeling for økonomisk statistikk holdt vi foredrag for Seksjon for arbeidsmarkedsstatistikk. På avdeling for næringsstatistikk holdt vi foredrag for Seksjon for inntekt- og lønnsstatistikk. På Forskningsavdelingen ble alle seksjoner invitert til foredrag. Utvalget vårt besto altså av i alt 8 seksjoner ved Statistisk sentralbyrå, fordelt på 6 foredrag.

Hvert av foredragene varte opptil 1 time, og var delt i tre sekvenser (se vedlegg 4). I første sekvens stilte vi seksjonene spørsmål om hva de visste om forløpsdatabasene i Statistisk sentralbyrå - FD-Trygd og NUBD. Hvor godt kjent er databasene hos potensielle interne brukere? Andre sekvens inneholdt informasjon om forløpsdatabaser generelt, NUBD og FD-trygd spesielt, med hovedvekt på FD-trygd. Formålet med denne sekvensen var å informere om og vise nytten av forløpsdatabaser. I tredje sekvens stilte vi igjen spørsmål til seksjonene: Har informasjonen økt interessen for forløpsdatabaser generelt og FD-Trygd spesielt? Kan de tenke seg prosjekter på seksjonene der forløpsdatabaser er relevant? Formålet med den siste sekvensen var å få en indikasjon på om foredragene førte til at seksjoner som ikke hadde så stor kjennskap til forløpsdatabaser var blitt mer interessert. På sikt kan økt bruk være en konsekvens.

Problemstillingen vi forsøker å få svar på ved hjelp av foredragene er "å vise nytten av og øke bruken av" forløpsdatabaser. Når alle foredragene var gjennomført ble informasjonen fra foredragene organisert slik at de kunne analyseres på en hensiktsmessig måte. Foredraget var planlagt på forhånd, og dette er i seg selv med på å strukturere materialet. Vi sammenliknet svarene fra de ulike seksjonene, og så etter likheter eller forskjeller: Var erfaringene de samme hos de ulike seksjonene?

Var det noen seksjoner eller grupper som skilte seg ut? Hva var hovedinntrykket etter foredragene? Var inntrykket at flere ønsker å bruke forløpsdatabaser etter presentasjonen vår?

3.2 Dokumentanalyse

Å vurdere eksisterende dokumentasjon i forhold til brukernes behov er viktig i seg selv, og også viktig i forhold til hovedproblemstillingen vår: Å øke bruken og vise nytten av forløpsdatabaser. Er dokumentasjonen god nok i dag? Vi valgte her å fokusere på forløpsdatabasen FD-Trygd. Spørsmål vi stilte oss var hvordan tilgjengeligheten er, hvordan saksbehandlingsprosessen fra innlevering av søknad fram til dataleveranse er, og om det er lett å finne fram til brukertjenester i SSB.

Vi gjennomgikk trykte dokumenter, dokumenter på Web og veiledning for utlevering av data på Web. Dokumentanalysen fikk oss til å stille spørsmål som var nyttige i prosessen med å lage spørreskjema til brukerne, det fikk oss til å sette søkelys på hvordan dagens saksbehandlingsprosess er og hvilke forbedringspotensialer som ligger der og det fikk oss til å se muligheter til forbedringer i ulike dokumenter og veiledninger. I kapittel 6 presenterer vi våre forslag til forbedringer og fremtidige utfordringer.

3.3 Spørreskjemaet

3.3.1 Datainnsamling

Utvalget til spørreskjemaet var aktuelle eksterne og interne brukere av forløpsdatabasen FD-Trygd. Vi trakk ut alle som hadde mottatt eller var i ferd med å motta data fra FD-Trygd, 13 eksterne og 12 interne. Spørreskjemaet besto både av lukkede og åpne spørsmål. Åpne spørsmål ble en løsning, fordi vi ønsket å få mest mulig informasjon ut av de få respondentene vi hadde til rådighet. Vi ønsket ikke å begrense svarene, fordi vi mente vi ville gå glipp av mye informasjon på denne måten.

Spørreskjemaene ble sendt ut 21. januar 2002. Etter at svarfristen var gått ut 5. februar, hadde 12 av de uttrukne svart. Nye skjema ble sendt ut mandag 11. februar. Etter puringen hadde 19 svart. Vi fikk 7 svar fra de eksterne mens alle interne svarte. 6 eksterne gikk altså til frafall. Det er usikkert om frafallet skyldes at noen av personene som har mottatt skjemaet har sluttet i jobbene sine, og at frafallet derfor skulle vært avgang.

Spørreskjemaene besto av 35 spørsmål. 13 av spørsmålene var åpne, og flere av disse var spørsmål der vi ønsket å få mye informasjon fra respondentene. Resten av spørsmålene var lukkede spørsmål. Spørreskjemaet var tredelt der første del omhandlet kjennskap til forløpsdatabasen FD-Trygd, andre del tilgjengelighet og tredje del bruk og nytte.

De første 8 spørsmålene omhandlet kjennskap til forløpsdatabasen FD-Trygd. Der fikk vi vite hvordan brukerne hadde fått kjennskap til FD-Trygd, hva som var relevant for dem der, om de visste om andre som kunne ha nytte av databasen og om de hadde noen forslag til hvordan man kan gjøre databasen mer kjent. Spørsmål 9-18 omhandlet tilgjengelighet. Brukerne fikk spørsmål om de visste hvem som var ansvarlig for brukertjenesten FD-Trygd, hvordan de fant fram til brukertjenesten, hvordan levering av data hadde vært og om de ville anbefale andre å bruke data fra FD-Trygd. De siste spørsmålene omhandlet bruk og nytte. Hvor mange ganger hadde brukerne søkt om data fra FD-Trygd, hvilke type data hadde de brukt, hadde de vært fornøyd med data og dataleveranser, hadde det vært noen mangler eller noe mer de ønsket seg, ville de søke om data fra FD-Trygd igjen ved en senere anledning.

3.3.2 Behandling av data fra spørreskjemaene

Utvalget til spørreskjemaene besto som tidligere nevnt av alle som har mottatt eller er i ferd med å motta data fra forløpsdatabasen FD-Trygd. Dette er derfor en total kartlegging av brukerne. Svarene på de lukkede spørsmålene ble sammenstilt slik at eventuelle tendenser i materialet bedre kom fram. Svar

på de åpne spørsmål ble analysert og strukturert på samme måte som data fra foredragene. Har brukerne samme oppfatninger om data fra FD-Trygd eller er oppfatningene forskjellige? Var det noen som skilte seg ut, hva var hovedinntrykket?

Da antall brukere var lite, og relativt kjent, valgte vi å innlede med å be svarpersonene om å føre opp navnet sitt. Dette fordi vi ønsket å ha mulighet til å snakke nærmere med noen av respondentene, hvis vi fant det nødvendig. Et problem blant de eksterne brukerne var at de aller fleste av de som svarte ikke førte opp navnet sitt. Dette har gjort det noe vanskeligere å tolke svar på enkelte spørsmål. Det er blant de eksterne at frafallet er høyt. Det er mulig at noe av frafallet kan skyldes vårt ønske om navn i spørreskjemaene. Det at spørreskjemaet ikke er anonymt kan muligens ha ført til at noen ikke har ønsket å svare. Det er også viktig å være klar over at vi ikke vet hvilke erfaringer omtrent halvparten av de eksterne brukerne har med forløpsdatabasen FD-Trygd. Våre svar baserer seg på de 7 som har svart på spørreskjemaet, og det er vanskelig å si om de som ikke har svart har liknende oppfatninger som de som har svart.

4. Vise nytten - øke bruken

Hovedformålet med foredraget var å informere om FD-Trygd og vise nytten av å bruke databasen. I tillegg ønsket vi å innhente opplysninger om hva og hvor mye potensielle interne brukere visste om forløpsdatabasene NUDB og FD-Trygd. Til slutt ønsket vi eksempler på hva forløpsdatabasene kunne brukes til på seksjonene, og om det i fremtiden ville være naturlig å ta i bruk databasene.

Vi ønsket en dialog, hvor vi skulle bidra med generell informasjon om forløpsdatabasene og de tilhørende skulle komme med spørsmål, eksempel på allerede bruk og mulig bruk av basene og innspill til forbedringer med hensyn på det tekniske, innhold og tilgjengelighet ved FD-Trygd. Vi ser oss fornøyd med deltakelsen og engasjement på foredragene.

De fleste seksjonene hadde en antakelse om hva FD-Trygd var, og noen personer hadde også brukt databasen. Det var få som visste hva NUDB var, og hvilke opplysninger man ville få fra utdanningsbasen i forhold til eksisterende data om utdanning.

På spørsmål fra prosjektgruppa til tilhørerne om hvilke områder/seksjoner FD-Trygd allerede hadde blitt tatt i bruk fikk vi følgende svar:

- Statistiske analyser, ved Seksjon for inntekts- og lønnsstatistikk, eksempelvis Statistiske Analyser (SA) for prosjektet Sosial rapportering av utsatte grupper
- Analyse av trygdekarriere, ved Seksjon for offentlig økonomi og personmodeller

4.1 Informasjon - foredraget

Informasjonsdelen av foredraget (se vedlegg 4) var bygd opp slik at vi ga en kort beskrivelse av hva forløpsdata er, noe de fleste tilhørende visste fra før. Deretter informerte vi hva NUDB er og noen eksempler på hva den nye basen kan gi av informasjon. Det er nok behov for mer informasjon til brukerne om NUDB, gjerne i god tid før opplysningene skal brukes i prosjekter eller statistikker.

Vi planla å bruke om lag 40 minutter av en time på å informere om FD-Trygd, der vi informerte om innhold (variable), årganger og viste eksempler på ulike uttakstabeller (avhengig av hvilken seksjon vi holdt foredrag for).

4.2 Kommentarer og dialog

De fleste seksjoner mente at de ville ta i bruk NUDB. Opplysninger om utdanning brukes innenfor mange områder/statistikker allerede. NUDB vil bedre utdanningsopplysningene og gi ny informasjon om utdanning.

Vi fikk kommentarer på at FD-Trygd har et enormt potensiale til analyse/forskning, og de spurte seg selv hvorfor ikke interne brukere mer tar i bruk allerede eksisterende databaser. Vi fikk tilbakemelding på at det er viktig at basen blir brukt, pga. nødvendig revisjonskontrollen av basen. I tillegg var det flere som lurte på hvorfor FD-Trygd ikke blir brukt til ordinær statistikkproduksjon ved Seksjon for levekårstatistikk.

Prosjekter som ble nevnt som mulige brukere av data fra forløpsdatabasen FD-Trygd:

- Fattigdomsprosjektet og
- Innvandring - grad av integrering, ved Seksjon for inntekts- og lønnsstatistikk
- Sentral sykefraværstatistikk, ved Seksjon for arbeidsmarkedsstatistikk

Ved seksjon for Arbeidsmarked ble det nevnt at NUDB vil bli brukt på prosjektet "Fra utdanning til arbeid".

For å kunne bruke informasjon fra eksisterende register, som tilleggsinformasjon til utvalgsundersøkelser, må man informere intervjuobjektet på forhånd. Det var derfor noen som arbeidet med utvalgsundersøkelser som nevnte at de hadde problemer med å få lov til å bruke registerinformasjon, som tilleggsinformasjon.

Det var flere seksjoner som kommenterte vanskelighetsgraden ved å bruke av data fra basen, fordi basen inneholder flere record per person (standard datadesign for forløpsdata). Man kan vrenge datasettet (det vil si at man får en record per person og med mange variable) og en slik behandling av datasettet vil gjøre det enklere å etterbehandle i vanlig SAS-programmering. De fleste trenger hjelp til å vrenge datasettet, og det ble hevdet at et slikt arbeide må gjøres av IT-personer og krever ressurser og tid.

En annen kommentar var at de ulike prosjektene/seksjonene ofte ikke har tid eller ressurser til å ta i bruk denne type informasjon. I den sammenheng kom også spørsmålet om hvor "ferske" årganger FD-Trygd inneholder, og at de ulike temaområdene har ulike årganger. Det ble spesielt kommentert at demografi-opplysninger som ikke følger med siste årgang på andre områder er verdiløst.

FD-Trygd basen vil dessuten bli mer interessant hvis familiebegrepet ville bli erstattet av husholdningsbegrepet. En slik overgang fra familiebegrep til husholdningsbegrep er ikke enkelt, og man må foreta noen valg på grunnlagsregisteret.

4.3 Oppsummering

Hovedformålet med foredraget var å informere generelt om forløpsdatabaser, med hovedvekt på FD-Trygd, og vise nytten av å bruke databasen. Det er for tidlig å se resultater av foredraget. Men vi fikk inntrykk at det var flere personer som fikk ny informasjon om forløpsdatabaser i SSB generelt, og FD-Trygd spesielt. I tillegg var det flere som mente at forløpsdatabaser var viktig for forskning og statistiske analyser.

5. Kjennskap, tilgjengelighet, bruk og nytte - brukernes erfaringer

Vi vil nå presentere resultatene av dataanalyser gjort på grunnlag av spørreskjemaundersøkelsen. Vi har valgt å presenterer resultatene i absolutte tall i stedet for prosentvise fordelinger. Dette på grunn av lite antall respondenter som gjør at prosentvise fordelinger ikke gir et riktig bilde av faktiske forhold. Vi har i tillegg valgt å benytte oss stort sett av bivariate frekvensfordelinger der vi ser på ulike forhold ut i fra verdiene for dikotomivariabelen intern/ekstern.

I alt var det 19 brukere av FD-Trygd (12 interne og 7 eksterne) som svarte på spørsmålene. Fire interne og to eksterne hadde valgt å fylle ut skjemaet sammen med en annen bruker (kollega), dermed fikk vi inn til sammen 16 utfylte skjema, ti fra interne og seks fra eksterne brukere. I de følgende analyser har vi tatt utgangspunkt i antall brukere, og ikke i antall skjema. For å kunne gi et helhetlig bilde av brukernes behov har vi i noen analyser supplert våre data med data fra brukertjenestes registreringer over ulike bestillinger.

5.1 Analyse av spørreskjemaene

5.1.1 Kjennskap til forløpsdatabasen FD-Trygd og det interessante med den

Brukerne av FD-Trygd har fått vite om databasen på ulike måter. Naturlig nok har de fleste interne brukere fått vite om FD-Trygd fordi de jobber i SSB. Fem av 12 interne brukere peker på sin arbeidsplass som eneste grunn til at de kjenner til basen, mens tilsvarende tall for eksterne brukere kun er en av syv. SSB's publikasjoner som eneste måte å få kjennskap til basen oppgis av tre interne og to eksterne brukere. To interne og to eksterne brukere oppgir SSB's publikasjoner som kilde til kjennskap til FD-Trygd i kombinasjon med andre alternativer. En intern bruker og to eksterne brukere har gitt uttrykk for at de fikk vite om databasen i forbindelse med andre prosjekter de jobbet med.

Hva fikk de vite som var interessant for sitt arbeidsområde gjennom de ulike kanalene? De interne delte seg her i tre grupper med litt ulik fokus. Den ene gruppen fokuserte på at dette var en base som inneholdt mye informasjon, en base som følger personer over tid, den gir muligheter til å gjøre forløpsanalyser og den inneholder individdata om trygd. Disse var opptatt av helheten ved basen og de generelle mulighetene en slik base kan gi. Den andre gruppen var mer opptatt av ulike variabler eller tema i basen FD-Trygd som passet godt for dem, fordi de dekket problemstillinger de jobbet med. En person var opptatt av basen fordi basen og innholdet i den er viktig for jobben de gjør, fordi jobben de gjør er å ha oversikt over slike baser.

De fleste eksterne fokuserte på muligheten en slik database ga til å kjøre forløpsanalyser og studere fenomener over tid generelt. En av de eksterne fokuserte mer konkret på et tema han var interessert i; nemlig trygdemedisinsk forskning.

5.1.2 Bruk og markedsføring

Det var stor enighet blant respondenter om at også andre institusjoner kan ha nytte av å ta i bruk databasen FD-Trygd. Vi stilte spørsmål om hvilke andre man mener kunne ha nytte av denne databasen. Blant de interne var det forskere, forskningsinstitusjoner og forvaltningen generelt som var gjennomgangstemaet. En nevnte at alle som er interessert i forløpsanalyser kan ha nytte av databasen, en nevnte alle som er interessert i trygdeproblemstillinger og en nevnte alle som driver med levekårsforskning. Blant de eksterne var det det samme mønsteret som viste seg: de fleste nevnte forskere, forskningsinstitusjoner og forvaltningen generelt. Men også ulike typer interesseorganisasjoner ble nevnt og mer spesifikt ble samfunnsmedisinske institutter ved universitetene i Norge nevnt.

Til tross for at både de interne og eksterne visste om mange som kunne ha nytte av å bruke en slik database, og også kom med gode forslag på slike institusjoner, visste de fleste ikke om disse

institusjonene allerede brukte basen eller ikke. I alt svarte syv ja, at de visste, en nei, at de ikke visste og ti var usikre. Vi spurte videre om de visste hvorfor de ikke bruker databasen. Seks svarte at de ikke visste og 11 svarte ikke på spørsmålet. Derimot svarte to eksterne brukere at de visste dette, og de pekte på manglende kjennskap og kunnskap som årsaker til å ikke ta i bruk databasen FD-Trygd.

Når det gjelder forslag til å gjøre databasen bedre kjent har seks personer (fem interne + en ekstern) forslag til dette, mens 11 personer ikke har noen forslag. Blant de interne ble følgende markedsføringsstrategier nevnt: avertere i egne SSB-publikasjoner, som Økonomiske Analyser og Samfunnsspeilet, holde innlegg på faglige konferanser, kontakte forskningsinstitusjoner med brosjyremateriell eller brev og informere om FD-Trygd gjennom NSD og Norges forskningsråd. En presiserte at vi burde fokusere på markedsføring mot institusjoner som benytter eller kan ha interesse for data fra FD-Trygd.

En av de eksterne foreslo å legge ut nøkkeltall på SSB's web-sider. En annen mente at man burde forsøke å spore informasjon i forskningsmiljøer som ikke i utgangspunktet assosieres med trygdeforskning. Politikere, byråkrater må bli klar over hvilke analysemuligheter som ligger i basen. Og hva med utenlandske forskningsmiljøer, også der er det et potensial.

5.1.3 Tilgjengelighet

Vi har forsøkt å måle tilgjengeligheten av databasen FD-Trygd med å stille spørsmål om respondentene visste hvem som er ansvarlig for brukertjenesten FD-Trygd, hvordan de fant frem til brukertjenesten, hvilke forventning de hadde til leveringstid og hvor lang tid det tok før de fikk innlevert dataene, ville de anbefale andre å bruke forløpsdatabasen FD-Trygd, m.m. Det er ikke lett å peke på klare tendenser eller rendyrkede mønstre, men vi finner en del forskjeller mellom interne og eksterne brukere.

Brukertjenesten og veiledninger

Mens ti av 12 interne brukere oppgir at de vet hvem som er ansvarlig for brukertjenesten FD-Trygd, oppgir halvparten (tre av seks) av eksterne brukere at de ikke vet hvem det er.

På spørsmålet om hvordan de fant frem til brukertjenesten FD-Trygd, svarer ni av de interne at de fant frem til basen via bl.a. sine kolleger. Ingen av de eksterne brukere oppgir det samme. I stedet har to eksterne brukere oppgitt personer og bekjente i SSB som sine kilder. Det var forventet at noen av de eksterne brukere skulle finne veien frem til FD-Trygd via sentral brukertjeneste eller biblioteket i SSB. Dette er ikke tilfelle. Derimot har en ekstern bruker kommet frem til FD-Trygd via sentralbordet til SSB.

Veiledninger for fremgangsmåter for uttak fra forløpsdatabasen FD-Trygd er ment å skulle hjelpe brukere til å finne lett frem til brukertjenesten FD-Trygd, og ikke minst være et hjelpemiddel for spesifisering av oppdraget slik at behandlingsprosessen reduseres noe. Disse veiledningene finnes på Internett, i dokumentasjoner og andre materiale fra SSB. Vårt datamateriale viser at de fleste brukere (ti personer) ikke hadde sett slike veiledninger. Fire interne og fire eksterne brukere hadde sett veiledninger for uttak fra databasen, en person hadde ikke svart spørsmålet. Tabell 5.1. viser at av de som hadde sett veiledningen, mente kun en person at det var vanskelig å finne frem til brukertjenesten. Resten svarer "ja" (lett å finne frem) på dette spørsmålet. Også tre av dem som ikke hadde sett veiledningen mente at det var lett å finne frem til brukertjenesten.

Tabell 5.1. Av de som hadde sett/ikke sett veiledningene og om det var lett/ikke lett å finne frem til ansvarlig for FD-Trygd.

Sett/ikke sett veiledningene	Var det lett å finne til ansvarlig for FD-Trygd?				
	I alt	Ja	Nei	Vet ikke	Uoppgitt
I alt	19	9	6	1	3
Ja	7	6	1	-	-
Nei	11	3	5	1	2
Uoppgitt	1	-	-	-	1

Fem som hadde svart nei på spørsmålet om de hadde sett veiledninger, hadde også svart nei på spørsmålet om det var lett å finne frem til brukertjenesten.

Blant de som hadde sett veiledninger på Internett, i dokumentasjonsnotater eller i annet materiale mente de interne at materialet var bra, eller "som ellers i Statistisk sentralbyrå". De eksterne som hadde sett på dette mener veiledningene er bra.

Levering av data

Et annet interessant spørsmål er om det å se på veiledninger påvirker spesifiseringen av oppdraget og dermed saksbehandlingstiden. Ut i fra datamaterialet har vi dessverre ikke noe grunnlag for å kunne svare på dette spørsmålet. I stedet kan vi se nærmere på forholdet mellom forventede og faktiske leveringstid fra ferdig spesifisert og godkjent bestilling til første hoveddataleveranse.

Faktisk leveringstid er avhengig av mye: Om kunden har spesifisert sine ønsker tilstrekkelig, kompleksiteten i oppdraget og hvor stor køen av oppdrag er på det enkelte tidspunkt m.m. I veiledninger står ikke noe om leveringstid av data. Likevel forventer vi at de som har sett (lest) veiledningene bør spesifisere sine søknader på tilfredstillende måte i samsvar med veiledningene, og dermed redusere leveringstiden. Derfor har vi valgt å se på forventet og faktisk leveringstid i lys av om respondentene har sett veiledningene for datauttak fra FD-Trygd. 16 respondenter har svart på spørsmålene om forventnings- og faktisk leveringstid.

Ved å sammenligne tallene på individnivå i tabellene 5.2. og 5. 3. kan vi si at de som hadde sett veiledningene hadde en lengre forventningstid enn de som ikke hadde sett dem.

Tabell 5.2. De som svarte at de hadde sett veiledningene, etter forventet og faktisk leveringstid (N=8).

Type bruker	Forventet leveringstid	Faktisk leveringstid
Interne brukere		
Under 1 mnd	1	-
1 - 2 mnd	3	-
Over 2 mnd	-	2
Uoppgitt	-	2
Eksterne brukere		
Under 1 mnd	-	2
1 - 2 mnd	2	-
Over 2 mnd	-	2
Uoppgitt	2	-

Av de som hadde sett veiledningene, forventet en intern og ingen eksterne brukere å få levert dataene under en måned. Tre interne og to eksterne brukere hadde en forventningstid mellom en og to måneder. Ingen hadde forventet at levering av data kan ta over to måneder. Likevel måtte to interne og to eksterne brukere faktisk vente over to måneder for å få levert dataene. Samtidig sier to eksterne brukere at de fikk dataene "raskt". I tabellen har vi tolket begrepet "raskt" som utlevering av data i løpet av under to uker.

Tabell 5.3. De som svarte at de ikke hadde sett veiledningene, etter forventet og faktisk leveringstid (N=8).

Type bruker	Forventet leveringstid	Faktisk leveringstid
Interne brukere		
Under 1 mnd	3	1
1 - 2 mnd	1	-
Over 2 mnd	-	2
Uoppgitt	-	1
Eksterne brukere		
Under 1 mnd	2	2
1 - 2 mnd	1	1
Over 2 mnd	-	1
Uoppgitt	1	-

Ser vi på tabell 5.3. oppdager vi at forventningene til de som ikke hadde sett veiledningene avviker i mer eller mindre grad fra forventningene til de som hadde sett veiledningene.

Mens det var kun en intern og ingen eksterne brukere med forventningstid under en måned, ser vi at det er tre interne og en ekstern med samme forventningstid. Og når vi sammenligner kolonnene for faktisk leveringstid i tabell 5.2. og 5.3., kommer det frem at alt i alt har de som ikke har sett veiledningene kortere forventningstid enn de som har sett veiledningene. Samtidig ser vi at de med lavere forventningstid har fått utlevert dataene raskere enn de med lengre forventningstid. Det vil si at forventningstid henger stort sett sammen med faktisk leveringstid. Det kan være slik at i etter tid har respondentene "justert" sine forventningstid i forhold til den faktiske tiden det tok det før de fikk utlevert dataene.

En del avvik i hver av de ovenstående tabellene, og i sammenlikningen mellom de to tabellene, kan muligens forklares med at, i hvert fall noen av respondentene, ikke har skilt mellom tidspunkt for levering av søknad om data og tidspunkt for godkjent bestilling. Vi må også understreke at kundens egen spesifisering av oppdraget blir en del av total leveringstid.

Vi stilte også spørsmål om de hadde opplevd noen problemer enten i forbindelse med spesifisering eller i forbindelse med uttak av data. I alt syv respondenter (fire interne, tre eksterne) svarte at de ikke hadde hatt problemer, mens ni hadde hatt problemer (fem interne, fire eksterne). Tre interne svarte ikke på spørsmålet. Det er også interessant å se på svarene på dette spørsmålet også i lys av om respondentene hadde sett veiledningen eller ikke.

Tabell 5.4. De som hadde sett/ikke sett veiledningene, etter problemer med spesifisering.

Sett/ikke sett veiledningene	Problemer med spesifisering?			
	I alt	Ja	Nei	Uoppgitt
I alt	19	9	7	3
Ja	7	4	3	-
Nei	11	5	4	2
Uoppgitt	1	-	-	1

Tabell 5.4 viser at i overkant av halvparten av de som hadde sett veiledningene, hadde problemer med å spesifisere og motta data. Det samme gjelder de som ikke hadde sett veiledningene. En kan spørre seg om veiledningene ikke er bra nok vedrørende spesifisering av oppdraget eller om veiledningene ikke leses nøye nok av søkere. Dette kommer vi tilbake til i kapittel 6.

De eksterne nevnte problemer i ulike faser av prosessen med å få til å jobbe med data. I spesifiseringsfasen ble nettopp spesifiseringen nevnt som vanskelig. Dette er uvant for de som ikke er kjent med forløpsanalyser. Et annet problem som er relevant for de som ikke er kjent med forløpsanalyser og som skal bruke data til tverrsnittsanalyser, er at kjøringene er store og tunge. Av problemer de eksterne nevnte var for det første problemer i spesifiseringsfasen. Vanskeligheter med avgrensning og spesifisering av variable og omfang. Et annet problem var selve leveringen, som to av de eksterne opplevde som meget dårlig.

Vi stilte også spørsmål om hvordan man ønsket å få data levert, hvis man kunne velge. Blant de interne ønsket flest (fire stykker) å få data fra FD-Trygd som en tradisjonell dataleveranse, eventuelt tverrsnittsutttak på bestemte tidspunkt. Men flere ønsket også å få data stilt opp etter samme type forløpsmodell som FD-Trygd er bygd opp etter, altså opplysninger som endres kontinuerlig over tid og som det dannes forløp på. En av de interne påpekte at det var avhengig av om han skulle gjøre forløpsanalyser eller tverrsnittsanalyser. Her er det altså formålet med datauttaket som i stor grad bestemmer hvordan man ønsker data. Blant de eksterne ønsker alle bortsett fra en å få data stilt opp etter samme type forløpsmodell som FD-Trygd er bygd opp etter. Dette skyldes nok at de fleste eksterne brukerne bruker forløpsmodellen i sine analyser.

På spørsmålet om de vil anbefale andre å bruke forløpsdatabasen, svarer 11 interne og fem eksterne ja. En nærmere analyse av datamaterialet viser at de (en intern og to eksterne) som ikke vil anbefale andre å bruke forløpsdatabaser er blant de som hadde ventet på utlevering av dataene fra FD-Trygd i over to måneder. Som vi allerede har vært inne på henger faktisk leveringstid nøye sammen med bl.a. tilstrekkelig spesifisering av oppdraget fra kundens side, kompleksiteten i oppdraget og hvor stor køen av oppdrag er på det enkelte tidspunkt.

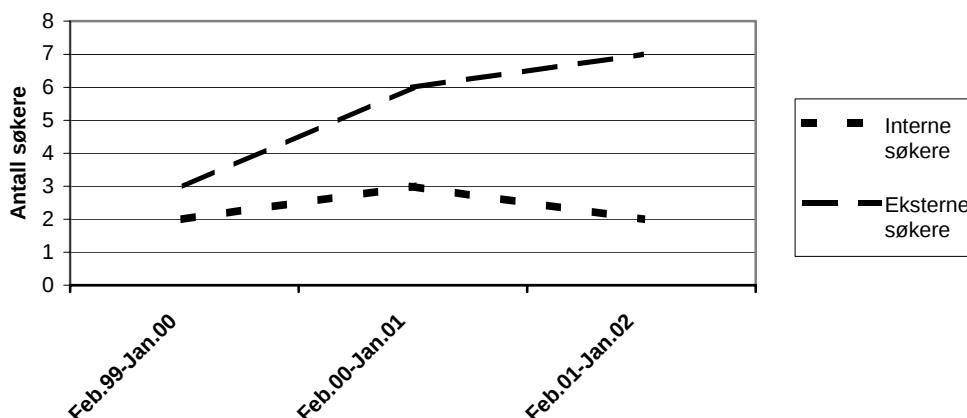
5.1.4 Bruk og nytte

Bruk av data og hovedproblemstillinger for prosjektene

FD-Trygd er for så vidt en "ung" database. Den har ikke eksistert i mer enn 3 år. Derfor er det ikke unaturlig at de fleste respondenter bare hadde søkt om data fra FD-Trygd en gang. Til sammen svarte 16 personer på spørsmålet. Ut i fra vårt datamaterial var det i alt to interne og fire eksterne førstegangssøkere i 1999. Året etter øker antall interne søkere til tre, mens antall eksterne søkere går ned til en førstegangssøker. Antall interne og eksterne førstegangssøkere øker igjen i 2001 da det var fire interne og to eksterne søkere.

For å kunne gi en bedre og bredere oversikt over førstegangssøkere, vil vi se på samtlige interne og eksterne førstegangssøkere fra FD-Trygd var klar til bruk. Dataene er hentet fra arkivet til brukertjenesten FD-Trygd, og er gjengitt i figur 5.1.

Figur 5.1 Antall førstegangssøkere for hvert år etter basens oppstart (N=23)



I perioden februar 1999 - januar 2000 var det i alt fem søkere, to interne og tre eksterne. I perioden februar 2000 - januar 2001 økte antall søkere med flere enn det som er vist i figuren. Men som vi ser i figuren øker også antall førstegangssøkere. Mens interne førstegangssøkere kun øker med en søker, øker antall eksterne søkere fra tre søkere i den første perioden til i alt seks søkere i den andre perioden. Figuren viser at antall interne førstegangssøkere i perioden februar 2001 - januar 2002 går tilbake til samme nivået som i den første perioden. I samme perioden øker antall eksterne førstegangssøkere med enda en søker i forhold til forrige perioden.

Det kan også nevnes at brukertjenesten i tillegg har fått syv førstegangssøknader knyttet til totalt 11 prosjekter via NSD: tre søknader i 1999, to i 2000 og to i 2001.

Figur 5.1. viser tydelig at mens det er flere eksterne førstegangsbrukere som tar i bruk data fra FD-Trygd, er antall interne førstegangsbrukere stabilt. Er det slik at stadig flere institusjoner får kjennskap til FD-Trygd og fordelene med å ta i bruk denne basen? Er den interne bruken av FD-Trygd begrenset til kun noen få avdelinger/seksjoner i byrået? Avslører denne tendensen ulike prioriteringer mellom SSB og miljøet utenfor med hensyn til forskningsfelt? Her er det mange interessante spørsmål som kan reises, men det er vanskelig å gi klare svar på dette ut fra vårt datamateriale.

For å kunne si noe om hvilke forskningsfelt som blir prioritert av SSB og andre institusjoner, kan vi se på typer problemstillinger brukerne forsøkte å belyse da de søkte om data fra FD-Trygd. La oss først se på hvilke problemstillinger de interne brukerne av FD-Trygd jobber med. Flere jobber med problematikken rundt enslige forsørgere og deres bruk av trygd og deltakelse på arbeidsmarkedet. Eldre arbeidstakere og førtidspensjonister er en annen gruppe det forskes på ved hjelp av forløpsdata. Personer som jobber med Folke- og bolig tellingen bruker også forløpsdata fra FD-Trygd. Andre problemstillinger er ekteskap i innvanderbefolkningen og kvinners yrkesdeltakelse og fruktbarhet.

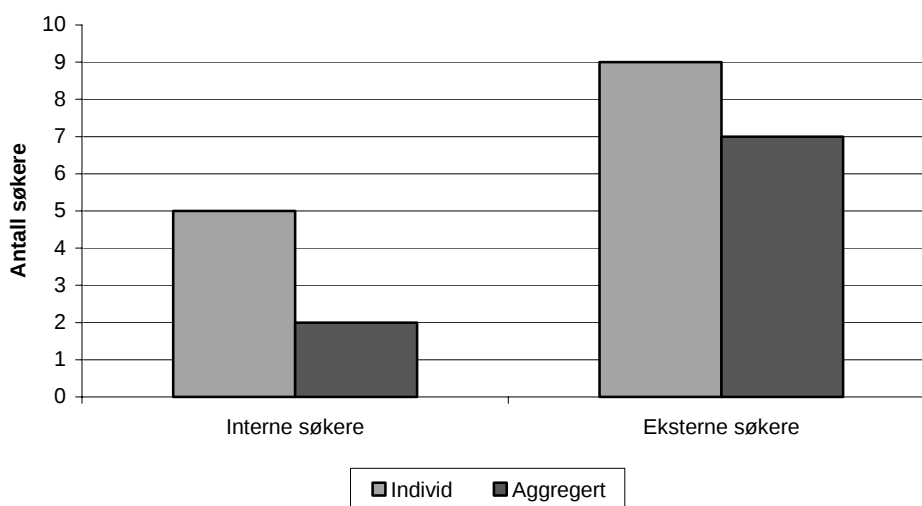
Så, hva med eksterne brukere? De eksterne brukere oppga følgende problemstillinger for prosjektene de jobbet med: Sammenhengen mellom uførepensjonering og utdanning, uførepensjonering og etnisitet, overgang fra arbeid til tidligpensjonering, effektevaluering av arbeid for sosialhjelp, overgang fra arbeidsledighet til jobb til pensjonering, bruk av ulike trygdeytelser i befolkningen.

En av de viktigste egenskapene ved forløpsdatabaser er at slike baser inneholder data på individnivå. Og det er nettopp denne egenskapen som gjør det mulig å kunne analysere "løpet". Samtidig er det fullt mulig å hente data på aggregert nivå fra forløpsdatabaser. Datamaterialet viser at 13 brukere (ni interne og fire eksterne) hadde søkt om data bare på individnivå. Kun en ekstern bruker hadde søkt om

data på bare aggregert nivå. Til sammen tre brukere (en intern og to eksterne) hadde søkt om data på både individ- og aggregert nivå. To brukere har ikke svart på spørsmålet.

Dersom vi tar utgangspunkt i samtlige førstegangssøkere etter basens oppstart, og ønsket type data blir bildet noe annet. Figur 5.2. gjengir et oversikt over interne og eksterne førstegangssøkerens ønske om type data med hensyn til nivå.

Figur 5.2 Førstegangssøkere på ulike datatyper m.h.t nivå (N=23)



Det er ingen tvil om at problemstillingen til et forskningsprosjekt bestemmer hvilken type data en søker på. Selv om de fleste førstegangssøkere har søkt på individdata, ser vi at også en god del har søkt om data på aggregert nivå. Fordelingen er slik at fem av sju interne, og ni av 16 eksterne førstegangssøkere har søkt på individdata. Derimot har to interne og sju eksterne førstegangssøkere søkt på aggregerte data.

5.1.5 Mangler, ønsker og eventuelle problemer

Over halvparten av respondentene i vår undersøkelse gir uttrykk for at utleverte data var i samsvar med deres ønsker, mens tre brukere gir uttrykk for det motsatte. Fra tabell 5.5. ser vi at andelen som svarer ja på spørsmålet er større blant eksterne enn interne brukere, tatt i betraktning at fem interne brukere ikke har svart på spørsmålet.

Tabell 5.5. Utleverte data i samsvar med brukernes ønsker, etter typer brukere.

Samsvar med søkernes ønsker	Type bruker		
	I alt	Intern	Ekstern
I alt	19	12	7
Ja	11	6	5
Nei	3	1	2
Uoppgitt	5	5	-

Av mangler oppgir den ene interne som har svart at datafila var vanskelig å "finne ut av", at det var lite selvforklarende. De to eksterne som oppgir mangler, nevnte at variabelen arbeidsmarked manglet.

Fem interne og to eksterne brukere oppga at de hadde bruk for variabler som ikke fantes i FD-Trygd, mens tre interne og fire eksterne mener at det ikke var noen variabler de hadde bruk for som ikke fantes i databasen. Hvis vi krysser svarene på spørsmålene om å få levert dataene i samsvar med egne

ønsker, og om det var noen variabler som søkerne hadde behov for, men som ikke fantes i FD-Trygd, finner vi at kun en av de som ikke fikk dataene i samsvar med sine egne ønsker, samtidig oppga at databasen manglet variabler vedkommende hadde behov for. De to andre ga uttrykk for at FD-Trygd ikke manglet noen variabler som de hadde behov for.

Tabell 5.6. Samsvar med søkerne ønsker, etter svare på spørsmålet om manglende variabler.

Samsvar med søkerne ønsker	Manglende variabler?			
	I alt	Ja	Nei	Uoppgitt
I alt	19	7	7	5
Ja	11	4	5	2
Nei	3	1	2	-
Uoppgitt	5	2	-	3

Ser vi på tabell 5.6. med utgangspunkt i svarene på spørsmålet om manglende variabler, finner vi ut at fire brukere melder om manglende variabler, mens de samtidig oppgir at de har fått dataene i samsvar med sine ønsker. To andre som opplyser at FD-Trygd mangler variabler de kunne ha behov for, har lot det være å svare på det andre spørsmålet.

Av manglende variabler oppga de interne manglende dato for endring av sivilstand, opplysninger fra barnetrygdregisteret, opplysninger om statsansatte og variabler utenom trygdevariabler. De eksterne nevner følgende mangler: arbeidsmarked, arbeidsledighetstrygd og mer generell sysselsettingsinformasjon.

Alle de eksterne brukere (sju stykker) sier at de fikk data for de årgangene de ønsket. Det samme sier åtte interne brukere. Om å kunne koble data fra FD-Trygd til andre datasett sier åtte brukere (fire interne og fire eksterne) at de hadde behov for en slik kobling. Samtlige mener at det var lett å kunne gjøre dette.

Vi stilte også spørsmål om respondentene ville søke om data fra forløpsdatabasen FD-Trygd ved en senere anledning. Åtte interne og fem eksterne brukere svarte ja på dette spørsmålet. Bare to interne brukere har svart nei, mens to eksterne brukere har sagt at de ikke vet.

De to eksterne som svarte nei oppga som grunn at de fikk dårlig oppfølging av bestilling, og at dette var synd fordi basen er god.

5.2 Oppsummering

I dette kapitlet har vi presentert resultatene fra analysen av spørreskjemaundersøkelsen til interne og eksterne brukere av FD-Trygd. Resultatene kan oppsummeres i tre hovedpunkter.

For det første tyder resultatene på at brukerne er bevisst sine ønsker og forventninger til både databasen - og brukertjenesten FD-Trygd. Interne og eksterne brukere gir uttrykk for sin tilfredshet eller misnøye med databasen og / eller forvaltningen av den. Selv om de fleste hadde positive erfaringer med bruk av FD-Trygd, bør en alltid tenke ut nye strategier for å imøtekomme brukernes behov.

For det andre viser datamaterialet at FD-Trygd er lite kjent både internt og eksternt. De fleste interne brukere fikk vite om FD-Trygd i kraft av sin stilling og sine arbeidsoppgaver, og ikke fordi de jobber i SSB. Det tyder på at en kan være ansatt ved en seksjon i SSB uten å vite noe om forløpsdatabaser, men når man begynner i en ny stilling ved en annen seksjon får kjennskap til FD-Trygd. Det kom frem at mens antall eksterne førstegangssøkere har økt i løpet av en tre års periode, har antall interne førstegangssøkere holdt seg mer eller mindre stabilt. Når det gjelder eksterne brukere så er det tydelig

at FD-Trygd er mer kjent i visse eksterne miljøer enn i andre. Det var stor enighet blant respondentene om at andre institusjoner og personer som ikke bruker/har brukt basen, kan ha stor nytte av å ta i bruk basen.

For det tredje kommer det frem at spesifisering av oppdraget fra brukernes side er et problem som en bør jobbe aktivt med. Noen av brukerne har hatt problemer med nettopp å kunne spesifisere oppdraget sitt godt nok. Dette kan skyldes at brukeren kan ha problemer med å spesifisere og eller avgrense problemstillingen sin. Dermed får brukeren problemer med å spesifisere eksakt hvilke variabler og fra hvilke tidsperioder hun/han ønsker seg når han/hun kontakter SSB.

Ikke alle brukere er tjent med mengden av data og opplysninger som finnes i SSB's forløpsdatabaser, der iblant FD-Trygd. Når en først kommer i kontakt med disse basene, oppdager en ofte at en kan ha behov for flere typer opplysninger enn det en hadde tenkt seg i utgangspunktet. 4 av respondentene i vår undersøkelse sier at kravene deres endret seg underveis etter mottak av data. Det vil si de ønsket å få opplysninger om nye variabler, nytt uttak m.m.

Problemer knyttet til spesifisering av oppdraget kan også henge sammen med kvaliteten på ulike dokumenter og veiledninger som er lagt ut på SSB's websider og ellers i papirutgaver. Vårt datamateriale viser at de fleste brukerne ikke har sett slike veiledninger. Mangle på klare retningslinjer kan være et annet problem i den sammenheng.

6. Forslag til forbedringer - fremtidige utfordringer

Prosjektgruppen har ikke til hensikt "å diagnostisere" SSB's forløpsdatabaser, og skrive ut "resepter" som skal bidra til å øke bruken av disse basene i enda større grad enn det som gjøres i dag. Derimot mener vi at det er hensiktsmessig å studere ulike sider ved bruken av forløpsdatabaser i dag for å kunne si noe om hvorfor og hvordan en kan benytte seg av de mangfoldige forsknings- og analysemuligheter slike databaser gir anledning til.

SSB bruker en god del ressurser for oppbygning, vedlikehold og forvaltning av forløpsdatabaser. Som nevnt tidligere i notatet, er SSB's forløpsdatabaser gull verdt i mange henseender. Vi mener at forløpsdatabaser blir brukt i en begrenset omfang i dag, og det er potensial for å kunne øke omfanget betraktelig. Derfor vil vi i det følgende først se på dagens situasjon i lys av det som vi gikk gjennom i de tre foregående kapitler for å deretter behandle spørsmålet om hva som kan forbedres, og eventuelt på hvilken måte.

6.1. Dagens situasjon, og hva som bør forbedres

Ut i fra diskusjoner i kapittel 2 finner vi dagens saksbehandling av innkomne søknader om data fra forløpsdatabasen FD-Trygd nokså uklar og komplisert. Prosessdiagram 2a og 2b (kapittel 2) viser kompleksiteten i saksbehandlingsprosessen, og hvor tungtvint det kan være å kunne spesifisere oppdraget. Det ble antydning at mangelfulle veiledninger og andre nødvendige dokumentasjoner i både SSB's Websider og papirpublikasjoner kan forårsake en del uklarheter hos brukerne. Prosjektgruppen kommer en forslag til veiledning på Web-siden ved bestilling av mikrodata.

En forbedring av eksisterende dokumentasjoner for forløpsdatabaser er i høyeste grad nødvendig. Dette kan begrunnes med bl.a. det faktum at bortimot halvparten av respondentene i vår undersøkelse hadde problemer med å spesifisere sitt oppdrag. Datamaterialet vårt tyder på at intern datalevering tar for så vidt lenger tid i forhold til eksterne datalevering. Det er i denne sammenheng mulig at interne kunder oftere stiller krav som impliserer en IT-utvikling på uttakssiden enn det som vanligvis tilbys de eksterne og som de eksterne kunder eventuelt må løse selv. Etter løsning av eventuelle IT-problemer vil det være mulig å tilby slike prosedyrer med raskere leveringstid. Det en fordel for de interne kunder å være oppmerksom på muligheten til å søke IT-støtte på utviklingssiden fremfor å løse det via brukertjenesten.

Våre analyser viser at ikke alle kjenner til SSB's forløpsdatabaser og fordelene med å ta i bruk disse. Av markedsføringsstrategier kan følgende nevnes: Annonser i SSBs egne publikasjoner som Økonomiske analyser og Samfunnsspeilet og DS-artikler som gir smakebiter på hva forløpsdatabasene kan gi av informasjon. I tillegg kan man være aktive på arenaer utenfor SSB, som innlegg på faglige konferanser, kontakte ulike forskningsinstitusjoner med brosjyremateriell eller brev eller å informere om FD-Trygd gjennom NSD og Norges forskningsråd. Utenlandske forskningsmiljøer er også en potensiell arena for markedsføring av SSB's forløpsdatabaser.

Samtidig bør en legge forholdene til rette slik at forløpsdatabaser blir lettere tilgjengelig for interne brukere. Av utfordringer i den sammenheng kan en nevne problemer med vanskelighetsgraden av bruk av data fra forløpsdatabaser og problemet med tid for å kunne gjennomføre dyptgående analysestudier.

I første del av kapittel 6 kommer vi med forslag til forbedringer av veiledning på Web og forenklinger av kompliserte saksbehandlingsprosesser. Siste del av kapitlet inneholder fremtidige utfordringer. Med fremtidige utfordringer mener vi aktuelle problemstillinger som har falt utenfor dette prosjektets rammer, og videre arbeid med forslag til forbedringer.

6.2 Forslag til forbedringer

6.2.1 Veiledning på Web

Det er mange som henvender seg til SSB for å bestille mikrodata om personer til forskning og planlegging. Det er strenge regler for hvem kan få slike data og hva slags type data man kan få utlevert. SSB har ikke per i dag en generell veiledning på Web, for hvilke betingelser vi har ved utlevering av mikrodata.

Som nevnt tidligere har FD-Trygd en veiledning for bestilling av data på sine internettsider (vedlegg 3). Denne siden inneholder mye av det vi mener er nødvendig for en slik veiledning, men siden har et forbedringspotensial. For eksempel står det på siden at "Statsforvaltningen og andre brukere henvender seg til Statistisk sentralbyrå", hvem er andre? og det burde vel stå offentlig forvaltning, siden kommune og fylkeskommune også kan få data fra FD-Trygd. I tillegg mener vi at siden er uryddig og lite pedagogisk skrevet.

Det er flere muligheter med en veiledning på internettsidene, det ene er å lage en spesifikk side, på informasjonssiden til FD-Trygd, ved bestilling av data fra FD-Trygd. Det andre er å lage en generell veiledning på internettsidene til SSB, med link til fagområder hvor man kan få mer informasjon om temaet og bestille mikrodata. I tillegg må man da ha en link på informasjonssiden til FD-Trygd, eventuelt andre temasider, til generell veiledning. Prosedyren ved bestilling av mikrodata om personer i SSB vil være lik for de fleste fagområder i SSB.

Vi foreslår en generell veiledning på Internett. Prosjektgruppen har samarbeidet med jurist, Thorleiv Valen, Sikkerhetsrådgiver, Egil Løksa og Sentral brukertjeneste, Halvor Strømme i arbeidet med å lage et forslag til generell veiledning på Web. I tillegg har vi kontaktet Web-redaktør Jan Bruusgaard og diskutert hvor en slik veiledning kan plasseres på Web-siden til SSB.

Vi anbefaler at veiledningen skal ligge på hovedsiden til SSB under punktet forskning (på venstresiden) og under punktet fra A til Å, med inngang både på forskning og planlegging.

Forslag til generell veiledning til Web-siden til SSB ved bestilling av mikrodata om personer til forskning og planlegging (se side 34).

Veiledningen er delt opp i fem temaer:

- Hvem kan søke
- Hvor skal man søke
- Hvilke data finnes
- Hvilke typer data kan utleveres
- Krav til søknad
- Personregistre i SSB - hvem svarer

Under punktet hvor skal man søke henviser vi forskere til NSD (med link) og offentlig forvaltning til SSB. Under punktet hvilke data finnes gir vi en definisjon på de ulike typer data. Punktet hvilke typer data kan utleveres gir vi informasjon om hvilke betingelser SSB krever ved utlevering, ved de ulike type data som finnes, med henvisninger (link) til ulike lover og Datatilsynet.

I punktet krav til søknad har vi listet opp hva en skriftlig søknad må inneholde og hvilke vedlegg som er nødvendig. I tillegg er det laget en link til et søknadsskjema som vi mener vil forenkle arbeidet hos saksbehandler i SSB. Et slikt søknadsskjema vil føre til at kunden må svare på alle spørsmål som SSB mener er nødvendig å hente inn informasjon om, for å utlevere data. Dette vil kunne standardisere søknadene om data, og man vil slippe skriftlige og muntlige henvendelser med mangelfulle opplysninger. Prosjektgruppen har dessverre ikke rukket å lage et slikt søknadsskjema.

I siste punkt hvem svarer ønsker vi en fullstendig oversikt over registre med mikrodata om personer og mailadresse til kontaktpersoner i SSB og eventuelle linker til informasjonssider på Web om temaområdet (f.eks. FD-Trygd). Søknader om utlevering av data bør sendes Sentral brukertjeneste eller direkte til fagseksjon/ansvarlig. Det bør derfor opprettes en mailadresse til Sentral brukertjeneste, Brukertjenesten FD-Trygd og eventuelt for andre områder som utleverer mye mikrodata (f.eks. NUDB).

Forslag til veiledning på Web-siden til SSB

På hovedsiden til Statistisk sentralbyrå - www.ssb.no, under punktet forskning på venstresiden:

Veiledning ved bestilling av mikrodata om personer til forskning og planlegging

Hvem kan søke

Hvor skal man søke

Hvilke data finnes

Hvilke typer data kan utleveres

Krav til søknad

Personregistre i SSB - hvem svarer

Veiledning ved bestilling av mikrodata om personer til forskning og offentlig planlegging

Hvem kan søke

Offentlig planlegger - offentlig forvaltning

Utlevering til offentlig planlegging kan bare skje til organer i stat eller kommune. Dersom en kommune benytter et privat konsultentselskap til å utføre deler av sitt planleggingsarbeid eller utarbeidelse av grunnlagsdokumenter for slik arbeid, kan kommunen utlevere de nødvendige data til konsultentselskapet etter retningslinjer gitt av SSB.

Forsker

Utlevering av person opplysninger kan skje til etablerte forskningsinstitutter, forskere med forskningskompetanse og hovedfagsstudenter som er under veiledning av en med forskerkompetanse ved undervisningsstedet. I vurderingen av søkeren vil det være et moment om søkeren har fått prosjektet vurdert og finansiert av etater som forvalter offentlige forskningsmidler.

Hvor skal man søke

Offentlig planlegger - offentlig forvaltning

Henvender seg til den aktuelle fagseksjonen eller Sentral brukertjeneste ved Statistisk sentralbyrå (se avsnitt: Personregistre i SSB - hvem svarer)

Forsker

Henvender seg til NSD dersom anonymiserte data er tilstrekkelig. SSB har inngått en avtale med Norsk Samfunnsfaglig Datatjeneste (NSD) om videreformidling av data fra SSB til forskere og hovedfagsstudenter. Avtalen går i korthet ut på at NSD mottar data fra SSB som NSD anonymiserer før utlevering. For nærmere opplysning om utlevering av personopplysninger til forskning, se [veiledning hos NSD/Datafaglig sekretariat](#).

Forskere som trenger aidentifiserte data henvender seg til aktuell fagseksjon eller Sentral brukertjeneste ved Statistisk sentralbyrå (se avsnitt: Personregistre i SSB - hvem svarer)

Hvilke data finnes

- a) *Identifiserbare opplysninger*: opplysninger hvor det er mulig å identifisere fysiske eller juridiske personer ved hjelp av identifikasjonsnøkkel (f.nr, org.nr, navn, eksakt adresse)
- b) *Aidentifiserte opplysninger*: opplysninger hvor data som brukes som identifikasjonsnøkkel er fjernet.

c) *Anonymiserte opplysninger*: opplysninger hvor så mange kjennemerker/kategorier er fjernet at en med rimelige midler ikke kan direkte eller indirekte identifisere fysiske eller juridiske personer.

Hvilke typer data kan utleveres

Identifiserbare opplysninger leveres normalt ikke ut til forskning eller planlegging.

Anonymiserte opplysninger

Det er en fordel om det ønskes anonymiserte data, siden det er få krav til utlevering av slike data. Kravene til anonymisering er imidlertid strenge. Det kreves skriftlig søknad til NSD eller SSB om utlevering av data, og instansene vil vurdere om søknaden tilfredsstiller kravene om anonymisering.

Avidentifiserte opplysninger

Dersom det ikke er mulig å gjennomføre prosjektet med anonymiserte opplysninger, kan det utleveres avidentifiserte data, dvs personopplysninger som kan knyttes til enkeltperson. Det kreves skriftlig søknad til SSB om hvilke kjennemerker som skal utleveres. Søknaden må inneholde opplysninger om utleveringen gjelder sensitive data eller ikke.

Sensitive personopplysninger er opplysninger om:

- a) rasemessig eller etnisk bakgrunn, politisk, filosofisk eller religiøs oppfatning,
- b) at en person har vært mistenkt, siktet, tiltalt eller dømt for en straffbar handling,
- c) helseforhold,
- d) seksuelle forhold,
- e) medlemskap i fagforeninger

Dersom ikke-sensitive data skal utlevers, må det foreligge kopi av bekreftet mottatt melding til Datatilsynet. Mottaker av personopplysninger (ikke-sensitive data) er underlagt bestemmelsene i personopplysningsloven (pol). Dette innebærer bl.a. at mottaker må ha selvstendig hjemmel for behandlingen (jf pol § 8 og § 9), og må melde behandlingen til Datatilsynet (jf pol § 31).

Dersom sensitive personopplysninger (jf pol § 2 nr 8) skal utleveres, må det foreligge kopi av mottakers konsesjon fra Datatilsynet som omfatter utlevering av de omsøkte opplysninger (jf pol § 33). Kopi av konsesjonssøknad og innvilget konsesjon skal oversendes SSB sammen med søknad. For mer informasjon om personopplysningsloven, kontakt Datatilsynet.

Etter prosjektslutt skal de utleverte avidentifiserte dataene returneres SSB eller slettes.

Lister med navn og adresse leveres ikke ut. Dersom et prosjekt er avhengig av å kontakte de registrerte (spørreundersøkelser), kan SSB under gitte omstendigheter yte bistand. Dette kan enten gjennomføres ved at SSB påtar seg hele undersøkelsen, eller ved at SSB står for utsendelse av spørreskjema hvor de som ønsker å delta returner skjemaet direkte til forskeren. Det gjelder særskilte betingelser for slike undersøkelser. Ta kontakt for nærmere informasjon i det enkelte tilfelle.

Utvalgsundersøkelser eller data fra andre offentlige organer

Dersom opplysningene som ønskes utlevert er hentet inn ved frivillige undersøkelser, kreves særskilt samtykke fra den registrerte før utlevering kan finne sted. Ved utlevering av anonymiserte opplysninger kreves ikke slikt samtykke.

Dersom opplysningene som ønskes utlevert er hentet fra andre offentlige organer, må SSB klarere utlevering med det avgivende organ (eksempelvis Rikstrygdeverket, Arbeidsmarkedsetaten, Skattedirektoratet). Aktuelt fagdepartement må alltid dispensere fra taushetsplikt før utlevering finner

sted hvis det dreier seg om opplysninger som er underlagt taushetsplikt etter særlovgivning (eksempelvis helseopplysninger).

Krav til søknad

Det må sendes skriftlig søknad om utlevering av data. Ved forskningsprosjekt skal prosjektbeskrivelse vedlegges søknaden.

Søknaden skal inneholde følgende forhold:

- navn, adresse og stilling på de som skal ha befatning med dataene
- navn på (og eventuelt finansiering av) prosjektet som dataene skal brukes i
- formål med datautleveringen
- beskrivelse av utvalgsriterier
- liste over alle variabler og periode som skal være med i uttaket
- eventuelt beskrivelse av hvordan data skal organiseres

Vedlegg til søknaden:

- eventuelt kopi av konsesjon eller kopi av bekreftet mottatt melding fra Datatilsynet.
- prosjektbeskrivelse.

Søknaden bør inneholde en så detaljert beskrivelse som mulig av hva man ønsker, formålet med datautleveringen og dokumentasjon av behovet for enkeltkjennemerker, med særlig vekt på behovet for identifiserende opplysninger. Som regel vil det være fornuftig å drøfte uttaket med SSB før søknaden skrives slik at søkeren får de riktige dataene og at man unngår unødige forsinkelser.

Søknadsskjema

Vurdering

Mottakers formål med bruk av opplysningene vil vurderes nøye før SSB gir et tilbud.

En viktig forutsetning for utlevering av opplysningene er at søknaden "finnes rimelig og ikke medfører uforholdsmessig ulempe for andre interesser".

Ved godkjent søknad sender SSB et pristilbud, samt skriftlig orientering om betingelsene for utlevering bl.a. taushetsregler og regler om tilbakelevering/sletting. Før arbeidet settes i gang må SSB ha mottatt signert taushetserklæring og utleveringsavtale og skriftlig bekreftelse på tilbudet.

Personregistre i SSB - hvem svarer

En av SSBs oppgaver er å gi opplysninger om statistikk til bruk for forskningsformål og for offentlig planlegging. I den sammenheng er det etablert en Sentral brukertjeneste som skal veilede og bistå ved bestilling av data. Sentral brukertjeneste har et særskilt ansvar for tverrgående prosjekter.

sentral.brukertjeneste@ssb.no

Ved utlevering av data fra et fagområde, eller de fleste av variablene i oppdraget inngår i et register, kontaktes fagseksjoner/personer direkte.

Arbeidsmarked:

Arbeidstakerregisteret
Registerbasert sysselsettingsfil
Arbeidskraftundersøkelsen
SOFA

Kontaktperson:?

Data fra Det sentrale folkeregister

Situasjonsfil

Loadregister

Fødelandsfil

Flyttefil

Kontaktperson:?

Folketellinger:

FoB 60

FoB 70

FoB 80

FoB 90

britt.elin.braten@ssb.no , tlf: 62 88 51 43**Inntekt:**

Selvangivelsesregisteret

Ligningsregisteret

Inntektsregisteret

Kontaktperson:?

Sosialhjelp og trygd:

Sosialhjelpsmottakere

FD-Trygdbrukertjenesten.fd-trygd@ssb.no, tlf: 21 09 47 94**Utdanning:**

Befolkningens høyeste utdanning

Nasjonal utdanningsdatabase

Kontaktperson:?

6.2.4 Retningslinjer for deling av oppdrag

En fordel med en sentral database som FD-Trygd og senere også NUDB har vært samling av relevante opplysninger innen et og samme system (Oracle database) og konsekvent bruk av samme modellering for revisjon av datamaterialet og konstruksjon av datastrukturen (forløpsmodell).

Dette impliserer også at gjenbruken av data fra flere fagseksjoner i SSB i hovedsak er konsentrert om én lokalisering av data, FD-Trygd, og én inngangsportal, brukertjenesten FD-Trygd.

Brukertjenesten FD-Trygd er imidlertid delvis eksternt finansiert via Norges forskningsråd og alle oppdrag på FD-Trygd er dermed definert som gratis. Internt i SSB har leverandører til FD-Trygd gjerne oppfattet dette som et konkurranseforhold til egne oppdragsmidler. Samtidig er det slik at brukertjenesten faktisk kan ha hatt problemer med å fordele oppdrag som andre fagseksjoner burde ha tatt hele eller deler av. Dette fordi det ikke finnes klare skillelinjer for denne deling. Det er derfor nødvendig å sette opp et forslag til en slik deling.

Rent umiddelbart er det to hovedregler som skiller seg ut. Det ene er kundens problemstilling/ formål, som bør ligge innenfor trygdeforskningen eller velferdsforskningen. Imidlertid, dersom kunden skal ha data som stammer helt og holdent fra annen fagseksjon enn eier av FD-Trygd, er det naturlig å henvise kunden til rette fagseksjon. Det er mulig at disse to hovedregler kan komme i konflikt, men vi henviser den diskusjonen til de fire seksjonsledere som leverer data til FD-Trygd.

Brukertjenesten FD-Trygd kan heller ikke ta det administrative ansvaret for betalte oppdrag og risikere å møte seg selv i døren over tid. Et konkret eksempel er dersom kunden krever supplerende opplysninger i form av ikke-fristilte årganger eller manglende variabler i FD-Trygd. Brukertjenesten kan ikke hente inn og behandle ulike datakilder for å dekke nødvendige årganger, blant annet pga. kompleksiteten i å stille sammen ulike datamodeller. Dette gjelder både definering av utvalget og hvilke data som skal leveres ut. Over tid kan man risikere at samme type oppdrag er betalt i en sammenheng og gratis i neste omgang, dvs. en inkonsekvent kundebehandling. Hvis bestillingen krever data som ikke finnes i FD-Trygd, bør den/de aktuelle fagseksjonen(e) selv utføre arbeidet, aller helst med en koordinering fra sentral brukertjeneste dersom oppdraget deles mellom fagseksjon og FD-Trygd. Denne avklaring kan også lette arbeidssituasjonen for brukertjenesten FD-Trygd.

Et annet forhold er hvis sentral brukertjeneste finner det enklere å hente data til egne oppdrag fra FD-Trygd enn direkte fra fagseksjon. Det blir da et forhold mellom Sentral brukertjeneste og fagseksjon å dele oppdragsmidlene.

Mange kunder forsøker også andre datakilder (eks. sykehusopplysninger innen trygdemedisinforskningen) før de eventuelt går til FD-Trygd som er en sentral kilde for velferdsforskningen. Gjennom forslaget til retningslinjer for oppdragsdeling blir FD-Trygd en mulig kilde også for andre fagseksjoner for å hente inn oppdrag man ellers ikke ville mottatt.

Forslag til retningslinjer for deling av oppdrag:

1. Kunden kan gå til FD-Trygd dersom prosjektets **problemstillinger** i hovedsak konsentreres om trygdeforskning og/ eller velferdsforskningen - dette er da et gratis oppdrag.
2. Dersom kunden kun skal ha data som leveres fra **andre fagseksjoner**, henvises kunden til rette fagseksjon eller sentral brukertjeneste (i tilfelle flere fagseksjoner involveres) - dette er da et betalt oppdrag
3. Dersom kunden i tillegg til FD-Trygd skal ha data som ikke er tilgjengelig, f.eks. nyere årganger eller andre variable, må det bli en **oppdragsdeling** mellom FD-Trygd og fagseksjon.
Fagseksjonen må da tilrettelegge data som tilhører eget fagområde, ikke bare utfyllende data - oppdraget bør koordineres av sentral brukertjeneste for å skille mellom betalt og gratis oppdrag
4. På egne oppdrag kan **sentral brukertjeneste** hente data direkte fra FD-Trygd i stedet for fagseksjon selv om oppdraget ikke faller inn under punkt 1 eller 3, dersom det er det mest formålstjenlige for brukertjenesten selv - dette impliserer eventuell deling av oppdragsmidler mellom SB og fagseksjon

6.2.2 Forenklet saksbehandlingsprosess

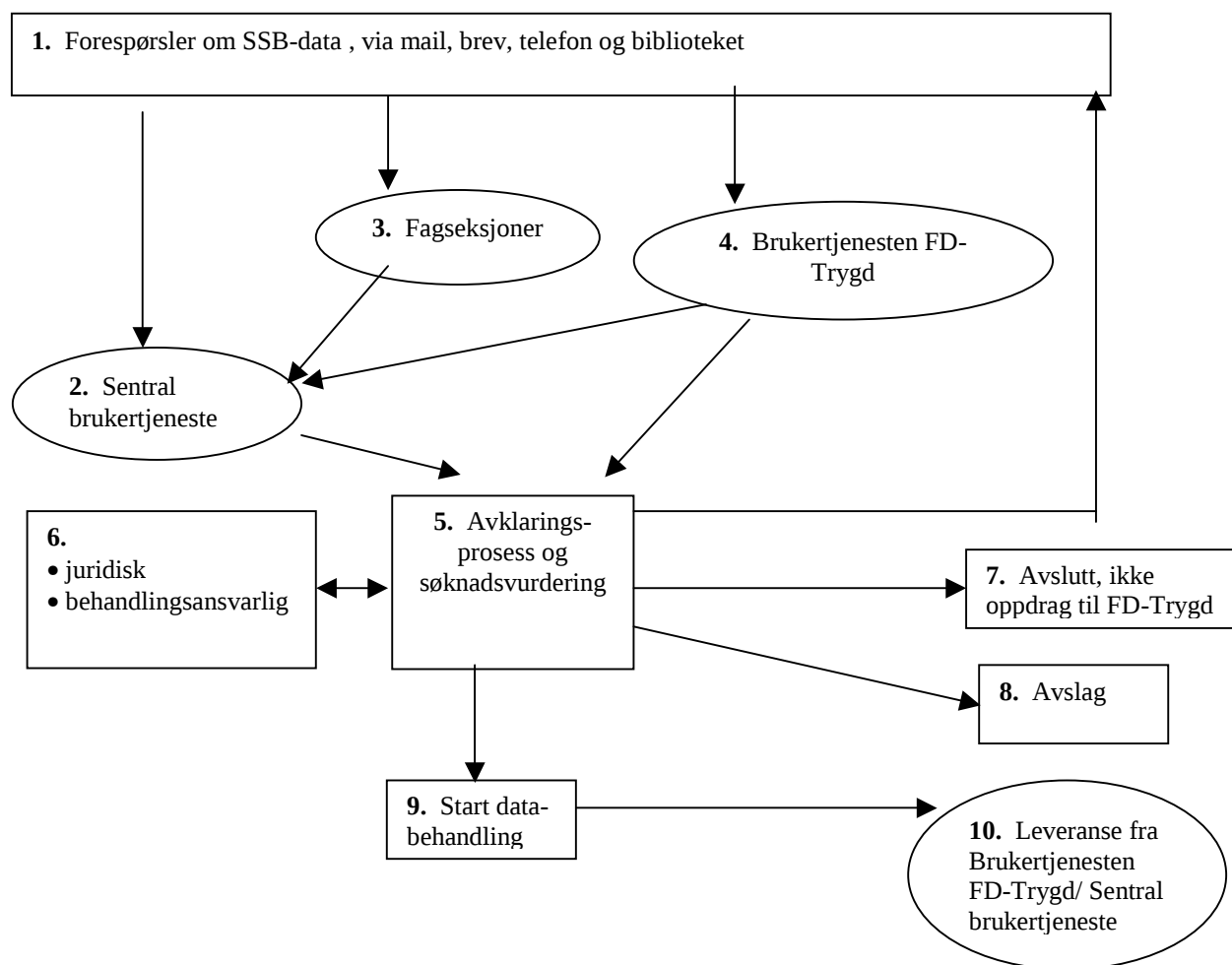
Målet med den generelle veiledningen på Web er at prosessen fra henvendelse fra en kunde til utlevering av data forenkles for både SSB og kunden. I tillegg vil retningslinjene for deling av oppdrag mellom Brukertjenesten FD-Trygd og fagseksjoner bidra til mindre saksbehandlingstid innad i SSB.

Prosessdiagram 3 nedenfor viser en forenklet saksbehandlingsprosess ved forespørsler av FD-Trygd data og andre datautleveringer fra SSB. Ved bruk av veiledningen på Web vil kunden som oftest henvende seg til riktig brukertjeneste i SSB, det vil si enten Sentral brukertjeneste eller Brukertjenesten FD-Trygd. Der hvor kunden ønsker data fra FD-Trygd og i tillegg data fra andre fagområder vil Sentral brukertjeneste utføre saksbehandling og leveranse av data. Dette fører til at Brukertjenesten FD-Trygd unngår bruk av tid til samarbeid med andre fagseksjoner og Sentral brukertjeneste. Det vil selvfølgelig være noen oppdrag som feilaktig havner hos Brukertjenesten FD-Trygd, men som videresendes direkte til Sentral brukertjeneste.

Etter fordeling av oppdrag mellom Sentral brukertjeneste og Brukertjenesten FD-Trygd vil prosessen være lik, derfor har vi tegnet boksene 2 og 4 til boks 5.

Som vi har nevnt tidligere er det vesentlig for saksbehandlingstiden at kunden beskriver formålet med dataleveransen så tidlig som mulig. Det vil si at en skriftlig henvendelse med vedlagt melding eller konsesjon fra Datatilsynet ved første henvendelse vil redusere saksbehandlingstiden. I forhold til prosessdiagram 2a og 2b vil det si at avklaringsprosessen og søknadsvurderingen (boks 5) foregår samtidig i den ideelle prosessen.

Prosesdiagram 3. Oversikt over prosessen ved forespørsler av FD-Trygd data til SSB og til leveranse av data.



6.3. utfordringer

6.3.1 Generelt dokumentasjonsnotat

Som et ganske komplisert og omfattende pilotprosjekt har prosjektet FD-Trygd i stor grad konsentrert dokumentasjonsarbeidet om å nedtegne den revisjonsprosess og omstrukturering av datamaterialet som ligger bak oppbyggingen av forløpsdatabasen. Å forklare modellen og hvorfor data er organisert som de er på det konkrete temaområde, med dateringer på mange analysevariabler i tillegg til start og stopp opplysninger på både hovedforløpsnivå og gruppeforløpsnivå, er en vanskelig sak nok i seg selv.

Det må skrives et dokumentasjon som er langt mer målrettet og konsentrert om produktet, resultatet uten å henviser til den revisjonsprosess som leder frem til databasen. For de fleste sluttbrukere tør vi hevde at denne informasjon er uten relevans. FD-Trygd har laget en metadatabase for generelle forløpsdatabaser og det gjenstår å definere ulike uttak til faste standardrapporter for dette formål. Her kan man da få dekket informasjonsbehovet til både interne statistikkprodusenter på forløpsprosjekter og til sluttbrukere av denne type data.

Et slikt generelt dokumentasjonsnotat bør inneholde:

- En teoretisk innledning om forløpsdatabaser, kan gjerne ta utgangspunkt i kapittel 2.1 i dette notat.
- En forklaring på datastruktur og modell, slik den er implementert på det enkelte temaområde.
- En tabell for mengden av individ ved utgangen av hvert år, på det enkelte temaområde.
- Variabelliste
- Veiledning til bestilling av mikrodata om personer til forskning og offentlig planlegging, jf forslag i kapittel 6.2.1

Det generelle dokumentasjonsnotatet bør legges på Web-siden til FD-Trygd.

I tillegg bør det lages enkeltstående notat om hvert temaområde, med grundigere innføring og flere kontrolltall på relevante variable.

6.3.2 Opplysninger om uttaksvalg

Målsettingen og ambisjonen er altså høy og resulterer i et komplisert datamateriale med mange opplysninger (records) pr. observasjon (forløpsstilfelle) og kan skremme "lurven" av brukere som er vant til mer tradisjonell programvare som SAS eller SPSS. Det er her vi tror at mer konsentrasjon om forholdet mellom databasen og uttakssiden kan bidra til større oppmerksomhet om og lette tilgjengeligheten til fenomenet forløpsdata.

Det må opplyses bedre om hvilke muligheter man har til forenklede uttak av datamengden. Opplysninger hentet ut kun på enten start- eller stopp-tidspunktet er ikke ueffekt for viderebehandling i andre program enn Oracle. Synliggjøring må til for å overkomme det hinder at ikke alle brukere trenger å forstå modellen til punkt og prikke for å ta den i bruk.

6.3.3 Veiledning på Web og deling av oppdrag

Forslaget til veiledning på Web (i kapittel 6.2.1) må sendes ut på høring til berørte seksjoner, det vil si seksjoner som leverer ut mikrodata. Seksjonene må selv føre på aktuelle registre og kontaktpersoner. I tillegg må det lages en mailadresse for Sentralbrukertjeneste og for Brukertjenesten FD-Trygd. Det bør også vurderes om Web-siden bør inneholde informasjon om utlevering av mikrodata til utenlandske forskere.

I tillegg bør det lages et standardisert søknadsskjema, som tar hensyn til de spørsmål som SSB mener er nødvendig å hente inn informasjon om, for å utlevere data. En slikt søknadsskjema må utformes av personer i SSB som er ansvarlige for juridiske spørsmål, med kjennskap til avtalen med Datatilsynet og i tillegg saksbehandlere i SSB som arbeider med utlevering av mikrodata.

Seksjonsledere som er ansvarlig for statistikk til FD-Trygd og Brukertjenesten FD-Trygd må møtes for å drøfte forslaget til retningslinjer.

1. Prosjektskriv

Til:	Johan-Kristian Tønder		
Fra:	Seksjonsleder Berit Otnes	Seksjon 350	Dato 27.11.01

Prosjektnavn Bruk av forløpsdatabaser i SSB		Produktnr. 0765
Oppdragsgiver SSB		
Ansvarlig (utførende) enhet Seksjon for levekårsstatistikk		
Prosjektleder Tonje M. Köber	Kvalitetslos Grete Olsen	

Se Statistisk sentralbyrås håndbøker 58 «Håndbok i prosjektstyring»

1. Bakgrunn

Her beskrives kort hvorfor/hvordan prosjektet har oppstått.

Prosjektet ble levert inn som forslag til kurset Prosjekt- og teamarbeid fra Seksjon for levekårsstatistikk. Bakgrunnen var at SSB's forløpsdatabaser inneholder en stor mengde av opplysninger som er viktig for samfunnsnyttig forskning og analyse, særlig i forbindelse med velferdsforskning. En har det inntrykket at det er liten bruk av eksisterende forløpsdatabaser i SSB, og dermed blir analysene i samfunns-/velferdsforskningen noe begrenset.

2. Formål

Forslag til målsettinger for prosjektarbeidet.

Formålet med dette prosjektet er å vise nytten av å ta i bruk data fra SSB's forløpsdatabaser med hovedvekt på FD-trygd, i tillegg Nasjonal utdanningsdatabase (Nudb).

Målet er at disse databasene skal synliggjøres for å bidra til å øke bruken av disse både internt og eksternt. Av de eksterne brukerne skal vi bare se på brukere som har mottatt data fra SSB via sentralbrukertjeneste, brukertjenesten FD-trygd eller i kombinasjon med andre fagseksjoner.

Erfaringene vil samles i et notat, som vil peke på tiltak for å bedre vise nytten av å ta i bruk denne typen data fra SSB. En vurdering av eksisterende dokumentasjon med tanke på brukernes behov inngår også i prosjektets rammer.

3. Faglige problemstillinger

Vektlegg kort faglige forutsetninger og problemstillinger.

Dette punktet utformes i detalj senere, i forbindelse med utformingen av milepælplanen.

For å kunne si noe om nytten av og ta i bruk data fra forløpsdatabasene må vi undersøke etterspørsel og behov hos brukerne. Brukerne kan deles inn i to grupper: potensielle brukere og aktuelle brukere (de som bruker forløpsdatabasene i dag). Disse kan igjen deles inn i to grupper: eksterne og interne brukere. Vi skal kartlegge brukerkravene til aktuelle interne og eksterne brukere og etterspørsel og behov hos potensielle interne.

Ved hjelp av å kartlegge disse ulike gruppene vil vi få innblikk i hva brukerne ønsker av informasjon fra forløpsdatabasene og hvordan de ønsker data behandlet og levert.

Vi skal intervju brukere ved hjelp av personlig intervju og spørreskjema. De potensielle brukerne vil vi møte i grupper og holde foredrag, og kartlegge gruppens informasjon og interesse for å bruke forløpsdatabaser før og etter foredraget.

4. Interessenter og deltakere Hvem er oppdragsgiver, og hvem er prosjektleder.

Hvem er øvrige interessenter i eller utenfor Statistisk sentralbyrå. Hvilke fagområder eller organisatoriske enheter må bidra.

Prosjektet er i utgangspunktet et prosjekt til utviklingsprogrammet Prosjekt- og teamarbeid i SSB.

Prosjektgruppen består av Tonje M. Köber (prosjektleder, S260), Hossein Moafi (S230), Elisabeth Rønning og Øyvind Sivertstøl (S350).

Følgende utgjør styringsgruppe for prosjektet: Berit Otnes (leder) og Elisabetta Vassenden.

Følgende referansegruppe for prosjektet er etablert: S320 Befolkning- og utdanning, Toril Vangen, S303 Person statistikk IT, Anders Akselsen, S310 Demografi- og levekårsforskning, Trude Lappegård, S510 Offentlig økonomi og personmodeller, Dag Rønningen og S260 Arbeidsmarkedsstatistikk, Magne Bråthen.

5. Ressursrammer Hvilke økonomiske, tidsmessige og personalmessige ressursrammer foreligger.

Prosjektet skal være ferdig 4. april 2002.

Prosjektarbeidet er beregnet til ett dagsverk per uke per deltaker i 25 uker frem til april, om lag 2,7 månedsverk totalt for prosjektgruppen.

6. Prosjektets prioritet ...også i forhold til andre oppgaver.

Prosjektet har høy prioritet for gruppen som deltar på kurset Prosjekt- og teamarbeid.

2. Fra miniregister/grunnlagsstatistikken til forløpstabell - et eksempel

Tabellen i avsnitt 2.4.3 angir hvordan "tabellene" ser ut for en enkelt månedsfil med en record pr. person (tverrsnitt) Etter at miniregistrene er opprettet og kontrollert lastes dataene til basen hvor de omorganiseres til forløpstabeller. Dette innebærer at alle månedsfilene settes sammen i en forløpsstruktur for alle årgangene fordelt etter temaområde. Det viktigste i denne prosessen er at alle records knyttes til ulike dateringer som legges inn i tabellene for å angi hendelsesforløpet.

Det er tre typer av dateringer, tilgang-, avgang- og hendelsesdateringer. Tilgang- og avgangsdatoer angir naturlig nok når en person begynte/sluttet å motta en ytelse. Hendelsesdateringer angir *når* det har skjedd endringer og *hva* som har endret seg. Datovariablene gjør at vi oppnår en stor gevinst i forløpsdatabasen: en record ligger i basen bare én gang og over tid så lenge det ikke skjer noe "nytt". Mao. en person som er registrert med tilgang til uførepensjon i januar, og hvor opplysningene er identiske på de påfølgende månedsfilene til f.eks. desember, blir liggende med én record for hele perioden frem til det skjer en endring eller avgang. I dette tilfellet vil 12 record fra RTV reduseres til en record i FD-Trygd uten at vi mister informasjon.

For å redusere abstraksjonsnivået kan et eksempel knyttet til pensjoner illustrere strukturen i et typisk forløp (jf. med vedlegg 3).

Tabellene nedenfor kan være et utsnitt fra forløpstabellen for uførepensjon. Person identifiseres med fødselsnummer i første kolonne. De fire neste kolonnene er dateringer for tilgang og avgang, forskjellen mellom dem beskrives nærmere nedenfor. *Regdato* (registreringsdato) er en hendelsesdatering. Denne er generell for alle forløpstabeller og vil fungere som en siste oppdaterings-/endringsdato for hele recorden sett som en helhet, uavhengig om det er tilgang, avgang eller endring. Denne har to formål. Den angir alltid siste registrerte hendelsesdato for recorden og den refererer til hvilke andre variabler som har endret seg. Temakode er i dette tilfellet *Pnsjkode* som er beskrevet tidligere. *Ufg* (uføregrad) er en hendelsesvariabel mens *ufgdato* er en hendelsesdato knyttet til eventuelle hendelser i uføregraden.

TILGANGSRECORD

-----dateringsvariabler-----					-temakode	-- variabler-----		
Fnr	Tilgdato	Avgdato	Pnsjtilg	Pnsjavg	Regdato	Pnsjkode	Ufg	Ufgdato
12345	199504		199504		199504	21	50	199504

Når en person registreres inn i en ordning vil det genereres en tilgangsrecord som vist ovenfor. *Pnsjkodens* 1. siffer = 2 og angir at dette er uførepensjon. Andre siffer =1 og angir at det er en tilgangsrecord. *Pnsjtilg* viser at personen mottar uførepensjon fra 199504. Merk at tilgangsdato, regdato og variabeldato normalt vil være identiske på en tilgangsrecord fordi alle datoer refererer til status ved oppstart. Kun tilgangsrecorden vil ligge i basen frem til det skjer en endring/avgang.

ENDRINGSRECORD

Anta at nye månedsfiler fra miniregisteret viser at personen har endret uføregrad fra 50 til 100%. Personen er altså fortsatt aktiv i ordningen men en variabel har endret seg og dette skal fremkomme i en endringsrecord.

Fnr	Tilgdato	Avgdato	Pnsjtilg	Pnsjavg	Regdato	Pnsjkode	Ufg	Ufgdato
12345	199504		199504		199607	22	100	199607

Pensjonskodens andre siffer er lik 2 og angir at det er en endringsrecord. Gitt at vi hadde tatt med alle variablene kan det være vanskelig å se hva som er endret. Vi kan da benytte *regdatoen* som angir den siste hendelsen. Ved å finne en variabeldato som har samme verdi som regdato vet vi at denne refererer til hvilken variabel som er endret. I dette tilfellet er regdato og *ufgdato* identiske og vi vet at

uføregraden er endret. Merk at tilgangsdatoen fortsatt er den samme fordi denne angir når tilfellet startet og avgangsdato er blank, altså vedkommende er fortsatt aktiv innenfor ordningen.

AVGANGSRECORD

Hvis en person har hatt avgang fra ordningen, dvs. kontroll av miniregisteret, viser at vedkommende ikke mottar ytelsen lenger og vi setter en avgangsdato. Merk at denne settes til siste måned tilfellet mottok ytelsen. I vårt tilfellet mottar personen uførepensjon siste gang i 199703.

Fnr	Tilgdato	Avgdato	Pnsjtilg	Pnsjavg	Regdato	Pnsjkode	Ufg	Ufgdato
12345	199504		199504	199703	199703	20	100	199607

Pensjonskodens andre siffer er nå 0 og indikerer at dette er en avgangsrecord. Regdato vil da være identisk med avgangsdato fordi dette er den sist registrerte hendelsen. Legg merke til at ufgdato fortsatt er den samme som på endringsrecorden. Den viser kun når det sist skjedde en endring i uføregrad. Men en endring kan også registreres på samme tidspunkt som en avgangsrecord, og vil da ha samme dato som regdato.

Pensjoner er delt inn etter grupper. Dette er utgangspunktet for at det er to tilgangs- og avgangsdatoer. *Tilgdato* refererer til startdato for tilfellet, dvs. når personen første gang mottok en av ytelsene definert under pensjoner. *Pnsjtilg* er en gruppetilgang, den sier hvilken type pensjon tilgangen gjelder.³ Anta at personen går fra uførepensjon og over på alderspensjon. En tilgangsrecord for alderspensjon for samme person kan da se slik ut.

Fnr	Tilgdato	Avgdato	Pnsjtilg	Pnsjavg	Regdato	Pnsjkode	Apd	Apddato
12345	199504		199704		199704	11	100	199704

Tilgdato viser samme tilgangsdato som for uførepensjon og indikerer altså at tilfellet har hatt en annen løpende ytelse før tilgang til alderspensjon. *Pnsjtilg* er nå 199704 og første siffer i pnsjkode =1 som er alderspensjon. Fordi alderspensjon normalt er siste ytelse man mottar settes avgdato når vedkommende dør og vil være identisk med pnsjavg. *Tilgdato* og *avgdato* viser da den sammenhengende perioden vedkommende mottok en pensjonsytelse. Dette kan selvsagt også inkludere flere ytelser definert under pensjoner.

Fnr	Tilgdato	Avgdato	Pnsjtilg	Pnsjavg	Regdato	Pnsjkode	Apd	Apddato
12345	199504	199806	199704	199806	199806	10	100	199704

Her er det viktig å være oppmerksom på følgende. Denne sammenhengen gjelder bare hvis det er kontinuitet mellom ytelsene. Hvis vedkommende ikke hadde gått rett over i alderspensjon og ikke mottok andre ytelser ville tilgdato og avgdato være identisk med pnsjtilg og pnsjavg på *avgangsrecorden for uførepensjon*. Tilsvarende ville tilgdato og pnsjtilg vært identisk på alderspensjon fordi dette ville blitt betraktet som et nytt tilfelle. Dette er av den enkle grunn at vi ikke kan sette noen dato hvis personen ikke finnes i noen av registrene for de ulike ytelsene.

³ Denne inndelingen vil variere noe etter temaområde, f.eks. har foreløpig uførestønad ingen undergrupper. Tilgdato vil da alltid være identisk med gruppetilgangs dato.

3. Veiledning til bestilling av data som ligger på Web-siden til FD-Trygd

Henvendelse ved bestilling av data fra FD-Trygd

I utgangspunktet skal det alltid foreligge en skriftlig søknad om datautlevering fra FD-trygd. En søknad bør inneholde:

navn, adresse og stilling på de som skal ha befatning med dataene
navn på (og eventuelt finansiering av) prosjektet som dataene skal brukes i
formål med datautleveringen
beskrivelse av utvalgsriterier
liste over alle variabler som skal være med i uttaket (jf. variabelliste)
beskrivelse av hvordan data skal organiseres.
evt. konsesjon ved søknad om utlevering av ikke-anonymiserte data
Søknaden bør inneholde en så detaljert beskrivelse som mulig av hva man ønsker, formålet med datautleveringen og dokumentasjon av behovet for enkeltkjennemerker, med særlig vekt på behovet for identifiserende opplysninger. Som regel vil det vil det være fornuftig å drøfte uttaket med SSB/NSD før søknaden skrives slik at søkeren får de riktige dataene og at man unngår unødige forsinkelser.

I de tilfellene der det kreves konsesjon kan forskere og studenter henvende seg til Datafaglig sekretariat ved NSD. Dette er et nasjonalt servicesenter der forskere og studenter kan få råd og veiledning i forbindelse med bruk av persondata. Sekretariatet rapporterer også overfor Datatilsynet hvilke forskningsprosjekter som har bruk for konsesjon. For forvaltningen og andre brukere anbefales det å ta direkte kontakt med Datatilsynet.

For service til forskere og studenter er det inngått en egen avtale mellom Statistisk sentralbyrå og Norsk Samfunnsvitenskapelig datatjeneste (NSD) som gir NSD adgang til å betjene disse miljøene.

Forskere/studenter henvender seg til:

NSD
Hans Holmboesgate 22
5007 Bergen
Tlf: 55 58 21 17

Statsforvaltningen og andre brukere henvender seg til:

Statistisk sentralbyrå
v/Øyvind Sivertstøl S 350
Kongensgate 6
P.B. 8131 Dep.
0033 Oslo
Tlf: 21 09 47 94
E-mail: oyvind.sivertstol@ssb.no

4. Foredrag

Foil 1:

BRUK AV FORLØPSDATABASER I SSB

Prosjekt i kurset Prosjekt- og teamarbeid.

Øyvind Sivertstøl og Tonje Köber

Foil 2:

Hva vet dere om de to forløpsdatabasene i SSB

FD-trygd

NUDB, utdanning

Brukes FD-trygd allerede på deres seksjon

Foil 3:

NUDB; Nasjonal utdanningsbase

Utdanningsdata fra og med 1970

Målet er og ta utgangspunkt i hvilken som helst gruppe, på et hvilket som helst tidspunkt og følge disse enten fremover eller bakover.

Klar til bruk høsten 2002

Foil 4:

Hva

Mikrodata, dvs. på individnivå

Som kan brukes til:

Tverrsnittsanalyser

Forløpsanalyser

Foil 5:

Forløpsanalyser

Kan vise hvor lenge (varighet) en person står registrert på en ordning(f.eks. uføretrygdet), hvor lang tid før overgang til annen ordning (f.eks. sysselsatt) og hvilke faktorer som påvirker overgangen fra en status til en annen.

Foil 6:

FD-trygd

Inneholder alle bosatte og ikke bosatte (døde og utvandret)

Første årgang er 1992

Bygger på administrative data fra RTV, AETAT og SSB

Foil 7:

Bakgrunn

Levering av tryggedata til folketellingen

Forskning

Oppstart i 1996, ferdig til bruk i februar 1999

Foil 8:

Temaområdene i FD-trygd

Bakgrunnsopplysninger

demografi (1992-1997)

inntekt og formue (1992-1999)
utdanning (1992-1997, avsluttet pga. NUDB)

Foil 9:

Temaområdene i FD-trygd

forts.

Trygdeopplysninger

folketrygdpensjoner (1992-2000)
grunn- og hjelpestønader (1992-2000)
foreløpig uførestønad (1992-1998)
enslige forsørgere (1992-1999)
økonomisk sosialhjelp (1992-1997)
fødsels- og sykepenges (1992-1999)
attføring (1992-1999)

Foil 10:

Temaområdene i FD-trygd

forts.

Arbeidsmarkedsdata

sysselsetting (1992-1998)
arbeidssøkere (1992-1999)
statsansatte (1992-1999)

Foil 11:

Sysselsettingsdata

Seksjon for arbeidsmarked

Arbeidstakerregisteret
Liknings- og momsregisteret

FD-trygd

Arbeidstakerregisteret med LTO-forhold
Liknings- og momsregisteret

Som er konsistentbehandlet med SOFA

Foil 12:

Hvordan ta i bruk/skaffe data fra FD-trygd

Bestemme utvalget og variabler selv vha egne administrative registre eller utvalgsundersøkelse
Utvalget bestemmes vha brukertjenesten v/FD-trygd, med ønskede variabler

Uttrekk og kobling utfører brukertjenesten, men bestiller må informere fagdirektør på avd. 300

Foil 13:

Eksempler på bruk av dataene

- Kartlegge og studere ulikheter i trygdekarrierer for grupper av personer.
- Fra trygd til arbeid, eller se på hvilke personer som kombinerer trygd og arbeid.
- Evaluere effekter av regelverksendringer knyttet til velferdsordninger.

Foil 14:

”Og hva så”?

- Spennende – økt lysten...
- Interessant for seksjonen?
- Noen prosjekt/områder FD-trygd eller NUDB kan brukes hos dere?
- Vil bruk av forløpsdatabaser forenkle arbeidet hos dere?

5. Spørreskjema

Kjennskap til data fra forløpsdatabasen FD-Trygd

NAVN:

Svarfrist: 5. februar

1. Hvordan fikk du vite om forløpsdatabasen FD-trygd?

- ☐ SSBs publikasjoner / Web
- ☐ Institusjonen du jobber på
- ☐ Annet, spesifiser:

2. Hva fikk du vite der som var interessant for deg/ditt arbeidsområde?

3. Vet du om andre institusjoner/bedrifter enn din arbeidsplass som kan ha nytte av denne databasen?

- ☐ Ja - *fortsett på spørsmål 4.*
- ☐ Nei - *gå til spørsmål 7.*

4. Hvilke andre mener du kan ha nytte av denne databasen? Spesifiser:

5. Bruker noen av disse forløpsdatabasen FD-Trygd?

- ☐ Ja - *fortsett til spørsmål 7*
- ☐ Nei - *gå til spørsmål 6*
- ☐ Vet ikke

6. Vet du hvorfor de ikke bruker forløpsdatabasen FD-Trygd?

- ☐ Ja - spesifiser, og gå videre til spørsmål 7:
- ☐ Nei - *gå til spørsmål 6*

7. Har du forslag til hvordan man kan gjøre denne databasen bedre kjent?

- ☐ Ja - *fortsett til spørsmål 8.*
- ☐ Nei - *gå til spørsmål 9.*

8. Beskriv hvordan man kan gjøre denne databasen bedre kjent:

Tilgjengelighet

9. Vet du hvem som er ansvarlig for brukertjenesten FD-Trygd?

- ☐ Ja - *fortsett til spørsmål 11*
- ☐ Nei - *gå til spørsmål 10*

10. Hvordan fant du fram til brukertjenesten FD-Trygd?

(Flere kryss er mulig)

- ☐ Via sentral brukertjeneste i SSB
- ☐ Via biblioteket i SSB
- ☐ Via andre seksjoner i SSB
- ☐ Via SSBs Internettsider

- ☐ Via sentralbordet til SSB
- ☐ Via personer / bekjente som jobber i SSB
- ☐ Via kolleger
- ☐ Andre, spesifiser:

11. Har du sett veiledninger for framgangsmåter for uttak fra forløpsdatabasen FD-Trygd, på Internett, i dokumentasjonsnotater eller i annet materiale?

- ☐ Ja - *fortsett til spørsmål 12*
- ☐ Nei - *gå til spørsmål 13*

12. Hva synes du om veiledningen(e)?

13. Var det lett å finne fram til ansvarlig for brukertjenesten FD-Trygd?

- ☐ Ja
- ☐ Nei

14. Hvilken forventning hadde du til leveringstid fra ferdig spesifisert og godkjent bestilling til første hoveddataleveranse?

Antall uker:

Ev. antall måneder:

15. Hvor lang tid tok det fra ferdig spesifisert og godkjent bestilling til første hoveddataleveranse?

Antall uker:

Ev. antall måneder:

16. Dersom du har mottatt data eller er i ferd med å spesifisere og motta data, hvilke problemer har du støtt på i forbindelse med å ta ut data fra forløpsdatabasen FD-trygd?

- ☐ Ingen problemer
- ☐ Problemer, spesifiser:

17. Dersom du skal benytte eller allerede har benyttet andre datakilder fra SSB enn de som er tilgjengelige innen forløpsdatabasen FD-trygd:

- a) ville det vært ønskelig om disse data også var stilt opp etter samme type forløpsmodell som FD-trygd er bygd opp etter? Vi tenker her på opplysninger som endres kontinuerlig over tid og som det kan dannes forløp på (ikke statistiske opplysninger)
- b) eller ville du foretrukket tradisjonell dataleveranse, eventuelt tverrsnittsutak på bestemte tidspunkt
- c) eller andre alternativ

Ønsket alternativ og begrunnelse:

18. Vil du anbefale andre å bruke forløpsdatabasen FD-Trygd?

- ☐ Ja
- ☐ Nei

Bruk og nytte

19. Hvor mange ganger har du søkt om data fra forløpsdatabasen FD-Trygd (vi er opptatt av antall hovedleveranser av data, ikke oppdateringer av årganger på et og samme prosjekt)?

Antall ganger:

20. Når søkte du om hovedleveranse av data første gang, og for hvilket prosjekt?

Måned, år:

Prosjekttittel:

21. Hva var hovedproblemstillingen for prosjektet?

22. Når søkte du om hovedleveranse av data sist gang, og for hvilket prosjekt?

Måned, år:

Prosjekttittel:

23. Hva var hovedproblemstillingen for prosjektet?

24. Hvilken type data har du brukt fra forløpsdatabasen FD-Trygd?

(Flere kryss er mulig)

- ☐ Data på individnivå
- ☐ Data på aggregert nivå
- ☐ Data på individ- og aggregert nivå

For spørsmål 24 til 35 ber vi deg svare for siste gang du fikk hovedleveranse av data til prosjekt fra FD-trygd hvis du har hatt flere uttak. Hvis du har fått oppdateringer av årganger etter dette, ber vi deg se bort fra dette og konsentrere deg om siste hovedleveranse av data på et prosjekt.

25. Var de utleverte data fra forløpsdatabasen FD-Trygd i samsvar med dine ønsker?

- ☐ Ja - fortsett til spørsmål 27
- ☐ Nei - gå til spørsmål 26

26. Hvilke mangler var det? Spesifiser:

27. Var det noen variabler du hadde bruk for som ikke finnes i forløpsdatabasen FD-Trygd?

- ☐ Ja - fortsett til spørsmål 28
- ☐ Nei - gå til spørsmål 29

28. Hvilke variabler? Spesifiser:

29. Fikk du data for de årgangene du ønsket?

- ☐ Ja
- ☐ Nei

30. Hadde du behov for å koble data fra forløpsdatabasen FD-Trygd sammen med andre data?

- ☐ Ja - *fortsett til spørsmål 31*
☐ Nei - *gå til spørsmål 33*

31. Var data fra forløpsdatabasen FD-Trygd lett å koble sammen med andre data?

- ☐ Ja - *fortsett til spørsmål 32*
☐ Nei - *gå til spørsmål 33*

32. Hva var problemet eller problemene? Spesifiser:

33. Endret brukerkravene dine seg underveis etter mottak av data - ønsket du noe mer, nye variabler, nytt uttak m.m ?

- ☐ Ja
☐ Nei

34. Vil du søke om å hente ut data fra forløpsdatabasen FD-Trygd ved en senere anledning?

- ☐ Ja - *Ferdig!*
☐ Nei - *gå til spørsmål 35*

35. Hvorfor ikke? Spesifiser:

TUSEN TAKK FOR HJELPEN!

6. Grunnlagstabeller - resultater fra spørreskjema

1. Hvordan fikk du vite om forløpsdatabasen FD-trygd?

	I alt	Interne	Ekterne
I alt	19	12	7
a) SSBs publikasjoner	5	3	2
b) Institus. du jobber på	6	5	1
a) og b)	2	2	-
c) Folketelling	1	1	-
c) Kirut	2	-	2
c) Kolleger	1	1	-
b) og c)	1	-	1
a), b) og c)	1	-	1

3. Vet du om andre institusjoner/bedrifter enn din arbeidsplass som kan ha nytte av denne databasen

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	17	10	7
Nei	2	2	-

5. Bruker noen av disse forløpsdatabasen FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	7	6	1
Nei	1	1	-
Vet ikke	10	4	6
Uoppgitt	1	1	-

6. Vet du hvorfor de ikke bruker forløpsdatabasen FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	3	1	2
Nei	5	3	2
Uoppgitt	11	8	3

7. Har du forslag til hvordan man kan gjøre denne databasen bedre kjent?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	6	5	1
Nei	11	7	4
Uoppgitt	2	-	2

9. Vet du hvem som er ansvarlig for brukertjenesten FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	13	10	3
Nei	5	2	3
Uoppgitt	1	-	1

10. Hvordan fant du fram til brukertjenesten FD-Trygd?

	I alt	Intern	Ekstern
I alt	25	19	7
a) Via sentral brukertjeneste i SSB	-	-	-
b) Via biblioteket i SSB	-	-	-
c) Via andre seksjoner i SSB	3	2	1
d) Via SSBs Internettsider	3	-	3
e) Via sentralbordet til SSB	1	-	1
f) Via personer / bekjente som jobber i SSB	6	4	2
g) Via kolleger	9	9	-
h) Andre, spesifiser	-	-	-
Uoppgitt	4	3	1

11. Har du sett veiledninger for framgangsmåter for uttak fra forløpsdatabasen FD-Trygd, på Internett, i dokumentasjonsnotater eller i annet materiale?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	7	4	3
Nei	11	7	4
Uoppgitt	1	1	-

13. Var det lett å finne fram til ansvarlig for brukertjenesten FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	9	5	4
Nei	6	4	2
Vet ikke	1	-	1
Uoppgitt	3	3	-

14. Hvilken forventning hadde du til leveringstid fra ferdig spesifisert og godkjent bestilling til første hoveddataleveranse?

	I alt	Intern	Ekstern
I alt	19	12	7
1 - 2 uker	4	2	
1 - 2 måned	7	4	3
Husker ikke	2	1	1
Uoppgitt	4	3	1

15. Hvor lang tid tok det fra ferdig spesifisert og godkjent bestilling til første hoveddataleveranse?

	I alt	Intern	Ekstern
I alt	19	12	7
Under 1 md.	3	1	2
1 måned	1	-	1
3 måneder	2	2	-
6 måneder	2	-	2
9 måneder	1	1	-
Noen md..	2	1	1
Raskt	1	-	1
Husker ikke	2	2	-
Uoppgitt	5	5	-

16. Dersom du har mottatt data eller er i ferd med å spesifisere og motta data, hvilke problemer har du støtt på i forbindelse med å ta ut data fra forløpsdatabasen FD-trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ingen probl.	7	4	3
Problemer	9	5	4
Uoppgitt	3	3	-

17. Dersom du skal benytte eller allerede har benyttet andre datakilder fra SSB enn de som er tilgjengelige innen forløpsdatabasen FD-trygd:

	I alt	Intern	Ekstern
I alt	19	12	7
a)	9	4	5
b)	5	4	1
c)	-	-	-
Vet ikke	1	1	-
Uoppgitt	4	3	1

18. Vil du anbefale andre å bruke forløpsdatabasen FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	15	10	5
Nei	1	1	-
Vet ikke	2	-	2
Uoppgitt	1	1	-

19. Hvor mange ganger har du søkt om data fra forløpsdatabasen FD-Trygd prosjekt)?

	I alt	Intern	Ekstern
I alt	19	12	7
Ikke fått data	1	1	-
1 gang	12	9	3
2 +	5	1	4
Uoppgitt	1	1	-

20. Når søkte du om hovedleveranse av data første gang, og for hvilket prosjekt?

	I alt	Intern	Ekstern
I alt	19	12	7
1998	1	1	-
1999	5	1	4
2000	4	3	1
2001	6	4	2
Uoppgitt	3	3	-

22. Når søkte du om hovedleveranse av data sist gang, og for hvilket prosjekt?

	I alt	Intern	Ekstern
I alt	18	12	6
2000	1	-	1
2001	5	2	3
2002	1	1	-
Uoppgitt	11	9	2

24. Hvilken type data har du brukt fra forløpsdatabasen FD-Trygd? (Flere kryss er mulig)

	I alt	Intern	Ekstern
I alt	19	12	7
Data på individnivå	12	13	4
Data på aggregert nivå	1	-	1
Begge deler	3	1	2
Uoppgitt	2	2	-

25. Var de utleverte data fra forløpsdatabasen FD-Trygd i samsvar med dine ønsker?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	11	6	5
Nei	3	1	2
Uoppgitt	5	5	-

27. Var det noen variabler du hadde bruk for som ikke finnes i forløpsdatabasen FD-Trygd?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	7	5	2
Nei	7	3	4
Uoppgitt	5	4	1

29. Fikk du data for de årgangene du ønsket?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	15	8	7
Nei	2	2	-
Uoppgitt	2	2	-

30. Hadde du behov for å koble data fra forløpsdatabasen FD-Trygd sammen med andre data?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	9	4	5
Nei	6	4	2
Uoppgitt	4	4	-

31. Var data fra forløpsdatabasen FD-Trygd lett å koble sammen med andre data?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	7	2	5
Nei	-	-	-
Uoppgitt	12	10	2

33. Endret brukerkravene dine seg underveis etter mottak av data - ønsket du noe mer, nye variabler, nytt uttak m.m ?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	5	3	2
Nei	10	6	4
Uoppgitt	4	3	1

34. Vil du søke om å hente ut data fra forløpsdatabasen FD-Trygd ved en senere anledning?

	I alt	Intern	Ekstern
I alt	19	12	7
Ja	13	8	5
Nei	2	2	-
Vet ikke	2	-	2
Uoppgitt	2	2	-

De sist utgitte publikasjonene i serien Notater

- | | | | |
|---------|---|---------|---|
| 2001/77 | G. Haakonsen: Beregninger av utslipp til luft av klimagasser. En gjennomgang av arbeidsprosess og dokumentasjon. 39s. | | framskrivinger basert på nasjonale anslag med modellene MODAG (1997-2005) og MSG (1995-2020). 35s. |
| 2002/1 | P. Scøning: Statistikk for 16 tettsteder og deres sentrumsarealer. Et innspill til programmet for utvikling av miljøvennlige og attraktive tettsteder i distriktene. 58s. | 2002/13 | H. Madsen og L. S. Stambøl: Kontrafaktiske regionale beregninger ved hjelp av SSBs modellsystem REGARD. Regionale beregninger basert på historiske tall på nasjonalt nivå kjørt bakover i tid på grunnlag av modellens basisår (her 1995). 55s. |
| 2002/2 | V.V. Holst Bloch: Arealbruksklassifisering av bebygde arealer. Revidert rutine for tilordning av arealbruksklasse til bygning. 58s. | 2002/14 | V. Hansen og H. Madsen: Månedlig og kvartalsvis elektrisitetsstatistikk. Dokumentasjon av produksjonsrutiner og systembeskrivelse. 41s. |
| 2002/3 | V.V. Holst Bloch: Metode og datagrunnlag for produksjon av arealstatistikk for tettstednære områder. Teknisk dokumentasjon. 29s. | 2002/15 | A. Rolland: Søkelys på det gode liv. 37s. |
| 2002/4 | G. Dahl og J Lajord: FD - Trygd: Konsistenskontroller. 21s. | 2002/17 | D.Rønningen og D. Fredriksen: Beskatningen av pensjonister. 41s. |
| 2002/5 | G. Dahl og C. Nordseth: FD - Trygd: Dokumentasjonsrapport. Inntekt og formue, 1992-1999. 41s. | 2002/18 | D. Rønningen: Overganger fra arbeidsmarkedet til trygd. En litteraturoversikt. 34s. |
| 2002/6 | G. Dahl og C. Nordseth: FD - Trygd: Dokumentasjonsrapport. Sysselsetting, 1992-1998. 107s. | 2002/19 | F. Gundersen og L. Solheim: Regionalisering av FoU-statistikken. 43. |
| 2002/7 | S. Lien og C. Nordseth: FD - Trygd: Dokumentasjonsrapport. Pensjonsgivende inntekt, 1992-1999. Omsorgspoeng, 1992-1997. 24s. | 2002/20 | L. Vågane: Omnibusundersøkelsen november/desember 2001. Dokumentasjonsrapport. 56s. |
| 2002/8 | V. Lund: Kostnadsindekser for lastebiltransport. Definisjoner og beregningsmetode. Vekter og representantvarer 2001. 47s. | 2002/21 | G. Claus, O. Haugen P. M. Holt og E. Knutsen: Regnskapsstatistikk. Næringsoppgaver for ikke-finansielle aksjeselskaper, 1999. Dokumentasjon. 34s. |
| 2002/9 | T.M. Normann: Rekruttering til erfaringskonferanse og undersøkelse om røykevaner blant kvinner i alderen 25-45 år. Dokumentasjonsrapport. 16s. | 2002/22 | M. Takle: Befolkningsstatistikk på rute-nett. Dokumentasjon. 35s. |
| 2002/10 | J. Holmøy: GERIX 1995-1999. Dokumentasjon, system, data, program. 47s. | 2002/23 | D. Roll-Hansen, S. Ferstad, M. Stålnacke, P. Tuhus og R. Nøtnæs: En spørreskjemametodisk gjennomgang av datainnsamling gjennom Grunnskolens informasjonssystem (GSI). 109s. |
| 2002/11 | T.M. Normann: Underøking om det lokale sjølvstyret. Dokumentasjonsrapport. 81S. | 2002/24 | T. P. Bøe og I. Håland: Dokumentasjon av arbeidskraftundersøkelsen (AKU) 85s. |
| 2002/12 | L.S. Stambøl: Regionale framskrivinger av sysselsetting og bruttoprodukt ved hjelp av SSBs modellsystem REGARD. Regionale | 2002/25 | A. Akselsen og T. Sandnes: FD - Trygd. Dokumentasjonsrapport. Stønader til enslig forsørger. 1992-2000. 46s. |