

# Ökad tillgänglighet till Sveriges officiella statistik

Professor Bo Sundgren  
December 1998

## Innehållsförteckning

<b>0</b>	<b>Inledning .....</b>	<b>8</b>
0.1	Uppdraget.....	8
0.2	Tillgänglighetsproblematiken.....	8
0.3	Uppsatsens uppläggning.....	12
<b>1</b>	<b>Historik och nuläge .....</b>	<b>16</b>
1.1	En tidig ansats (1970): det arkivstatistiska systemet (ARKSY) .....	17
1.2	Tydligare krav från statsmakterna och tunga användare (1990) .....	22
1.3	En samlad attack för att lösa dokumentationsproblemet .....	24
1.3.1	SCBDOK-modellen: begrepp och dokumentationsmallar .....	24
1.3.2	Kvalitetsbegrepp och kvalitetsdeklarationer .....	26
1.3.3	Översiktliga produktbeskrivningar.....	29
1.3.4	Mot ett integrerat metadatasystem.....	30
1.4	Tillgänglighetsprojektet.....	33
1.5	Databasprojektet och Sveriges statistiska databaser .....	35
1.5.1	Innehåll: makrodata, mikrodata och metadata .....	36
1.5.2	Tillgänglighet via Internet .....	38
1.5.3	Användarstyrd statistikframställning: möjligheter och begränsningar .....	39
1.5.4	Elektronisk publicering .....	40
1.5.5	Internet som ingång till all officiell statistik i Sverige och internationellt.....	41
1.5.6	Rutiner för fortlöpande aktualisering av Sveriges statistiska databaser .....	41
1.5.7	Några viktiga egenskaper hos Sveriges statistiska databaser .....	42
<b>2</b>	<b>Vad menas med tillgänglighet? .....</b>	<b>46</b>
2.1	Fysisk och innehållsmässig tillgänglighet.....	46
2.2	Olika tillgänglighetsprofiler.....	47
2.3	Tillgängliggörandeprocessen ur statistikanvändarens synpunkt.....	49

2.4	Önskemål och klagomål beträffande den officiella statistikens tillgänglighet .....	54
2.5	Åtgärdsområden.....	55
2.6	Tillgänglighetens plats i ”det goda statistiska systemet” .....	57
<b>3</b>	<b>Hur kan man öka tillgängligheten? .....</b>	<b>61</b>
3.1	Kunskapsbaserad dokumentation.....	61
3.1.1	Behovet av dokumenterad kunskap vid användning av statistiska data .....	61
3.1.2	Kvalitetsfrågor.....	64
3.2	Teknik och presentation .....	67
3.2.1	Strategi för utvecklingsarbete i en snabbt föränderlig teknisk miljö .....	68
3.2.2	Internet-baserade system .....	69
3.2.3	Standardkomponenter och standardiserade gränssnitt.....	70
3.2.4	Söksystem .....	73
3.2.5	Databashanterare .....	75
3.2.6	Presentationsteknik .....	78
3.3	Ekonomi: finansiering och prissättning .....	80
3.3.1	Officiell statistik – en kollektiv nytta.....	80
3.3.2	Marginalkostnadsprincipen för prissättning av informationstjänster .....	81
3.3.3	Bastjänster och tilläggstjänster .....	83
3.3.4	Internet-baserade informationstjänster .....	84
3.4	Effektivare hantering av statistiksekretessen .....	85
3.4.1	Avsiktliga och oavsiktliga röjanden .....	85
3.4.2	Det lagstadgade skyddet för statistiksekretessen.....	87
3.4.3	Avvägningen mellan statistiksekretess och statistiktillgänglighet .....	89
3.4.4	Behövs en tillgänglighetsprincip för hantering av statistiksekretessen? .....	90
3.4.5	Olika sätt att hantera statistiksekretessen i särskilt känsliga fall.....	91
3.4.6	Varför hanteras inte statistiksekretessen effektivare?.....	93
3.4.7	Möjliga åtgärder .....	94
3.5	Lagar och regler.....	97
3.5.1	Tryckfrihetsförordningen och Sekretesslagen.....	98
3.5.2	Lagen om den officiella statistiken.....	98

3.5.3	Förordningen om den officiella statistiken.....	99
3.5.4	Föreskrifter och allmänna råd om offentliggörande av officiell statistik.....	99
3.5.5	Personregister för officiell statistik.....	102
3.5.6	Begreppet ”Sveriges officiella statistik”.....	103
3.6	Organisationsfrågor.....	107
3.6.1	Centraliserat eller decentraliserat system?.....	107
3.6.2	Det svenska statistiska systemet, dess omvärld och delsystem.....	111
<b>4</b>	<b>Önskvärda åtgärder.....</b>	<b>120</b>
4.1	Lätt att söka och komma åt all officiell statistik, som är relevant för ett visst problem.....	122
4.1.1	Fullständiga dokumentationer och kvalitetsdeklarationer....	122
4.1.2	Tekniska åtgärder för att förbättra sökbarhet och åtkomstmöjligheter.....	122
4.1.3	Marginalkostnadsbaserad prissättning.....	125
4.1.4	Begreppet ”officiell statistik” preciseras, och enhetliga publiceringsregler föreskrivs.....	126
4.1.5	Skärpt tillsynsverksamhet.....	127
4.2	Lätt även för icke-specialister att tolka officiell statistik och använda sig av den på rätt sätt.....	127
4.2.1	Skräddarsydd information om statistiken för olika användarkategorier.....	127
4.2.2	Datorstödd framställning av användar- och problemanpassad dokumentation.....	128
4.3	Möjligt för kvalificerade statistikanvändare att använda den officiella statistikens datamaterial på egen hand.....	129
4.3.1	Fullständig dokumentation av dokumentationsregister, så att de kan återanvändas utan förstahandskunskap.....	130
4.3.2	Standardiserade tekniska lösningar för att göra datamaterialen tillgängliga.....	130
4.3.3	Bastillgänglighet till lågpris för ”gör-det-självaren”, prisvärda tilläggstjänster till marginalkostnad.....	130
4.3.4	Skydd för statistiksekretessen som inskränker tillgängligheten minimalt.....	130
4.3.5	Tillgänglighetskraven på officiell statistik föreskrivs gälla även observationsregistren.....	131
4.3.6	Offensiv bevakning av tillgängligheten.....	131

---

4.4	Tillgängligheten till den officiella statistikens datamaterial inskränks inte mer än nödvändigt för att säkra statistiksekretessen.....	132
4.4.1	Kombinera legala, administrativa och tekniska åtgärder.....	132
4.4.2	Minska tolkningsutrymmet och inför tillgänglighetsprincip.....	133
4.4.3	Bästa kända tekniska lösningar för att maximera tillgängligheten inom given statistiksekretess.....	133
4.4.4	Registerfunktion inom databasverksamheten.....	134
4.5	Möjligt att komma åt den officiella statistiken genom bastjänster till låga priser i kombination med valfria tilläggstjänster .....	134
4.5.1	Infrastruktur som möjliggör flexibla kombinationer av billiga bastjänster och prisvärda tilläggstjänster .....	135
4.5.2	Marginalkostnadsbaserad prissättning .....	135
	<b>Referenser.....</b>	<b>141</b>

# Ökad tillgänglighet till Sveriges officiella statistik

## 0 Inledning

### 0.1 Uppdraget

I samband med statistikreformen gav statsmakterna SCB i uppdrag att öka tillgängligheten till Sveriges officiella statistik, inklusive de bakomliggande observationsregistren och deras dokumentation. Tillgängligheten skulle bl.a. ökas genom att en databasverksamhet byggdes upp vid SCB, omfattande den officiella statistiken och ett antal särskilt uppräknade register.

På uppdrag av *Utredningen om utvärdering av statistikreformen* (Fi 1998:05) redogör jag i denna uppsats för en del erfarenheter från arbetet med att bygga upp den nämnda databasverksamheten, som numera går under beteckningen "Sveriges statistiska databaser". Jag diskuterar också några idéer om hur tillgängligheten till Sveriges officiella statistik ytterligare skulle kunna förbättras för olika kategorier av användare.

De uppfattningar och förslag, som jag för fram i denna uppsats, står för mig personligen.

### 0.2 Tillgänglighetsproblematiken

Det har länge stått klart, att Sverige har en internationellt sett ovanligt rik tillgång på statistiska grunddata. En viktig anledning till detta är att Sverige har världens äldsta officiella statistik, ursprungligen baserad på ett unikt folkbokföringssystem. En annan anledning, som i modern tid blivit allt viktigare, är att det svenska samhället och den svenska förvaltningen är uppbyggd kring välutformade administrativa system, idag datoriserade, som också lämpar sig väl som datakälla för officiell statistik.

Så tidigt som på 1960-talet demonstrerade Svein Nordbotten<sup>1</sup> på ett övertygande sätt, hur man redan med dåtidens teknik – stordatorer med mindre kapacitet än dagens minsta persondatorer i kombination med

<sup>1</sup> Se t.ex. Nordbotten Svein: "A Statistical File System", Statistisk Tidskrift, No. 2, Stockholm 1966 sid 99-109.

magnetbandsteknik för datalagring – på ett tämligen enkelt sätt skulle kunna drastiskt öka forskares, utredares och andra kvalificerade statistikanvändares möjligheter att bearbeta de väldiga volymerna av statistiska grunddata i den svenska statistikens arkiv. Med lämpligt utformade tekniska system, som var möjliga att utveckla redan vid denna tid, skulle användarna själva kunna bestämma vilka statistiska bearbetningar och analyser, som skulle utföras på arkiverade grunddata, utan att behöva vara begränsade till förutbestämda aggregeringsnivåer, tabellformat och liknande. En återanvändare av statistiska grunddata skulle kunna ompröva många av de antaganden och designbeslut, som den ursprungliga statistikproducenten gjort, och skulle därigenom med stor flexibilitet kunna använda statistiska grunddata för nya ändamål.

Sådana var möjligheterna på 1960-talet, och sådana är naturligtvis möjligheterna ännu mera idag, när vi har tillgång till en teknik, som är åtskilliga tiopotenser effektivare. Det som på 1960-talet tog timmar, dagar och veckor av datorbearbetningar, klaras idag av på minuter och sekunder.

Ändå måste vi idag, liksom på 1960-talet, konstatera – för min del med stor bedrövelse – att dessa unika möjligheter, som vi besitter i Sverige, att återanvända den officiella statistikens arkiverade grunddata, utnyttjas i mycket liten utsträckning. Vad beror det på?

Det magra utnyttjandet av statistiska grunddata beror knappast på bristande intresse från de närmast berörda statistikanvändarna: forskare, utredare m.fl. Dessa kategorier av statistikanvändare ser utan tvekan den stora potentialen i statistiska grunddata, men måhända har de varit alltför modesta, när de uttryckt sina önskemål gentemot statistikansvariga. Det är åtminstone mitt intryck. Kanske beror detta delvis på att statistikanvändarna sist och slutligen är mest intresserade av sakfrågorna inom sitt ämnesområde och inte vill ägna alltför stor del av sin tid åt ”teknikaliteter”, som de – med rätta – förväntar sig att statistikproducenten till stor del skall kunna hantera åt dem.

Därmed har vi kommit till nästa hinder mot tillgängligheten till statistiska grunddata. Statistikansvariga och statistikproducenter har aldrig visat någon särskilt aktiv entusiasm för att verkligen ställa statistiska grunddata till statistikanvändarnas förfogande för deras egna bearbetningar och analyser. Producenterna är i och för sig inte negativa till att utnyttja de tekniska möjligheterna, men man föredrar att göra det i egen regi, hellre än att låta statistikanvändarna göra det själva. Och då tycker statistikanvändarna ofta att det blir för krångligt och dyrt, och framför allt tycker man inte, att man får den frihet, som man skulle vilja ha, att i många och snabba interaktioner med datamaterialen själva kunna testa och formulera nya statistiska hypoteser.

Den dominerande attityden hos statistikansvariga och statistikproducenter skulle kunna beskrivas som "den upplyste despotens" eller "den gode monopolistens". Man tycker helt enkelt att man som statistikansvarig producent bäst kan värna om statistikbearbetningarnas kvalitet och datamaterialens sekretess. I den mån man bedriver uppdragsverksamhet, kan det även finnas en begriplig frestelse att snegla mot de intäkter man kan få genom att utföra tjänster åt statistik-användarna hellre än att låta dessa göra jobbet själva.

Fram till 1990-talets början accepterade statistik användarna i stor utsträckning statistikansvarigas förklaringar, att det nog ändå var bäst att låta dem som hade förstahandskunskapen om de statistiska datamaterialen göra även de specialbearbetningar, som användarna önskade sig, om inte annat så av sekretessskäl.

Omkring 1990 kunde man se helt nya förutsättningar hos de professionella statistik användare, som vi här talar om: forskare, utredare m.fl. Användarna hade numera i stor utsträckning och i flera avseenden lika goda statistiska och metodologiska kunskaper som statistikproducenterna, och genom teknikens snabba framsteg hade de fått tillgång till minst lika goda tekniska förutsättningar som dessa.

Några kategorier av statistik användare, främst utredare och analytiker inom regeringskansliet började nu artikulera tillgänglighetskraven tydligare. Man ville bl.a. använda de statistiska grundmaterialen i de simuleringsmodeller, som man börjat bygga upp för att kunna testa möjliga utfall av olika reformer, som man övervägde i de olika departementen. Forskarsamhället uttryckte liknande önskemål om flexibel tillgång till grunddata för statistiska bearbetningar och analyser, men till skillnad från departementstjänstemännen hade forskarna inga maktmedel att åberopa.

Påtryckningarna från statistik användarna hade denna gång viss effekt. En statistikreform utreddes och genomfördes. Den s.k. Genomförandekommittén fokuserade bl.a. på tillgänglighetsfrågorna. SCB gavs i uppdrag att bygga upp en central databasverksamhet, som skulle underlätta tillgängligheten till den officiella statistiken och ett antal bakomliggande observationsregister. Sveriges statistiska databaser sattes i drift fr.o.m. 1997.

Inte minst genom att Sveriges statistiska databaser är baserade på den nya Internettekniken, gör de det oerhört mycket enklare, flexibla och billigare för alla kategorier av statistik användare att komma åt den officiella statistiken. De mera avancerade statistik användarna (forskare, utredare m.fl.) kan dessutom relativt enkelt bearbeta statistiken vidare enligt egna önskemål och med hjälp av egna program, om man så vill.

Däremot har den centrala databasverksamheten ännu inte åstadkommit särskilt mycket för att öka tillgängligheten till statistiska grund-



data, det som egentligen var forskarnas och vissa utredares och departementstjänstemäns huvudönskemål. Sveriges statistiska data-baser skall visserligen tillgängliggöra 17 explicit uppräknade observationsregister, men dels är detta bara en del av de grunddata, som skulle kunna tillgängliggöras, och dels har inte ens arbetet med att verkligen tillgängliggöra dessa register ännu kommit särskilt långt. Detta kan förklaras med att prioriteringar varit nödvändiga i uppbyggnaden av Sveriges statistiska databaser, och då har "man" (statistikansvariga och statistikproducenter) klart prioriterat tillgängliggörandet av aggregerade data, "statistik", på bekostnad av grunddata och dokumentation. Denna prioritering kan naturligtvis försvaras med att det är så många fler statistik användare, som använder sig av "färdig statistik", än som har nytta av ökad tillgång till statistiska grunddata. Icke desto mindre hade det säkert varit möjligt att göra mera för tillgängligheten till grunddata, utan att det behövde ha inverkat negativt på tillgängligheten till aggregerade material.

Attityderna vad gäller tillgänglighetsfrågorna hos statistikansvariga och statistikproducenter har sålunda förändrats klart till det bättre, vad gäller tillgängligheten till den färdiga statistiken. När det gäller tillgängligheten till statistiska grunddata måste man tyvärr konstatera att attitydförändringarna fortfarande går mycket långsamt. Här är det alltså viktigt att statistik användarna och statsmakterna även fortsättningsvis utövar – och enligt min mening ytterligare förstärker – ett väl artikulerat efterfrågetryck på statistikansvariga och statistikproducenter.

Som jag närmare kommer att utveckla i den här uppsatsen, bör det ingå i "spelreglerna" för den officiella statistiken, att inte bara den färdigaggregerade statistiken, utan även bakomliggande observationsregister och dokumentation, tillgängliggörs för statistik användarna på bästa möjliga sätt inom ramen för givna anslag och tillgängliga metoder. Här finns oerhört mycket att göra. Allt som behöver göras behöver inte heller kosta några astronomiska belopp. Det handlar väl så mycket om vilja, standardisering och förmåga att på ett konstruktivt sätt utnyttja tillgängliga metoder och tekniska lösningar.

Samtidigt som jag känner ett stort engagemang för att aktivt verka för att de professionella statistik användarnas enligt min mening mycket berättigade önskemål och krav på ökad tillgänglighet till statistiska grunddata faktiskt tillgodoses, så vill jag betona att mitt engagemang för statistik kvaliteten och statistik sekretessen inte är mindre. Min poäng är, att det finns mycket stora möjligheter att drastiskt öka statistik användarnas tillgång till grunddata för statistiska ändamål, utan att man på något sätt behöver tumma på kvalitet och sekretess. Kvalitets- och sekretesskraven kan säkras, samtidigt som statistik användarna ges ökad flexibilitet och ökade möjligheter att statistiskt bearbeta grunddata enligt egna

önskemål. Förmyndarmentaliteten – i och för sig ofta väl-menande – som ibland kan skönjas hos statistikansvariga och statistikproducenter bör nu definitivt förvisas till historien. Statistik-användarna bör ges effektiv tillgång till alla statistiska data i den officiella statistiken och till all den kunskap och dokumentation, som behövs för att använda den officiella statistikens data på bästa sätt för den mångfald av statistikbehov, som nu och framdeles föreligger hos statistikanvändarna.

### 0.3 Uppsatsens uppläggning

Min uppsats utvecklar det tema kring den officiella statistikens tillgänglighet, som jag introducerade i föregående avsnitt. Kapitel 1 ger en historisk tillbakablick och en redogörelse för några viktiga tillgänglighetsrelaterade projekt, som genomförts under 1990-talet.

I kapitel 2 diskuteras begreppet "tillgänglighet" lite mera ingående, och olika tillgänglighetsdimensioner och tillgänglighetskomponenter identifieras. Det betonas att olika kategorier av användare har olika tillgänglighetsbehov och därmed olika tillgänglighetsprofiler. Begreppet "tillgänglighet" relateras till andra viktiga begrepp i "det goda systemet" för officiell statistik, t.ex. kvalitet.

Processen att tillgängliggöra statistik analyseras ur ett användarperspektiv. Processens huvudflöde är följande:

1. Användaren har ett problem och behöver statistik som kan vara relevant för att belysa problemet.
2. Med hjälp av tillgängliga sökverktyg identifierar användaren successivt och interaktivt potentiellt relevanta statistiska data och bedömer med hjälp av tillgänglig dokumentation, huruvida de olika datamaterialen är så intressanta att de skall hämtas fram.
3. Identifierade statistiska datamaterial, som användaren bedömer vara av intresse för problemet, lokaliseras och görs fysiskt tillgängliga för användaren tillsammans med erforderlig dokumentation.
4. Därefter tolkar användaren framtagna statistiska data med hjälp av den tillgängliga dokumentationen och eventuell expertis.
5. Vissa användare gör egna bearbetningar och analyser på basis av framtagna statistiska data.
6. Resultaten av tolkningar och analyser sammanställs till någon form av utredning eller beslutsunderlag.
7. Användningen av statistiken kan leda till att nya statistikbehov identifieras, och att delar av den skisserade processen upprepas.

Några typiska klagomål från användarna beträffande den svenska officiella statistikens tillgänglighet identifieras också i kapitel 2 och utgör

grund för diskussionerna i kapitel 3 kring vad som kan göras för att öka tillgängligheten till Sveriges officiella statistik. De tyngsta klagomålen avser dokumentationen av statistiken och datamaterialen, sök- och åtkomstprocessernas användarvänlighet, kostnaderna för olika typer av tjänster i sammanhanget, samt hanteringen av statistiksekretessen.

I kapitel 3 diskuteras vad som har gjorts och vad som kan göras för att komma till rätta med de olika problemen kring tillgängligheten till den officiella statistiken och de bakomliggande datamaterialen. Sex viktiga åtgärdsområden identifieras och diskuteras:

1. Kunskapsbaserad dokumentation
2. Teknik och presentation
3. Ekonomi: finansiering och prissättning
4. Effektivare hantering av statistiksekretessen
5. Lagar och regler
6. Organisationsfrågor

I kapitel 4 formuleras ett antal förslag till konkreta åtgärder för att öka tillgängligheten till Sveriges officiella statistik. Varje åtgärdsförslag hör hemma i något av de åtgärdsområden, som diskuterades i kapitel 3, och ger bidrag till ett eller flera delmål, som i sin tur bidrar till det övergripande målet om ökad tillgänglighet. De delmål som avses är:

1. Lätt att söka och komma åt all officiell statistik, som är relevant för ett visst problem
2. Lätt även för icke-specialister att tolka statistiken och använda sig av den på rätt sätt
3. Möjligt för kvalificerade statistikanvändare att använda den officiella statistikens datamaterial på egen hand
4. Tillgängligheten till den officiella statistikens datamaterial inskränks inte mer än nödvändigt för att säkra statistiksekretessen
5. Möjligt att komma åt den officiella statistiken genom bastjänster till låga priser i kombination med valfria tilläggstjänster

Några viktiga åtgärder som föreslås är:

- All officiell statistik och tillhörande observationsregister dokumenteras och kvalitetsdeklarerar fullständigt enligt fastlagda regler och riktlinjer
- En registerfunktion inrättas inom den centrala databasverksamheten för bearbetningar och samkörningar av identifierade mikrodata för statistiska ändamål
- Begreppet ”officiell statistik” definieras så, att reglerna beträffande den officiella statistikens tillgänglighet blir tillämpliga även för observationsregister, som ligger till grund för officiell statistik
- Datamaterial på olika aggregeringsnivåer, som ligger till grund för officiell statistik, skall tillgängliggöras enligt samma regler som gäller för den officiella statistiken och dess observationsregister
- All officiell statistik tillgängliggörs på ett standardiserat sätt genom Sveriges statistiska databaser och Svenska statistiknätet
- Nya metoder och verktyg utnyttjas för effektiv hantering av statistiska data och metadata och för problembaserade sökningar på många olika sätt
- Infrastrukturen för att tillgängliggöra den officiella statistiken och dess datamaterial byggs upp med hjälp av standardkomponenter och standardiserade gränssnitt, så att billiga bastjänster och dyrare tilläggstjänster lätt kan fogas in i strukturen och väljas eller väljas bort av statistikanvändarna
- Tjänsterna kring den officiella statistiken och dess data utformas så, att grundtillgänglighet säkerställs med billiga bastjänster, varefter användaren kan välja mellan självbetjäning eller tilläggstjänster prissatta enligt den marginalkostnadsprincip, som formuleras i Riksrevisionsverkets (RRV) förslag till huvudprincip för prissättning av informationstjänster<sup>2</sup>
- Skyddet för statistiksekretessen utformas så, att statistikanvändarnas tillgång till datamaterialen inskränks minimalt: bästa kända metoder, tekniker och programvaror utnyttjas för att maximera tillgängligheten inom given statistiksekretess
- Det centrala tillsynsorganet för den officiella statistiken ges en starkare ställning och åläggs att utöva en skärpt tillsyn särskilt vad gäller (i) de skärpta föreskrifterna kring dokumentation och kvalitetsdeklaration, (ii) tillgängliggörandet av den officiella statistiken och de bakomliggande observationsregistren genom databasverksamheten (iii) enhetliga prissättningsprinciper baserade på RRV:s marginal-

<sup>2</sup> Riksrevisionsverket: ”Principer för prissättning av informationstjänster”, regeringsuppdrag, RRV1995:64.

kostnadsbaserade huvudprincip för prissättning av informations-tjänster

Den som vill ha en mera fullständig översikt över de tillgänglighets-höjande åtgärder som föreslås i uppsatsen kan ha nytta av den tabellariska sammanställningen i figur 24 a och b.

# 1 Historik och nuläge

Sverige har en lång och ärorik historia på den officiella statistikens område. Tidigare än i något annat land i världen, började statsmakterna i Sverige att systematiskt bygga upp objektiva informationsunderlag för sina beslut i form av statistik. Redan från början spelade administrativa processer (folkbokföringen) en viktig roll i statistikframställningen.

Den långa statistiktraditionen i kombination med välorganiserade och datagenererande förvaltningsprocesser har givit Sverige ett omfattande och rikt historiskt dataarkiv, innehållande långa tidsserier av socio-demografiska och ekonomiska data.

Denna rikedom förpliktat. Forskare från hela världen blickar med avund på de svenska dataarkiven och det informationskapital, som ligger där, till stora delar outnyttjat. Även sentida material, som ur teknisk synpunkt är mera lättillgängliga utnyttjas mycket sparsamt. Vad beror det på?

Administrativa och statistiska datamaterial från de senaste tre decennierna är i mycket stor utsträckning elektroniskt lagrade och förhållandevis lätt åtkomliga för datorbearbetningar, även om det idag, p.g.a. den mycket snabba tekniska utvecklingen, kan vara vissa svårigheter att hitta lämpliga datorer och program för att bearbeta de äldsta materialen. Men de huvudsakliga hindren för ett effektivt utnyttjande av det slumrande informationskapitalet är inte längre av teknisk natur. Det handlar mer om vår oförmåga att utnyttja tekniken på ett bra sätt och om hinder av administrativ och organisatorisk natur. Mera konkret handlar hindren bl.a. om

- inflexibla informationssystem
- bristande standardisering av komponenter och gränssnitt
- bristande dokumentation och bristande motivation att göra något åt det
- bristande motivation att utnyttja möjligheterna att öka tillgängligheten till känsliga data inom ramen för gällande statistiksekretess
- oklara och ibland olämpliga prissättningsprinciper

Man kan i förbigående notera, att framställningen av officiell statistik under större delen av de senaste 40 åren varit centraliserad till SCB. De ovan antydda problemen har funnits under hela denna tidsperiod och har sålunda ingenting med den partiella decentralisering att göra, som genomfördes i början av 1990-talet. Problemen finns i ungefär lika hög grad inom en central statistikorganisation som SCB som inom ett mera

decentraliserat system. Teoretiskt skulle man tro, att det vore lättare att åtgärda problemen inom ett centralt system än i ett decentraliserat, men så tycks inte vara fallet i praktiken. Det verkar behövas en mera kraftfull och tydlig extern styrning från statsmakterna och statistikanvändarna för att komma till rätta med problemen, oavsett graden av centralisering/decentralisering i det statistiska systemet.

## 1.1 En tidig ansats (1970): det arkivstatistiska systemet (ARKSY)

Det har dock inte saknats ansatser att komma tillrätta med tillgänglighetsproblemen. Redan under 1960-talet, när datorer just börjat användas i statistikproduktionen, tog den dåvarande verksledningen vid SCB, generaldirektören Ingvar Ohlsson och planeringschefen, sedermera överdirektören, Lennart Fastbom initiativ till införandet av ett s.k. arkivstatistiskt system (ARKSY) för att möjliggöra flexibla och mera användaranpassade statistikbearbetningar. Inspirationen kom från Svein Nordbotten, då avdelningschef vid statistiska centralbyrån i Norge, sedermera professor i informationsvetenskap vid Universitetet i Bergen och chef för FN:s statistiska kontor i New York.

Figur 1 åskådliggör det arkivstatistiska systemet. Sedan primärdata till statistiken inhämtats från uppgiftslämnare (personer, företag och/eller myndigheter), och indata i vederbörlig ordning kodats, registrerats, granskat och rättats, skulle de lagras och organiseras på ett standardiserat sätt, i form av s.k. flata filer, enhetligt dokumenterade i en s.k. variabelkatalog. De flata filerna, innehållande vad som senare kom att benämnas mikrodata, skulle användas på två sätt:

- dels omedelbart, som underlag för den reguljära publiceringen av standardiserade statistiska redovisningar, s.k. statistiska meddelanden (SM)
- dels vid behov, ad hoc, och vanligen på uppdragsbasis, som underlag för mera skräddarsydda statistikprodukter, speciellt anpassade till olika användarbehov

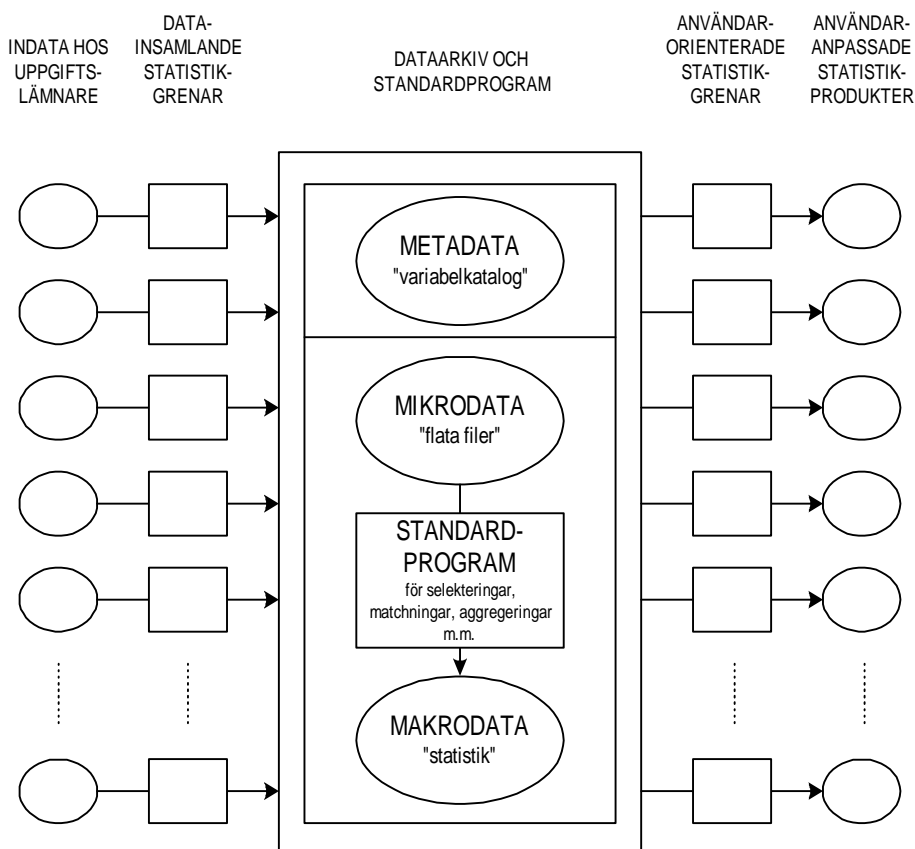
I bägge fallen skulle de erforderliga bearbetningarna (selekteringar, matchningar, aggregeringar m.m.) utföras på ett standardiserat sätt med hjälp av standardprogram. Utgående från mikrodata i de flata filerna, och med hjälp av databeskrivningar, s.k. metadata, skulle man med hjälp av de standardprogramstyrda men ändå flexibla bearbetningsprocesserna framställa såväl standardiserad som mera användar-

anpassad statistik, s.k. makrodata, och tillhörande beskrivningar, meta-data.

Det är värt att notera, att magnetband var den enda praktiskt tillgängliga tekniken för lagring av stora datamängder vid denna tidpunkt. Ända kan man idag konstatera, att det arkivstatistiska konceptet som sådant ligger mycket nära modern databasteknik, med relationstabeller motsvarande de flata filerna och SQL-språket (Structured Query Language) motsvarande standardprogrammen. Man kan t.o.m. hävda att det arkivstatistiska systemets dataarkiv har stora likheter med ett modernt datalager, s.k. "data warehouse".

Den dåvarande verksledningen vid SCB gick så långt att man utarbetade ett detaljerat förslag till radikal omorganisation av SCB enligt den arkivstatistiska modellen. De flesta statistikbyråer i världen hade dittills, och har fortfarande, en s.k. "stovepipe organisation", där statistiken produceras, "från ax till limpa", inom tämligen självständiga och parallellt arbetande små organisatoriska enheter, s.k. statistikgrenar. Vid en statistisk centralbyrå kan det typiskt finnas ett par hundra statistikgrenar, grupperade hierarkiskt till statistikprogram, sektioner, enheter, avdelningar eller vad nu benämningarna kan vara. Organisationen är i grunden datainsamlingsbaserad, och samordningen mellan statistikgrenarna är vanligen svag. Personalen på statistikgrenarna har ett starkt kunskapsmonopol; de är de enda som har förstahandskunskap om de data och processer, som statistiken baseras på.





**Figur 1. Schematisk modell av det arkivstatistiska systemet (ARKSY) som det konceptualiserades under åren kring 1970.**

Statistikanvändare har ofta behov av att kombinera statistik från olika statistikgrenar. Även de som primärt är ute efter ekonomisk statistik, har t.ex. inte sällan ett starkt behov av kompletterande bakgrundsdata från befolkningsstatistiken. En datainsamlingsbaserad statistikgrensorganisation av traditionellt slag är dåligt anpassad till användar- och användningsstrukturerna. Få har en överblick över all statistik, som i princip är tillgänglig vid en statistisk centralbyrå, och de som har en viss överblick, t.ex. någon form av informationstjänst, har i allmänhet en alltför ytlig kunskap, bl.a. därför att det är personalen på statistikgrenarna som har monopol på förstahandskunskaperna om sina respektive statistikprodukter och bakomliggande data och processer. De flesta försök till ökad användaranpassning i en sådan organisation stupar på detta och på det faktum att den erforderliga samordningen mellan statistikgrenarna blir mycket komplicerad, tungrodd och resurskrävande

och inte heller är särskilt populär hos de traditionellt mycket självständiga statistikgrenarna.

I den tilltänkta, arkivstatistiskt inspirerade organisationen skulle de traditionella statistikgrenarna delas upp i indata- och utdataorienterade processer, sammankopplade i ett "många-till-många"-mönster via ett centralt "clearing-house", en s.k. databascentral, svarande för vården av dataarkivet (mikrodata, makrodata, metadata) och de standardiserade bearbetningsverktygen. De indataorienterade processerna skulle svara för att producera mikrodata och tillhörande metadata till dataarkivet. Eventuellt skulle de också kunna svara för den rutinmässiga framställningen av resultatrapporter (SM) från respektive datainsamling; det skulle då ske med hjälp av standardverktygen.

De utdataorienterade processerna skulle struktureras efter användare och användningar. Tack vare den standardiserade lagringen av data och metadata i dataarkivet skulle den användarorienterade personalen själv snabbt och effektivt kunna framställa skraddarsydda statistikprodukter för olika behov med hjälp av standardverktygen.

Verksledningens förslag till ny organisation möttes av sådana högljudda protester från stora delar av personalen och från de fackliga organisationerna, att det aldrig officiellt lades fram. (Ett komplett utkast på flera hundra sidor finns dock tillgängligt för den intresserade.) Förslaget ansågs bygga på en alldeles för "mekanisk" syn på statistikproduktionen; det beaktade inte betydelsen av personalens unika kompetens och behovet av "handpåläggning" i de olika produktionsmomenten. Kort sagt skulle det leda till ett totalt kvalitetsmässigt haveri för den officiella statistiken.

Ambitionerna att med hjälp av ett arkivstatistiskt system öka användaranpassningen av och tillgängligheten till den officiella statistiken och bakomliggande datamaterial drabbades ungefär samtidigt av ytterligare ett bakslag. I spåren av Folk- och bostadsräkningen 1970 (FoB 70) och den integritetsdebatt, som då utbröt, utbreddes sig en stark rädsla, inte minst hos SCB:s verksledning, att på ett alltför aktivt sätt använda sig av mikrodata i den sekundära statistikproduktionen, d.v.s. efter det att de primära statistikresultaten framställts och publicerats. Man beförde att det skulle kunna föranleda ytterligare opinions-stormar, vilka i sin tur skulle kunna försämra uppgiftslämnarklimatet och öka bortfallet i de statistiska undersökningarna.

Av dessa skäl tonade SCB:s verksledning ner det arkivstatistiska konceptet och speciellt de användar- och mikrodataorienterade delarna av det. Organisationen förblev i stort sett oförändrad. Dock fortsatte man att satsa på det tekniska utvecklingsarbete, som ARKSY-modellen förutsatte, men nu med huvudinriktningen att rationalisera de interna produktionsprocesserna vid SCB. Sålunda standardiserades datalag-

ringen, först i form av ”flata filer”, senare som relationsdatabaser, och en hel familj av användarvänliga och flexibla standardprogram, den s.k. TAB68-familjen, utvecklades och infördes i stor skala med pedagogiskt understöd i form av en standardiserad systemutvecklingsmetod, SCB:s systemeringsmodell. Ämnespersonal utan omfattande ADB-kunskaper kunde därefter, i stor utsträckning på egen hand, utan omfattande insatser av IT-specialister, både utveckla statistikproduktionssystem och sköta dem driftmässigt.

Statistikproduktionen ökade kraftigt i omfattning under denna tid, samtidigt som antalet anställda kontinuerligt minskade, särskilt bland s.k. ”smör-och-bröd”-programmerare och bland icke-handläggare. I ett senare skede kunde även kostnaderna för datordrift skäras ner kraftigt, efter ”outsourcing” och införande av persondatorer i nätverk.

En mera användar- och tillgänglighetsorienterad del av det arkivstatistiska konceptet blev dock förverkligad. Vid mitten av 1970-talet öppnade SCB sin första externt riktade databasservice. Den baserades enbart på aggregerade data, d.v.s. makrodata, och bestod innehållsmässigt av en regionalstatistisk databas (RSDB) och en tidsseriedatabas (TSDB), bägge hanterade av den egenutvecklade, multidimensionella databashanteraren AXIS. Databaserna var tillgängliga för externa användare via telefonförbindelse med SCB:s stordator. Systemet var avancerat för sin tid, bl.a. genom sin metadatahantering och den multidimensionella datamodellen. Det är fortfarande i drift, men det kommer under 1999, efter mer än 20 års trogen tjänst, helt att ersättas av Sveriges statistiska databaser, det moderna Internetbaserade systemet, som sattes i drift den 1 januari 1997.

En annan effekt av teknikutvecklingen kring ARKSY-modellen, som onekligen kom användarna till godo, var att många uppdrag därefter snabbt och förhållandevis billigt kunde utföras med hjälp av de framtagna standardprogrammen, utan tidsödande planerings- och programmeringsarbete. Användarna fick därigenom en indirekt förbättring av den tillgängligheten till mikrodata, även om denna tillgänglighet med dagens mått mätt ter sig otillräcklig, dyr och inte särskilt användarvänlig.

## 1.2 Tydligare krav från statsmakterna och tunga användare (1990)

Det arkivstatistiska utvecklingsarbetet under början av 1970-talet hade i vissa avseenden drivits fram av starka användarkrav från de samhällsplanerare på central, regional och lokal nivå, som då hade en mycket stark ställning och numerär i den svenska förvaltningen. I minst lika hög grad var dock utvecklingen resultatet av ett starkt och visionärt engagemang hos SCB:s dåvarande verksledning. När verksledningen fick "kalla fötter" p.g.a. integritetsdebatten och det interna motståndet, tappade utvecklingen mot ökad tillgänglighet för externa användare momentum, helst som det externa trycket minskade i takt med att de nämnda samhällsplanerarna successivt förlorade terräng på alla nivåer inom förvaltningen.

Tjugo år senare, omkring år 1990, vaknade dock vissa grupper av statistikanvändare till liv igen och krävde allt ihärdigare, och med viss makt bakom orden, att SCB skulle öka tillgängligheten till den officiella statistiken och, inte minst, till de bakomliggande mikrodatamaterialen. Det var bl.a. Finansdepartementets personal, som hade upptäckt den moderna teknikens möjligheter, och som hade börjat bygga avancerade simuleringsmodeller i PC-miljö för att snabbt kunna ta fram ändamålsenliga beslutsunderlag inför viktiga politiska beslut kring reformer av olika slag, bl.a. skattereformen. Modellerna krävde effektiv och flexibel hantering av mycket stora mängder av statistiska mikrodata, inte bara i form av engångsbearbetningar, utan alltmer i höginteraktiva statistiska och analytiska processer, där användaren testar och modifierar sina förutsättningar och hypoteser gång på gång och med krav på mycket korta väntetider mellan iterationerna.

Även inom forskarsamhället hade man vid den här tidpunkten lärt sig använda den moderna teknikens möjligheter på egen hand, och man förväntade sig alltmer otåligt att få en ändamålsenlig tillgänglighet till SCB:s mikrodata inom ramen för statistiksekretessen. Forskarna har inte ensamma så stora möjligheter att påverka SCB:s verksamhet, men tillsammans med tongivande tjänstemän på Finansdepartementet bildade de en ganska effektiv påtryckargrupp. Situationen förstärktes också av att SCB vid denna tidpunkt hade en generaldirektör, Sten Johansson, som själv var forskare, och som hade stor förståelse för forskarnas behov och även för liknande behov hos andra analyserande statistikanvändare, t.ex. vid privata eller offentliga utredningsorgan och inom förvaltningen.

Uttalade klagomål, krav och önskemål kom som sagt från tunga statistikanvändare: finansdepartementet, forskarsamfundet, utredare och massmediarepresentanter. Några viktiga synpunkter var:

- Det är svårt att få reda på vilken statistik som faktiskt finns, och som skulle kunna vara av intresse för att belysa ett visst problem. Inte ens på SCB kan man ge bra svar på sådana frågor, eftersom man har en alltför datainsamlingsorienterad organisation. Det saknas en god överblick över hela statistikfältet, och en sådan överblick är ofta nödvändig, eftersom ett visst problem ofta belyses bäst av en kombination av statistik från flera olika ämnesområden.
- När man som användare, tillsammans med SCB, lyckats identifiera intressant statistik för att belysa ett visst problem, så visar det sig ofta svårt att rent tekniskt få fram statistiken, särskilt om man vill ha tillgång till den i datorläsbar form för egna fortsatta bearbetningar och analyser. Det beror på att SCB använder sig av gammalmodiga, stordatorbaserade och illa dokumenterade system, som inte är gjorda för snabba, oförberedda uttag.
- De omoderna, inflexibla och ofullständigt dokumenterade systemen gör också att uttag och bearbetningar blir oerhört tids- och arbetskrävande och, inte minst, kostsamma.
- Som kvalificerad statistikanvändare har man ofta både vilja, kunskaper och tekniska möjligheter att själv bearbeta statistiken och de bakomliggande datamaterialen för olika vidareanvändningar och analyser. SCB verkar inte särskilt beredvilligt att konstruktivt medverka till sådana lösningar utan vill hellre lösa användarnas problem inom ramen för sin egen uppdragsverksamhet. SCB "överutnyttjar" därvid vissa kvalitets- och sekretessargument:
  - (i) de efterfrågade bearbetningarna och analyserna kräver sådana kunskaper om datamaterialen och deras kvalitet, som bara finns hos SCB:s personal
  - (ii) förfördelad statistik kan inte lämnas ut till statistikanvändarna, och det kan inte heller mikrodata, inte ens om de aidentifierats, eftersom det för sådana data alltid finns en risk att en användare genom matematiska härledningar kan röja känslig information om enskilda objekt (personer eller företag) i statistiken

Det var ett antal viktiga klagomål, krav och önskemål från statistikanvändarna, som ledde statsmakterna och SCB att initiera projekt och verksamheter som

- Tillgänglighetsprojektet
- Framtagandet av ett dokumentationssystem för statistiska undersökningar (SCBDOK)

- Framtagandet av ett kvalitetsredovisningssystem
- Omläggning av SCB:s produktionssystem från stordator till datorbaserad client/serverteknik; det s.k. plattformbytet
- Uppbyggnaden av Sveriges statistiska databaser för spridning av statistik via Internet

## 1.3 En samlad attack för att lösa dokumentationsproblemet

### 1.3.1 SCBDOK-modellen: begrepp och dokumentationsmallar

Redan några år tidigare hade Sten Johansson själv upptäckt ett avgörande hinder för tillgängligheten till SCB:s data: dokumentationen – eller snarare bristen på dokumentation. Han uppdrog åt två experter, Bengt Rosén och Bo Sundgren, den ena professor i statistik och den andra i informationsbehandling, att utforma en begreppsapparat och ett beskrivningssystem för statistiska undersökningar i syfte att möjliggöra en så fullödig dokumentation av SCB:s datamaterial ur såväl statistisk som informationsteknisk synpunkt, att datamaterialen skulle kunna återanvändas av forskare om 200 år, d.v.s. av personer som inte skulle ha någon förstahandskunskap om vare sig datamaterialen som sådana, eller om de bakomliggande statistiska undersökningarna, eller ens om det samhälle och den tid, där undersökningarna genomfördes.

Arbetet resulterade i en sammanhängande och detaljerad beskrivningsmodell för statistiska undersökningar (i vid mening) och arkiverade mikrodatamaterial, s.k. slutliga observationsregister.<sup>3</sup> Beskrivningsmodellen integrerade produktionsstatistiska, metodstatistiska och informationstekniska begrepp och lades till grund för ett dokumentationssystem, SCBDOK, baserat på dokumentationsmallar, som implementerades med hjälp av en vanlig ordbehandlare.

På en övergripande nivå strukturerar en SCBDOK-mallen dokumentationen i ett antal dokumentationsavsnitt, svarande mot de typiska momenten i en statistisk undersökning. Mallen fungerar därvid som en checklista för vad man bör överväga att ta upp i dokumentation. För att

<sup>3</sup> Rosén, Bengt & Sundgren, Bo: *"Dokumentation för återanvändning av mikromaterial från SCB:s undersökningar"*, Stockholm 1991. Finns även i engelsk översättning.

göra tröskeln för dokumentationsarbetet så låg som möjligt, kan merparten av informationen ges i form av löpande text, på samma sätt som när man framställer vilket dokument som helst med hjälp av sin ordbehandlare. På vissa punkter, t.ex. variabel- och postbeskrivningar, skall informationen tillhandahållas enligt mera formaliserade delmallar, med stöd av mallar och s.k. makros i ordbehandlaren.

Kriteriet för vad som måste behandlas i en SCBDOK-dokumentation är i princip enkelt och klart. Med hjälp av dokumentationen skall en framtida (åter)användare av ett arkiverat observationsregister från en statistisk undersökning kunna tolka och analysera datamaterialet på ett självständigt och statistiskt ansvarsfullt sätt, även om han eller hon inte har möjlighet att konsultera några som själva medverkat i den bakomliggande undersökningen. Användaren måste givetvis ha erforderliga statistiska metodkunskaper eller kunna rådfråga sådan expertis, men i övrigt skall dokumentationen ge all den information om datamaterialet, som användaren behöver.

Figur 2 åskådliggör huvuddragen i version 2.0 av SCBDOK-mallen. Mera detaljerade instruktioner, checklistor och delmallar finns i den handbok som åtföljer dokumentationssystemet.<sup>4</sup>

<sup>4</sup> "Att dokumentera statistiska undersökningar – Användarhandbok för SCBDOK version 2.0", SCB 1996.

<p><b>0 Administrativa uppgifter</b></p> <p>0.1 Benämning  0.2 Statistikområde  0.3 Ansvarig myndighet, person etc  0.4 Kostnad  0.5 Syfte och historik  0.6 Användare och användningsområden  0.7 Uppgiftsskyldighet  0.8 Sekretessregler och gallringsföreskrifter  0.9 EU-reglering</p>	<p><b>1 Översikt</b></p> <p>1.1 Undersökningens uppläggning  1.2 Innehåll: statistiska storheter  1.3 Utflöden: statistik och mikrodata  1.4 Tidsplaner  1.5 Dokumentation</p>
<p><b>2 Uppgiftsinsamling</b></p> <p>2.1 Ram och ramförfarande  2.2 Urvalsförfarande  2.3 Mätinstrument  2.4 Insamlingsförfarande  2.5 Databeredning</p>	<p><b>3 Observationsregister</b></p> <p>3.1 Mål- och observationsobjekt  3.2 Variabelförteckningar  3.3 Fysisk organisation  3.4 Erfarenheter från senaste undersökningsomgången</p>
<p><b>4 Statistisk bearbetning och redovisning</b></p> <p>4.1 Skattningar: antaganden och beräkningsformler  4.2 Redovisningsförfaranden</p>	<p><b>5 Databehandlingssystem</b></p> <p>5.1 Systemöversikt och systemflöde  5.2 Bearbetningar  5.3 Databasmodell  5.4 Databastabeller  5.5 Databastillbehör  5.6 Rapporter  5.7 Övriga datamängder</p>
<p><b>6 Loggbok</b></p>	

Figur 2. Huvudstrukturen i dokumentationsmallen SCBDOK 2.0.

### 1.3.2 Kvalitetsbegrepp och kvalitetsdeklarationer

På basis av samma begrepps- och beskrivningsmodell, som låg till grund för dokumentationssystemet SCBDOK, utarbetades sedermera också en standard för kvalitetsdeklarationer.<sup>5</sup> En kvalitetsdeklaration har ett delvis

<sup>5</sup> Standarden är utförligt beskriven och exemplifierad i "*Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik SCB*", Meddelanden i



annat syfte än en observationsregisterdokumentation, vilket återspeglas i den dokumentationsmall, som ligger till grund för kvalitetsdeklarationen. Det handlar i många stycken om samma information som i SCBDOK-mallen men presenterad på ett annat sätt, för andra användningar och användare.

Figur 3a åskådliggör strukturen i en standardiserad kvalitetsdeklaration för officiell statistik enligt MIS 1994:3. Figur 3b åskådliggör en engelsk översättning av kvalitetsbegreppets struktur.<sup>6</sup> Det handlar om en reviderad version, där författarna även tagit hänsyn till olika synpunkter som inkommit, i första hand från Eurostat.

Innehåll	Tid
Statistiska storheter	Referenstid
Objekt och population	Framställningstid
Variabler	Punktlighet
Statistiska mått	Frekvens
Redovisningsgrupper	Jämförbarhet över tiden
Jämförbarhet med annan statistik	
Tillförlitlighet	Tillgänglighet
Tillförlitlighet totalt	Spridningsformer
Osäkerhetskällor	Presentation
Täckning	Dokumentation
Urval	Primärmaterial
Mätning	Upplysningar
Bortfall	
Bearbetning	
Modellantaganden	
Redovisning av osäkerhetsmått	

Figur 3a. Kvalitetsbegreppets struktur för officiell statistik.

samordningsfrågor (MIS) 1994:3, utarbetad av Chris Denell, Eva Elvers och Bengt Rosén, samtliga vid SCB.

<sup>6</sup> Den i samarbete med Eurostat föreslagna strukturen beskrivs och kommenteras i Eva Elvers och Bengt Rosén: "Quality Concept for Official Statistics", R&D Report 1997:2, SCB.

<p><b>Contents</b></p> <p>Statistical target characteristics  Units and population  Variables  Statistical measures  Study domains  Reference time</p> <p>Comprehensiveness</p>	<p><b>Timeliness</b></p> <p>Frequency  Production time  Punctuality</p>
<p><b>Accuracy</b></p> <p>Overall accuracy  Sources of inaccuracy  Sampling  Coverage  Measurement  Non-response  Data processing  Model assumptions</p> <p>Presentation of accuracy measures</p>	<p><b>Coherence, especially comparability</b></p> <p>Comparability over time  Comparability over space  Coherence in general</p>
	<p><b>Availability and clarity</b></p> <p>Forms of dissemination  Presentation  Documentation  Access to microdata  Information services</p>

**Figur 3b. Quality concept for official statistics**

SCBDOK-mallen är i första hand till för mycket kvalificerade och självständiga statistikanvändare, t.ex. forskare, som inte utan vidare vill acceptera de antaganden om det ena och det andra, t.ex. bortfallets effekter, som den ursprungliga statistikproducenten gjort, utan som kanske vill pröva alternativa antaganden. En kvalitetsdeklaration är däremot i första hand till för en användare, som inte går tillbaka till bakomliggande observationsregister för att göra nya bearbetningar, utan en användare, som utgår från den statistik på aggregerad nivå, makrodata, som den ursprungliga producenten publicerat. Även denna typ av användare har naturligtvis behov av god information om olika förutsättningar och omständigheter kring statistiken och dess fram-

tagande, men informationen behöver inte vara lika djupgående och detaljerad som för en användare, som vill tränga djupt in i underliggande mikrodata. Däremot väger pedagogiska aspekter tungt för användaren av publicerad statistik, som inte sällan har ont om tid, och som kanske inte har så goda statistiska metodkunskaper.

En kvalitetsdeklaration av publicerad statistik kan alltså vara något ytligare och mera översiktlig än en observationsregisterdokumentation, men i gengäld måste den vara mera pedagogiskt utformad. Den fokuserar också på makrodata, medan observationsregisterdokumentationen fokuserar på mikrodata.

### 1.3.3 Översiktliga produktbeskrivningar

För vissa syften behövs mera kortfattade, översiktliga och lättillgängliga beskrivningar av den officiella statistiken. För att tillgodose dessa behov, har en standardiserad mall för s.k. produktbeskrivningar tagits fram. Enligt föreskrifter från SCB,<sup>7</sup> baserade på förordningen om den officiella statistiken,<sup>8</sup> skall en sådan produktbeskrivning tas fram av respektive statistikansvariga myndigheter för varje statistikprodukt, som ingår i den officiella statistiken, och den skall aktualiseras en gång om året.

SCB har utarbetat en dokumentationsmall för produktbeskrivningarna.<sup>9</sup> Även denna mall är baserad på den begrepps- och beskrivningsmodell, som ligger bakom dokumentationssystemet SCBDOK.

Figur 4 åskådliggör produktbeskrivningsmallen. Som synes har den stora strukturella likheter med mallen för kvalitetsdeklarationer, bortsett från att ett avsnitt 0 med administrativa uppgifter lagts till. Av anvisningarna till produktbeskrivningsmallen framgår dock, om man jämför med MIS 1994:3, att det ställs lägre krav på omfattning och djup i en produktbeskrivning jämfört med vad som gäller för en fullödig kvalitetsdeklaration.

<sup>7</sup> "Föreskrifter och allmänna råd om offentliggörande av officiell statistik", SCB, Meddelanden i samordningsfrågor (MIS) 1995:5, utarbetade av Per Samuelson, Bo Sundgren

<sup>8</sup> "Förordningen om den officiella statistiken", SFS 1992:1668, ändrad i SFS 1994:1108, omtryck.

<sup>9</sup> "Produktbeskrivningsmall 1997 för officiell statistik – Med anvisningar", SCB 1997, reviderad version 1998, utarbetad av Chris Denell, Gunilla Lundholm, Per Nyström, Bo Sundgren m.fl.

<p><b>0 Administrativa uppgifter om produkten</b></p> <p>0.1 Benämning  0.2 Statistikområde  0.3 Ansvarig myndighet, person etc  0.4 Kostnad  0.5 Syfte och historik  0.6 Användare och användningsområden  0.7 Uppgiftsskyldighet  0.8 Sekretessregler och gallringsföreskrifter  0.9 EU-reglering</p>	<p><b>1 Innehåll</b></p> <p>1.1 Statistiska storheter  – objekt och population  – variabler  – statistiska mått  – redovisningsgrupper</p> <p>1.2 Jämförbarhet med annan statistik</p>
	<p><b>2 Tid</b></p> <p>2.1 Referenstid  2.2 Framställningstid  2.3 Punktlighet  2.4 Frekvens  2.5 Jämförbarhet över tiden</p>
<p>3 Uppläggning och tillförlitlighet</p> <p>3.1 Tillförlitlighet totalt  3.2 Osäkerhetskällor:  – täckning  – urval  – uppgiftsinsamling/mätning  – bortfall  – bearbetning  – modellantaganden  3.3 Redovisning av osäkerhetsmått</p>	<p>4 Tillgänglighet</p> <p>4.1 Statistik: publicering och spridning (inkl elektroniska media och databaser)  4.2 Primärmaterial: lagring och åtkomstmöjligheter  4.3 Presentation  4.4 Dokumentation  4.5 Övriga upplysningar</p>

Figur 4. Produktbeskrivningsmall för officiell statistik.

### 1.3.4 Mot ett integrerat metadatasystem

Dokumentationsarbete upplevs ofta av dem som måste göra det, delvis obefogat, som tråkigt och tidsödande. Det är därför viktigt, att det läggs upp på ett sådant sätt, att ansträngningarna kan minimeras, och att de som dokumenterar känner sig så välmotiverade som möjligt.

Ett sätt att minimera ansträngningarna är att se till att dokumentationsarbetet blir så integrerat med andra arbetsuppgifter som möjligt,

och att en dokumentation som utarbetats för ett visst syfte också så långt som möjligt utnyttjas för andra syften, d.v.s. man skall aldrig behöva dokumentera samma sak två gånger.

Sålunda bör dokumentationsarbetet påbörjas redan när man börjar utforma en statistisk undersökning, så att den fullständiga dokumentationen växer fram mer eller mindre som en biprodukt av planerings- och genomförandeprocesserna. Ett och samma dokument kan användas först som ett underlag för ett designbeslut och senare, när det diskuterats färdigt och "frysts", infogas som ett bidrag till dokumentationen av det färdiga systemet.

Med hjälp av ordbehandlare, mallar och andra IT-baserade verktyg kan man stödja dokumentationsarbetet och säkerställa att man inte behöver göra om en dokumentationsinsats, som man redan gjort i ett annat sammanhang. När man genomför en ny omgång av en regelbundet återkommande statistisk undersökning, skall man naturligtvis inte behöva dokumentera om sådant som inte förändrats sedan föregående undersökningsomgång. Delar av en observationsregisterdokumentation kan också återanvändas (kopieras) för motsvarande delar av en kvalitetsdeklaration av den statistik som tagits fram ur observationsregistret. En produktbeskrivning består i mångt och mycket av utdrag ur och sammanfattningar av mera fullständiga observationsregisterdokumentationer och kvalitetsdeklarationer. Etc.

Figur 5 åskådliggör vilka logiska samband som finns mellan några viktiga metadataobjekt.<sup>10</sup>

Den begreppsmässiga och tekniska utvecklingen på metadataområdet har kommit ganska långt inom SCB. Resultaten av dessa insatser har även fått ett visst genomslag i det internationella utvecklings- och standardiseringsarbetet på detta område. Ändamålsenliga datorbaserade verktyg utvecklas för att stödja och "knyta ihop" olika delar av dokumentations- och metadataarbetet. Detta arbete har delvis skett inom ramen för ett forskningsprojekt<sup>11</sup> inom EU:s fjärde

<sup>10</sup> En dokumentation består av metadata. Metadata kan vara mer eller mindre strukturerade och formaliserade. Beskrivningar i form av löpande text är exempel på ostrukturerade metadata. Med hjälp av checklistor och mallar kan beskrivningarna ges en viss struktur. Genom att mera i detalj styra hur en mall skall fyllas i, kan man framtvinga en hårt strukturerad beskrivning, s.k. formaliserade metadata. Formaliserade metadata består i stor utsträckning av koder av olika slag och behövs i första hand för styrningen av datorprogram. Människor föredrar i allmänhet att få tillgång till metadata i form av lätt strukturerade texter i naturligt språk.

<sup>11</sup> Projektet heter "Integrated MetaInformation Management" (IMIM). Projektkoordinator är Bo Sundgren och projektledare Lars Rauch.

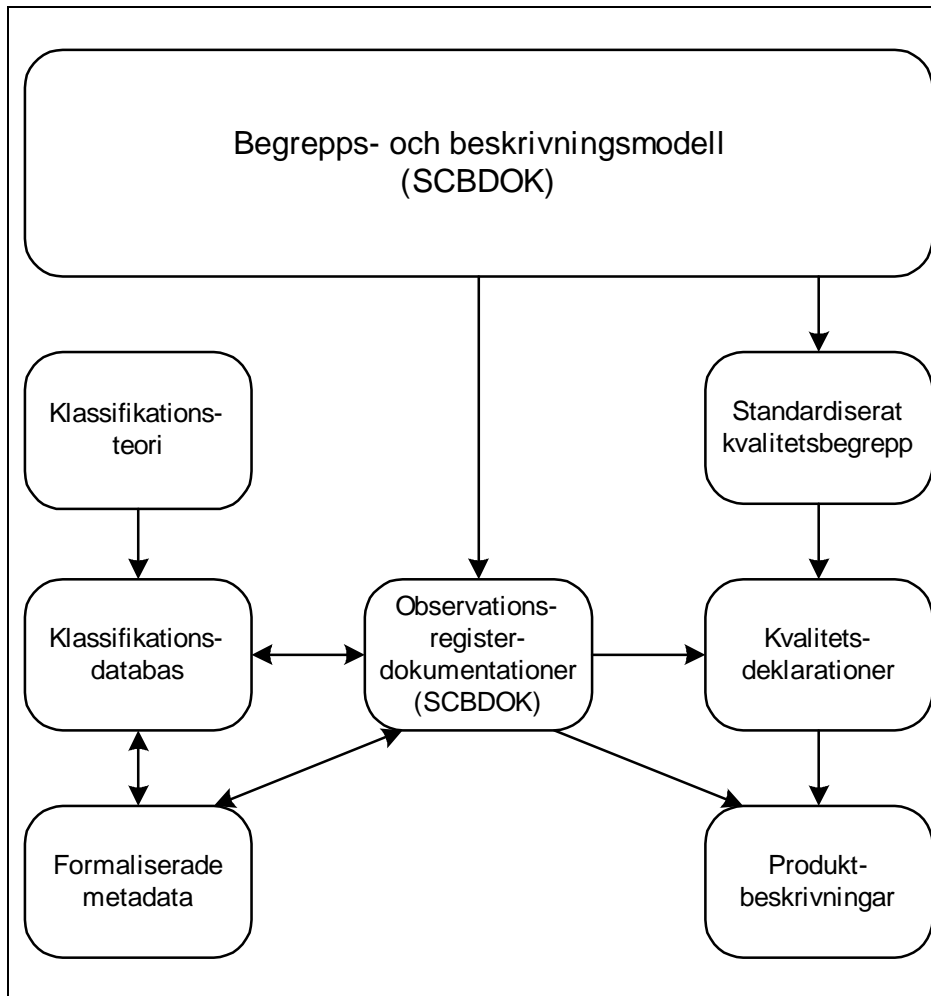
ramprogram för forskning och utveckling, där SCB svarat för projektledningen.

Utvecklingen på dokumentations- och metadataområdet har alltså kommit ganska långt, och den går vidare i rask takt. Den svaga punkten är det systematiska genomförandet av de uppnådda resultaten. Detta är tyvärr inte ovanligt när det gäller statistiskt utvecklingsarbete, och som alltid är det i första hand statistikanvändarna som blir lidande. Den enda av ovan beskrivna dokumentationsinsatser, som blivit någorlunda systematiskt och konsekvent genomförd i praktiken, är framställningen av årligen aktualiserade produktbeskrivningar. Detta är i och för sig utmärkt, men produktbeskrivningarna är alldeles för ytliga – och ibland dessutom av tveksam kvalitet – för att duga för de kvalificerade statistikanvändare, som vill ha tillgång till mikrodata för egna bearbetningar, eller som vill göra mera avancerade analyser av den publicerade statistiken.

Fullständiga observationsregisterdokumentationer föreligger hittills bara för en handfull produkter i den officiella statistiken.

En vanlig ”ursäkt” från de statistikansvariga – inte minst inom SCB – är att man inte har tid att genomföra det erforderliga dokumentationsarbetet. För att undersöka om detta kunde vara en rimlig förklaring gjordes ett experiment. Ett antal produkter från olika avdelningar inom SCB dokumenterades enligt SCBDOK-modellen. Tidsåtgången mättes. Den varierade givetvis, beroende på produkternas omfattning och komplexitet, men i genomsnitt handlade det om ca en personmånads arbetsinsats per produkt. Om man räknar med att det finns ca 240 produkter i den officiella statistiken, så skulle en fullständig dokumentation av dessa kräva 20 personår, d.v.s. knappt 2% av den volym arbete som utförs vid SCB under ett år.<sup>12</sup> Det är knappast orimligt att en arbetsinsats av denna omfattning skall kunna rymmas inom befintliga anslag.

<sup>12</sup> Andra statistikansvariga myndigheter än SCB är i och för sig ansvariga för ungefär hälften av de 240 produkterna, men eftersom SCB-anställda i stor utsträckning svarar för produktionen av även dessa produkter, kan det vara rätt att räkna med arbetsinsatsen för dem också, när man ställer den i relation till mängden av allt arbete som utförs vid SCB under ett år. Å andra sidan handlade det här om en ”förstagångsdokumentation” av samtliga produkter, och den utfördes i flera fall av personer, som inte hade någon som helst förkunskap om de produkter, som de skulle dokumentera. Aktualisering av befintliga dokumentationer, utförd av personer som ”kan” sina produkter bör ta betydligt kortare tid. I ett fortvarighetstillstånd bör dokumentationsarbetet sålunda inte kräva mer än ca 1% av statistikproducentens resurser.



Figur 5. En översikt över olika objekt på metadataområdet och deras samband med varandra.

## 1.4 Tillgänglighetsprojektet

Som ett resultat av den ovan refererade kritiken från statsmakterna och tunga statistikanvändare (se avsnitt 1.2) tillsatte SCB ett särskilt verksamhetsprojekt med uppgift att se vilka åtgärder av olika slag, som skulle kunna vidtagas för att drastiskt öka tillgängligheten till statistiska datamaterial.

De huvudsakliga klagomålen var alltså:

- att SCB använder gammalmodig stordator teknik, klumpig, dyr och föga användarvänlig
- att datamaterialen är dåligt dokumenterade och därför svåra att hitta och utnyttja
- att SCB:s medarbetare, med hänvisning till kvalitets- och sekretessproblem, i alltför stor utsträckning vill utföra ad hoc bearbetningar av mikrodata själva, på uppdragsbasis, i stället för att låta användarna göra bearbetningar och analyser på egen hand, enligt självbetjäningsprincipen

Projektet skulle föreslå lösningar på de nämnda problemen, och även utföra en del praktiska experiment i samarbete med externa användare för att testa de föreslagna lösningarna.

I kort sammandrag blev resultaten av Tillgänglighetsprojektet följande:

- En helt ny modell för statistikproduktionen togs fram, helt baserad på persondatorer i nätverk och databasorienterade client/server-system. Försök visade, att även mycket stora databaser skulle kunna hanteras på detta sätt. Ett nytt projekt, Plattformsprojektet, initierades för att förverkliga den nya produktionsmodellen. Omläggningen skulle vara klar före år 2000, så att SCB då skulle kunna göra sig av med stordatorn och därmed som en bieffekt bli av med en stor del av den s.k. 2000-problematiken.
- SCBDOK-modellen för dokumentation av observationsregister testades praktiskt tillsammans med både statistikproducenter och externa statistik användare. Testerna utföll till användarnas belåtenhet. Producenterna hade dock vissa tidigare refererade kritiska synpunkter beträffande de erforderliga arbetsinsatserna (se avsnitt 1.3), och man trodde inte – i motsats till användarna själva – att dokumentationerna skulle vara tillräckligt användarvänliga.
- Ett alternativt, mera variabelorienterat dokumentationsverktyg, DAISY, testades. Systemet hade utvecklats vid en ämnesenhet på SCB och var skräddarsytt för ett tiotal observationsregister, som man hanterade på enheten ifråga. Testen skedde i samarbete med Socialstyrelsen, som var intresserad av att använda produkten, förutsatt att man fick generalisera och vidareutveckla den för sina egna behov. Ämnesenheten på SCB, som "ägde" verktyget, var dock inte intresserad av ett samarbete med Socialstyrelsen på dessa villkor. Man vill hellre vidareutveckla verktyget själv men hade inte resurser för detta, varför projektet rann ut i sanden.



- Olika metoder att hantera sekretessproblematiken på ett konstruktivt sätt prövades. Genom förbudet mot olovlig identifiering,<sup>13</sup> som just hade lagstadsrats av riksdagen, hade s.k. bakvägsidentifiering kriminaliserats, vilket ökade möjligheterna för en statistikansvarig myndighet att lämna ut avidentifierade mikrodata till forskare och andra seriösa statistik användare, t.ex. utredare vid privata och offentliga utredningsorgan. Tillgänglighetsprojektet prövade dessa möjligheter genom att anskaffa programvaran SuperCross från ett programvaruföretag i Australien, Space Time Research, en programvara som visade sig kunna hantera mycket stora mängder avidentifierade mikrodata på ett mycket effektivt, flexibelt och säkert sätt. På en vanlig PC kunde man t.ex. på någon minut framställa en godtycklig statistisk tabell över hela den svenska befolkningen, ca 9 miljoner poster.

## 1.5 Databasprojektet och Sveriges statistiska databaser

Genom 1990 års statistikutredning<sup>14</sup> och 1993 års kommitté med uppgift att förbereda ändringar i ansvarsfördelningen för den statliga statistiken, den s.k. Genomförandekommittén,<sup>15</sup> genomfördes en reform med syfte att öka användarinflytandet över styrningen av den officiella statistiken i Sverige genom att i viss utsträckning decentralisera beställansvaret för officiell statistik till ett antal sektormyndigheter, som vid sidan av SCB skulle fungera som statistikansvariga myndigheter. SCB skulle behålla statistikansvaret för ungefär hälften av statistikprodukterna, som ingår i Sveriges officiella statistik. Genom statistikreformen gavs SCB ett ökat ansvar för bl.a. tillgänglighetsfrågorna. Sålunda fick SCB ett särskilt ansvar för att med anslagsmedel bygga upp en databasservice omfattande ett antal centrala databaser och register.

För att utföra uppdraget att bygga upp en databasservice tillsatte SCB först en databasutredning och därefter ett databasprojekt, vars resultat redovisades i slutet av 1994,<sup>16</sup> respektive i slutet av 1996.<sup>17</sup>

<sup>13</sup> Lagen om den officiella statistiken, 10§ och 15§, 1992.

<sup>14</sup> Utredningens förslag redovisades 1992 i betänkandet "*Effektivare statistikstyrning. Den statliga statistikens styrning och samordning*", SOU 1992:48.

<sup>15</sup> Genomförandekommitténs förslag redovisades 1994 i betänkandet "*Ändrad ansvarsfördelning för den statliga statistiken*", SOU 1994:1.

<sup>16</sup> "*Den framtida databasverksamheten vid SCB – Slutrapport*", SCB 1994.

<sup>17</sup> "*Slutrapport från databasprojekten*", SCB 1997.

Verksamheten sattes i reguljär drift 1 januari 1997 och gavs därvid namnet "Sveriges statistiska databaser". Ett nytt databasprojekt tillsattes samtidigt med uppgift att fullborda utvecklingen av den nya databasservicen fram till år 2000.

### 1.5.1 Innehåll: makrodata, mikrodata och metadata

Sveriges statistiska databaser startade alltså 1 januari 1997. Databaserna innehåller tre olika typer av data:

- makrodata: det som vi kallar "statistik" i dagligt tal, d.v.s. aggregerade data om populationer och klasser av objekt, t.ex. summor, frekvenser, medelvärden
- mikrodata: data om enskilda objekt, t.ex. företag, individer, händelser, som ligger till grund för statistik, och som kan aggregeras till statistik på ett flexibelt sätt
- metadata: "data om data", d.v.s. data som definierar och beskriver mikrodata och makrodata

En mera omfattande definition av "statistik" eller "statistiska data" skulle kunna vara

- makrodata med bakomliggande mikrodata och tillhörande metadata

SCB har som mål att Sveriges statistiska databaser skall innehålla

- all officiell statistik, som SCB är statistikansvarig myndighet för<sup>18</sup>
- ett antal av statsmakterna uppräknade mikrodatamaterial (observationsregister), som ligger till grund för officiell statistik, och som statistik användarna ofta vill bearbeta och samköra

Någon allmänt accepterad definition av "officiell statistik" finns inte. Statsmakterna har definierat begreppet genom explicit uppräknade av 22

<sup>18</sup> Anledningen till begränsningen till den statistik som SCB svarar för är att det i enlighet med statistikutredningen inte råder något tvång för statistikansvariga myndigheter att använda Sveriges statistiska databaser för publicering och spridning. Alla statistikansvariga myndigheter är dock välkomna att använda Sveriges statistiska databaser som spridningskanal. Man kan diskutera om det inte ur samordnings- och tillgänglighetssynpunkt skulle vara en fördel om alla statistikansvariga myndigheter hade skyldighet sprida all officiell statistik, som de svarar för, genom Sveriges statistiska databaser. Databaserna skulle naturligtvis inte behöva vara den enda spridningskanalen, men den skulle vara obligatorisk.

ämnesområden; se rutan till vänster i figur 6. Inom vart och ett av ämnesområdena finns ett antal explicit angivna statistikprodukter.<sup>19</sup>

En mera beskrivande definition av begreppet "officiell statistik" skulle kunna vara följande:

- Med officiell statistik avses statistik av allmänt intresse, som uppfyller av statsmakterna fastställda krav på objektivitet, allmänintresse och kvalitet.

Ämnesområden i officiell statistik	Observationsregister
<ul style="list-style-type: none"> <li>• Arbetsmarknad</li> <li>• Befolkning</li> <li>• Bostäder och byggande</li> <li>• Energi</li> <li>• Finansmarknad</li> <li>• Handel med varor och tjänster</li> <li>• Hälso- och sjukvård</li> <li>• Inkomster och inkomstfördelning</li> <li>• Jord- och skogsbruk, fiske</li> <li>• Kultur och fritid</li> <li>• Levnadsförhållanden</li> <li>• Medborgarinflytande</li> <li>• Miljövard och naturresurshushållning</li> <li>• Nationalräkenskaper</li> <li>• Näringsverksamhet</li> <li>• Offentlig ekonomi</li> <li>• Priser och konsumtion</li> <li>• Rättsväsende</li> <li>• Socialförsäkring</li> <li>• Socialtjänst</li> <li>• Transporter och kommunikation</li> <li>• Utbildning och forskning</li> </ul>	<p>Följande register tillgängliggörs för statistikbearbetningar genom SCB:s databasservice (SFS 1995:1060):</p> <ul style="list-style-type: none"> <li>• Registret över totalbefolkningen, RTB</li> <li>• Befolkningsförändringar</li> <li>• Folk- och bostadsräkningar, FoB</li> <li>• Fastighetstaxeringar, FTR</li> <li>• Utbildning, UREG</li> <li>• Arbetskraftsundersökningar</li> <li>• Kontrolluppgifter</li> <li>• Centrala företags- och arbetsställeregistret, CFAR</li> <li>• Utrikeshandel</li> <li>• Bilregistret</li> <li>• Lantbruksregistret, LBR</li> <li>• Inkomst- och förmögenhet, IoF</li> <li>• Årlig sysselsättning, ÅRSYS</li> <li>• Utbildningsanalys</li> <li>• För brott lagförda personer</li> <li>• Levnadsförhållandena i samhället, ULF</li> <li>• Mervärdesskatt</li> </ul>

**Figur 6. Översikt över innehållet av makrodata (till vänster) respektive mikrodata (till höger) i Sveriges statistiska databaser.**

Till den officiella statistiken bör man strängt taget också räkna de mikrodata (s.k. observationsregister) som ligger till grund för den officiella statistiken. Ett antal av dessa har redan av statsmakterna givits en speciell ställning, genom att de skall hållas tillgängliga för statistikbearbetningar genom SCB:s databasservice; se rutan till höger i figur 6. På sikt borde alla observationsregister, som ligger till grund för officiell statistik, kunna tillgängliggöras på likartat sätt. Inom SCB bedrivs f.n. ett verksprojekt, det s.k. registerprojektet, som bl.a. förväntas förbättra förutsättningarna för snabba, flexibla och kostnadseffektiva statistikbearbetningar av observationsregistren, såväl bearbetningar av enstaka register som kombinationsbearbetningar (samkörningar) av flera register samtidigt. Forskarna i SCB:s vetenskapliga råd har uttalat sig för att SCB borde inrätta en särskild registerfunktion med uppgiften att underlätta bearbetningar och samkörningar av observationsregister för

<sup>19</sup> "Förordningen (1994:1108) om den officiella statistiken", bilaga.

statistiska ändamål. Givetvis skulle alla sådana bearbetningar ske under beaktande av statistiksekretessen.

En viktig egenskap hos Sveriges statistiska databaser är att de, förutom data, innehåller en rad olika definitioner, beskrivningar, förklaringar och kvalitetsdeklarationer av data, s.k. metadata, ”data om data”. En del metadata lagras i fritextform och hanteras med hjälp av s.k. sökmotorer och fritexthanterare. Andra data lagras på i princip samma sätt som data, d.v.s. i tabellstrukturerade relationsdatabaser.

Exempel på olika slag av metadata ges i figur 7.

<b>Exempel på beskrivande och förklarande metadata i fritextform i Sveriges statistiska databaser</b>	<b>Exempel på formaliserade och strukturerade metadata i Sveriges statistiska databaser</b>
<ul style="list-style-type: none"> <li>• Produktbeskrivningar och kvalitetsdeklarationer: översiktliga beskrivningar av alla statistikprodukter, som ingår i Sveriges officiella statistik (SOS)</li> <li>• Fullständiga beskrivningar (enligt SCBDOK) av observationsregister och produkter, som ingår i SOS</li> <li>• Beskrivningar av arkiverade datamaterial (ARKEN)</li> <li>• Beskrivningar av standardiserade klassifikationer (KDB)</li> </ul>	<ul style="list-style-type: none"> <li>• Variabler</li> <li>• Värdemängder och värden</li> <li>• Tidsskalor</li> <li>• Fotnoter</li> <li>• Korrespondensregler mellan olika klassifikationer (KDB)</li> <li>• Tekniska metadata</li> </ul>

**Figur 7. Exempel på olika slag av metadata i Sveriges statistiska databaser.**

De mest ingående och utförliga beskrivningarna och förklaringarna av innehållet i Sveriges statistiska databaser ges i dokumentationer, som utformats enligt den s.k. SCBDOK-modellen, en beskrivningsmodell för observationsregister och statistiska undersökningar (se avsnitt 1.3). En SCBDOK-beskrivning av en statistikprodukt skall vara så ingående och utförlig, att även en statistikanvändare, som inte själv har någon förstahandserfarenhet av undersökningen, skall kunna återanvända alla statistiska data, som genererats av undersökningen, kanske årtionden efter det att undersökningen genomförts, när de som ursprungligen genomförde undersökningen inte längre finns tillgängliga.

### 1.5.2 Tillgänglighet via Internet

Statistiken (makrodata) i Sveriges statistiska databaser och tillhörande beskrivningar (metadata) tillgängliggörs för statistikanvändarna i första hand via Internet. Det sker på flera olika sätt. Utgångspunkten är att Internetanvändaren har tillgång till ett abonnemang på Sveriges statistiska databaser (kostar f.n. 6.000 kr per juridisk person och år; gratis för

ett 70-tal bibliotek) samt en vanlig s.k. webbläsare (Internet Explorer eller Netscape). Via Internetgränssnittet kan användaren navigera sig fram till önskad statistik med hjälp av olika sökverktyg, och därefter kan han eller hon betrakta statistiken på skärmen och, alltefter egna önskemål, skriva ut den eller ladda ner den i önskat format till den egna datorn för vidare bearbetningar.

### 1.5.3 Användarstyrd statistikframställning: möjligheter och begränsningar

Alternativt kan användaren ladda ner en särskild programvara, PC-AXIS, en s.k. statistikbläddrare, med vars hjälp man kan "vända och vrida" på statistisk information och redigera och bearbeta den, tills dess att man har fått den presenterad just så, som man vill ha den. Med hjälp av PC-AXIS kan man också snabbt och enkelt komma åt ett antal färdigframställda statistikmaterial, som ofta efterfrågas.

En användare som vill ha mera skräddarsydda statistikuttag, kan antingen själv, eller med SCB:s hjälp, göra mera komplicerade bearbetningar direkt mot de relationsdatabaser, där statistiken ligger lagrad. Även de bakomliggande observationsregistren skulle på detta sätt kunna göras flexibelt tillgängliga för statistikbearbetningar, men den rätten förbehålls tills vidare SCB-personalen. Som skäl för denna begränsning brukar anföras

- att alla, icke helt fördefinierade och av statistikproducenten sekretessgranskade bearbetningar av mikrodata måste granskas av SCB, innan de får genomföras; s.k. skadeprövning enligt sekretesslagen
- att produktbeskrivningar, kvalitetsdeklarationer och annan dokumentation inte ännu håller en sådan kvalitet, att statistikanvändarna ännu kan få göra vilka statistikbearbetningar som helst på egen hand, utan att först konsultera ansvarig statistikproducent

Ett ytterligare skäl för begränsningen, som ibland framkastas av vissa statistikanvändare, är att skräddarsydda bearbetningar av mikrodata för olika statistikanvändares räkning utgör en viktig del av SCB:s uppdragsverksamhet. Om dessa uppdrag plötsligen föll bort, genom att användarna kan göra dem själva, skulle det åtminstone kortsiktigt ge SCB problem med att finansiera den berörda personalen. Samma kritiker brukar också påpeka, att det kan finnas ett visst egenintresse hos statistikproducentens experter att slå vakt om det kunskapsmonopol man *de facto* har, så länge inte dokumentationen av statistiken och bakomliggande data och processer är tillräckligt bra och komplett.

Det är naturligtvis mycket svårt att objektivt avgöra huruvida det verkligen ligger något i ovan refererade anklagelser, eller om det bara rör sig om orättvisa misstänkliggöranden. Bäst vore enligt min mening att se till att ännu större ansträngningar görs för att eliminera de kvar-varande hindren för maximal tillgänglighet på användarnas villkor, inom ramen för gällande statistiksekretess. Mycket kan fortfarande göras. Mer om detta senare.

#### 1.5.4 Elektronisk publicering

Sveriges statistiska databaser lämpar sig även som bas för en rad andra spridningsformer, t.ex. CD-ROM, disketter och publikationsframställning. När det gäller statistikpubliceringen, har SCB nyligen startat ett projekt, Publiceringsprojektet, som syftar till att standardisera och automatisera produktionen av pressmeddelanden och s.k. Statistiska Meddelanden (SM), genom vilka den officiella statistiken rutinmässigt publiceras och sprids.

De nya publiceringsrutinerna, som nu tas fram, baseras på att all officiell statistik i första hand publiceras genom att den läggs in i Sveriges statistiska databaser, och där görs tillgänglig vid en viss tidpunkt. Med den databaslagrade statistiken och tillhörande metadata som grund, kan standardiserade, elektroniska pressmeddelanden och SM framställas med hjälp av fördefinierade mallar och program. Tryckta versioner framställs och skickas till dem som beställt dem, men någon reservupplaga trycks inte. Om någon statistikanvändare senare skulle vilja ha ett tryckt exemplar, kan han eller hon själv trycka ut det på sin egen skrivare via Internet eller beställa ett exemplar från SCB, som då trycker upp det enligt principen för s.k. Print On Demand.

Bara besparingseffekten av att inte behöva trycka upp och lagerhålla restupplagor för en eventuell, framtida efterfrågan beräknas bli så stor att den kan finansiera gratis tillgängliggörande av motsvarande elektroniska produkter, s.k. webb-SM. De prenumererade eller spontanbeställda tryckta exemplaren kan å andra sidan behöva bli något dyrare än idag.

### 1.5.5 Internet som ingång till all officiell statistik i Sverige och internationellt

Internet håller snabbt på att bli den naturliga ingången till all officiell statistik:

- Korta nyhetsbulletiner på statistikansvariga myndigheters hemsidor
- Rutinmässig statistikrapportering med korta kommentarer i form av elektroniska pressmeddelanden och webb-SM
- Flexibel tillgänglighet till detaljerad statistik och tidsserier, inklusive omfattande och ingående beskrivningar och förklaringar, via Sveriges statistiska databaser
- Länkar till relaterad statistik från andra statistikproducenter inom och utom Sverige

För att knyta ihop den officiella statistiken, oavsett vem som är ansvarig för den, utvecklas f.n. det svenska statistiknätet efter en del svenska och utländska förebilder såsom det svenska miljödatanätet och FedStatnätet i USA.

Tillgången till internationell statistik kommer att underlättas av att många nationella statistikbyråer världen över, liksom flera internationella organisationer, är i full färd med att bygga upp system liknande Sveriges statistiska databaser och Svenska statistiknätet.

### 1.5.6 Rutiner för fortlöpande aktualisering av Sveriges statistiska databaser

För att ett system som Sveriges statistiska databaser skall fungera effektivt, inte bara tillfälligt, utan även uthålligt och långsiktigt, måste man lägga stor vikt vid att databaserna förs med indata (mikrodata, makrodata och metadata) på ett rutinmässigt och automatiserat sätt. Alla erforderliga indata och aktualiseringar skall helst flöda in från de bakomliggande statistikproduktionssystemen till databaserna via standardiserade gränssnitt och utan att det behöver ske någon ”handpåläggning”.

Sveriges statistiska databaser är utformade för att kunna fungera på detta sätt, men ännu så länge kommer merparten av data- och metadatainflödet från SCB:s gamla, stordatorbaserade databaser. Det beror på att merparten av SCB:s produktionssystem fortfarande går i stordatormiljön. Vid millenniumskiftet skall emellertid alla produktionssystem vid SCB vara stordatorberoende, och produktionen kommer i stället att ske i en modern, PC-baserad client/servermiljö med relationsdatabaser. En standardiserad s.k. databaskoppling har utformats för

att möjliggöra automatisk överföring av statistiska data och metadata från de nya statistikproduktionssystemen till Sveriges statistiska databaser via ett s.k. SQL-gränssnitt.

Även andra statistikansvariga myndigheter än SCB kan utnyttja den standardiserade databaskopplingen och SQL-gränssnittet för att på samma sätt överföra statistik och metadata från sina produktionssystem till Sveriges statistiska databaser. Detta gäller vare sig en statistikansvarig myndighet väljer att köpa själva statistikproduktionen från SCB eller någon annan statistikproducent eller väljer att sköta statistikproduktionen i egen regi. Kostnaderna för eventuellt nödvändiga anpassningsåtgärder och manuella ingrepp ligger på respektive statistikansvarig myndighet, liksom de ligger på SCB:s produktansvariga statistikprogram, när det är SCB som är den statistikansvariga myndigheten.

Databasservicen, Sveriges statistiska databaser, svarar för att den inkommande statistiken och tillhörande metadata är utformade i enlighet med fastställda standarder. När en viss statistikproduktion för första gången ansluts till Sveriges statistiska databaser, är det vanligen nödvändigt med visst expertstöd från databaspersonalen, för att statistiska data och metadata skall bli strukturerade på ett bra sätt, men därefter handlar de centrala insatserna i samband med data/metadainflödet till databaserna mest bara om att ”pricka av” indata och godkänna eventuella förändringar i innehåll och struktur.

### 1.5.7 Några viktiga egenskaper hos Sveriges statistiska databaser

Figurerna 8a–8f sammanfattar några viktiga egenskaper hos Sveriges statistiska databaser.

Figur 8a illustrerar problematiken med att databaserna skall tillgodose många olika användares olika, och ibland motstridiga behov. För en forskare eller en kvalificerad utredare är det t.ex. viktigt att kunna göra egna bearbetningar och analyser, helst på basis av mikro-data, på den egna datorn. Det kräver bl.a. god tillgång till metadata och ingående kvalitetsinformation. För en journalist är det viktigt med snabbhet och extrem användarvänlighet. För en aktör på finans-marknaden är det helt avgörande att ny statistik levereras snabbt och senast samtidigt med att någon annan aktör får tillgång till den. Etc.

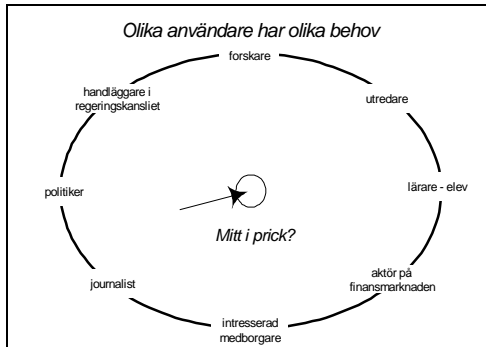
Ett system, som skall tillgodose en så komplicerad och motsägelsefylld kravbild, får inte utformas som en kompromiss mellan de olika kraven. Alla skulle bli missnöjda med en sådan kompromiss. I stället måste man bygga systemet, så att det består av olika delar och gräns-



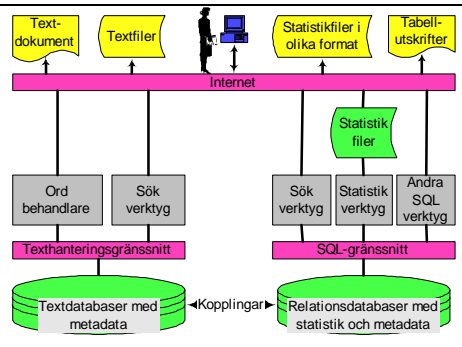
snitt, som i kombination med varandra ger erforderlig mångsidighet och flexibilitet.

Figur 8b visar de viktigaste delarna och gränssnitten i Sveriges statistiska databaser. Tanken är att denna moduluppbyggda arkitektur, byggd på standarder, skall möjliggöra den flexibilitet, som nyss beskrevs.

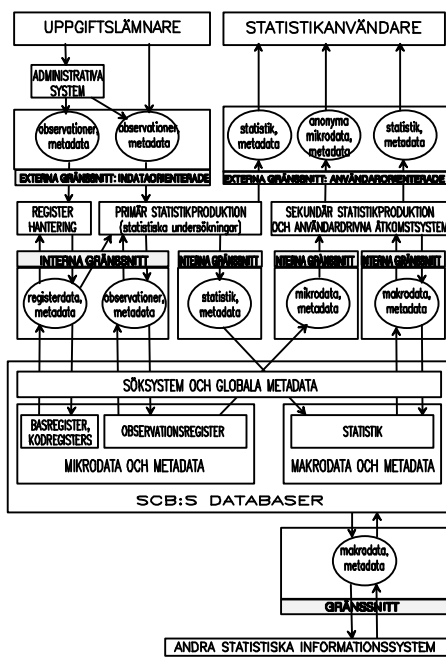
Figur 8e åskådliggör den s.k. servicetrappan, en marknadsföringsfilosofi, som också bidrar till mångsidighet och flexibilitet till gagn för olika kategorier av statistikanvändare. För dem som inte har så gott om pengar, men som kanske har mera gott om tid, finns på steg 1 i servicetrappan självbetjäningstjänster, som har ett mycket lågt pris eller t.o.m. är gratis. På steg 2 återfinns olika standardtjänster som tillhandahålls mot avgifter enligt prislista, och på steg 3 finns skräddarsydda tjänster, som offereras kunder individuellt.



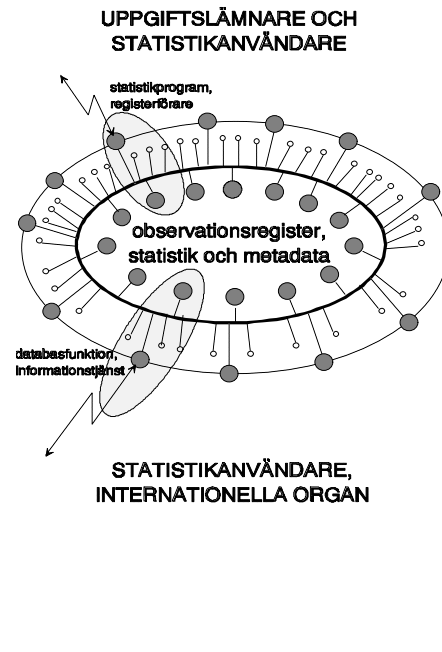
Figur 8a. Hur tillgodose olika behov?



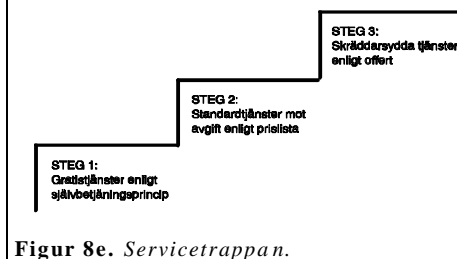
Figur 8b. Sveriges statistiska databaser.



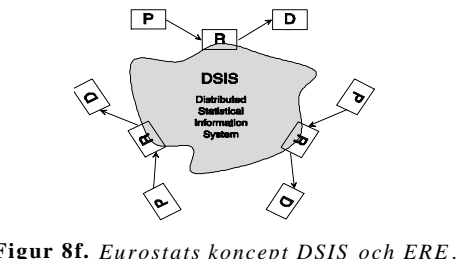
Figur 8c. Processer och gränssnitt.



Figur 8d. Samspelet mellan produktions-systemen och databasservicen.



Figur 8e. Servicetrappan.



Figur 8f. Eurostats koncept DSIS och ERE.

Figur 8d åskådliggör schematiskt<sup>20</sup> hur Sveriges statistiska databaser samspelar med de produktionssystem, som genererar statistiska data och metadata till databaserna, och hur såväl produktionssystemen som spridningsdatabaserna samspelar med omvärlden. Den inre ovalen representerar det inre nätverket hos statistikproducenten<sup>21</sup> med i princip en client/serverkonfiguration per statistikprodukt eller grupp av statistikprodukter (statistikprogram); de små ringarna symboliserar klienter och de större servrar; f.n. finns ett hundratal servrar och ett tusental klienter i detta nätverk.

Den yttre ovalen i figur 8d representerar ett yttre nätverk, som står i förbindelse med omvärlden, dels för indataöverföringar från uppgiftslämnare och administrativa system, dels för spridning av statistik via databasservicen. Mellan det yttre och det inre nätet finns s.k. brandväggar, som säkerställer att obehöriga inte kommer åt det inre nätet. Sveriges statistiska databaser finns i två fysiskt skilda men innehållsmässigt lika kopior, en på en inre server i det inre nätet och en på en yttre server.

Figur 8c illustrerar på en något mera detaljerad nivå de olika delsystemen och gränssnitten i den arkitektur för insamling, produktion och spridning av statistik, som nu håller på att förverkligas.

Figur 8f åskådliggör den motsvarande arkitektur, Distributed Statistical Information Services (DSIS), som EU:s statistikorgan, Eurostat, anbefaller för samspelet mellan nationella statistikbyråer och Eurostat. De nationella statistiksystemen representeras där av (P→R→D)-flöden, där

- P står för "production"
- R står för "reference"
- D står för "dissemination"

De nationella systemen knyts ihop med varandra och med Eurostat via sina respektive referenskomponenter till "the European Reference Environment" (ERE) eller det europeiska statistiska systemet.

Sveriges statistiska databaser utgör den naturliga referenskomponenten i det svenska statistiska systemet och erkänns som sådan av Eurostat.

<sup>20</sup> Den tekniska lösningen ser givetvis inte ut exakt som denna starkt förenklade bild.

<sup>21</sup> Med "statistikproducenten" avses f.n. SCB, men konceptet kan tillämpas även hos andra statistikproducenter och, efter viss vidareutveckling, även på det svenska statistiska systemet som helhet.

## 2 Vad menas med tillgänglighet?

Innan vi ger oss in på att diskutera, vad som kan göras för att öka tillgängligheten till den officiella statistiken, kan det vara på sin plats att kortfattat genomlysja själva begreppet ”tillgänglighet”. Vad betyder det egentligen?

För att officiell statistik skall anses vara tillgänglig, måste man

- kunna lokalisera och komma åt statistiken, på rimlig tid och till rimlig kostnad
- kunna tolka statistiken och använda den för egna bearbetningar och analyser

### 2.1 Fysisk och innehållsmässig tillgänglighet

De två ovan formulerade tillgänglighetskraven svarar mot två huvuddimensioner i tillgänglighetsbegreppet:

- fysisk tillgänglighet
- innehållsmässig tillgänglighet

Ju enklare, snabbare och billigare man kan lokalisera och komma åt statistiken, desto högre grad av fysisk tillgänglighet har den.

Ju enklare, snabbare och säkrare man kan tolka statistiken och bedöma dess kvalitet och användbarhet för de bearbetningar och analyser man vill göra, desto högre grad av innehållsmässig tillgänglighet har den.

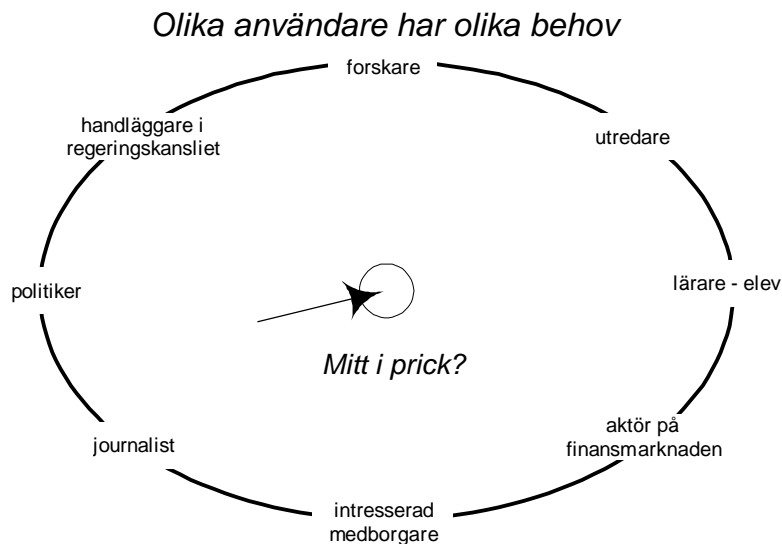
Fysisk tillgänglighet är en nödvändig förutsättning för innehållsmässig tillgänglighet, i den meningen att ett statistiskt datamaterial naturligtvis måste vara lagrat och fysiskt tillgängligt på något sätt, för att det överhuvudtaget skall vara meningsfullt att tala om datamaterialets innehållsmässiga tillgänglighet. Statistikanvändare, t.ex. användare av Sveriges statistiska databaser, uttrycker ofta detta på så sätt, att man kräver ett utökat innehåll i databaserna, varvid man menar att flera statistiska datamaterial från den officiella statistiken borde vara fysiskt och innehållsmässigt tillgängliga i databaserna.

## 2.2 Olika tillgänglighetsprofiler

Tillgänglighet betyder olika saker för olika användare i olika situationer. Annorlunda uttryckt: olika komponenter i tillgänglighetsbegreppet betyder olika mycket för olika användare i olika situationer.

Betrakta figur 9, som är en upprepning av figur 8a. Där finns ett antal typiska kategorier av statistik användare. De prioriterar olika tillgänglighetskomponenter på olika sätt. Exempel:

- För en forskare är det viktigt med en omfattande och djup dokumentation av statistikens innebörd och kvalitet, så att han eller hon kan göra komplicerade analyser och tolka resultaten på ett korrekt sätt. Särskilt intressant för forskaren är god tillgänglighet till väldokumenterade mikrodata. Forskaren kan vilja ompröva de antaganden, som den ursprungliga statistikproducenten gjort, och behöver därför ha detaljerad information om dessa antaganden. För forskaren är det viktigt att det är billigt att komma åt data, och att data är så väl dokumenterade att han eller hon själv kan göra så mycket som möjligt av arbetet, utan att behöva köpa dyra tjänster av statistikproducenten. Tidsåtgången är vanligen av mindre betydelse för forskaren. Däremot är det viktigt för forskaren att kunna arbeta mycket interaktivt med data.
- En handläggare i regeringskansliet, som förbereder en ny reform, har krav på tillgängligheten, som i mångt och mycket liknar forskarens. Dock har handläggaren mycket ont om tid. Däremot är låga kostnader inte lika viktigt som för forskaren.



**Figur 8a. Hur tillgodoses olika behov?**

- En aktör på finansmarknaden har extremt höga krav på att den senaste statistiken skall vara omedelbart tillgänglig, och att han eller hon också uppmärksammas på att de nya uppgifterna kommit. Det handlar ofta om nyckeltal. Deras egentliga innebörd kan vara svår att beskriva, men den är heller inte av något större intresse för denna typ av statistikanvändare, som är mest intresserad av att tydligt se förändringar i förhållande till tidigare uppgifter av samma slag. Det är alltså viktigt att statistiken görs tillgänglig tillsammans med andra relevanta uppgifter, så att relevanta jämförelser lätt kan göras över tiden, mellan länder och i andra avseenden. Ovanliga eller oväntade uppgifter skall helst vara tydligt signalerade. Den här typen av användare har extremt ont om tid men har i gengäld mycket gott om pengar.
- För lärare, elever, journalister, politiker och intresserade medborgare är den pedagogiska tillgängligheten av största betydelse, d.v.s. det skall vara lätt att förstå den huvudsakliga innebörden av statistiken. Även för dessa kategorier av statistikanvändare är ofta jämförelser av stort intresse. För privatpersoner och skolfolk är låga kostnader mycket väsentligt; helst skall statistiken vara gratis. För journalisten är det ovanliga intressant, liksom uppgifter som gör det lätt att identifiera "vinnare" och "förlorare". Politikern vill att det skall vara lätt att sätta informationen i relation till aktuella stridsfrågor i samhällsdebatten.

Det är instruktivt att på detta sätt skissera typiska "tillgänglighetsprofiler" för olika kategorier av användare. Man upptäcker snabbt, att ett system för tillgängliggörande av officiell statistik måste vara mycket mångsidigt och flexibelt. Det duger inte med halvdåliga kompromisser, utan att var och en av de sinsemellan mycket olika användarna skall få känslan av att systemet är skraddarsytt just för honom eller henne. Det är idealet, och det krävs en hel del av systemkonstruktören, för att man skall komma i närheten av detta ideal.

## 2.3 Tillgängliggörandeprocessen ur statistikanvändarens synpunkt

Låt oss analysera steg för steg, delprocess för delprocess, hur en användare skaffar sig och använder sig av statistisk information:

1. Användaren har ett problem eller någon annan typ av arbetsuppgift eller behov, där han eller hon tror eller vet, att statistisk information skulle kunna vara till nytta. Det kan röra sig om några olika typer av behovssituationer:
  - a. Användaren befinner sig i en tämligen unik situation. Användaren vet inte exakt vilken statistisk information, som skulle kunna vara till nytta, men tror ändå att det finns användbar statistik. Användaren behöver alltså hjälp med att ta reda på om det finns någon relevant statistisk information för att belysa en viss frågeställning. Det kan handla om såväl färdigframställd statistik som statistik, som kan framställas på begäran från tillgängliga mikrodata.
  - b. Användaren befinner sig i en återkommande situation och anser sig veta, vilken statistisk information, som han eller hon behöver. Här handlar det om att snabbt och effektivt hitta den informationen.
  - c. Användaren befinner sig i ett pågående förlopp, där han eller hon vill bli uppmärksam på, när ny statistisk information av visst slag föreligger. En aktör på finansmarknaden vill t.ex. få reda på senaste uppgiften om konsumentprisindex, så snart den föreligger. Användaren vill alltså inte själv behöva ta reda på om viss statistisk information föreligger, utan informationskällan, "systemet", skall automatiskt skicka ut informationen till användaren. Om en viss fördröjning kan accepteras, kan automatisk utsändning av elektronisk post vara en lösning på problemet, men ibland är kraven på snabbhet betydligt större, och då måste andra metoder användas.

2. Situationen 1a leder naturligen till en interaktiv process, där användaren successivt, och med informationssystemets hjälp, ringar in ett antal statistiska datamaterial (makrodata och/eller mikrodata), som kan vara av intresse i sammanhanget. Här krävs sökhjälpmedel av olika slag: systematiska innehållsförteckningar, sökmotorer m.m. Det behövs också metadata, som såväl översiktligt som mera i detalj beskriver de tillgängliga statistiska materialen, så att användaren kan avgöra hur användbara de olika datamaterialen är.
3. När användaren identifierat vissa statistiska datamaterial, som han eller hon vill titta på och eventuellt bearbeta vidare, måste datamaterialen och tillhörande metadata lokaliserats och göras fysiskt tillgängliga för användaren. Det fysiska tillgängliggörandet kan ske på några principiellt olika sätt:
  - a. Data och metadata överförs till användarens egen dator, där han eller hon kan bearbeta dem vidare på egen hand. Här är det viktigt att data och metadata kan tillgängliggöras i sådan form, att användaren kan använda dem på sin egen dator med sina egna program.
  - b. Data och metadata ligger kvar hos producenten, och användaren kan titta på dem och eventuellt bearbeta dem vidare där. Eventuellt kan vissa slutresultat överföras till användaren, t.ex. för utskrift. Här är det viktigt att producenten kan erbjuda en mångfald av användarvänliga och prisvärda bearbetningsmöjligheter, och att resultaten av de centralt utförda bearbetningarna kan tillgängliggöras för användaren på ett ändamålsenligt sätt. De flesta användare och användningar kräver idag, att resultaten tillgängliggörs i elektronisk form, så att de lätt kan kombineras och integreras med annan information, som användaren använder sig av.
  - c. Man kan också tänka sig olika kombinationer av a och b, där vissa bearbetningar kan göras centralt, medan andra kan göras lokalt hos användaren.
4. Därefter måste användaren kunna tolka den framtagna statistiken. Nu handlar det om innehållsmässig tillgänglighet. Återigen kan några olika typiska situationer föreligga:
  - a. Om det rör sig om statistiska data, som är helt nya för användaren, kan han eller hon behöva åtskillig hjälp med att tolka data. Om användaren är okunnig och hjälpbehoven komplicerade, kan det vara nödvändigt, att en mänsklig expert finns att tillgå, men allt oftare kräver statistikanvändarna, att det skall finnas ett datoriserat stöd i form av beskrivningar och förklaringar, d.v.s. metadata, och verktyg av olika slag, t.ex. dynamiska länkar och s.k. expertsystem, som gör det möjligt för användaren själv att snabbt och kostnadseffektivt tolka data. Hur avancerade metadata och verktyg, som behövs, är naturligt-



- vis också beroende av hur komplicerad den tilltänkta användningen är. En forskare behöver naturligtvis mycket mera ingående meta-information än en journalist. Journalisten kan å andra sidan behöva få metainformationen presenterad på ett mera pedagogiskt sätt, än vad som är nödvändigt för forskaren.
- b. Om det rör sig om statistisk information av ett slag, som användaren är väl förtrogen med sedan tidigare, d.v.s. ett mer eller mindre regelbundet återkommande informationsbehov, behöver användaren inte alltid ha tillgång till så mycket metainformation. Däremot är det i en sådan situation ofta viktigt för användaren att bli uppmärksam på viktiga förändringar i förhållande till tidigare perioder, jämförelser med andra regioner, branscher eller liknande. Det är också viktigt för användaren att bli uppmärksam på fel och störningar av olika slag, som gör att föreliggande data på ett eller annat slag avviker från "det normala".
  - c. I ännu mera strukturerade situationer, t.ex. hos aktören på finansmarknaden, där användaren fortlöpande använder sig av vissa väl kända nyckeltal, är det inte så mycket fråga om att förstå den exakta innebörden av informationen som att snabbt bli uppmärksam på oväntade mönster i data, vilka i sin tur kräver handling av data-användaren.
5. Vissa användare nöjer sig inte med att bara tolka de statistiska data, som han eller hon fått fram från det statistikproducerande systemet. Det kan ofta vara aktuellt att göra egna, mer eller mindre djupgående analyser, ibland genom att kombinera statistik från olika källor med varandra och med egna data av olika slag. Härvidlag krävs, att det är möjligt att tekniskt och innehållsmässigt jämföra data från olika källor med varandra, vilket i sin tur kan kräva ganska ingående kunskaper om de olika datamaterialens kvalitet.
  6. Resultaten av tolkningar och analyser sammanställs till någon form av utredning eller beslutsunderlag, som i sin tur används i någon verksamhet.
  7. Tolkning, analys och användning av statistiska data leder ofta till följdfrågor, vilka i sin tur leder till upprepningar av stegen 1 – 4 ovan. Användaren kan också vara intresserad av att få tips om liknande analyser, som gjorts av andra användare.

I tabell 10 sammanfattas "tillgängliggörandeprocessen" för statistiska data, sedd ur användarens synpunkt, och där anges några viktiga, tillgänglighetsskapande hjälpmedel, som behövs i de olika delprocesserna.

Delprocess	Tillgänglighetsskapande hjälpmedel och åtgärder
Användaren behöver statistisk information i sin verksamhet och identifierar den. (Steg 1 och 2.)	<ul style="list-style-type: none"> <li>All statistik och alla slutliga observationsregister i den officiella statistiken är tillgängliga genom en databastjänst.</li> <li>Sökmetoderna är intuitiva, lättlärd (cf Internet, Windows), och enhetliga, oavsett var data ligger fysiskt lagrade.</li> </ul>
<ul style="list-style-type: none"> <li>Unik situation. Finns det några relevanta statistiska datamaterial?</li> </ul>	<p>Sökhjälpmedel som associerar problem eller frågeställning med potentiellt relevanta statistiska datamaterial (statistik eller observationsregister). Exempel:</p> <ul style="list-style-type: none"> <li>systematiska innehållsförteckningar för tillgängliga statistiska datamaterial "var de än finns"</li> <li>sökmotorer baserade på innehållsbeskrivningar</li> </ul> <p>Beskrivningar som hjälper användaren att bedöma om ett föreslaget datamaterial är av intresse eller ej.</p>
Återkommande situation. Användaren vet vilka statistiska datamaterial som är relevanta.	Sökhjälpmedel som gör det möjligt för användaren att på ett enkelt sätt peka ut de statistiska datamaterial, som han eller hon vill ha tag i.
<ul style="list-style-type: none"> <li>Kontinuerligt pågående verksamhet där användaren vill ha ny statistisk information så snart den föreligger.</li> </ul>	<p>Verktyg som skickar prenumererad statistisk information till prenumererande användare, så snart den föreligger:</p> <ul style="list-style-type: none"> <li>med elektronisk post (måttliga krav på snabbhet)</li> <li>med s.k. "data shooting" (höga krav på snabbhet)</li> </ul>
<p><b>Identifierade statistiska material lokaliseras och görs fysiskt tillgängliga för användaren. (Steg 3.)</b></p>	<ul style="list-style-type: none"> <li>Lätt att "hämta hem" statistik och andra tillgängliga till den egna miljön.</li> <li>Metadata följer med data, så att användaren lätt kan arbeta vidare med data med egna program.</li> </ul>
<ul style="list-style-type: none"> <li>Data och metadata överförs till statistikanvändarens egen dator.</li> </ul>	<ul style="list-style-type: none"> <li>Standardiserade lagringsformat och effektiva överföringsprocedurer för data och metadata.</li> <li>Verktyg som gör det möjligt för användaren att transformera data och metadata för effektiv användning i egna databaser och program.</li> </ul>
<ul style="list-style-type: none"> <li>Data och metadata ligger kvar hos producenten.</li> </ul>	<ul style="list-style-type: none"> <li>Verktyg som gör det möjligt för användare att fjärrstyra bearbetningar av datamaterialen hos producenten, på ett sätt som så långt som möjligt ger användaren samma möjligheter, som om data och metadata låg lokalt hos användaren.</li> <li>Verktyg som gör det möjligt för användare att få resultaten av de fjärrstyrda bearbetningarna överförda till den egna lokala miljön på ett sådant sätt att de där kan presenteras och bearbetas vidare enligt användarens önskemål.</li> </ul>
<p><b>Användaren tolkar framtagen statistik och presenterar tolkningar och analyser. (Steg 4 och 6.)</b></p>	<ul style="list-style-type: none"> <li>Fullständiga beskrivningar av innehåll och kvalitet</li> <li>Pedagogiskt utformade översikter: språk, grafik m.m.</li> <li>Varningar om begränsningar och fallgropar.</li> <li>Mer hjälp genom länkar och expertsystem.</li> </ul>

<ul style="list-style-type: none"> <li>• Komplicerad användning, t.ex. forskning, utredning.</li> </ul>	<ul style="list-style-type: none"> <li>• Både översiktlig och detaljerad dokumentation av alla förhållanden, som är av betydelse för statistiskt korrekt användning och analys, och som gör det möjligt för användaren att ompröva producentens antaganden.</li> <li>• Verktyg, t.ex. dynamiska länkar, som gör det möjligt att "gräva djupare" i data och metadata vid behov.</li> <li>• Expertsystem som stöd för tolkning och analys</li> <li>• Experthjälp hos producenten som "sista utväg".</li> </ul>
<ul style="list-style-type: none"> <li>• Pedagogisk och journalistisk användning.</li> </ul>	<ul style="list-style-type: none"> <li>• Presentationsmetoder och översiktliga beskrivningar, som gör det möjligt för användaren att snabbt förstå och tolka statistiken på ett rimligt sätt.</li> </ul>
<ul style="list-style-type: none"> <li>• Återkommande analyser av likartat slag.</li> </ul>	<p>Information, presentationsmetoder och verktyg, som</p> <ul style="list-style-type: none"> <li>• framhäver förändringar och skillnader, och som ger underlag för att tolka och förklara dessa</li> <li>• underlättar jämförelser av olika slag, t.ex. över tiden, mellan regioner, mellan branscher.</li> </ul>
<ul style="list-style-type: none"> <li>• Handlingsinriktad användning på basis av nyckeltal och signaler av olika slag.</li> </ul>	<p>Information, presentationsmetoder och verktyg, som</p> <ul style="list-style-type: none"> <li>• enkelt och tydligt framhäver förändringar och ovanligheter</li> <li>• passar väl ihop med användarens beslutsmodeller och beslutsverktyg, och som</li> <li>• ger underlag för ett snabbt agerande</li> </ul>
<p><b>Användning i kombination med statistiska och andra data från egna och andras källor. (Steg 5.)</b></p>	<p>Information och verktyg som underlättar</p> <ul style="list-style-type: none"> <li>• bedömningen av olika datas jämförbarhet</li> <li>• jämförelser mellan olika data från olika källor</li> </ul>
<p>Uppföljning och fördjupning av gjorda analyser.</p>	<p>Information och verktyg som hjälper användaren att</p> <ul style="list-style-type: none"> <li>• minnas vad han eller hon gjort tidigare</li> <li>• upprepa "liknande" bearbetningar och analyser på ett förenklat sätt jämfört med att behöva definiera dem helt från grunden igen</li> <li>• få referenser till andra användares analyser av liknande frågeställningar</li> </ul>
<p><b>Hela processen. (Steg 1 – 7.)</b></p>	<p>Åtgärder för att bemöta vanligt förekommande kritik mot statistikansvariga för att vara alltför restriktiva, dyra, förmyndaraktiga och monopoliserande.</p>
<ul style="list-style-type: none"> <li>• Effektiv hantering av statistiksekretessen.</li> </ul>	<ul style="list-style-type: none"> <li>• Kostnadseffektiva och säkra metoder och verktyg, som så lite som möjligt inskränker tillgängligheten till datamaterialen för statistiska bearbetningar enligt steg 1– 7 ovan.</li> <li>• Metoder som fokuserar på den kombinerade effekten av tekniska, metodologiska, legala och administrativa åtgärder.</li> <li>• Lagstiftning som tydligare framhåller den statistikansvariges plikt att tillgängliggöra datamaterial, som ingår den officiella statistiken.</li> </ul>

<ul style="list-style-type: none"> <li>• Rimliga kostnader. Möjlighet att välja passande pris/prestandanivå.</li> </ul>	<ul style="list-style-type: none"> <li>• Produktion av officiell statistikanslagsfinansieras, liksom infrastrukturen för tillgängliggörandet. Endast kostnader därutöver avgiftsfinansieras (marginalkostnadsprincip).</li> <li>• Avgiftsfinansierade tilläggstjänster paketeras så att användaren inte behöver köpa en mera omfattande tjänst än vad han eller hon vill.</li> </ul>
<ul style="list-style-type: none"> <li>• Kunskapen om datamaterialen dokumenterad och tillgänglig. Inget otillbörligt kunskapsmonopolhos statistikansvarig eller producent.</li> </ul>	<ul style="list-style-type: none"> <li>• Lagstiftning som tydligare framhåller den statistikansvariges plikt att dokumentera datamaterial och processer, som ingår den officiella statistiken.</li> <li>• Kostnaderna för dokumentationsarbetet skall anses ingå i anslaget för produktion av officiell statistik</li> <li>• Dokumentationen möjliggör självbetjäning</li> </ul>

**Figur 10. Översikt över krav på tillgänglighet och tillgänglighetsskapande åtgärder och hjälpmedel i olika delar av tillgängliggörandeprocessen, sedda ur användarsynpunkt.**

## 2.4 Önskemål och klagomål beträffande den officiella statistikens tillgänglighet

Av diskussionerna i kapitel 1 och 2.1 – 2.3 framgår att användarna av officiell statistik bl.a. har framfört följande önskemål och klagomål beträffande den officiella statistikens tillgänglighet:

1. All officiell statistik bör vara lätt tillgänglig på ett enhetligt sätt, oavsett vem som är statistikansvarig. Det skall gå snabbt och lätt att identifiera och lokalisera statistik och observationsregister som är relevanta i en viss användningssituation.
3. Alla statistik användare, även oerfarna och tillfälliga, skall kunna tolka statistiken och använda sig av den på ett korrekt sätt.
4. Kvalificerade användare skall kunna analysera och vidarebearbeta statistiska datamaterial självständigt och på egen hand, utan att vara tvungna att konsultera eventuella experter hos producenten. All erforderlig kunskap om statistik och bakomliggande data och processer, som man kan behöva för att använda de statistiska datamaterialen på ett statistiskt ansvarsfullt sätt i olika situationer, måste därför vara väl dokumenterad och lätt tillgänglig för användarna.
5. Tillgängligheten till statistiska data skall inte inskränkas mer än vad som är absolut nödvändigt för att uppnå lagstadgad statistiksekretess.
6. Ett antal bastjänster, som är nödvändiga för att tillgängliggöra den officiella statistiken skall tillhandahållas till låg marginalkostnad. Bastjänsterna skall vara tillräckliga för de användare, som är villiga att göra en del arbete själva. Därutöver skall det finnas en konkur-

rensutsatt marknad, där statistikanvändaren kan köpa olika tilläggs-tjänster, om han eller hon så önskar.

## 2.5 Åtgärdsområden

För att komma till rätta med ovanstående önskemål och klagomål krävs aktiva åtgärder av olika slag. De faller inom följande åtgärdsområden:

1. Kunskapsbaserad dokumentation
2. Teknik och presentation
3. Ekonomi: finansiering och prissättning
4. Effektivare hantering av statistiksekretessen
5. Lagar och regler
6. Organisationsfrågor

Matrisen i figur 11 åskådliggör måldimensionen horisontellt och åtgärdsdimensionen vertikalt. De viktigaste kombinationerna av möjliga åtgärder och effekter är markerade med "x" i respektive celler i matriserna. Dessa kommer att diskuteras mera i detalj i kapitel 3 och 4.

Åtgärds- område	Mål				
	1. Lätt att söka och komma åt all officiell statistik, som är relevant för ett visst problem	2. Lätt även för icke-specialister att tolka officiell statistik och använda sig av den på rätt sätt	3. Möjligt för kvalificerade statistik-användare att använda den officiella statistikens datamaterial på egen hand	4. Tillgängligheten till den officiella statistikens datamaterial inskränks inte mer än nödvändigt för att säkra statistik-sekretessen	5. Möjligt att komma åt den officiella statistiken genom bastjänster till låga priser i kombination med valfria tilläggstjänster
<i>DOK:</i> <i>Kunskaps-</i> <i>baseraddoku-</i> <i>mentation</i>	x	x	x		
<i>TEK:</i> <i>Teknik och</i> <i>presentation</i>	x	x	x	x	x
<i>EKO:</i> <i>Ekonomi:</i> <i>finansiering</i> <i>och prissätt-</i> <i>ning</i>	x		x		x
<i>SEK:</i> <i>Hantering av</i> <i>statistik-</i> <i>sekretessen</i>			x	x	
<i>LAG:</i> <i>Lagar och</i> <i>regler</i>	x		x	x	
<i>ORG:</i> <i>Organisation</i>	x		x	x	

Figur 11. Matris för att ange åtgärder för att bättre tillgodose vanligt förekommande önskemål och klagomål beträffande den officiella statistikens tillgänglighet.

## 2.6 Tillgänglighetens plats i ”det goda statistiska systemet”

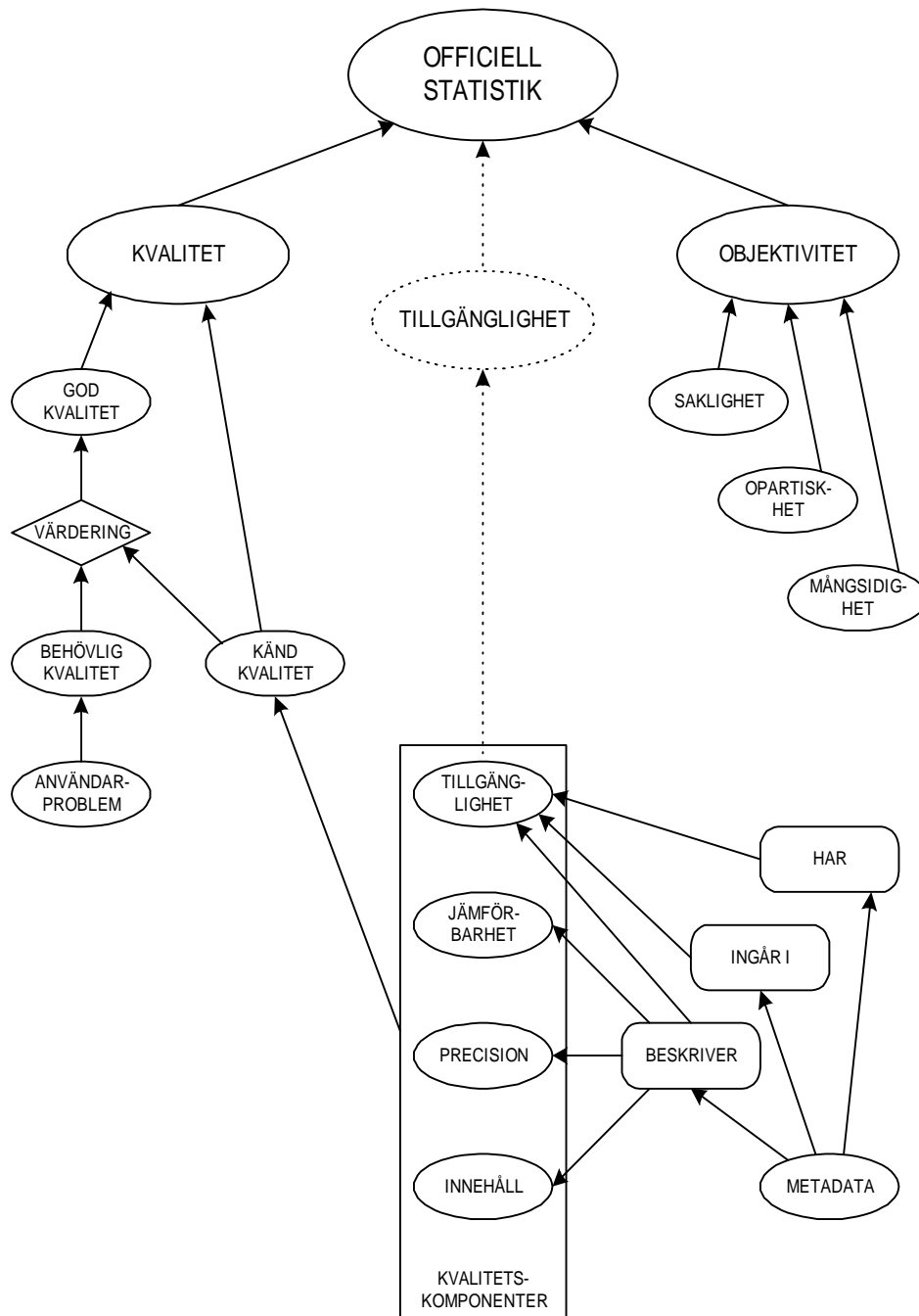
Begreppet ”tillgänglighet” har nära och delvis lite komplicerade relationer till andra centrala begrepp och värden i ett system för officiell statistik. Se figur 12.

Enligt SCB:s standard för kvalitetsdeklarationer,<sup>22</sup> är ”tillgänglighet” en av fyra huvuddimensioner (komponenter) i kvalitetsbegreppet.

Alla kvalitetskomponenter beskrivs med hjälp av metadata. Dessa metadata ingår i tillgänglighetsdimensionen för den statistik, som de beskriver. Dessutom har de själva en viss fysisk och innehållsmässig tillgänglighet.

Begreppet ”kvalitet” används på ett par principiellt olikartade sätt i dagligt tal. Å ena sidan frågar vi oss vilken kvalitet en viss produkt har. En kvalitet (= egenskap) behöver då inte nödvändigtvis vara ”bättre” eller ”sämre” än en annan, men olika kvaliteter passar för olika ändamål, och för en köpare gäller det att jämföra produktens faktiska kvaliteter (= egenskaper) med de kvaliteter (= egenskaper) som behövs i den tänkta användningen av produkten. I andra sammanhang talar vi om kvalitet som liktydigt med ”god kvalitet” i en mera absolut mening, oberoende av användning. Man kan t.ex. kvalitetsdeklarera en produkt på en skala från 1 till 5, där 1 är ”mycket dålig” och 5 är ”mycket bra”.

<sup>22</sup> Standarden är utförligt beskriven och exemplifierad i *”Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik”*, SCB, Meddelanden i samordningsfrågor (MIS) 1994:3, utarbetad av Chris Denell, Eva Elvers och Bengt Rosén, samtliga vid SCB.



Figur 12. Tillgänglighetens plats i "det goda systemet" för officiell statistik.



När det gäller statistik, så är det ytterst användaren i användningssituationen, som avgör om statistikens kvalitet är bra, eller snarare ändamålsenlig, för användningen ifråga. Statistik som är lämplig för en användning kan vara olämplig för en annan. Bedömningen går i princip till så, att användaren jämför statistikens faktiska (kända) kvaliteter (egenskaper) enligt metadata i kvalitetsdeklarationen med önskvärda kvaliteter (egenskaper) hos statistiken enligt användarens egen värdering av vad förhandenvarande problem eller frågeställning kräver. Både den av användaren värderade kvaliteten och den faktiska, i kvalitetsdeklarationen beskrivna kvaliteten ingår i kvalitetsbegreppet för officiell statistik.

Tillgängligheten är en så fundamental egenskap hos den officiella statistiken, att man kan fråga sig, om den inte borde vara en helt fristående värderingsdimension vid sidan av (snarare än som en del av) kvaliteten. Detta är markerat med en streckad linje i figur 12.

Redan idag behandlar lagstiftningen<sup>23</sup> såväl "tillgänglighet" som "objektivitet" som fristående dimensioner i kraven på officiell statistik.

Ett väl fungerande system för officiell statistik skall stödja

- rationella beslut i samhälle, näringsliv och privatliv
- demokratiskt beslutsfattande

Detta sker genom att systemet för officiell statistik producerar och tillgängliggör beslutsunderlag som präglas av

- objektivitet (saklighet, opartiskhet, mångsidighet)
- kvalitet (väl beskriven och ändamålsenlig för många, viktiga användningar)
- god tillgänglighet för alla potentiella användare

Ett bra statistiskt system värnar också om

- statistiksekretessen
- uppgiftslämnarförtroendet
- effektiviteten

Statistiksekretess och uppgiftslämnarförtroende har en positiv påverkan på statistikens kvalitet. Statistiksekretessen har dock även en negativ inverkan på tillgängligheten, som också är en kvalitetsaspekt.

God effektivitet i det statistiska systemet har positiv inverkan på flera kvalitetsaspekter, t.ex.

- relevansen: mera statistik kan produceras för samma summa pengar, vilket ökar chanserna att relevant statistik finns att tillgå, när den behövs

<sup>23</sup> Lagen om den officiella statistiken, 2§.

- noggrannheten: effektivare urvalsdesign, bortfallsförebyggande åtgärder, bortfallshantering och kontroller av primärdata ger bättre precision i statistiken
- tillgängligheten: statistiken produceras, publiceras och sprids snabbare och säkrare och blir därmed aktuellare och punktligare

## 3 Hur kan man öka tillgängligheten?

En hel del har redan gjorts för att komma till rätta med de inledningsvis beskrivna problemen kring den officiella statistikens tillgänglighet, men en hel del återstår också att göra. I den fortsatta framställningen behandlas några strategiskt viktiga områden:

- Kunskapsbaserad dokumentation
- Teknik och presentation
- Ekonomi: finansiering och prissättning
- Statistiksekretessen
- Lagar och regler
- Organisationsfrågor

### 3.1 Kunskapsbaserad dokumentation

Den som vill kunna

- hitta relevant statistik för ett visst problem
- förstå vad statistiken betyder och bedöma dess kvalitet
- analysera statistiken vidare

måste besitta eller kunna skaffa sig kunskaper av olika slag. Omfattningen och djupet av de erforderliga kunskaperna är beroende av vad statistiken skall användas till.

#### 3.1.1 Behovet av dokumenterad kunskap vid användning av statistiska data

För att hitta statistik, som kan vara relevant för ett visst problem, måste den potentielle statistikanvändaren ha överblick över vilken statistik som finns tillgänglig. Om vi begränsar oss till svensk officiell statistik är detta problem relativt hanterbart. Det torde räcka med att ha tillgång till en uppräknig av de ca 200 statistikprodukter fördelade på 22 ämnesområden, som f.n. utgör den officiella statistiken, tillsammans med översiktliga beskrivningar av varje statistikprodukt av ungefär den omfattning som ingår i Produktbeskrivningarna i Sveriges statistiska databaser. Söksystem av olika slag, t.ex. strukturerade, menystyrda söksystem och fritextsökningssystem, underlättar för användaren.

För att rätt tolka den statistik man funnit och för att förstå vad den betyder, vilka slutsatser man kan dra av den, och vilka slutsatser man inte kan dra, behövs mera kunskap. Exempelvis måste man förstå innebörden av de begrepp (statistiska storheter: objekttyper, populationer, variabler, mätförfaranden, mått), som statistiken är uppbyggd av. Man måste också ha en viss förståelse för statistiska metoder och kvalitetsbegrepp och kunskap om den faktiska kvaliteten hos den statistik man försöker tolka, jämfört med den kvalitet, som den tilltänkta användningen kräver.

Om de produktbeskrivningar och kvalitetsdeklarationer som obligatoriskt skall åtfölja officiell statistik<sup>24</sup> är välgjorda, räcker de en bra bit för att ge statistikanvändaren den kunskap om statistikens innehåll och kvalitet som är nödvändig, när man skall tolka statistiken och använda sig av den som informationsunderlag för problemlösning och beslutsfattande.

För mera djupgående analyser av officiell statistik räcker det inte med att förstå begreppsdefinitioner och att ta del av kvalitetsdeklarationer för de aggregerade slutprodukterna, själva statistiken. Man också ha en hel del kunskap om de underliggande datamaterialen och de processer, genom vilka data har uppkommit, samlats in, preparerats och bearbetats. Man måste kort sagt veta, hur bakomliggande statistiska undersökningar och administrativa processer är utformade, och hur de faktiskt genomförts. Full kunskap om dessa saker har i bästa fall några få personer, de som deltagit i utformningen och genomförandet av ifrågavarande processer, om ens de. Efter viss tid har med stor sannolikhet även denna kunskap förflyktigats, och man riskerar betydande informationskapitalförstöring, om inte kunskapen dessförinnan dokumenterats på ett personoberoende sätt.

De största kraven på kunskaper om de data och processer, som ligger bakom den officiella statistiken, föreligger när en användare vill återanvända gamla observationsregister från en eller flera undersökningar för att bearbeta och kombinera dessa data på nya sätt. Det är denna typ av återanvändning, som motiverar, att vi överhuvudtaget arkiverar gamla data. Lite tillspetsat kan man säga att all arkivering av data är bortkastad och inte borde göras, om man inte säkerställer, att den kunskap, som finns i huvudena hos dem, som en gång planerade och

<sup>24</sup> Regleras genom "Föreskrifter och allmänna råd om offentliggörande av officiell statistik", Meddelanden i samordningsfrågor 1995:5, SCB. Se särskilt 4-5 §§.

genomförde undersökningarna, också dokumenteras och lagras. Detta är en uppgift som är långtifrån trivial.<sup>25</sup>

En stor del av de grunddata, som den officiella statistiken i Sverige grundas på, närmare bestämt mer än 95%, genereras inte av statistiska undersökningar i traditionell mening utan av administrativa processer. Statistikerna som producerar den officiella statistiken har sällan den förstahandskunskap om de administrativa processerna, som de har om de mätförfaranden och datainsamlingsprocesser, som ingår i mera traditionellt utformade statistiska undersökningar. Det är sålunda en ännu mera komplicerad uppgift att dokumentera administrativa processer och av dessa genererade administrativa data, som även används för statistiska ändamål. Ofta kombineras dessutom administrativa data och data som genererats genom regelrätta statistiska undersökningar.

Kvalificerade statistikanvändare, som forskare och professionella utredare, kräver idag en tillgänglighet till den officiella statistiken, som innebär, att de själva kan genomföra självständiga analyser av såväl publicerad statistik som bakomliggande observationsregister, utan att de nödvändigtvis måste söka reda på och rådfråga de statistikproducenter, om dessa överhuvudtaget står att återfinna, som genomfört de bakomliggande statistiska undersökningarna. Denna kategori av statistikanvändare vill också numera utnyttja de tekniska möjligheterna att bearbeta och analysera statistiska data på ett mycket interaktivt sätt utan att behöva hamna i byråkratiska processer av olika slag mellan de olika iterationerna. De vill själva kunna formulera nya frågor till data-materialen, allteftersom de får svar på tidigare frågor. De vill kunna formulera hypoteser, testa dessa, omformulera hypoteserna, bygga modeller, teorier, etc. De vill veta, vilka antaganden och modeller, som den ursprungliga producenten utgått ifrån, så att de själva kan välja, om de vill acceptera dessa förutsättningar, eller om de vill formulera sina egna, alternativa antaganden och modeller som grund för (åter) användningen av de statistiska grundmaterialen.

Man kan mycket grovt särskilja två motsatta attityder till den ovan antydda problematiken kring kunskap och tillgänglighet:

1. Bara statistikproducenten har erforderlig kunskap för att rätt tolka och analysera statistiken. Alla statistikanvändare bör därför vända sig till statistikproducenten med sina frågor och uppdrag. Endast med stor försiktighet kan statistiken överhuvudtaget användas för andra syften än de för vilka den i första hand framtagits.
2. Statistiken skall vara så väl dokumenterad att "vem som helst" kan använda den på egen hand, för sina egna syften.

<sup>25</sup> Se Rosén, Bengt & Sundgren, Bo: " *Dokumentation för återanvändning av mikromaterial från SCB:s undersökningar*", Stockholm 1991.

Det säger sig självt, att statistikproducenter har en tendens att inta attityder liknande (1) ovan, medan kvalificerade statistikanvändare hävdar ståndpunkter liknande (2). Statistikproducenten har ett visst övertag i denna dragkamp, eftersom statistikproducenten har ett kunskapsmonopol *de facto*; det är bara statistikproducenten, som har direktkunskap om statistikens innehåll och kvalitet, liksom om de bakomliggande data-materialen och processerna. Det krävs aktiva och kvalificerade arbetsinsatser av statistikproducenten för att på ett bra sätt dokumentera denna kunskap och göra den tillgänglig för statistikanvändarna.

En inte alldeles ovanlig replik från statistikproducenten i detta sammanhang är, att man i och för sig mycket gärna skulle vilja göra dessa insatser för att dokumentera och dela med sig av sina kunskaper, men att man tyvärr inte har "resurser" för att göra det. Ett förslag i detta sammanhang kunde vara, att statsmakterna i samband med att man beviljar anslag för produktion av officiell statistik mycket tydligt klargör, att det är en prioriterad uppgift, att inom ramen för anslaget genomföra ett gediget dokumenteringsarbete enligt standarder; dessa standarder måste då också fastställas, och de måste ligga på en ambitionsnivå motsvarande den i avsnitt 1.3 behandlade SCBDOK-modellen.

Statistikproducentens "kunskapsmonopol" är sålunda i praktiken ett ganska allvarligt hot mot statistikens tillgänglighet.

### 3.1.2 Kvalitetsfrågor

Statistikkvalitetens komplexitet anförs inte sällan (företrädesvis av statistikproducenter) som ett argument mot att göra statistiken alltför lättillgänglig. Bara den som har producerat statistiken anses rätt kunna uttolka och analysera den. Innan en statistikanvändare tillåts ta del av en statistikuppgift, bör han eller hon åtminstone tvingas ta del av en hel del varningar och s.k. metodbeskrivningar. Kort sagt: det finns ibland en viss förmyndarmentalitet hos statistikproducenten.

Å andra sidan finns naturligtvis också de som tar alltför lättvindigt på kvalitetsfrågorna. En representant för Eurostat utbrast nyligen ungefär följande: Vi på Eurostat tror inte på värdet av att statistiken åtföljs av en massa förklaringar, fotnoter och s.k. metadata. Våra användare på kommissionen vill bara ha siffrorna, de är inte intresserade av det andra, det bara förvillar dem.

En bra medelväg finns kanske antydd i följande slagord från Nederländernas statistiska centralbyrå: "Use tools, not rules" – "Använd verktyg, inte regler" i något mindre poetisk svensk översättning.

I stället för att förhindra en statistikanvändare att ta del av en statistikuppgift, innan han eller hon tagit del av – eller åtminstone klickat sig

förbi – en varningstext, så kan en statistikproducent, som lever enligt maximen ”use tools, not rules”, t.ex. förse statistiken på Internet med klickbara länkar, som gör det möjligt för en intresserad användare att få en relevant förklaring till snart sagt allt, inklusive de enskilda siffrorna, som finns på webbplatsen. En sådan kvalitetsmedvetenhetsstrategi bygger på användarnas naturliga nyfikenhet snarare än på producentens förmyndarmentalitet och kan säkert vara väl så effektiv.

Officiell statistik i allmänhet, och officiell statistik på Internet i synnerhet, är av s.k. general purpose karaktär, d.v.s. statistikproducenten måste räkna med att den används av många olika typer av användare och för många olika syften. Användningarna kan t.o.m. ligga årtionden fram i tiden, när samhället och användarnas referensramar är helt annorlunda jämfört med hur det var när statistiken producerades.

Alla som funderar över begreppet statistikkvalitet kommer snart till insikten att det är ett komplicerat begrepp. I dagligt tal menar man med kvalitet ofta ”god kvalitet”, och man ”varudeklarerar” en produkts kvalitet genom att poängsätta den i ett antal dimensioner (t.ex. 1=dålig, ..., 5=bra). Begreppet statistikkvalitet innehåller också dimensioner eller aspekter, som kan värderas ungefär på detta sätt; man kan t.ex. ange urvalsstorlekar, bortfallsandelar, varianser, och i en någorlunda absolut och objektiv mening kan man därvid anse att stora urval, små bortfallsandelar och små varianser innebär ”bättre kvalitet” än motsatsen.

Begreppet statistikkvalitet består emellertid också av dimensioner eller aspekter, som inte alls låter sig kvantifieras eller ens på något annat sätt värderas i en absolut och objektiv mening. Det gäller t.ex. s.k. innehålls- eller relevansaspekter.

En statistikuppgift kan vara mer eller mindre relevant för en användning och inte alls relevant för en annan. Ytterst är det användaren själv, i en viss användningssituation, som måste avgöra vilken statistik, som är relevant för den användningen.

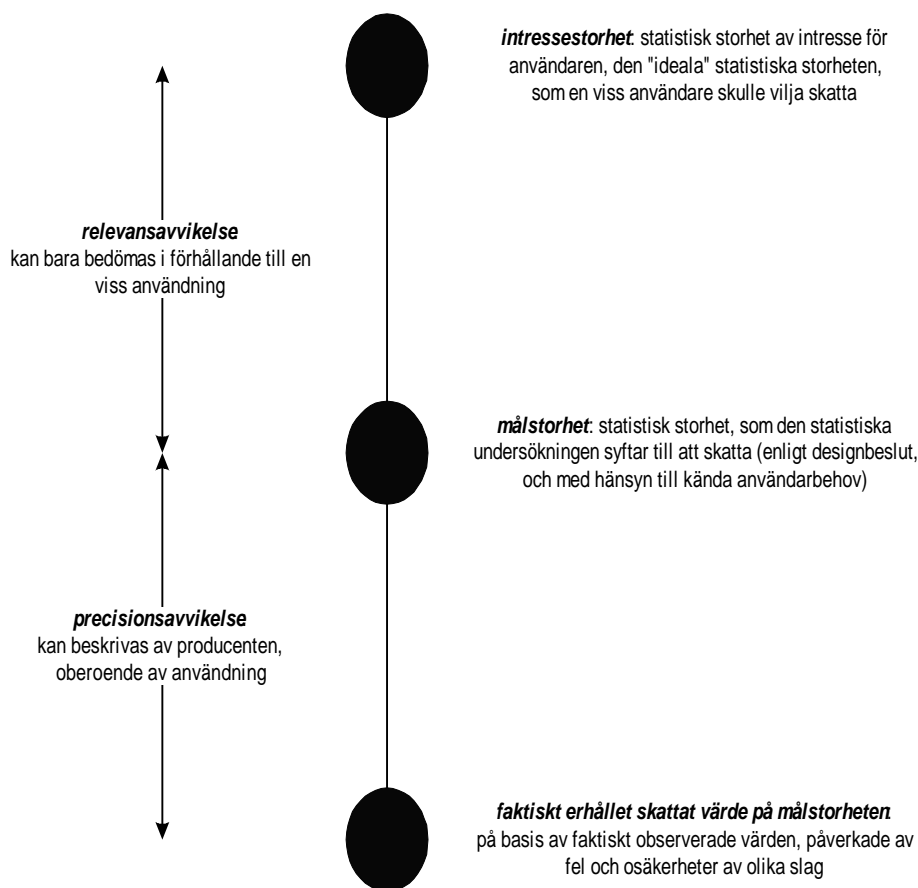
Statistikproducenten kan på olika sätt underlätta för användaren att bedöma statistikens relevans för olika ändamål. I första hand gäller det för statistikproducenten att informera om vad statistiken betyder, d.v.s. hur de ingående statistiska storheterna definieras, hur mätningar och observationer utförts, under vilka förutsättningar, som skattningarna av de statistiska storheterna beräknats, m.m. Användaren kan då jämföra

- de *faktiskt föreliggande skattningarna* av de av statistikproducenten *faktiskt begagnade statistiska storheterna*; med
- de *ideala skattningar av ideala statistiska storheter*, som användaren egentligen skulle behöva för att på ett helt relevant sätt belysa användarens föreliggande frågeställning.

Figur 12 illustrerar kvalitetsproblematiken. Den faktiskt föreliggande statistikens egenskaper skiljer sig från den ideala statistik, som använ-

daren skulle vilja ha tillgång till i en viss användningssituation, för ett visst ändamål, i en rad olika avseenden eller kvalitetsdimensioner. I figuren har kvalitetsdimensionerna (och därmed avvikelserna mellan den faktiskt föreliggande och den önskade statistiken) grovt sorterats i två nivåer:

- (i) innehålls- eller relevansnivån
- (ii) noggrannhets-, tillförlitlighets- eller precisionsnivån



Figur 12. Kvaliteten hos statistiska data, sedd ur användarens synpunkt.

Vad gäller statistikens noggrannhet eller precision kan statistikproducenten ge ganska detaljerad och exakt information, oberoende av vad statistiken skall användas till. Statistiker brukar oftare tala om osäkerhet och fel i statistiken än om noggrannhet och precision, och faktum är, att det lättare att ange bristerna i statistiken och att uppskatta dessa bristers betydelse för de framtagna skattningarna av statistiska storheter än att



mera direkt och "positivt" beskriva statistikens kvalitet i dessa avseenden.

När det gäller statistikens relevans kan statistikproducenten bara tillhandahålla vissa pusselbitar, som har med definitioner, mätmetoder och förutsättningar att göra. Användaren måste själv kombinera denna producentorienterade information med sin egen kännedom om det föreliggande statistikändamålets behov. De storheter, som statistik-användaren egentligen skulle vilja ha skattningar av, brukar kallas för intressestorheter, medan de storheter, som statistikproducenten försökt skatta så bra som möjligt, kallas för målstorheter. Det föreligger vanligen en diskrepans, sett ur användarens synpunkt, mellan intressestorheterna och målstorheterna. Sedan föreligger det dessutom ytterligare en diskrepans mellan de "sanna" värdena på målstorheterna och de mer eller mindre osäkra skattningar av målstorheterna, som härletts ur de faktiskt registrerade värden på observerade storheter, vilka uppnåtts med hjälp av de mätmetoder och mätinstrument, som statistikproducenten använt sig av. Denna diskrepans beror i sin tur på olika brister och ofullkomligheter i observations- och mätförfarandena.

### 3.2 Teknik och presentation

Man kan lugnt konstatera att utvecklingen på IT-området har gått snabbare, mycket snabbare, än vad någon kunnat drömma om. Jag har ibland roat mig med att titta tillbaka på "visioner", som jag blivit ombedd att formulera vid olika tidpunkter. När de presenterades blev dessa visioner i och för sig uppskattade men ofta behandlade med ett visst kritiskt överseende såsom varande alltför djärva och orealistiska. I efterhand har jag regelmässigt kunnat konstatera, att mina visioner snarare varit alltför försiktiga.

Ingenting tyder på att IT-utvecklingen skulle komma att gå långsammare under de kommande 5-10 åren. Vi kan alltså se fram mot mängder av nya möjligheter. Vilka är dessa möjligheter, och hur tar vi bäst vara på dem?

En första viktig fråga är hur man skall förhålla sig den snabba utvecklingen. Frågan är inte trivial, vilket belyses av följande paradox. Många av de tekniska problem man har i ett utvecklingsprojekt tenderar att få enkla och billiga standardlösningar, bara man väntar något. För ett problem, som just nu skulle kräva en ganska resurskrävande egenutveckling, kan det t.ex. mycket väl finnas ett billigt och väl fungerande kommersiellt standardprogram om ett halvår eller ett år. När programmet kommer ut på marknaden, har man kanske bara kommit halv-

vägs i sin egenutveckling av motsvarande funktionalitet, om man nu satsade på det.

Å andra sidan: om man aldrig satsar på att lösa problem på egen hand, utan bara väntar på att den tekniska och kommersiella utvecklingen skall lösa problemen åt en, så kommer man inte heller framåt. Man måste alltså ha ganska mycket fingertoppskänsla (och en hel del tur) för att kombinera egenutveckling och inväntande av färdiga standardlösningar på ett bra sätt.

### 3.2.1 Strategi för utvecklingsarbete i en snabbt föränderlig teknisk miljö

En vettig strategi för att främja utvecklingsarbetet i en teknisk miljö med extremt hög förändringstakt kan vara,

- att satsa de egna resurserna på att lösa problem som verkar vara unika, eller åtminstone ganska speciella, för det egna tillämpningsområdet,
- medan man inväntar kommersiella standardlösningar av problem, som man tror är vanliga för många tillämpningsområden.

När det gäller att öka tillgängligheten till den officiella statistiken bör vi fokusera vårt eget tekniska utvecklingsarbete på att

- ta fram en övergripande strategi och modell för tillgängliggörandet av officiell statistik, tillhörande dokumentation och bakomliggande datamaterial
- avgränsa och utarbeta lösningar av problem, som vi inte förväntar oss att andra aktörer, t.ex. kommersiella programvarutillverkare, skall lösa åt oss
- definiera enkla och tydliga gränssnitt, som gör det möjligt för egenutvecklade och kommersiellt tillgängliga komponenter att samverka med varandra på ett smidigt sätt

Observera att dessa principer kan tillämpas både i smått och i stort. En viss funktion i ett tekniskt system för att tillgängliggöra officiell statistik kan förverkligas genom ett program, som byggs upp av en kombination av kommersiellt tillgängliga och egenutvecklade komponenter, vilka samverkar via väldefinierade och väletablerade tekniska gränssnitt. Ett helt system för tillgängliggörande av officiell statistik – t.ex. det tekniska och organisatoriska system som förverkligar Sveriges statistiska databaser – kan på motsvarande sätt byggas upp av tekniska och organisatoriska delfunktioner, vilka samspelar med varandra via tekniska och mänskliga gränssnitt.

Genom att på detta sätt systematiskt arbeta med moduler och gränssnitt uppnår man också att systemet blir ”öppet”, d.v.s. man kan lätt byta ut en otillräcklig eller föråldrad modul mot något bättre, och man kan bygga ut systemet och dess funktionalitet genom att lägga till helt nya moduler. Detta är ett sätt att skapa flexibilitet i förhållande till nya och förändrade användarönskemål, samtidigt som man upprätthåller goda möjligheter att smidigt och fortlöpande anpassa ett system till den snabba och ständigt pågående tekniska utvecklingen.

### 3.2.2 Internetbaserade system

De inslag i den tekniska utvecklingen, som betytt särskilt mycket för utvecklingen av Sveriges statistiska databaser, är

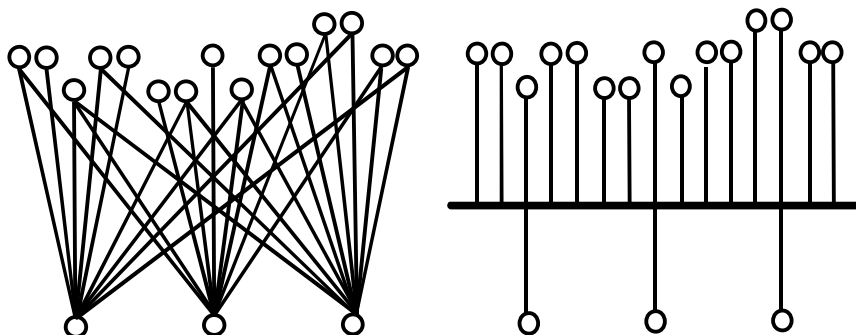
- Internetteknikens genombrott
- databastekniken med standardiserade gränssnitt (SQL etc) mot andra systemfunktioner
- intuitivt lättbegripliga *de facto* standarder för användargränssnitt (Windows, Netscape/Explorer)
- kraftfulla sökmotorer för att söka i stora datamängder bestående av mer eller mindre ostrukturerad fri text

Fortfarande måste man programmera rätt mycket själv för att på ett bra sätt utnyttja och kombinera dessa tekniker med varandra. Här kommer emellertid att hända en hel del under de närmaste åren. Microsoft visade t.ex. nyligen upp hur man på ett mycket enkelt sätt kommer att kunna bygga upp avancerade, Internetbaserade applikationer med hjälp av databashanteraren SQL Server 7.0 i kombination med de olika komponenterna i Office 2000 (Word, Excel, Outlook etc) inom ramen för utvecklingsmiljön Visual Studio 6.0, som i sig innehåller Visual Basic 6.0 och ett antal kraftfulla utvecklingsverktyg. Med några enkla ”drags and drops” kan man t.ex. snabbt förverkliga en applikation, där ett e-mail skickas till en statistikanvändare, så snart viss statistik, som användaren sagt sig vara intresserad av, finns tillgänglig. När användaren öppnar sitt e-mail, innehåller det inte bara den efterfrågade statistiken i någon form, som användaren önskat sig, t.ex. i ett Excel-ark, utan via Excel-arket kan användaren gå in i den bakomliggande databasen, på producentens Internetserver, och dynamiskt hämta ner andra data eller få sig en deldatabas nerladdad på den egna datorn.

### 3.2.3 Standardkomponenter och standardiserade gränssnitt

Frågan om standardkomponenter och – framför allt – standardiserade gränssnitt är en fråga av största betydelse, och det av flera skäl:

1. Ett standardiserat gränssnitt ger många kombinationsmöjligheter och därmed flexibilitet. Betrakta figur 13.



**Figur 13. Standardiserade gränssnitt minskar komplexiteten och ökar flexibiliteten.**

I vänstra delen av figur 13 visas, hur man kan kombinera olika komponenter (t.ex. programvaror) med varandra genom direkt kommunikation mellan komponenterna. Man måste då utforma kommunikationen på ett unikt sätt för varje par av komponenter, som skall kunna kommunicera med varandra. Om  $m$  komponenter skall kunna kommunicera med  $n$  andra komponenter, måste man alltså utforma  $(m \times n)/2$  "kommunikationsprotokoll" (överenskommelser om hur kommunikationen mellan olika komponenter skall gå till).

I högra delen av figur 13 visas, hur kommunikationen kan förenklas med hjälp av ett standardiserat gränssnitt. Alla komponenter, som kan kommunicera med gränssnittet, kan därmed också kommunicera med alla andra komponenter, som kan kommunicera med gränssnittet. Det räcker sålunda i detta fall med att utforma  $(m + n)$  kommunikationsprotokoll för att  $m$  komponenter skall kunna kommunicera med  $n$  andra komponenter.

2. Genom att introducera ett standardiserat gränssnitt i ett system (eller mellan system) kan man alltså uppnå
  - (a) en viss flexibilitet med mindre komplexitet; och/eller
  - (b) större flexibilitet vid viss komplexitet.

Tack vare det standardiserade gränssnittet blir systemet också inkrementellt utbyggbart: man kan i små steg bygga ut systemet och öka dess flexibilitet multiplikativt; när man lägger till en ny komponent, ser man bara till att denna komponent kan kommunicera med det standardiserade gränssnittet – och därmed kan det också kommunicera med alla andra komponenter, som redan kan kommunicera med gränssnittet. Om man inte har ett standardiserat gränssnitt, måste man däremot etablera lika många nya kommunikationsprotokoll, som det finns komponenter, som det nytillagda komponenten skall kunna kommunicera med.

3. Genom att bygga ett system runt några väl valda gränssnitt, kan man snabbt utveckla en hållbar systemstomme, som kan sättas i drift för nyttiga ändamål. Systemet kan därefter vidareutvecklas kontinuerligt under lång tid, genom att man tillför nya komponenter runt de givna gränssnitten. Man kan även lägga till nya gränssnitt, utan att man därför behöver ta bort eller ändra de gamla.
4. Även ett standardiserat gränssnitt kan någon gång behöva ändras, och detta är naturligtvis ett mera djupgående ingrepp i systemet. Man bör dock lägga märke till, att det är betydligt lättare att byta från en standard till en annan än att göra motsvarande ändring i ett system, där det inte finns någon standard alls.
5. Standardiserade gränssnitt möjliggör decentraliserade system och decentralisering av själva systemutvecklingen. Olika system eller delsystem kan utvecklas självständigt av olika aktörer. Ändå kan de kommunicera och samspeja med varandra, bara de utformas så, att de kan kommunicera med de standardiserade gränssnitten. Detta är en egenskap av stor betydelse i ett statistiskt system med flera statistiskansvariga myndigheter och flera statistikproducenter. Med hjälp av ett litet antal väl valda, standardiserade gränssnitt, kan man uppnå den (bl. a. ur användarsynpunkt) erforderliga samordningen inom ett sådant system, och man kan göra det flexibelt och inkrementellt utbyggbart. Systemet blir inte heller så sårbart för om en eller annan aktör i systemet skulle fallera.

Man brukar tala om system, som på detta sätt byggts kring standardiserade gränssnitt, som "löst sammanfogade system" (eng: "loosely coupled systems") – till skillnad från "hårt integrerade, monolitiska system" (grek: "monolit" = "en sten").

Standardiserade gränssnitt kan återfinnas på olika nivåer i ett informationssystem. Det kan t.ex. röra sig om användargränssnitt, kommunikationsgränssnitt, databasgränssnitt och gränssnitt mellan programmoduler på högre eller lägre nivå. Principerna för, och fördelarna med, standardiserade gränssnitt är dock desamma, oavsett systemnivå.

Några konkreta exempel på standardiserade gränssnitt är:

- Internet som sådant
- HTML, XML, SGML
- Webbläsare som Internet Explorer och Netscape
- Databasgränssnittet SQL
- EDIFACT/GESMES (Generic Statistical Message) för statistiska data och metadata
- Active X, CORBA, OLE, ODBC

Följande figurer i föreliggande rapport (avsnitt 1.5) innehåller exempel på standardiserade gränssnitt: 8b, 8c, 8d, 8f.

Ett problem med standarder i allmänhet, och med standardiserade gränssnitt i synnerhet, är att det ofta finns konkurrerande standarder för samma sak. Vissa standarder är formella standarder, fastställda av internationella standardiseringsorgan, andra är kommersiellt etablerade *de facto* standarder. I den bästa av världar är den officiella standarden också den kommersiellt dominerande *de facto* standarden, men detta lyckliga tillstånd uppkommer tyvärr sällan. Ett typiskt mönster är i stället, att en marknadsledande leverantör (tidigare IBM, numera Microsoft) framgångsrikt etablerar kommersiella *de facto* standarder, medan den dominerande leverantörens konkurrenter sluter sig samman och stöder en officiell standard, som tas fram av en internationell standardiseringsgrupp. Bägge grupperingarna har goda affärsmässiga skäl för sitt agerande.

Frågan är vad en standardanvändare (systembyggare) gör i en situation som denna. Den enda helsäkra lösningen för maximal flexibilitet är att implementera bägge standarderna parallellt, men detta är naturligtvis ofta alltför komplicerat och dyrbart. Det ”politiskt korrekta” är att välja den officiella standarden, men vad hjälper det att vara politiskt korrekt, om man därmed inte får tillgång till de marknadsdominerande systemkomponenterna? Den pragmatiskt korrekta lösningen är inte sällan att välja den marknadsdominerande leverantörens *de facto* standard. Man får därigenom tillgång till denna leverantörs samtliga produkter – och dessutom, vilket ibland kan vara ännu viktigare, även många konkurrenters produkter, eftersom få leverantörer vågar vara inkompatibla med den marknadsledande leverantörens produkter. Man kan också vara säker på, att den dag den marknadsledande leverantören förlorar sin marknadsledande position, och/eller det dyker upp en ny och helt överlägsen produkt från en konkurrerande leverantör, så uppkommer omedelbart en livskraftig marknad för att anpassa eller konvertera den gamla standarden till en ny.

Det kan i detta sammanhang vara instruktivt att jämföra med situationen på videomarknaden för ett antal år sedan. En numera praktiskt

taget bortglömd standard, som heter Betamax, ansågs av dåtidens videoexperter vara helt överlägsen. Den konkurrerande standarden VHS blev emellertid den som lyckades etablera sig som den kommersiellt dominerande *de facto* standarden, trots att den enligt de flesta experterna var klart sämre. De kunder som satsade på Betamax fann ganska snart att de hade ett litet och snabbt krympande utbud av videofilmer att välja mellan, d.v.s. deras valfrihet (flexibilitet) blev starkt beskuren.

### 3.2.4 Söksystem

I Sveriges statistiska databaser finns flera olika sätt att söka reda på det man är intresserad av och att associera från något man hittat till något annat, som man kommer på att man skulle vilja titta närmare på. Några sök- och associationsmöjligheter åskådliggörs i figur 14.

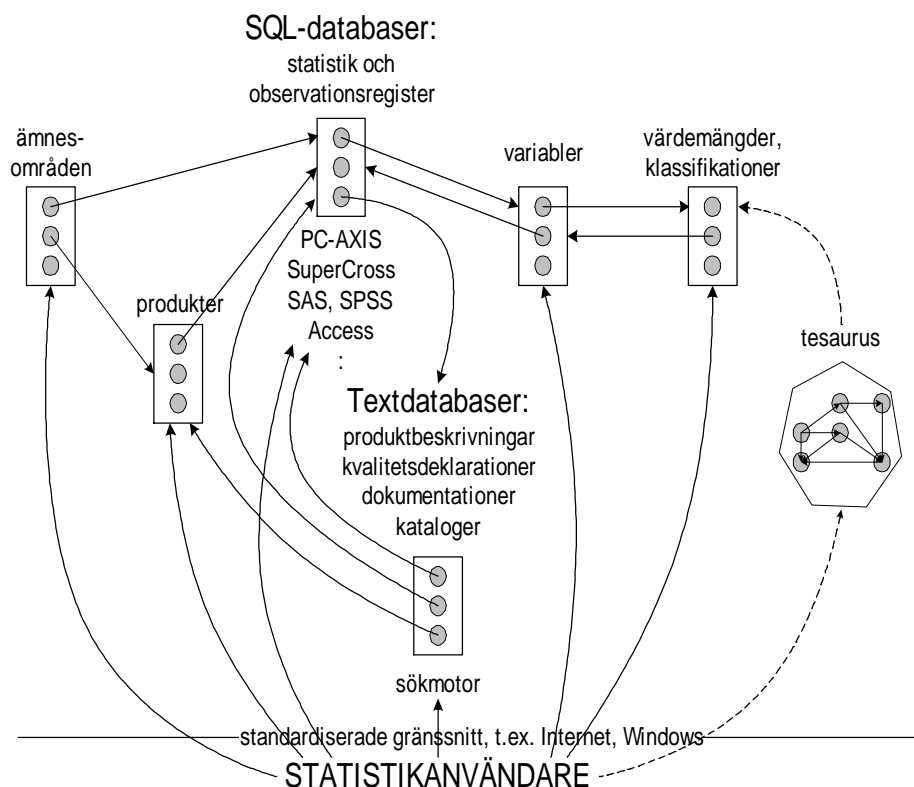
Av figur 14 framgår bl.a. att man kan hitta intressanta statistikmaterial och observationsregister

- antingen genom fritextsökning med hjälp av en sökmotor
- eller genom systematisk sökning via ämnesområden och produkter

När man hittat och kanske även bearbetat ett statistikmaterial eller observationsregister (med hjälp av programvaror som PC-AXIS, SuperCross, SAS, SPSS), vill man kanske veta mer om någon variabeldefinition eller annan kvalitetsaspekt. Man kan då länka sig till beskrivningar av variabler och klassifikationer eller till någon mera fullständig dokumentation av fritextkaraktär.

Det finns önskemål från användarna om ytterligare sökmöjligheter. En sådan möjlighet är antydd i figuren genom en streckad linje via en tesaurus.<sup>26</sup> En tesaurus ger mera avancerade möjligheter att associera begreppsmässigt närliggande termer med varandra. En användare är kanske intresserad av statistik om kylskåp. Ordet "kylskåp" förekommer kanske inte i de texter som beskriver tillgänglig statistik. Däremot finns kanske statistik om vitvaror, som i sig inkluderar kylskåp. I en tesaurus kan "kylskåp" vara associerat med "vitvaror" och därigenom möjliggöra att en användare hittar kylskåpen i statistiken.

<sup>26</sup> Denna funktion finns inte ännu i Sveriges statistiska databaser, men den efterfrågas ofta av statistikanvändare.



Figur 14. Några sökmöjligheter i Sveriges statistiska databaser.

En annan önskvärd sökmöjlighet är att man via det svenska statistiknätet, som knyter ihop Sveriges statistiska databaser med andra statistikkällor med hjälp av webblänkar, på ett enkelt sätt skall kunna utföra sökningar efter all statistik i statistiknätet, oavsett var den fysiskt ligger lagrad. Detta kräver ett visst organiserat utbyte av metadata mellan de olika system som deltar i det svenska statistiknätet, så att en sökmotor kan söka på all statistik, oavsett var den finns. Inom ramen för Miljödatanätet finns ett exempel på, hur man kan lösa denna typ av problem.

Givetvis krävs mera arbete med att bygga upp och underhålla sökmöjligheter och sökhjälpmedel ju mer användarvänliga och kraftfulla lösningar man önskar sig. Detta arbete kostar pengar. Men samtidigt går den tekniska utvecklingen snabbt framåt, så att man relativt billigt kan skaffa sig allt bättre standardverktyg för uppgifter av detta slag. Allt underhåll av metadata, som kräver manuella insatser, kommer dock att förbli dyrbart och tidsödande. Man bör därför undvika sökmetoder som bygger på manuella insatser, t.ex. manuell klassificering. I stället bör man satsa på automatiska och självlärande system.

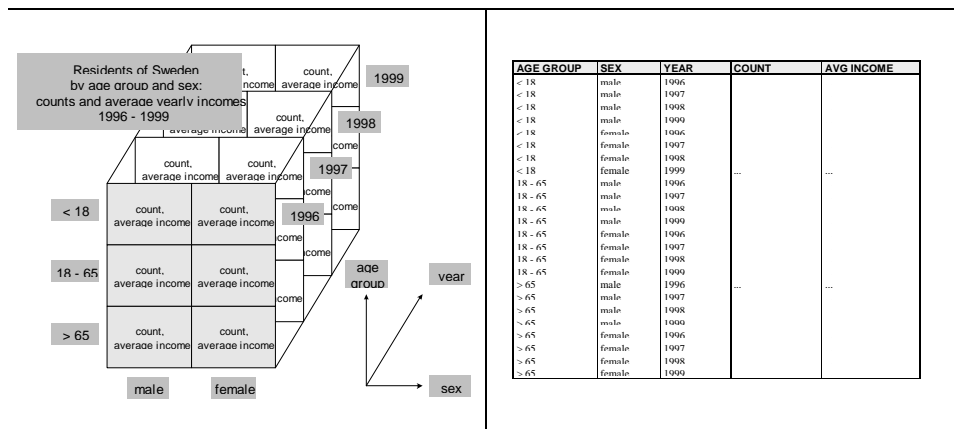


### 3.2.5 Databashanterare

I Sveriges statistiska databaser ingår

- dels relationsdatabaser med SQL-gränssnitt,
- dels fritextdatabaser.

I relationsdatabaser lagras alla slags data, både ”siffror” och verbala beskrivningar, i hårt strukturerad form – närmare bestämt i form av tvådimensionella tabeller (se figur 15, högra delen).<sup>27</sup>



Figur 15. En multidimensionell modell (till vänster) och en relationsdatamodell (till höger) för lagring av statistisk information.

Även fritextdatabaser kan vara mer eller mindre strukturerade. Textformateringspråk som HTML, XML och SGML (och för den delen även vanliga ordbehandlare som Microsoft Word) ger en rad möjligheter att strukturera text med hjälp av märkkoder. Sådan uppmärkning underlättar givetvis sökningar, bl.a. genom att de ökar möjligheterna att skilja mellan det som är väsentligt och det som är mindre väsentligt. Samtidigt kräver uppmärkningen mer eller mindre omfattande manuella insatser.

Även helt ostrukturerad text är sökbar med hjälp av moderna sökmotorer av det slag som berördes i avsnittet om sökning ovan.

<sup>27</sup> Figur 15 är hämtad från en översiktlig beskrivning av Sveriges statistiska databaser, som presenterats internationellt; se Sundgren, Bo: "Sweden's Statistical Databases: an infrastructure for flexible dissemination of statistics", Report to the UN/ECE Conference of European Statisticians, June 1997.

Relationsdatabaser är i praktiken sökbara endast på ett mycket strukturerat sätt, med hjälp av sökverktyg, som är baserade på SQL-språket (Structured Query Language). Man kan visserligen lagra data och metadata i textform i en relationsdatabas, men denna lagringsform lider av åtminstone tre brister:

1. Den minsta enheten som lagras i en relationsdatabas måste lagras "odelbart", d.v.s. ostrukturerat, in en enda tabellcell. Detta innebär att den textmassa, som lagras i en cell inte kan struktureras på något sätt, vare sig för att underlätta läsbarhet eller för att effektivisera sökningar.
2. "Fritextsökning" i en relationsdatabas blir, i den mån den alls kan utföras, i allmänhet mycket ineffektiv och långsam.
3. Om texterna är av omfattande och mycket varierande storlek, kan de vara svåra att hantera på ett effektivt sätt i en relationsdatabashantare. Omfattande textdata kan även påverka effektiviteten i bearbetningar av "sifferdata" på ett negativt sätt.

Inom databasområdet är relationsdatabaser för närvarande den helt dominerande standarden, såväl teorimässigt som kommersiellt. Ingenting inom IT-området gäller dock för evigt, eller, annorlunda uttryckt, evigheten inom IT-området är högst 5 år, och det finns redan alternativ eller komplement till relationsdatamodellen, som kan komma att bli kommersiellt intressanta under de närmaste åren. I det här aktuella sammanhanget tänker jag främst på två möjligheter.

Den ena möjligheten är s.k. multidimensionella databaser (se figur 15, vänstra delen). En multidimensionell databas är som klippt och skuren för att lagra statistiska data. Var och en av de  $n$  dimensionerna i en multidimensionell databas representerar en klassificering av de observationer man samlat in. Ett statistiskt observationsmaterial kan t.ex. som i figur 15 vara klassificerat efter kön, åldersgrupp och tidpunkt/tidsintervall. Varje cell i den  $n$ -dimensionella struktur, som spänns upp av de klassificerande variablerna, innehåller observerade eller skattade värden på en eller flera statistiska storheter för den delpopulation som motsvarar cellen ifråga.

*Exempel.* Cellen <male, 18-65, 1997> i figuren ovan motsvarar delpopulationen "män som var i åldern 18-65 år vid någon viss tidpunkt under 1997". För denna grupp objekt innehåller cellen ifråga observerade eller skattade värden på de statistiska storheterna

- antal personer
- medelinkomsten för dessa personer

Den multidimensionella datamodellen ligger uppenbarligen nära det som skulle kunna kallas en "naturlig" datamodell för statistiska data. Samma data kan visserligen logiskt ekvivalent lagras i en traditionell relationsdatamodell (se figur 15, högra delen), men denna modell motsvarar inte lika väl statistikanvändarens "mentala bild" av statistiken.<sup>28</sup>

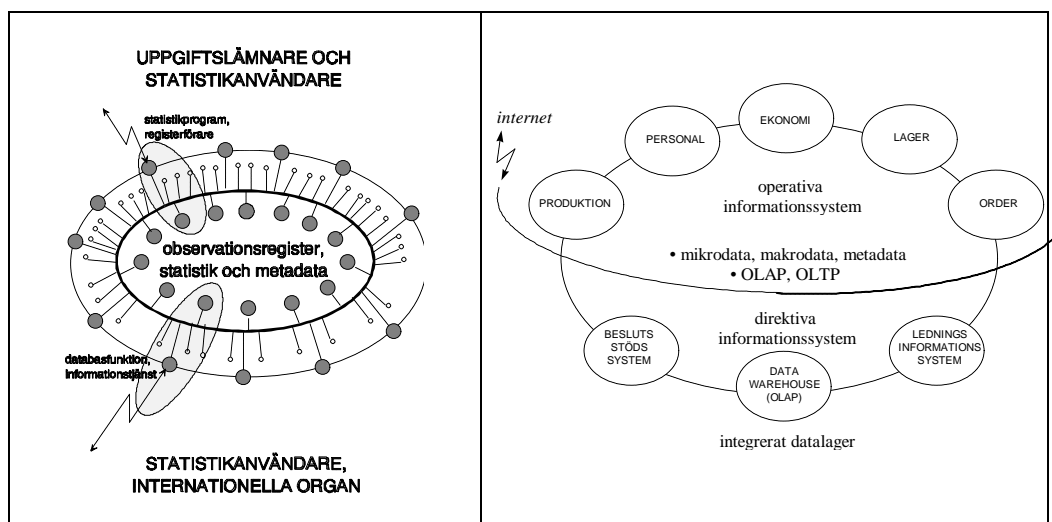
Det finns redan programvaror, som bygger på renodlade multidimensionella datamodeller. Det finns också programvaror, som bygger på hybridmodeller, där den grundläggande lagringen av data sker på vanligt sätt i en relationsdatamodell, samtidigt som vissa data permanent eller vid behov (även) lagras multidimensionellt, bl.a. för att snabba upp statistikuttag ur databasen.

Multidimensionell lagring av data blir alltmer vanlig i samband med s.k. OLAP-tillämpningar, vilka förekommer i anslutning till s.k. data-lager, "data warehouses". OLAP står för "On Line Analytical Processing", som bl.a. förekommer, när ett företag vill analysera stora mängder data, som genererats av de dagliga, operativa processerna, i syfte att planera verksamheten mera övergripande och på längre sikt, eller när man skall fatta något strategiskt beslut, t.ex. en stor investering eller en etablering på en ny marknad.

Motsatsen till OLAP är OLTP, "On Line Transaction Processing", d.v.s. system som direkt stöder den dagliga, operativa verksamheten i ett företag.

Figur 16 åskådliggör parallelliteten mellan å ena sidan (figuren till vänster) statistiska system av den typ som förekommer i samband med Sveriges statistiska databaser och å andra sidan (figuren till höger) OLAP-system i företagsmiljö.

<sup>28</sup> Som en kuriositet i sammanhanget kan man konstatera, att databashanteraren AXIS för lagring och spridning av statistik och metadata, som SCB utvecklade i början av 1970-talet, bygger på en multidimensionell datamodell. I dagens läge är det dock orealistiskt att utveckla och underhålla programvaror av detta komplicerade slag, utan den kommersiella utvecklingen bör inväntas.



Figur 16. Samspelet mellan operativa system och datalager/OLAP-system i statistikproduktionsmiljö (till vänster) och i företagsmiljö (till höger).

Utvecklingen av produktions-, analys- och spridningssystem för officiell statistik bör alltså framdeles kunna dra nytta av den parallellt pågående utvecklingen av metoder och verktyg för datalager och analysystem i företagsmiljö. Alltfler programvaruleverantörer kommer nu med verktyg anpassade för denna växande marknad.

En annan utvecklingslinje av intresse i detta sammanhang är utvecklingen av s.k. objektorienterade databaser. I en objektorienterad databas kan man lagra många olika typer av objekt, t.ex. mer eller mindre strukturerade data, mer eller mindre strukturerade texter, multimediaobjekt som bilder och/eller ljud. Man kan också på ett kraftfullt sätt bearbeta dessa olika typer av data, var för sig eller tillsammans, och man kan på ett förhållandevis enkelt och flexibelt sätt utveckla och anpassa applikationer till nya behov.

Objektorienterade databaser kan visa sig bli användbara, bl.a. för att hantera stora och komplicerade metadatabaser, fristående eller integrerade med statistikdatabaser.

### 3.2.6 Presentationsteknik

De traditionella presentationsteknikerna för statistisk information är

- tabeller
- grafik
- texter

Tabellerna och grafiken presenterar "siffrorna", och texter av olika slag förklarar och kommenterar statistikens innebörd, förutsättningar och begränsningar, t.ex. genom

- ledtexter (rubriker, förspalter, överspalter, diagramtexter)
- kommentarer
- fotnoter

Statistiken presenteras traditionellt på ett seriellt sätt, på sida efter sida i ett tryckt medium.

Med de elektroniska medierna för lagring och spridning av statistik har bl.a. följande möjligheter tillkommit:

- länkar ger möjlighet att organisera statistiken i semantiska nätverk, som ger möjlighet att associera och söka i många dimensioner samtidigt, oavsett hur statistiken rent fysiskt ligger lagrad
- animationer ger nya möjligheter att illustrera flöden och förändringar i statistiken
- guider och expertsystem kan ge användarna kraftfullt stöd i deras arbete med att tolka och analysera statistiken på egen hand
- parallella processer ger möjligheter att titta på flera statistiska material samtidigt
- simuleringar ger möjligheter att illustrera effekterna av olika förutsättningar m.m.

Det finns fortfarande förvånansvärt mycket att göra, när det gäller de traditionella presentationsteknikerna. Exempelvis finns det inte någon allmänt accepterad teori och metodik för hur statistiska tabeller skall byggas upp och presenteras.

Men de stora presentationstekniska utmaningarna finns förstås i anslutning till de nya möjligheter som de elektroniska medierna erbjuder. Uppenbarligen finns det mycket stora behov och möjligheter att förbättra många aspekter av statistikens tillgänglighet med hjälp av nya presentationstekniker. Under de senaste åren har datorn blivit ett nyttigt och lättanvänt arbetsverktyg mycket tack vare en kreativ utveckling av grafiska användargränssnitt. På analogt sätt bör vi kunna hitta nya vägar för att snabbt och effektivt söka oss fram till relevant statistik samt få hjälp med att tolka och analysera den. Detta utvecklingsarbete befinner sig ännu i sin linda.

### 3.3 Ekonomi: finansiering och prissättning

#### 3.3.1 Officiell statistik – en kollektiv nytthet

Officiell statistik är på det hela taget ett bra exempel på en kollektiv nytthet. Kostnaden för att producera den officiella statistiken är praktiskt taget densamma, oavsett hur många som konsumerar den. När den väl producerats för någon konsument, är marginalkostnaden för att ställa den till ytterligare konsumenters förfogande låg, ibland nära noll.

I ett demokratiskt samhälle är det naturligt och rationellt att genom det politiska systemet fatta beslut om huruvida en kollektiv nytthet skall produceras eller inte. Om man beslutar, att den skall produceras, är det dessutom naturligt och rationellt att finansiera den kollektivt, d.v.s. genom anslag från skattemedel, åtminstone till större delen. Alternativet vore att fördela kostnaderna på dem som faktiskt utnyttjar den kollektiva nyttheten, men en sådan princip leder lätt till ett suboptimalt underutnyttjande av nyttheten ifråga.

Å andra sidan är det sällan optimalt att låta en kollektiv nytthet vara helt gratis. Det vore i och för sig optimalt, om marginalkostnaden för att låta ytterligare en konsument konsumera den vore exakt noll, men så är sällan fallet. Även om de flesta kostnaderna för produktion och distribution av nyttheten ifråga är helt oberoende av hur många som konsumerar den, så finns det nästan alltid några kostnader som ökar med antalet konsumenter, proportionellt eller språngvis. Dessa kostnader bör konsumenterna stå för. Det är den s.k. marginalkostnadsprincipen, som enligt ekonomisk teori leder till optimalt resurs-utnyttjande.

De flesta kostnaderna för att producera och tillgängliggöra en viss statistikprodukt är helt oberoende av, hur många som konsumerar statistikprodukten ifråga. Det gäller t.ex. kostnaderna för att

- utforma och konstruera ett system för att producera statistiken ifråga
- samla in data till statistiken
- granska och rätta data som samlats in
- genomföra bearbetningar som leder fram till den förutbestämda statistikprodukten (t.ex. ett antal statistiska tabeller och/eller ett eller flera slutliga observationsregister)
- utforma och konstruera ett system för att tillgängliggöra och vidarebearbeta den förutbestämda statistikprodukten

Emellertid finns det också kostnader, om än i de flesta fall av betydligt lägre storleksordning, som ökar med antalet användare. Det gäller t.ex. kostnader för att

- producera (ytterligare) exemplar av en statistisk publikation (utöver originalet)<sup>29</sup>
- ge (betydligt) fler användare tillgång till en databas (med given servicenivå)<sup>30</sup>
- skräddarsy särskilda varianter av statistikprodukten för vissa användare
- utföra vidarebearbetningar och analyser av statistiken för en viss användares räkning

Rörliga kostnader av detta slag bör bekostas av användarna. Annars uppkommer köer eller ransoneringsbehov, om inte statsmakterna tillåter anslagsöverdrag. Underdebitering leder liksom överdebitering till en suboptimal användning av statistikprodukten ifråga.

Givetvis vore det ofta orationellt att tillämpa millimeterrättvisa vid debiteringen av de rörliga kostnaderna. Bl.a. måste man ta hänsyn till kostnaderna för själva kostnadsfördelnings- och debiteringsförfarandet. Om en viss typ av tjänst har en marginalkostnad, som ligger mycket nära noll, kan det ibland vara lämpligt att låta den vara gratis, samtidigt som man tar ut ett något för högt pris på en annan tjänst, så att man totalt sett får balans mellan (marginal)kostnader och (marginal)intäkter. En förutsättning är då, att man på goda grunder kan anta, att efterfrågan på gratistjänsten inte växer så våldsamt, att totalkostnaden för denna tjänst trots allt blir mycket stor.

### 3.3.2 Marginalkostnadsprincipen för prissättning av informationstjänster

Man kan också uttrycka den här förespråkade marginalkostnadsprincipen för prissättning av officiell statistik på följande sätt:

- De fasta kostnaderna för att producera och tillgängliggöra en viss basversion av den officiella statistiken och bakomliggande datamaterial, inklusive kostnaderna för erforderlig infrastruktur, skall vara helt anslagsfinansierade och inte till någon del debiteras statistik-användaren.
- De rörliga kostnaderna för att tillgängliggöra och vidareförädla basmaterialen (mikrodata, makrodata och metadata), som ingår i den

<sup>29</sup> Här avses t.ex. papperskostnader och andra rörliga kostnader för att trycka ytterligare exemplar.

<sup>30</sup> Här avses t.ex. kostnader för ökad datorkapacitet för att kunna upprätthålla en given svarstid gentemot användarna.

officiella statistiken, skall i sin helhet debiteras användarna genom någon lämplig form av avgifter.

De här föreslagna principerna för prissättning av olika tjänster i samband med produktion och tillgängliggörande av officiell statistik motsvarar mycket väl den huvudprincip, som föreslås i ett av Riksrevisionsverket (RRV) på regeringens uppdrag utarbetat förslag till prissättning av informationstjänster inom den offentliga sektorn.<sup>31</sup>

Det är angeläget att rationella och enhetliga prissättningsprinciper tillämpas konsekvent i hela det svenska statistiska systemet.<sup>32</sup> Det är också mycket angeläget att Sverige hävdar samma principer i internationella sammanhang, där man ofta återfinner en brist på samordning, för att inte säga en stor aningslöshet, när det gäller prissättning av informationstjänster. Även internationella statistikorganisationer, finansierade av medlemsländerna, kan inte sällan helt ogenerat tänka sig att exploatera sina *de facto* monopol vad gäller tillgången till statistik genom att ta ut betydligt högre priser än de rörliga marginalkostnaderna för tillgängliggörandet. Man ser det som en välkommen extra inkomstkälla.

Principlös prissättning och datamonopol (*de facto*) är i själva verket mycket allvarliga hot mot statistikens tillgänglighet. Om en statistikansvarig myndighet utnyttjar sin monopolställning och sina förstahandskunskaper om de statistiska datamaterialen till att ta betalt för data och dokumentation, trots att man redan fått betalt genom anslag, leder det inte bara till ett resursslöseri i form av underutnyttjande av den officiella statistiken. Det förhindrar eller bromsar åtminstone onödigtvis den utveckling, som annars skulle kunna komma statistikanvändarna till godo, genom att olika innovatörer och entreprenörer i samhället i konkurrens med varandra och med de etablerade statistikproducenterna skulle kunna ta fram nya och intressanta informationstjänster med anknytning till den officiella statistiken.

<sup>31</sup> "Principer för prissättning av informationstjänster", regeringsuppdrag, RRV 1995:64.

<sup>32</sup> Helst borde samma principer också tillämpas för närbesläktade informationstjänster i samhällets regi, t.ex. geografisk information (kartor och koordinater). Här tycks man dock ibland ha funnit anledning att göra undantag från RRV:s huvudprincip för prissättning av informationstjänster, vilket i sin tur helt naturligt lett till ett enligt min mening beklagligt underutnyttjande av t.ex. tillgänglig geografisk information



### 3.3.3 Bastjänster och tilläggstjänster

Annorlunda uttryckt kan man säga, att de statistikansvariga myndigheternas anslagsfinansierade monopol bör begränsas till

- produktion och tillgängliggörande av vissa basprodukter inom den officiella statistiken: fördefinierade statistiska sammanställningar och bakomliggande observationsregister med tillhörande dokumentation och metadata
- tillhandahållande av ett antal bastjänster, som gör det möjligt att överhuvudtaget få tillgång till den officiella statistiken och bakomliggande datamaterial

I övrigt bör samhället välkomna framväxten av en fri och konkurrensutsatt marknad för informationstjänster, s.k. tilläggstjänster, med anknytning till den officiella statistiken och bakomliggande datamaterial. Marknaden skall inte subventioneras, men den skall heller inte motarbetas, t.ex. genom att statistikansvariga myndigheter tillåts monopolisera eller begränsa datatillgången genom omotiverade restriktioner,<sup>33</sup> eller genom att tilläggstjänster subventioneras till den statistikansvariga myndighetens fördel med hjälp av anslagsmedel.<sup>34</sup>

Sammanfattningsvis bör nödvändiga bastjänster kring tillgängliggörandet av officiell statistik tillhandahållas av det svenska statistiska systemet till marginalkostnader, som, inte minst tack vare den snabba tekniska utvecklingen under senare år, bör kunna hållas mycket låga. Tilläggstjänster, som ger statistik användarna mervärden utöver bastjänsterna, men som användarna fritt kan välja att begagna sig av eller avstå ifrån (t.ex. genom att man kan skapa sig motsvarande mervärde på egen hand), skall produceras och tillhandahållas på en fri och konkurrensutsatt marknad, där statistikansvariga myndigheter också kan vara aktörer, men i så fall på samma villkor som andra aktörer.

<sup>33</sup> Exempel på motiverade restriktioner är vad som krävs för att upprätthålla statistiksekretessen. Men restriktionerna får inte ges en sådan utformning och tillämpning, att den statistikansvarige går längre i sin restriktivitet än vad som är absolut nödvändigt av sådana skäl.

<sup>34</sup> Däremot kan man tänka sig, att statsmakterna använder anslagsmedel för att finansiera uppbyggnaden av en infrastruktur, med utgångspunkt i vilken alla aktörer på "statistik tjänstmarknaden" kan utveckla och tillhandahålla tilläggs-tjänster på lika villkor.

### 3.3.4 Internet-baserade informationstjänster

Internet-teknikens genombrott har gjort marginalkostnadsprincipen extra attraktiv. Internet-tekniken gör det nämligen möjligt att tillgängliggöra producerad statistik till mycket låg marginalkostnad. I själva verket är marginalkostnaden så låg, att många argumenterar för att statistiken på Internet borde vara helt gratis.

Det finns flera goda argument för att låta Sveriges officiella statistik vara gratis på Internet. Produktionen av officiell statistik är helt finansierad av statsmakterna. Traditionellt anses detta innebära, att statsmakterna har betalt till och med framställningen av ett tryckoriginal. Reproduktion och spridning av tryckta statistikpublikationer brukar man däremot åtminstone till viss del finansiera genom att man sätter ett pris på publikationerna. Kostnaderna för traditionell reproduktion och spridning av tryckta publikationer är ganska stora, och även om man tillämpar marginalkostnadsprincipen, kan priset för en statistikpublikation bli högt, ibland alltför högt.

Om man i stället sprider statistiken helt elektroniskt, via Internet, blir motsvarande marginalkostnader mycket låga. Man kan då fråga sig om man inte kunde anse att denna låga kostnad i själva verket skulle kunna anses ingå i anslaget från statsmakterna. Då skulle statistikproducenten slippa besväret och kostnaderna för att ta betalt för statistiken. Vidare skulle användningen av statistiken säkert öka, och en ökad användning av officiell statistik ligger väl både i effektivitetens och demokratins intresse?

Det finns emellertid också argument mot "nolltaxa" för officiell statistik. Allmänt sett leder ett för lågt pris till överkonsumtion och slöseri med resurser. Även om styckekostnaden per statistikuppgift är extremt lågt, så kan kraftigt ökade konsumtionsvolymen trots allt leda till betydande kostnader, även om man använder Internetteknik. För att tillgängliggöra statistik via Internet, måste statistikproducenten lagra statistiken på servrar, som ansluts till Internet. Om alltför användare kopplar upp sig mot servrarna och gör alltmer avancerade bearbetningar på dessa, uppstår förr eller senare trängsel, varvid statistikproducenten måste skaffa fler eller större servrar, vilket i sin tur leder till kostnader. Om tillgängliggörandet av statistiken är helt anslagsfinansierad, måste statistikproducenten då begära ökade anslag, vilket dels tar tid, dels inte säkert beviljas, vilket i så fall leder till att alla statistikanvändare tvingas acceptera längre kötider och sänkt servicenivå.

Ett pris som täcker marginalkostnaden löser ovanstående problem på ett smidigt sätt. Om användningen av statistikproducentens tjänster ökar, så ökar också intäkterna, och serverkapaciteten kan byggas ut i takt med

ökad efterfrågan, utan att statsmakterna behöver involveras i en långdragen och osäker budgetprocess.

Ett pris ger också värdefull återföring av information till statistikproducenten. Det kanske visar sig att vissa statistikanvändare är beredda att betala ett högre pris för ytterligare tjänster, även tjänster som är förenade med ganska höga marginalkostnader. "Nolltaxa" skulle göra det svårt, kanske omöjligt, för statistikproducenten att tillhandahålla sådana tjänster.

### 3.4 Effektivare hantering av statistiksekretessen

Varje seriös statistikanvändare och statistikproducent är fullt på det klara med statistiksekretessens betydelse. Om en uppgiftslämnare misstänker, att de uppgifter han eller hon lämnar ifrån sig för statistiska ändamål även kan komma att användas andra ändamål, t.ex. för att kontrollera uppgiftslämnarens laglydnad eller allmänna leverne, så kommer uppgiftslämnaren att lämna felaktiga eller snedvridna svar eller inga svar alls. Statistik baserad på sådana uppgifter blir mer eller mindre värdelös, rentav skadlig.

#### 3.4.1 Avsiktliga och oavsiktliga röjanden

Det mest näralliggande brottet mot statistiksekretessen är att en statistikproducent själv använder sig av insamlade uppgifter för icke-statistiska ändamål, medan uppgifterna fortfarande är knutna till identifierade personer eller företag, eller att en statistikproducent lämnar ut sådana uppgifter till andra (t.ex. polisen), som i sin tur använder dem för icke-statistiska ändamål. Sådana handlingar innebär avsiktliga röjanden av sekretesskyddade statistikuppgifter och är kriminella. I Sverige är vi dessbättre förskonade från denna typ av missbruk av statistikuppgifter insamlade för officiell statistik, och uppgiftslämnarna i Sverige har i konsekvens därmed ett högt förtroende för producenterna av officiell statistik.

Så långt föreligger alltså inget större problem kring statistiksekretessen. Det som leder till problem är följande. Även sedan insamlade statistikuppgifter avidentifierats och summerats (aggregerats) till statistik, kan de under vissa förutsättningar åter knytas till de personer eller företag, som lämnat dem. Det gäller t.ex. följande statistiska tabeller:

**Tabell 1a. Befolkningen i Knypplinge efter ålder och kön.**

	Män	Kvinnor
< 18 år	5	3
18 - 65 år	37	39
> 65 år	19	27

**Tabell 1b. Antal dömda för brott i Knypplinge efter ålder och kön.**

	Män	Kvinnor
< 18 år	5	0
18 - 65 år	8	5
> 65 år	3	1

**Figur 17. Statistiska tabeller som möjliggör bakvägsidentifiering.**

Om man kombinerar de var för sig oskyldiga tabellerna, kan man sluta sig till att alla pojkar under 18 år i Knypplinge är dömda för brott, och om jag nu råkar känna en pojke under 18 år i Knypplinge, så får jag alltså reda på att denne är dömd för brott, om jag inte redan kände till det. Jag kan också sluta mig till, att de flickor under 18 år jag eventuellt känner i Knypplinge, inte är dömda för något brott, men detta är troligen mindre känslig information.

Denna typ av röjanden av sekretessbelagda statistikuppgifter skiljer sig från de inledningsvis i detta avsnitt nämnda exemplen genom att de är indirekta och oavsiktliga. Man kallar dem därför även oavsiktliga röjanden (eng: "inadvertent disclosures"). Enligt svensk lag är det straffbelagt att utnyttja sådana eventuellt existerande möjligheter att forcera statistiksekretessen, genom att kombinera officiell statistik med annan information.<sup>35</sup>

Det har under de senaste 30 åren bedrivits oerhört mycket forskning kring denna problematik. Till viss del har forskningen gått ut på att försöka hitta metoder för att förhindra oavsiktliga röjanden av det slag som åskådliggörs i ovanstående enkla exempel. I större utsträckning har dock forskningen gått ut på att visa, att oavsett vilka raffinerade skyddsåtgärder, som vidtas, så finns det ännu mera raffinerade metoder för att bryta sig igenom statistiksekretessen. En något nedslående slutsats, som jag för övrigt själv drog redan 1975 i en uppsats, som jag skrev tillsammans med Edmund Rapaport,<sup>36</sup> är att det praktiskt taget

<sup>35</sup> Lagen om den officiella statistiken, 10§. Brottet kallas "olovlig identifiering" (samma lag, 15§) och straffas med böter eller fängelse i högst ett år, om gärningen inte är belagd med straff i brottsbalken eller datalagen.

<sup>36</sup> Rapaport, Edmund & Sundgren, Bo: "Output Protection in Statistical Data Bases". Conference of the International Statistical Institute (ISI), Warszawa, 1975. Finns också tillgänglig som rapport från SCB.

alltid, under vissa förutsättningar, är teoretiskt möjligt att röja en eller annan mer eller mindre känslig uppgift om en enskild person eller ett enskilt företag ur statistik, som överhuvudtaget behandlar någon typ av känslig information.

De viktigaste förutsättningarna för att jag på detta sätt skall kunna bryta igenom statistiksekretessen är:

- att jag har ett starkt motiv att vilja röja känsliga uppgifter om enskilda;
- att jag är beredd att begå en brottslig gärning;
- att jag har erforderliga resurser (knskaper, pengar, datorkraft);
- att jag utöver statistiken själv har tillgång till viss ytterligare information från andra källor, t.ex. min egen lokalkännedom och uppgifter från offentliga administrativa register.

Det är sålunda omöjligt att ge en 100-procentig garanti för att det är helt omöjligt för någon illvillig statistik användare att med hjälp av statistiken röja känsliga förhållanden om enskilda, med mindre än att man helt upphör att producera och sprida statistik om förhållanden, som på något sätt kan upplevas som känsliga eller kontroversiella.

I en viss mening måste vi alltså välja mellan pest och kolera. Om vi vill ha någon intressant statistik överhuvudtaget (vilket är nödvändigt om vi vill ha en demokratisk process baserad på objektiva informationsunderlag snarare än okontrollerbara slagord), måste vi vara beredda att ta vissa risker vad gäller statistiksekretessen. Men för att statistiken verkligen skall vara intressant och rättvisande, har vi inte råd att ta risker som äventyrar förtroendet för statistikproducenten hos uppgiftslämnarna.

En rimlig upplösning av detta dilemma står att finna i en lämplig kombination av lagstiftning, förtroendeskapande åtgärder gentemot uppgiftslämnarna och tekniska/metodologiska insatser.

### 3.4.2 Det lagstadgade skyddet för statistiksekretessen

Den svenska lagstiftningen ger ett mycket gott skydd för statistiksekretessen. Enligt grundlagen måste kräver all sekretessbeläggning av handlingar hos myndigheter att man kan hänvisa till ett lagstadgat explicit undantag från offentlighetsprincipen. För statistiksekretessen

sker detta genom Sekretesslagens 9 kapitel 4§, vars innehåll sammanfattas nedan:<sup>37</sup>

<p><b>Sekretesslagen (9kap, 4§)</b> enligt riksdagsbeslut</p>
<p><i>Statistiksekretessen:</i></p> <ul style="list-style-type: none"> <li>• Sekretess gäller för sådan särskild verksamhet hos myndighet, som avser framställning av statistik, för uppgift, som avser enskilda personliga eller ekonomiska förhållanden, och som kan hänföras till den enskilde. Sekretessen gäller i 70 år för uppgift om enskilda personliga förhållanden och annars i 20 år.</li> <li>• Med undantag från ovanstående gäller, att följande uppgifter får lämnas ut, om det står klart att den som uppgiften rör eller någon honom närstående inte lider skada eller men: <ul style="list-style-type: none"> <li>– uppgift i företagregister</li> <li>– uppgift som avser avliden</li> <li>– uppgift som behövs för forskningsändamål</li> <li>– uppgift som avser personal- och lönestatistik</li> <li>– uppgift som inte genom namn eller annan identitetsbeteckning eller liknande är direkt hänförlig till den enskilde</li> </ul> </li> </ul>

Figur 18. Statistiksekretessens innebörd enligt Sekretesslagen.

Ytterligare skydd för statistiksekretessen stadgas i Lagen om den offentliga statistiken, 2§, 9§, 10§ och 15§ (se vidare avsnitt 3.5). Enligt 10§ får inte uppgifter i den officiella statistiken sammanställas med andra uppgifter i syfte att utröna enskildas identitet. Denna paragraf förbjuder alltså bakvägsidentifiering (via statistik eller avidentifierade mikrodata) av det slag, som exemplifierades med tabellerna i början av detta avsnitt. I 15§ anges straffsatsen för brott mot förbudet av sådana s.k. olovliga identifieringar.

Huvudsyftet med statistiksekretessen är att säkerställa förutsättningarna för att statistikansvariga myndigheter skall kunna framställa objektiv officiell statistik av god kvalitet, genom att uppgiftslämnarna kan lita på att de uppgifter de lämnar till den officiella statistiken inte används för andra än statistiska ändamål. Exempelvis måste de kunna lita på att en uppgift de lämnat till den officiella statistiken inte som sådan kan användas mot dem i någon administrativ process, som förorsakar dem skada eller men.

<sup>37</sup> Se även Samuelson, Per: "Statistikrätt: Den rättsliga regleringen av den statliga statistiken", Fritzes 1994.

Däremot kan det inte anses ingå i statistiksekretessen att förhindra att den *statistik*, som framställs med utgångspunkt i en uppgift från en viss uppgiftslämnare tillsammans med uppgifter från andra uppgiftslämnare, i ett senare skede leder till beslut hos statistikanvändare, t.ex. statsmakterna, som uppgiftslämnaren anser vara till skada eller men för denne.

*Exempel 1:* En statistisk undersökning kan tänkas visa att fastighetsägare generellt sett har en mycket god ekonomisk situation. Det kan inte anses vara ett brott mot statistiksekretessen, om statsmakterna bl.a. på basis av statistiken ifråga, beslutar om en höjning av fastighetsskatten.

*Exempel 2:* Av officiell statistik framgår att en viss kommun är osedvanligt brottsbelastad. Det kan inte anses vara ett brott mot statistiksekretessen, om Rikspolisstyrelsen därför beslutar att tilldela kommunen ifråga fler poliser, trots att detta troligen är till skada eller men för ett antal invånare i kommunen, som lämnat uppgifter till statistiken ifråga.

### 3.4.3 Avvägningen mellan statistiksekretess och statistiktillgänglighet

Avvägningen mellan statistiksekretess och statistiktillgänglighet är inte okomplicerad, och en statistikansvarig myndighet är måhända inte i alla lägen den lämpligaste att ensam göra denna avvägning. Låt oss något analysera vari svårigheterna består.

Om en uppgift, som lämnats till den officiella statistiken, är knuten till en uppgiftslämnare, som är entydigt, eller nästan entydigt, knuten till en enskild med hjälp av namn, personnummer eller liknande, torde knappast någon, inte ens en forskare eller annan seriös statistikanvändare, förvänta sig att den statistikansvarige lämnar ut uppgiften i den form den föreligger. Enligt lagen skulle den statistikansvarige kunna göra detta, om det står klart att ett utlämnande inte skulle vara till skada (= ekonomisk skada) eller men (= integritetskränkande) för uppgiftslämnaren eller denne närstående.

Harmlösa uppgifter av det slag som skulle möjliggöra utlämnande av dem, även om de kan knytas till enskild uppgiftslämnare, är måhända i sig själva sällan så särskilt intressanta för statistiska bearbetningar och analyser. Däremot har forskare och analytiker ofta ett behov av att komplettera egna uppgiftsinsamlingar, som kan innehålla mer eller mindre känsliga uppgifter, med bakgrundsdata från den officiella statistiken. För att kunna göra detta på individnivå, måste de berörda individerna (t.ex. personer eller företag) vara entydigt identifierade. Om bakgrundsuppgifterna då är harmlösa, d.v.s. om det står klart att ett

utlämnande inte skulle vara till skada eller men för uppgiftslämnaren eller denne närstående, kan den statistikansvarige lämna ut ifrågasvarande uppgifter till forskaren ifråga. Ett annat alternativ är att den statistikansvarige kan åta sig att hos sig berika forskarens känsliga data med bakgrundsdata från den officiella statistiken. Detta förutsätter att forskaren får och är villig att lämna över sitt material till den statistikansvariga myndigheten för påförande av bakgrundsdata. Det förutsätter också att denna bearbetning kan göras till en för forskaren rimlig kostnad.

#### 3.4.4 Behövs en tillgänglighetsprincip för hantering av statistiksekretessen?

Om en forskare eller analytiker vill utnyttja identifierade uppgifter hos den statistikansvarige, som är känsliga i den meningen, att det *inte* står klart att ett utlämnande *inte* skulle vara till skada eller men för uppgiftslämnaren eller denne närstående, är det enbart den senare möjligheten, som står öppen, d.v.s. forskaren måste uppdra åt den statistikansvarige att genomföra de bearbetningar och sambearbetningar, som han eller hon vill och får genomföra. Återigen är det då viktigt för statistik användaren, att den statistikansvarige kan åta sig sådana uppdrag på ett snabbt och smidigt sätt och till låg kostnad. Detta senare krav på den statistikansvarige skulle rentav kunna formuleras som en *tillgänglighetsprincip för hantering av statistiksekretessen*, avsedd att kompensera för de olägenheter, som statistiksekretessen förorsakar statistik användarnas berättigade krav på tillgänglighet (se figur 19).

*Tillgänglighetsprincip för hantering av statistiksekretessen:*

- *Hanteringen av statistiksekretessen skall vara så effektiv, så att den inte, t.ex. av tids- eller kostnadsskäl, onödigtvis inskränker tillgängligheten till statistiken för statistik användarnas egna bearbetningar och analyser.*

**Figur 19.** Tillgänglighetsprincip för hantering av statistiksekretessen.



### 3.4.5 Olika sätt att hantera statistiksekretessen i särskilt känsliga fall

Ett svårt fall föreligger, när forskaren vill genomföra mycket interaktiva analyser av känsliga uppgifter hos den statistikansvarige, uppgifter som eventuellt också måste kombineras på individnivå med känsliga uppgifter hos forskaren. Forskaren kan då inte i förväg specificera exakt vilka bearbetningar han eller hon vill göra, vilket skulle möjliggöra att bearbetningarna uppdrogs till den statistikansvarige enligt ovan, utan forskaren måste själv betrakta resultatet av en bearbetning, innan nästa bearbetning kan specificeras. Detta leder till en mycket komplicerad process, där forskaren och den statistikansvarige hela tiden måste ha mycket täta kontakter med varandra. Ett sådant förfarande upplevs ofta av forskare som alltför klumpigt, dyrt och tidsödande och därför helt otillfredsställande ur tillgänglighetssynpunkt. Vad kan man göra i en sådan situation?

#### *Forskare "svärs in" vid forskararbetsplatser hos den statistikansvarige*

En möjlighet, som utnyttjas ibland, både i Sverige och utomlands, är att forskaren "svärs in" hos den statistikansvarige, d.v.s. temporärt behandlas som en anställd hos den statistikansvariga myndigheten, med de plikter och privilegier, som detta för med sig. Vid SCB har t.ex. av och till funnits ett antal "forskararbetsplatser" i anslutning till vissa statistikprodukter, t.ex. ULF (undersökning om levnadsförhållanden). Denna lösning förutsätter, att forskaren utan alltför stora problem under en viss tid kan flytta sitt arbete rent fysiskt till den statistikansvarige och passar därför inte dem, som har sin ordinarie forskarplats i någon annan del av landet.

#### *Elektroniskt tillgängliga forskararbetsplatser med fjärrstyrning*

En elegantare lösning bygger på elektronisk kontakt mellan forskaren och den statistikansvarige. En sådan lösning förutsätter bl.a., att man, ungefär som i Internetsammanhang, kan skapa någon form av elektronisk "brandvägg" mellan den statistikansvariga myndighetens register och databaser, å ena sidan, och den externa forskaren, å den andra, så att bara harmlösa data kan passera brandväggen. I korta drag skulle ett sådant system fungera på följande sätt:

1. Forskaren specificerar med hjälp av program och metadata på sin egen PC, eventuellt i kombination med metadata från den statistikansvarige, som naturligtvis är harmlösa och får passera brandväggen, en bearbetning av identifierade, känsliga data hos den statistikansvarige. Specifikationen skickas till den statistikansvarige via brandväggen och kontrolleras automatiskt, eller i särskilt komplicerade fall, delvis manuellt hos den statistikansvarige, så att det kan konstateras, att den specificerade bearbetningen inte hotar statistiksekretessen. Det handlar väsentligen om att kontrollera, att resultatet av den specificerade bearbetningen kan lämnas ut. Eventuellt kan denna kontroll utföras först efter det att bearbetningen faktiskt utfört, och resultatet är klart att leverera ut till forskaren via brandväggen.
2. Bearbetningen utförs automatiskt hos den statistikansvarige, på dennes datorer.
3. Resultatet kontrolleras och skickas elektroniskt till forskaren via brandväggen.
4. Forskaren begrundar resultatet och utför eventuellt egna analyser på sin egen PC.
5. Forskaren formulerar, på basis av sina egna analyser, en ny fråga, som skickas till den statistikansvarige och bearbetas där enligt pkt 1-3 ovan.

Med den teknik, som idag är tillgänglig, kan det ovan beskrivna förfarandet göras minst lika säkert som det betydligt klumpigare förfarandet med en forskararbetsplats, som är fysiskt placerad hos den statistikansvarige. I bägge fallen bygger skyddet på en kombination av legala, avtalsmässiga och fysiska åtgärder. I lösningen med en forskararbetsplats hos den statistikansvarige, måste någon övervaka, att den betrodde forskaren verkligen håller sig till det överenskomna avtalet och inte gör något annat än det som är tillåtet enligt lag och avtal. I den elektroniska lösningen sköts övervakningen till största delen helt automatiskt, och forskaren ges i praktiken inga möjligheter att bryta mot lag och avtal. Den elektroniska lösningen är självklart både smidigare och billigare för alla inblandade.

Det ovan beskrivna elektroniska förfarandet blir naturligtvis ännu enklare att implementera och övervaka, om de data, som forskaren specificerar sina bearbetningar mot är avidentifierade. I en sådan situation kan den statistikansvarige t.o.m., om forskaren så skulle önska, och efter vissa kontroller, som i huvudsak kan utföras automatiskt, överföra avidentifierade data till forskaren för lokala bearbetningar hos denne. Förbudet mot olovlig identifiering ger ett legalt skydd mot att forskaren genom mer eller mindre komplicerade operationer försöker "bakvägsidentifiera" de avidentifierade uppgifterna, om han eller hon

mot förmodan skulle ha något incitament att göra detta, och i kombination med den förebyggande, huvudsakligen automatiska sekretessgranskningen hos den statistikansvarige, ger detta ett mycket kraftfullt skydd för statistiksekretessen, fullt i klass med, ja troligen betydligt bättre än motsvarande traditionella, huvudsakligen manuellt administrerade skyddet mot röjanden hos den statistikansvarige.

#### *Avidentifiering och uttunning av känsliga data*

Avidentifiering av känsliga data är en utmärkt metod i många fall, och den kan ytterligare förstärkas vad gäller skyddet av statistiksekretessen, genom att man uttunnar informationsinnehållet i materialet, t.ex. genom att man redovisar kvantitativa variabelers värden i intervall i stället för genom exakta tal, och genom att man förgrovar klasserna för kvalitativa variabler. Dock är avidentifiering inte användbar, när man vill kombinera olika datamaterial med varandra på individnivå, innan man framställer statistiken. Sedan man genomfört den erforderliga samkörningen av datamaterialen, kan man dock göra en avidentifiering för att ytterligare minska riskerna för brott mot statistiksekretessen i de fortsatta bearbetningarna.

### 3.4.6 Varför hanteras inte statistiksekretessen effektivare?

Det finns alltså redan idag utmärkta tekniska metoder för att, inom ramen för en minst lika betryggande säkerställd statistiksekretess som tidigare, ge forskare och analytiker nästan lika effektiva, flexibla och billiga bearbetningsmöjligheter, som om de känsliga statistikmaterialen faktiskt låg hos dem själva. Varför utnyttjas då inte dessa möjligheter i någon större utsträckning?

Svaret på frågan finns kanske i följande omständigheter:

1. Oavsett hur omfattande åtgärder man vidtar till skydd för statistiksekretessen, så kvarstår alltid en om än aldrig så liten risk, att skyddet bryts. En illvillig person, extern forskare eller anställd hos den statistikansvarige, kan bryta mot lagar och avtal och kan rentav fysiskt bryta igenom aldrig så väl genomtänkta skydd. Det kan visserligen vara svårt att föreställa sig, att någon skulle ha något rimligt motiv att på detta sätt begå brottsliga gärningar, men teoretiskt kan man aldrig vara helt säker. Om inte annat kan en illvillig journalist, utsänd av en skandaltidning, vilja skapa sensation och upprördhet hos läsarna; det säljer bra.

2. Ett brott mot statistiksekretessen är katastrofalt för en statistikansvarig myndighet, eftersom det undergräver förtroendet för myndigheten hos uppgiftslämnarna och därmed allvarligt skadar den officiella statistikens trovärdighet och kvalitet. Det är därför föga förvånande om en statistikansvarig myndighet i avvägningen mellan sekretess och tillgänglighet har en tendens att prioritera sekretessen. Ett alltför frikostigt tillgängliggörande kan innebära katastrof för den statistikansvarige, om statistiksekretessen på något sätt äventyras. Ett alltför snålt tillgängliggörande medför i värsta fall viss kritik från en mycket begränsad grupp av statistikanvändare, som knappast kan uppstå någon stark opinion för sin sak.
3. Den statistikansvariga myndigheten litar mera på egna anställda än på externa forskare och analytiker.
4. Bearbetningar hos den statistikansvariga myndigheten för forskares och analytikers räkning genererar icke obetydliga intäkter, särskilt om de är arbetskrävande, och bidrar därmed till att säkra anställningstryggheten hos myndigheten ifråga, åtminstone kortsiktigt.
2. Ovanstående omständigheter har inte bara direkt betydelse, i samband med att ett konkret utlämnande- eller tillgängliggörandeärende skall avgöras hos den statistikansvariga myndigheten. De har också en allmänt tillbakahållande effekt hos myndigheten på sådant utvecklingsarbete, som skulle kunna bidra till ökad tillgänglighet inom ramen för bibehållen statistiksekretess.

### 3.4.7 Möjliga åtgärder

#### *Införande av tillgänglighetsprincip i kombination med kryptertolkningsutrymme*

De ovan uppräknade svårigheterna vid avvägningen mellan skyddet för statistiksekretessen och främjandet av statistik tillgängligheten skulle emellertid inte föreligga om

- (i) den ovan formulerade *tillgänglighetsprincipen för hantering av statistiksekretessen* tillämpades och övervakades strikt, så att hanteringen av statistiksekretessen i praktiken verkligen var så effektiv, så att den inte, t.ex. av tids- eller kostnadsskäl, onödigtvis inskränker tillgängligheten till statistiken för statistikanvändarnas egna bearbetningar och analyser
- (ii) lagstiftningen var glasklar och inte lämnade något utrymme för olika tolkningar hos den statistikansvariga myndigheten

Det är särskilt följande formuleringar i sekretesslagen som öppnar för ett betydande tolkningsutrymme:

- ”..., uppgift som behövs för forskningsändamål...,”
- ”..., uppgift, som inte genom namn, annan identitetsbeteckning eller därmed jämförbart förhållande är direkt hänförlig till den enskilde,”
- ”får dock lämnas ut, om det står klart att uppgiften kan röjas utan att den som uppgiften rör eller någon honom närstående lider skada eller men.”

Kursiveringarna i citaten är mina och pekar ut följande tolkningsproblem:

1. Vad är forskningsändamål?
2. Vad är det som kan vara jämförbart med namn eller annan identitetsbeteckning? Måste det vara ett begrepp, eller kan det vara en mer eller mindre långsökt kombination av egenskaper, som var för sig inte är identifierande, men som tillsammans kan vara det?
3. När är en uppgift direkt hänförlig till den enskilde? Kan ”hänförandet” bestå av en kortare eller längre logisk/matematisk deduktionskedja, och kan ”hänförandet” då anses vara ”direkt”?
4. Vad menas med ”skada eller men”? Med skada avses ”ekonomisk skada”, och med ”men” avses i första hand integritetsinskränkningar.<sup>38</sup> Bedömningen skall enligt förarbetena utgå från den berörda personens upplevelse, men den måste i viss mån kunna korrigeras på grund av gängse värderingar i samhället; det räcker inte att en person tycker, att det i största allmänhet är obehagligt, att andra vet något om honom eller henne.
2. När kan man anse att ”det står klart” att en uppgift kan röjas, utan att det leder till skada eller men? Är det t.ex. tillräckligt om uppgiften inte direkt kan hänföras till uppgiftslämnaren, utan att någon begår brottsliga gärningar, t.ex. genom att fysiskt forcera ett skalskydd, eller genom att bryta mot förbudet mot olovlig identifiering? (I så fall borde alla uppgifter om enskilda utan ytterligare överväganden kunna lämnas ut efter avidentifiering.)

Erfarenheter från SCB:s verksamhet bekräftar, att tolkningsramarna är vida, och praxis har följaktligen varierat en hel del över tiden. Efter integritetsdebatter, t.ex. i samband med FoB70 och senare i samband med Metropolitprojektet, har utlämnandeärenden hanterats extra restriktivt. När statsmakterna tillsammans med tunga forskarintressen har artikulert legitima statistikanvändarintressen på ett tydligt sätt, t.ex. i budgetpropositioner och regleringsbrev, har SCB gjort extra ansträng-

<sup>38</sup> Se Samuelson, a.a.

ningar att öka tillgängligheten till mikrodata för statistiskt bruk utan att därför göra avkall på statistiksekretessen. I början av 1990-talet bedrev SCB sålunda ett särskilt projekt, Tillgänglighetsprojektet, som hade till uppgift att ta fram tekniska lösningar på problemet att göra mikrodata tillgängliga för statistikbearbetningar på användarnas villkor, utan att man därför inkräktade på statistiksekretessen. Ett antal lösningar togs fram och prövades med goda resultat, men efter en tid "lätade trycket" från statsmakterna och användarna, och då rann denna verksamhet mer eller mindre ut i sanden. Lösningarna finns naturligtvis fortfarande kvar som praktiskt användbara möjligheter, men möjligheterna utnyttjas inte av de statistikansvariga i någon större omfattning, trots att det föreligger ett stort intresse för denna typ av lösningar från forskare och utredare.

Anpassningar åt bägge hållen (extremt restriktiva tolkningar och något mera frikostiga) har sålunda rymts inom en oförändrad lagstiftning, genom att man utnyttjat det ovan beskrivna tolkningsutrymmet. Av praxis kan man dra slutsatsen att statsmakterna måste vara mycket tydliga och pådrivande för statistik användarnas intressen, om en statistikansvarig myndighet skall anstränga sig för att utnyttja alla möjligheter, som faktiskt finns, att öka tillgängligheten till statistiska mikrodata. Det är både bekvämare och riskfriare, ja kanske t.o.m. lönsammare för den statistikansvarige, att avstå från dessa möjligheter.

Statsmakterna skulle t.ex. årligen kunna avkräva statistikansvariga myndigheter redogörelser för vad de gjort under det gångna året för att öka tillgängligheten till den officiella statistiken – och då särskilt tillgängligheten till mikrodata för statistiska bearbetningar och analyser på användarnas egna villkor och inom ramen för statistiksekretessen.

#### *Inrättande av central registerfunktion inom databasverksamheten*

En annan åtgärd för att öka tillgängligheten till mikrodata för statistiska bearbetningar och analyser kunde vara, att statsmakterna uppdrar åt SCB att ge det pågående Registerprojektet en sådan inriktning, att det inte bara underlättar för SCB att internt utföra olika bearbetningar och samkörningar av SCB:s register, utan att man också bygger upp en externt riktad registerfunktion, förslagsvis inom ramen för den databas-service, som redan byggts upp och fungerar kring Sveriges statistiska databaser.

En registerfunktion skulle kunna finansieras enligt förebild från databasverksamheten, d.v.s. statsmakterna beviljar ett anslag för uppbyggandet och vidmakthållandet av själva infrastrukturen, medan marginalkostnaderna för uttag och bearbetningar bekostas av användarna. Infrastrukturen skulle utnyttja alla tekniska möjligheter, så att

användarna, om de så önskar, själva kan utföra så mycket som möjligt av arbetet med de statistiska bearbetningarna och analyserna, ungefär på det sätt som kortfattat beskrevs tidigare i detta avsnitt. Kostnaderna för mikrodatabearbetningarna skulle därigenom kunna hållas på en mycket låg nivå, genom att SCB-personal inte behöver engageras i någon större utsträckning. En lösning av detta slag förutsätter dock också att detaljerade dokumentationer (metadata och processdata) och kvalitetsdeklarationer av tillgängliga mikrodata och bakomliggande processer ingår i infrastrukturen.

En externt riktad registerfunktion av ovan beskrivet slag har efterlysts av bl.a. SCB:s vetenskapliga råd och av Databasrådet, en extern referensgrupp för Sveriges statistiska databaser.

### 3.5 Lagar och regler

Den rättsliga regleringen av tillgängligheten till Sveriges officiella statistik sker bl.a. genom följande lagar, förordningar, föreskrifter, riktlinjer m.fl. typer av styrningsinstrument:

- Tryckfrihetsförordningen
- Sekretesslagen
- Lagen om den officiella statistiken
- Förordningen om den officiella statistiken
- Föreskrifter och allmänna råd om offentliggörande av officiell statistik
- Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik

De olika styrningsinstrumenten är av olika dignitet. Tryckfrihetsförordningen, som stadfäster offentlighetsprincipen, är grundlag, och Sekretesslagen är, liksom Lagen om den officiella statistiken, lagar beslutade av Sveriges riksdag. Förordningen om den officiella statistiken är utfärdad av regeringen, medan de övriga ovan nämnda instrumenten har utfärdats av SCB genom serien Meddelanden i samordningsfrågor (MIS). Därutöver publicerar SCB som ett resultat av sitt generella metodarbete fortlöpande rapporter och promemorior av relevans för tillgänglighetsfrågorna. Det sker t.ex. genom serien R&D Reports, gröna serien (statistiska metoder) och gula serien (statistiska informationsbehandlingsmetoder).

### 3.5.1 Tryckfrihetsförordningen och Sekretesslagen

Offentlighetsprincipen enligt Tryckfrihetsförordningen behandlades liksom Sekretesslagen i avsnitt 3.4. Statistiksekretessen enligt Sekretesslagen sammanfattades i figur 18.

<p><b>Lagen om den officiella statistiken</b> enligt riksdagsbeslut</p>
<p><i>Framställning av officiell statistik</i></p> <ul style="list-style-type: none"> <li>• Lagen tillämpas när statliga myndigheter under regeringen framställer officiell statistik.</li> <li>• Definition av officiell statistik (1§): Officiell statistik är sådan statistik för samhällsplanering, forskning, allmän information och internationell rapportering, som som en myndighet framställer enligt föreskrifter som regeringen meddelar.</li> <li>• Krav på objektivitet, tillgänglighet och statistiksekretess (2§)</li> </ul> <p><i>Uppgiftsskyldighet</i></p> <ul style="list-style-type: none"> <li>• Specificering av fall där uppgiftsskyldighet föreligger (3-8§).</li> <li>• Tvångsmedel och straff (11-14§)</li> </ul> <p><i>Uppgiftsskydd</i></p> <ul style="list-style-type: none"> <li>• Hänvisningar till sekretesslagen och datalagen (9§).</li> <li>• Förbud mot olovlig identifiering (10§): Uppgifter i den officiella statistiken får inte sammanföras med andra uppgifter i syfte att utröna enskildas identitet.</li> <li>• Straff vid brott mot förbudet mot olovlig identifiering (15§)</li> </ul>

Figur 19. Sammanfattning av innehållet i Lagen om den officiella statistiken.

### 3.5.2 Lagen om den officiella statistiken

Innehållet i Lagen om den officiella statistiken sammanfattas i figur 19. Ur tillgänglighetssynpunkt är det särskilt 2§, 9-10§ samt 15§ som är av intresse.



### 3.5.3 Förfordningen om den officiella statistiken

Innehållet i Förfordningen om den officiella statistiken sammanfattas i figur 20. Där preciserar regeringen genom explicit uppräknig, vilken statistik som skall ingå i Sveriges officiella statistik. I övrigt är det 8-9§ som är av intresse. Där befullmäktigas SCB att meddela föreskrifter om hur

- den officiella statistiken skall dokumenteras, kvalitetsdeklareras och hållas allmänt tillgänglig (8§)
- den officiella statistiken skall offentliggöras i serien Sveriges officiella statistik (9§)

### 3.5.4 Föreskrifter och allmänna råd om offentliggörande av officiell statistik

Med stöd av dessa befullmäktiganden har SCB utfärdat närmare föreskrifter och allmänna råd i *"Föreskrifter och allmänna råd om offentliggörande av officiell statistik"*, vars innehåll sammanfattas i figur 21.

Av figur 21 framgår bl.a. att den statistikansvarige fritt kan välja mellan att publicera officiell statistik

- på traditionellt sätt, i tryckt form; och/eller
- på ADB-medium

Oavsett vilken form den statistikansvarige väljer, ställs i princip samma krav på offentliggörandet och på åtföljande dokumentation.

Man kan fråga sig, om det inte är dags att kräva att all officiell statistik skall offentliggöras genom Sveriges statistiska databaser. Därutöver skulle det naturligtvis stå den statistikansvarige fritt att dessutom offentliggöra genom andra ADB-medier och genom traditionella tryckta publikationer.

Av figur 21 framgår också, att det inte är obligatoriskt att ta fram fullständiga dokumentationer av de observationsregister, som ligger till grund för (och utgör ett resultat av) den officiella statistiken. Enligt *"Föreskrifter och allmänna råd om offentliggörande av officiell statistik"* bör den statistikansvarige framställa sådan dokumentation, medan man skall framställa kvalitetsdeklaration och produktbeskrivning. "Börkravet" bör skärpas till ett "skallkrav" för att underlätta för kvalificerad statistik användare att analysera och vidarebearbeta statistiken på egen hand, i enlighet med vad som diskuterats tidigare i denna uppsats.

**Förordningen om den officiella statistiken**

utfärdad av regeringen

*Den officiella statistikens omfattning och innehåll:*

- örordningen tillämpas när statliga myndigheter under regeringen framställer officiell statistik enligt lagen om officiell statistik (1§).
- Uppräkning av den officiella statistiken fördelad på 22 ämnesområden och med angivande av ansvarig myndighet för respektive produkt, som ingår i den officiella statistiken (2§+bilaga).
- Väsentliga förändringar av officiell statistik skall prövas av regeringen (2§).
- Individbaserad statistik skall, när inte särskilda skäl talar däremot, vara könsuppdelad (10§)

*Uppgiftslämnandet:*

- Myndighet som samlar in uppgifter till den officiella statistiken skall eftersträva att uppgiftslämnandet inte blir onödigt betungande (3§).
- SCB får lämna ytterligare föreskrifter om uppgiftsskyldighet (4§).
- Myndighet är skyldig att samråda med SCB, (i) innan den börjar samla in uppgifter till den officiella statistiken; och (ii) vid överlämnande av sådana uppgifter mellan myndigheter (5§).
- Vid inhämtande av uppgifter från näringsidkare och kommuner föreligger samrådsskyldighet med organisationer, som företräder dessa uppgiftslämnare (5§).
- Myndighet som samlar in uppgifter till den officiella statistiken från någon annan än statlig myndighet skall informera om (6§)
  - ändamålet med uppgiftsinsamlandet
  - grunder för uppgiftsskyldighet, om uppgiftsskyldighet föreligger
  - att uppgiftslämnandet är frivilligt, om uppgiftsskyldighet inte föreligger
  - huruvida samråd har skett med den organisation som företräder uppgiftslämnaren
  - vad som gäller om uppgiftsskydd
  - vad som gäller om uppgifternas bevarande
  - eventuella påföljder om uppgifterna inte lämnas
  - andra förhållanden som är av betydelse i sammanhanget.
- Myndighet som samlar in uppgifter för annat ändamål än officiell statistik, men som också skall användas för officiell statistik, skall informera om detta (7§)

*Den officiella statistikens tillgänglighet:*

- Den officiella statistiken skall dokumenteras, kvalitetsdeklareras och hållas allmänt tillgänglig enligt föreskrifter som SCB meddelar (8§).
- Myndighet skall offentliggöra den officiella statistiken i serien Sveriges officiella statistik enligt föreskrifter som SCB meddelar (9§).

**Figur 20. Sammanfattning av innehållet i Förordningen om den officiella statistiken.**

**Föreskrifter och allmänna råd om offentliggörande av officiell statistik**

enligt beslut av SCB (Meddelanden i samordningsfrågor 1995:5)

*Offentliggörande och prissättning av officiell statistik*

- Officiell statistik skall offentliggöras i serien Sveriges officiella statistik (SOS) i tryckta publikationer eller på ADB-medium (1§).
- Officiell statistik skall offentliggöras i SOS så snart den är framställd (2§).
- Vissa bibliotek skall få den officiella statistiken gratis. I övrigt skall endast marginalkostnader tas ut, t.ex. kostnader för mångfaldigande och distribution (3§).

*Dokumentation av officiell statistik*

- När officiell statistik offentliggörs skall den även omfatta de förklaringar som behövs för att användarna skall kunna förstå statistiken (4§).
- Officiell statistik skall åtföljas av en kvalitetsdeklaration och en produktbeskrivning, som fylls i enligt formulär som SCB tillhandahåller (5§). Även datamaterial som ligger till grund för den officiella statistiken bör dokumenteras. Dokumentationen bör ha en sådan omfattning och detaljeringsgrad att den möjliggör framtida användning av bevarade datamaterial, s.k. observationsregister. Produktbeskrivningar och kvalitetsdeklarationer skall offentliggöras och hållas tillgängliga på samma sätt som den officiella statistiken. Detta sker genom att uppgifterna lagras i den gemensamma databas, som handhas av SCB:s databasverksamhet. Riktlinjer för kvalitetsdeklaration av officiell statistik ges i MIS 1994:3. Riktlinjer för hur man dokumenterar observationsregister tillhandahålls också av SCB.

- Myndigheter som ansvarar för officiell statistik skall årligen till SCB lämna förteckningar över dels planerad, dels genomförd publicering av officiell statistik (11§).

*Formalia-*

- Vid offentliggörande i SOS skall texten ”Sveriges officiella statistik” samt utgivande myndighet, utgivningsort och årtal anges (6§).
- Tryckta publikationer i SOS skall innehålla uppgift om tidigare publicering i SOS, om det inte är uppenbart att detta är onödigt (7§).
- Tryckta publikationer i SOS skall innehålla översättningar till engelska (8§). Översättningar till engelska bör även finnas i de redovisningar som görs på ADB-medium.
- I SOS ingår publikationsserien Statistiska meddelanden (SM), för vilken SCB är ansvarig utgivare (9§). Även andra statistikansvariga myndigheter kan publicera i SM-serien. En myndighet kan också ge ut en egen serie, där publikationer i SOS ingår.

*Offentliggörande på ADB-medium-*

- Vid offentliggörande av officiell statistik på ADB-medium skall ett standardiserat format användas (10§). Standardiserade format är nödvändiga bl.a. för att man skall kunna integrera statistik från olika ämnesområden i databaser och för en effektiv internationell rapportering.

Figur 21. Sammanfattning av innehållet i MIS 1995:5.

### 3.5.5 Personregister för officiell statistik

Lagen (SFS 1995:606, ändrad t.o.m. SFS 1998:1433) respektive Förordningen (SFS 1995:1060, ändrad t.o.m. SFS 1998:1330) om vissa personregister för officiell statistik reglerar, hur personuppgifter i personregister får hanteras för framställning av officiell statistik.

I bilagan till förordningen anges för ett antal statistikansvariga myndigheter

- vilka personregister som respektive myndighet får ha för officiell statistik
- vad som är ändamålet för respektive register
- vilken information som respektive register får innehålla
- med vilka andra register respektive register får sambearbetas

Av speciellt intresse, när det gäller tillgängligheten till datamaterialen i Sveriges officiella statistik är punkt 1.15 i bilagan till förordningen. Den lyder som följer:

”1.15 Register för databasservice

#### Ändamål

Registren får utgöra grund för Statistiska centralbyråns (SCB) databasservice enligt 2 och 3 §§ förordningen (1988:137) med instruktion för centralbyrån.

#### Innehåll

Registren för databasservice består av kopior av följande personregister

- Registret över totalbefolkningen (RTB),
- Registret över befolkningsförändringar vid SCB,
- personregister som inrättats för folk- och bostadsräkningar,
- Fastighetstaxeringsregistret vid SCB,
- Utbildningsregistret vid SCB,
- registret Arbetskraftsundersökningarna (AKU),
- Registret över kontrolluppgifter,
- Centrala företags- och arbetsställeregistret (CFAR) vid SCB,
- Utrikeshandelsregistret vid SCB,
- Lantbruksregistret (LBR) vid SCB,
- Inkomst och förmögenhetsregistret (IoF),
- Årliga sysselsättningsregistret (ÅRSYS) vid SCB,
- Registret över utbildning-arbetsmarknad (UA-registret),
- För brott lagförda personer,
- Levnadsförhållandena i samhället (ULF),
- Mervärdesskatteregistret vid SCB,
- Bilstatistikregistret vid SIKÄ.

#### Sambehandling

Ett register för databasservice får sambehandlas med ett annat sådant register.”

Här finns sålunda den legala grunden för möjligheterna att inom ramen för SCB:s databasservice sambehandla den officiella statistikens observationsregister med varandra för statistikändamål. Förteckningen ovan kan antas omfatta de register, som av statistikanvändarna bedöms mest angelägna att få tillgång till för statistikbearbetningar. I den mån ytterligare behov uppkommer, kan förordningen i vederbörlig ordning kompletteras.

### 3.5.6 Begreppet ”Sveriges officiella statistik”

Begreppet ”Sveriges officiella statistik” har stor betydelse för att hålla samman det svenska statistiska systemet. Det är den delmängden av den statliga statistiken, som större delen av den ovan sammanfattade lagstiftningen tar sikte på, och det är den statistiken, som alla aktörer i det svenska samhället, de vanliga medborgarna likaväl som professio-

nella statistik användare i näringsliv och förvaltning, i första hand har att förlita sig som en gemensam och objektiv referensram och informationstillgång av god kvalitet och hög tillgänglighet.

Det finns goda skäl att något diskutera begreppet "Sveriges officiella statistik", inte minst ur tillgänglighetssynpunkt. I samband med att statsmakterna och statistik användarna artikulera allt tydligare krav på den officiella statistikens tillgänglighet, har det av och till funnits tendenser hos statistikansvariga att mildra effekterna av dessa krav genom att helt enkelt snäva in tolkningen av vad som skall räknas som officiell statistik.

Möjligheterna för statistikansvariga att göra vidare eller snävare tolkningar av vad som skall räknas som officiell statistik har sin grund i att begreppet inte är särskilt skarpt definierat i lagstiftningen. I de lagar och förordningar, som vi nyss gått igenom, finns två principiellt olika definitionsansatser:

- en intentionell, d.v.s. syftesorienterad ansats
- en extensionell, d.v.s. explicit uppräknande ansats

#### *En intentionell, syftesorienterad definitionsansats*

Den intentionella definitionsansatsen finns i "Lagen om den officiella statistiken" (jfr figur 19). I lagens två första paragrafer sägs sålunda:

- 1 § ... Officiell statistik är sådan statistik för samhällsplanering, forskning, allmän information och internationell rapportering som en myndighet framställer enligt föreskrifter som regeringen meddelar.
- 2 § Den officiella statistiken skall vara objektiv och allmänt tillgänglig. Den skall framställas och offentliggöras med beaktande av skyddet för enskilda.

Den första paragrafen uttrycker den officiella statistikens syfte. Den andra paragrafen ställer krav på objektivitet, allmän tillgänglighet och skydd för statistiksekretessen. Man kan sålunda bl.a. notera att kravet på tillgänglighet är ett krav av mycket hög dignitet i lagstiftningen.

I Förordningen om den officiella statistiken ställs ytterligare krav på statistikens tillgänglighet:

- 8 § En officiella statistiken skall dokumenteras, kvalitetsdeklareras och hållas allmänt tillgänglig enligt föreskrifter som Statistiska centralbyrån meddelar.
- 9 § En myndighet skall offentliggöra den officiella statistiken i serien Sveriges officiella statistik enligt föreskrifter som Statistiska centralbyrån meddelar.

Det kan möjligen förvåna, att det inte ställs några uttryckliga krav på den officiella statistikens kvalitet utöver att den skall vara objektiv, vilket naturligtvis i sig är ett mycket starkt krav, och att den skall vara dokumenterad, kvalitetsdeklarerad och allmänt tillgänglig.

#### *En extensionell, explicit uppräknande definitionsansats*

Den extensionella definitionsansatsen finns i "Förordningen om den officiella statistiken" (jfr figur 20). Där sägs i andra paragrafen:

2 § Officiell statistik är sådan statistik som förtecknas i bilagan. I bilagan anges också den myndighet som svarar för statistiken. Om en myndighet föreslår väsentliga förändringar i den officiella statistiken, skall förslaget prövas av regeringen.

I den nämnda bilagan räknas sedan samtliga statistikprodukter upp, som ingår i den officiella statistiken. De är sorterade i 22 ämnesom-råden. Av uppräknningen framgår inte, vilken omfattning och detaljeringsgrad, som statistikprodukterna skall ha. Sålunda anges varken vilka populationer eller vilka variabler, som skall undersökas. Dock skall alla förslag till väsentliga förändringar i statistiken prövas av regeringen.

#### *Möjliga åtgärder för att precisera innebörden av "officiell statistik"*

Ett par åtgärder som skulle kunna vidtas för att precisera innebörden av begreppet "Sveriges officiella statistik" och samtidigt skärpa kraven på dess tillgänglighet är följande:

1. Lagen skulle kunna säga, att all statistik, som framställs med hjälp av anslag från statsmakterna, skall vara officiell statistik, om det inte är uppenbart, att statistiken saknar intresse utanför myndigheten som framställer den.
2. Lagen skulle kunna tydliggöra, att med "officiell statistik" avses inte bara summerade sammanställningar av det slag, som brukar redovisas i traditionella statistikpublikationer, utan även de observationsregister och redovisningar på lägre aggregeringsnivåer, som ligger till grund för sådana summeringar.

Observationsregister och detaljerade redovisningar måste naturligtvis tillgängliggöras på ett sådant sätt att ett gott skydd för statistiksekretessen upprätthålls, t.ex. genom att man vidtar sådana åtgärder för effektivare hantering av statistiksekretessen som föreslogs i avnitt 3.4 i föreliggande rapport.

Den officiella statistiken skulle sålunda bestå av

- makrodata: skattningar av statistiska storheter, d.v.s. det som i dagligt tal kallas "statistik" eller aggregerade uppgifter
- mikrodata: de observationer, organiserade i observationsregister, som ligger till grund för framställningen av den officiella statistikens makrodata
- metadata: dokumentation, som beskriver den officiella statistiken och de bakomliggande datamaterialen och processerna på ett så fullständigt sätt, att även statistikanvändare, som inte själva deltagit i framställningen av den officiella statistiken, självständigt kan (åter) använda såväl mikrodata som makrodata på ett statistiskt ansvarsfullt sätt

För att ytterligare tydliggöra begreppet "Sveriges officiella statistik", och för att säkerställa den officiella statistikens användbarhet och kvalitet, skulle man kunna utöka listan med t.ex. följande krav och egenskaper:

3. Den officiella statistikens olika delar skall statistiskt belysa viktiga samhällsområden, deras tillstånd och utveckling. Sammantaget skall den officiella statistiken ge en sammanhängande statistisk beskrivning av det svenska samhället som helhet.
4. Statsmakterna fastställer vilka samhällsområden, som skall beskrivas med hjälp av officiell statistik, och (efter förslag från statistikansvarig myndighet) vilka statistikprodukter som skall ingå i den officiella statistiken för respektive samhällsområde. Samma statistikprodukt kan ingå i den officiella statistiken för flera samhällsområden.
5. Den officiella statistiken skall ha både djup och bredd, så att den täcker såväl specialisters behov av djup och högkvalitativ information om viktiga samhällsområden som generalisters (och allmänhetens) behov av bred och lättillgänglig information.
6. Officiell statistik skall uppfylla vissa objektivitets- och kvalitetskriterier, som fastställs av statsmakterna efter förslag från den centrala statistikansvariga myndigheten (SCB).
7. Officiell statistik och bakomliggande observationsregister (mikrodata) får bara användas för statistiska ändamål, men skall å andra sidan på ett så smidigt och effektivt sätt som möjligt tillgängliggöras för statistisk användning och statistiska bearbetningar hos statistikanvändarna. Den som är statistikansvarig för officiell statistik skall med lämpliga juridiska, administrativa och tekniska åtgärder säkerställa att statistiksekretessen upprätthålls, när statistiken och bakomliggande datamaterial tillgängliggörs för statistikanvändarna. Vare sig statistiksekretessen eller prissättningsinstrumentet får dock onödigtvis begränsa tillgänglighöret. Tvärtom skall den statistikansvarige utnyttja bästa kända metoder för att på alla sätt underlätta för statistikanvändarna att på det sätt de själva önskar få genomföra statistiska bearbetningar med bibehållen statistiksekretess.
8. Officiell statistik finansieras normalt med anslag från statsmakterna och tillgängliggörs sedan den producerats för användarna enligt marginalkostnadsprincipen, d.v.s. statistikproducenten tar bara betalt för de kostnader för spridning m.m., som inte täcks av anslaget för produktionen. Prissättningen bör sålunda följa huvudprincipen i



RRV:s förslag till prisättning av informationstjänster. Prissättningen av officiell statistik och av produkter och tjänster baserade på officiell statistik övervakas av tillsynsorganet för det svenska statistiska systemet och av riksrevisionsverket.

9. Statistik som finansierats på annat sätt kan i efterhand av statsmakterna definieras som officiell statistik. Den skall därefter prissättningsmässigt behandlas på samma sätt som om den hade varit anslagsfinansierad, och den skall även i övrigt uppfylla alla de krav, som gäller för officiell statistik.

## 3.6 Organisationsfrågor

### 3.6.1 Centraliserat eller decentraliserat system?

Med bra tekniska lösningar (baserade på Internet etc), bra standarder och bra krav på den officiella statistiken blir frågan om en mer eller mindre (de)centraliserad organisation av den officiella statistiken mindre betydelsefull ur tillgänglighetssynpunkt. Dock måste en effektiv övervakning ske, och det övervakande organet får inte ha sådana egenintressen, som lätt kan komma i konflikt med rollen att bevaka statistikens tillgänglighet för användarna. Övervakaren av den officiella statistiken skall stå på statistikanvändarnas och uppgiftslämnarnas sida.

#### *Argument för en centraliserad statistikorganisation*

Några huvudargument, som brukar anföras till förmån för en centraliserad organisation av statistikproduktionen, är:

- knappa och svårindelbara resurser (t.ex. datorkraft och metodkompetens)
- stordriftsfördelar, särskilt i datainsamlingsledet
- professionalism
- innehållsmässig samordning och standardisering
- samkörningsmöjligheter
- statistiksekretessen

När det gäller datorkraften, är denna som bekant inte längre vare sig knapp eller svårindelbar. Även metodkompetens (främst statistiska metoder, men även andra typer av metodkompetens som har betydelse för statistikproduktionen) borde idag kunna tillhandahållas i tillräcklig omfattning av den högre utbildningen. Däremot kan små organisationer (eller organisationer där statistiken utgör en liten del av verksamheten) ha svårt att "bära" de specialister som skulle behövas; de kan t.o.m. ha

svårt att hålla sig med den upphandlingskompetens, som behövs för att köpa in kompetensen, om man inte har råd att hålla sig med den i egen regi. Man måste också beakta att en grupp specialister måste ha en viss kritisk massa för att på sikt vidmakthålla sin kompetens.

En centraliserad organisation av statistiken erbjuder vissa stordriftsfördelar, särskilt i datainsamlingsledet. Vissa moment i datainsamlingen är fortfarande mycket resurskrävande. Främst gäller det direktinsamling av data med hjälp av intervjuare. Det finns visserligen numera en fungerande kommersiell marknad för datainsamling, men den officiella statistiken ställer höga krav på kvaliteten i datainsamlingen, och det är inte så lätt för t.ex. en liten statistikansvarig myndighet att vara en kompetent beställare gentemot kommersiella aktörer; det är lätt hänt att man alltför ensidigt fokuserar på priset och därmed missar viktiga kvalitetsaspekter. En liten statistikorganisation har inte heller råd att organisera vissa typer av datainsamling i egen regi.

Det enklaste sättet att drastiskt sänka kostnaderna för datainsamling är att återanvända data, som redan finns tillgängliga i elektronisk form, t.ex. hos företagare, som är uppgiftslämnare, eller hos andra myndigheter, som samlar in data för administrativa ändamål. Även när man samlar in data på sådana sätt, är det emellertid nödvändigt att besitta kompetens, som kan avgöra, om dessa data kan användas för statistiska ändamål, och vilka kvalitetsåtgärder, som i så fall måste vidtas.

Ett återkommande argument för en centraliserad statistikorganisation är behovet av professionalism i produktionen av officiell statistik. Det finns visserligen god utbildning i de flesta av de metoddiscipliner, som är relevanta för statistikproduktionen (statistik, informationsbehandling, mätteknik, m.fl.), men det är först när dessa metoder integreras med varandra, som man kan tala om en metodik för statistikproduktion, och det är strängt taget bara inom en stor, praktiserande statistikorganisation, som en sådan metodik kan växa fram och utvecklas.

Många samhällsfrågor är så pass komplexa, att de behöver belysas av en kombination av statistik från flera olika områden, t.ex. befolkningsstatistik och ekonomisk statistik. För att kunna kombinera statistik från olika områden med varandra, och för att kunna jämföra statistik över tiden och mellan olika regioner och länder, är det viktigt, att man använder sig av samma definitioner. Detta försöker professionella statistikorganisationer uppnå genom att etablera och använda sig av standarddefinitioner av objekttyper, variabler, klassifikationer m.m. Man eftersträvar en övergripande innehållsmässig samordning av den officiella statistiken, så att den "hänger ihop" begreppsmässigt; i bästa fall kan man på detta sätt etablera konsistenta statistiska modeller för större eller mindre områden av samhället. Nationalräkenskaperna är kanske det bäst kända exemplet på en sådan modell.

Samordning, standardisering och modellbyggande kräver uppenbarligen ett visst mått av centralisering.

Innehållsmässig samordning är som sagt en nödvändig förutsättning för att man skall kunna kombinera statistik från olika områden. Är det en tillräcklig förutsättning? Knappast. Även om tekniken har blivit oerhört mycket billigare och kraftfullare, så ställs man fortfarande inför en rad praktiska, tekniska problem, när man vill samköra olika datamaterial med varandra. Även ett visst mått av teknisk samordning är sålunda nödvändig för att möjliggöra sådana statistikbearbetningar. Återigen är standardisering, i detta fall teknisk standardisering av gränssnitt, lagringsformat m.m., ett viktigt medel, och det behövs professionella organ, som kan driva på framväxten och tillämpningen av sådana standarder.

Även om väldefinierade (innehållsmässiga och tekniska) standarder och gränssnitt på sikt gör det allt lättare att samköra statistiska data från olika organisationer, så kommer det säkert under överskådlig tid att vara något enklare att göra samkörningar inom en organisation än mellan organisationer. Värnandet om statistiksekretessen är ytterligare ett argument för att se till att vissa typer av samkörningar görs inom en och samma organisation.

#### *Argument för en decentraliserad statistikorganisation*

Några huvudargument för en decentraliserad statistikorganisation är:

- bättre anpassning av statistiken till användarnas behov
- mera optimal prioritering genom att statistiken vägs mot andra behov inom ett sakområde – inte mot annan statistik
- flera inslag av tävlan och konkurrens – till förmån för nytänkande och effektivitet

Jag skulle vilja lägga till ytterligare ett argument:

- en decentraliserad statistikorganisation synliggör ett antal problem, som finns i varje statistiskt system, och påkallar därmed på ett starkare sätt lösningar på problemen, än vad som kanske skulle ske inom ramen för en enda, mera monolitisk organisation

Vad jag menar är att problem, som t.ex. sammanhänger med statistik-användarnas behov av innehållsmässig och teknisk samordning m.m., uppenbarligen föreligger vare sig man har en centraliserad eller en decentraliserad statistikorganisation. I en centraliserad statistikorganisation finns problemen inom en och samma organisationen och är därmed i teorin – och kanske även i praktiken – lättare att lösa. Erfarenheter från centrala statistikorganisationer visar dock med stor tydlighet, att

dessa förhållandevis bättre förutsättningar i en centraliserad organisation ingalunda är någon garanti för att problemen verkligen blir lösta. För att man skall lösa dem på ett bra sätt, krävs i själva verket inomorganisatoriska åtgärder, som är ganska likartade de mellanorganisatoriska åtgärder som behövs för att lösa motsvarande problem i ett decentraliserat statistiskt system. I bägge fallen handlar det bl.a. om att etablera och tillämpa standarder och att få olika system och delsystem att kommunicera och samspela med varandra på ett smidigt sätt.

*Syntes: balanserad kombination av decentrala och centrala komponenter inom ramen för ett integrerat statistiskt system*

Valet mellan en centraliserad och en decentraliserad organisation har vissa likheter med valet mellan ett hårt integrerat, skräddarsytt programvarusystem och ett system där man på ett lösare sätt, med hjälp av standardiserade gränssnitt, har sammanfogat flera självständiga programvaror till en helhet; se figur 13 i avsnitt 3.2.3.

Varje väl fungerande statistiskt system behöver sålunda ha vissa egenskaper och funktioner, vare sig det statistiska systemet är centraliserat eller decentraliserat. Detta gäller för att ett statistiskt system generellt skall fungera väl, och det gäller speciellt för att systemet skall fungera väl ur tillgänglighetssynpunkt.

Man kan diskutera huruvida olika grader av centralisering eller decentralisering gör det lättare eller svårare att säkerställa att systemet verkligen får dessa önskvärda egenskaper och funktioner. Decentralisering synliggör behoven, medan centralisering möjligen ökar chanserna att egenskaperna och funktionerna uppstår spontant. I bägge fallen behövs i alla fall likartade åtgärder, även om åtgärderna i det ena fallet huvudsakligen är inomorganisatoriska, medan de i det andra fallet är mellanorganisatoriska eller överorganisatoriska.

Det svenska statistiska systemet kan beskrivas som en balanserad kombination av decentrala och centrala komponenter inom ramen för ett integrerat statistiskt system.

### 3.6.2 Det svenska statistiska systemet, dess omvärld och delsystem

När man diskuterar det svenska statistiska systemet och dess organisation, måste man ta hänsyn till ett antal aktörer, system och delsystem (se även figur 22):

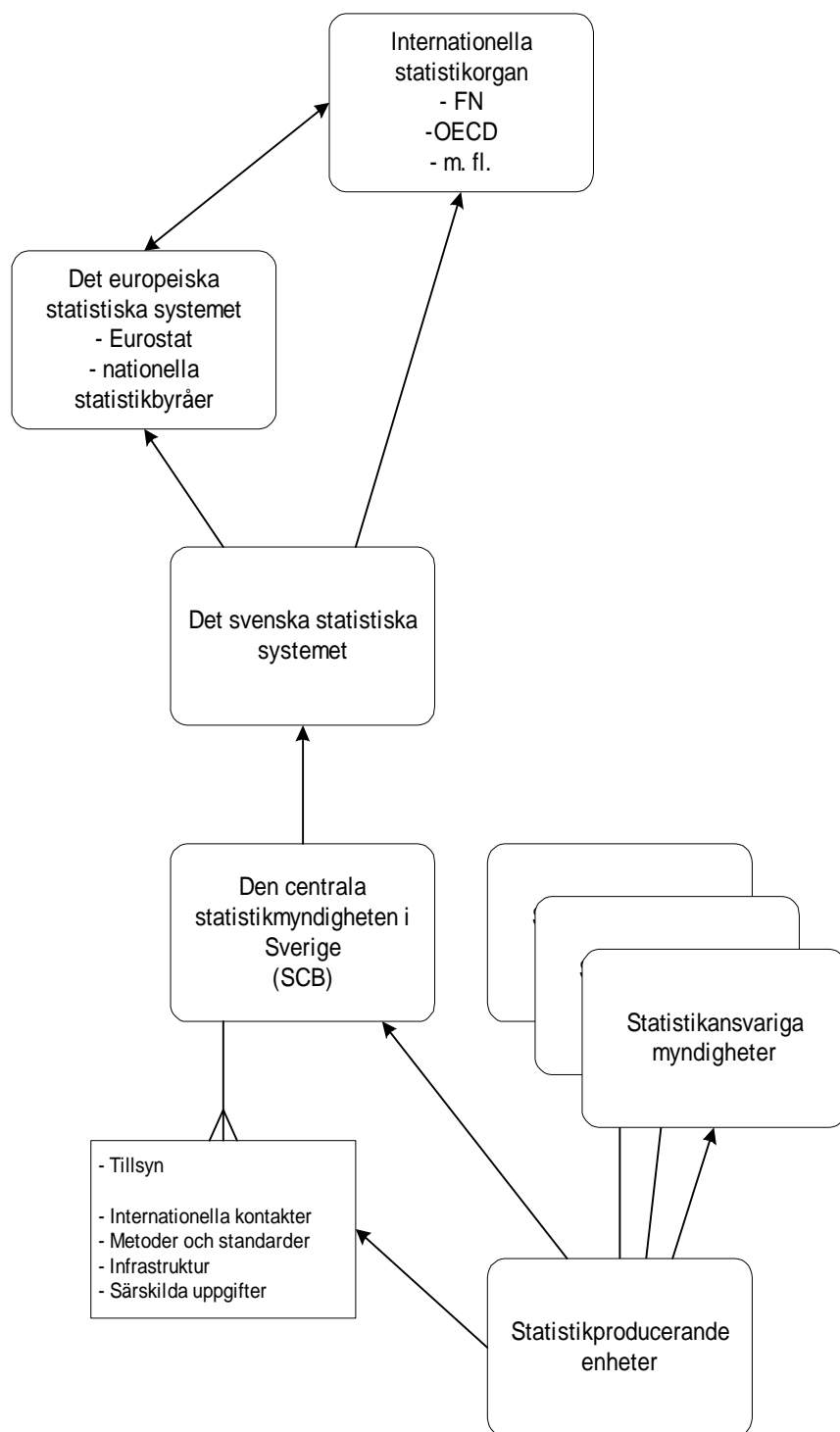
- Internationella statistikorgan (FN, OECD m.fl.)
- Det europeiska statistiska systemet, representerat av Eurostat
- Det svenska statistiska systemet som helhet
- Den centrala statistikmyndigheten i Sverige (SCB)
- Statistikansvariga myndigheter
- Statistikproducerande enheter

#### *Internationella statistikorgan*

På den internationella arenan är idag Eurostat den dominerande kraften, vad gäller inflytande över den svenska statistiken. Man har både pengar och makt. OECD har pengar men ingen större makt över statistiken. FN har varken pengar eller makt.

#### *Det europeiska statistiska systemet*

Eurostat har ägnat mycket tid och resurser åt att skissera det europeiska statistiska systemet. Eurostats nuvarande filosofi illustreras av figur 8f i avsnitt 1.5. Två nyckelbegrepp i filosofin är



Figur 22. Det svenska statistiska systemet: omgivning och delsystem.

- DSIS, som står för "Distributed Statistical Information System" eller "Distributed Statistical Information Services"
- ERE, som står för "European Reference Environment"

Man tänker sig det europeiska statistiska systemet som ett distribuerat informationssystem med Eurostat i en samordnande och i viss mån styrande roll. De nationella statistiksystemen beskrivs som <P, R, D>-tripler, där

- P står för "production"
- R står för "reference"
- D står för "distribution"

Det är i första hand genom R-komponenterna, som de nationella statistiska systemen knyts ihop med varandra och med det europeiska statistiska systemet. Mera konkret står R-et för referensdatabaser liknande Sveriges statistiska databaser, innehållande i första hand makrodata och metadata. Eurostat har av och till också uttryckt önskemål om att för det europeiska statistiska systemets räkning få tillgång till mikrodata, något som Sverige ställt sig avvisande till.

P- och D-komponenterna i de nationella statistiksystemen ses mera som lokala angelägenheter. Dock har Eurostat ibland gjort ansatser att vilja standardisera även produktionsmetoderna för officiell statistik, d.v.s. inte bara innehållet i statistiken utan också processerna och metoderna för hur statistiken skall tas fram. Man har t.ex. tidigare ställt sig negativ till den skandinaviska ansatsen att i stor utsträckning basera officiell statistik på administrativa data, och man har rentav ibland velat tvinga medlemsländerna att i stället använda sig av urvalsundersökningar, vilket skulle bli mycket dyrare. På senare tid synes Eurostat dock ha svängt i denna fråga och ser fördelarna, inte minst ur kostnadsynpunkt, med att använda administrativa data i statistikproduktionen.

I praktiken har utvecklingen av ett europeiskt statistiskt system enligt Eurostats intentioner ännu inte kommit särskilt långt. Eurostat har visserligen drivit en lång rad kostsamma projekt inom databas- och metadataområdena, men resultaten av dessa ansträngningar kan i stort sett bedömas som undermåliga. Det beror bl.a. på att samordningen inom Eurostat är mycket dålig för att inte säga obefintlig, och att man i alltför stor utsträckning lagt ut arbetet på kommersiellt arbetande konsultfirmor, som saknar erforderlig kompetens, och som fått en alltför svag styrning från Eurostat. Medlemsländerna har först under senare tid – och efter en hel del bråk – fått ett visst inflytande över utvecklingsarbetet och konsultinsatserna. Detta inflytande är dock fortfarande alltför svagt, och den professionella kompetensen hos de nationella statistikbyråerna tas dåligt tillvara.

En mycket kompetent och informativ utvärdering av Eurostats insatser på detta område ges i en rapport av Rodrigo Magalhães, forskare vid London School of Economics.<sup>39</sup>

#### *Det svenska statistiska systemet*

Det svenska statistiska systemet består för närvarande av följande komponenter

- en central statistikmyndighet (SCB)
- statistikansvariga myndigheter (varav SCB är en)
- producerande enheter (varav SCB är en)

Man kan som tidigare nämnts beskriva systemet som en balanserad kombination av decentrala och centrala komponenter inom ramen för ett integrerat statistiskt system.

#### *Decentrala komponenter*

De decentrala komponenterna i det svenska statistiska systemet utgörs av

- ett antal decentrala statistikansvariga myndigheter med ansvar för officiell statistik, som de också är den dominerande användaren av, och som faller inom ett avgränsat samhällsområde
- operatörer, som på marknadsmässiga villkor fungerar som producer av olika delprodukter i produktionsprocesser för officiell statistik; i första hand avses produktionsmoment, som inte kräver expertkunskaper, eller som kräver expertkunskaper, som inte finns i tillräcklig utsträckning hos statistikansvarig
- operatörer, som på marknadsmässiga villkor förädlar officiell statistik; kan t.ex. avse anpassning av officiell statistik för särskilda användarbehov

<sup>39</sup> "DSIS Impact Project – Final Report" (draft), London 1998.



### *Decentrala statistikansvariga myndigheter*

Kraven på en decentral statistikansvarig myndighet kan sammanfattas på följande sätt:

- Myndigheten svarar för den officiella statistiken inom ett avgränsat samhällsområde, där myndigheten ifråga också har ett sakansvar. Myndigheten är den dominerande användaren av statistiken ifråga, och därigenom får man ett starkt och direkt användarinflytande.
- Statistikansvaret måste klart skiljas från sakansvaret, så att myndigheten inte kan misstänkas för att otillbörligt påverka statistiken eller för att använda insamlade data för icke-statistiska ändamål.
- Den statistikansvariga myndigheten måste ha tillräckliga kontakter med andra användare av statistiken ifråga, så att även deras behov blir rimligt tillgodosedda. Särskild uppmärksamhet måste ägnas allmänintresset för statistiken ifråga.
- Den statistikansvariga myndigheten svarar för att de allmänna kraven på officiell statistik uppfylls, d.v.s. att den uppfyller kraven på objektivitet och kvalitet, att standarder och dokumentationskrav följs, och att statistiken publiceras och sprids på ett lättillgängligt sätt, samordnat med annan officiell statistik, bl.a. genom Sveriges statistiska databaser och svenska statistiknätet, etc.
- Den statistikansvariga myndigheten producerar statistiken ifråga i egen regi eller helt eller delvis med hjälp av andra producenter, t.ex. SCB. I det senare fallet har den statistikansvariga myndigheten beställarsvaret.

### *Operatörer på statistikproduktionsmarknaden*

På en väl fungerande statistikproduktionsmarknad finns ett antal operatörer, som på marknadsmässiga villkor och i konkurrens med varandra, åtar sig olika arbetsmoment i statistikproduktionen.

SCB är f.n. den dominerande operatören på statistikproduktionsmarknaden. Även SCB använder dock andra operatörer för avgränsade moment i statistikproduktionen, t.ex. datainsamling och datordrift.

Operatörer, som åtar sig uppdrag avseende större eller mindre delar av statistikproduktionsprocessen, måste förbinda sig att följa de lagar, standarder och riktlinjer, som gäller för den officiella statistiken i Sverige.

En statistikansvarig myndighet, som beställer statistikproduktion, har ansvaret för att alla inblandade operatörer följer lagar, standarder och riktlinjer för den officiella statistiken och måste bevaka detta bl.a. i samband med anbudsinfördran.

*SCB:s roller*

SCB har en tredubbel roll i det svenska statistiska systemet:

- rollen som den centrala statistikmyndigheten
- rollen som den största statistikansvariga myndigheten (primus inter pares)
- rollen som statistikproducerande enhet, för egen räkning och för andra statistikansvariga myndigheter

*Rollen som den centrala statistikmyndigheten*

Rollen som central statistikmyndighet innefattar ett antal uppgifter:

- att vara den naturliga centralpunkten för den officiella statistiken i Sverige och det svenska statistiska systemet med ansvar för bl.a.
  - (i) samordning och övervakning av Sveriges officiella statistik och det svenska statistiska systemet som helhet
  - (ii) övergripande kontakter med internationella statistikorgan
  - (iii) utveckling av metoder och standarder för det svenska statistiska systemet
  - (iv) infrastruktur för det svenska statistiska systemet
  - (v) produktionsmoment, som av särskilda skäl, t.ex. sekretesskäl eller kompetenskrav, bör utföras av den centrala statistikmyndigheten

Av de fem ovan nämnda delrollerna för SCB som central statistikmyndighet är det väl egentligen bara delrollen (i) som är kontroversiell. Jag återkommer till en mera utförlig diskussion av den rollen.

De övriga delrollerna, (ii) – (v), ligger naturligt på SCB, eftersom SCB är den i särklass största statistikansvariga myndigheten, och eftersom uppgifterna ifråga utförs bäst i en allsidig statistikproduktionsmiljö, i nära samarbete med erfarna och professionella statistikansvariga och statistikproducenter. Det gäller t.ex. den starkt tillämpningsorienterade metodutvecklingen liksom utvecklingen av en praktiskt fungerande infrastruktur för det svenska statistiska systemet med hjälp av system som Sveriges statistiska databaser, Svenska statistiknätet, m.fl.

### *Övergripande kontakter med internationella statistikorgan och utländska statistikbyråer*

Det internationella samfundet vill ha en central kontaktpunkt i statistikfrågor i varje medlemsland. Arbetsuppgifterna kan sammanfattas i följande punkter:

- Internationell kontaktpunkt för det svenska statistiska systemet.
- Tar emot internationella studiebesök.
- Deltar aktivt i internationella möten på statistikområdet.
- Deltar aktivt i internationellt standardiserings- och metodutvecklingsarbete.
- Tillhandahåller standardiserade rutiner för internationella statistikleveranser.

### *Metoder och standarder för det svenska statistiska systemet*

Arbetsuppgifterna kan sammanfattas i följande punkter:

- Utvecklar och anpassar metoder för statistikproduktion, där olika discipliner integreras (statistisk metodik, IT-metoder, sociologi, ekonomi m.fl.). Behovet av praktiskt förankrade teorier och metoder för statistikproduktion täcks inte av berörda discipliner på universitet och högskolor.
- Etablerar och sprider kunskap om "bästa metoder" (eng: "current best methods"), d.v.s. standardmetoder för statistikproduktion.
- Tillhandahåller metodkompetens och information i metod- och standardfrågor till statistikansvariga myndigheter och producenter.

### *Infrastruktur för det svenska statistiska systemet*

- Standarder för dokumentation, kvalitetsdeklarationer, publicering, data- och metadata-utbyte, klassifikationer, variabeldefinitioner m.m.
- Sveriges statistiska databaser för publicering och spridning av officiell statistik.
- Registerfunktion för flexibla bearbetningar och samkörningar av observationsregister.
- Standardiserade verktyg för olika moment i statistikproduktionen.
- Standardiserade gränssnitt.

*Särskilda produktionsmoment*

Det kan vara lämpligt, att den centrala statistikmyndigheten har ett särskilt ansvar för produktionsmoment, som av särskilda skäl, t.ex. sekretessskäl eller kompetenskrav, bör koncentreras till en organisation. Det kan t.ex. gälla

- Känslig datainsamling.
- Datainsamling vid uppgiftsplikt.
- Bearbetning och samkörning av känsliga mikrodata.

*Rollen som samordnare och övervakare*

SCB:s roll som samordnare och övervakare av Sveriges officiella statistik och det svenska statistiska systemet kan uppfattas som kontroversiell av det skälet, att SCB, samtidigt som man har ansvaret för denna "polisfunktion", också är det främsta "övervakningsobjektet". Tillsynsverksamheten har visserligen organiserats som en relativt fristående enhet under SCB:s chefsjurist, och jag tror inte att någon hittills anklagat detta Tillsynskansli för att på något sätt ha favoriserat SCB framför andra statistikansvariga myndigheter. Men det finns ändå tydliga nackdelar med den nuvarande organisationen.

En principiell invändning, som man kan göra, är att policyfrågor i tillsynsverksamheten behandlas på samma sätt som andra policyfrågor inom myndigheten SCB, d.v.s. de bereds i verkets direktion, innan beslut fattas av generaldirektören. I verksdirektionen sitter samtliga ämnesavdelningschefer, vilka i högsta grad representerar den statistikansvariga myndigheten SCB och statistikproducenten SCB. Däremot är inte några andra statistikansvariga myndigheter representerade vid den slutliga beredningen av sådana ärenden. (Mera informellt sker dock en hel del sådana kontakter.) En bättre ordning vore att policyfrågor i tillsynsverksamheten hanterades av generaldirektören ensam, eventuellt efter beredning i något organ, där samtliga statistikansvariga myndigheter är representerade. (Dessutom borde statistikanvändarna och uppgiftslämnarna vara representerade.)

En mera praktisk invändning mot den nuvarande ordningen är att tillsynsverksamheten möjligen tenderar att bli mera tandlös än den skulle behöva vara. Det är ofrånkomligt med en viss grad av "självcensur", genom att de tillsynsansvariga i sin roll som SCB-anställda är mycket väl medvetna om, att den statistikansvariga myndigheten SCB inte alltid är något mönster och föredöme, när det gäller att uppfylla kraven på den officiella statistiken, t.ex. vad beträffar kvalitetsdeklarationer och dokumentation.

Fördelen med den nuvarande ordningen är å andra sidan att tillsynsverksamheten genom sin närhet till praktiskt fungerande statistikproduktion får en god känsla för vad som är rimligt och önskvärt att kräva av de övervakade myndigheterna. Man kan på det sättet i bästa fall få en god förankring för de regler och påbud, som man går ut med. Man har också god tillgång till expertis av olika slag, som kan hjälpa till med erforderliga beredningsinsatser.

#### *Rollen som den ledande statistikansvariga myndigheten*

SCB:s roll som den ledande statistikansvariga myndigheten i Sverige innefattar

- ansvar för officiell statistik, som används av många användare och inom flera samhällsområden, en central kärna av officiell statistik med stort allmänintresse
- ansvar för omfattande och djupgående användarkontakter, bl.a. via programråd
- eget producentansvar och/eller beställaransvar gentemot andra producenter och delproducenter
- ansvar för att uppfylla alla krav på officiell statistik

#### *Rollen som producerande enhet*

SCB:s roll som producerande enhet innefattar

- produktion av officiell statistik, som SCB är statistikansvarig myndighet för
- produktion på uppdragsbasis av officiell statistik, som andra statistikansvariga myndigheter svarar för
- andra statistikproduktionsorienterade uppdrag för offentliga och privata kunders räkning

## 4 Önskvärda åtgärder

Matrisen i figur 23 (som är en upprepning av figur 11 i avsnitt 2.5) åskådliggör hur åtgärder inom vart och ett av de sex åtgärdsområden, som vi diskuterat ingående i kapitel 3, kan bidra till bättre måluppfyllelse i var och en av de fem måldimensioner, som tillsammans utgör målet "ökad tillgänglighet till Sveriges officiella statistik". De viktigaste kombinationerna av möjliga åtgärder och effekter är markerade med "x" i respektive cell i matrisen. Här sammanfattas nu innebörden av varje "x" i form av ett eller flera förslag till åtgärder. Förslagen identifieras med hjälp av "koordinaterna" för respektive cell i matrisen. Exempelvis är TEK.1 ett förslag till åtgärd inom området "teknik och presentation" med syfte att bidra till mål 1, att göra all officiell statistik tillgänglig på ett enhetligt sätt.

Åtgärds- område	Mål				
	1. Lätt att söka och komma åt all officiell statistik, som är relevant för ett visst problem	2. Lätt även för icke-specialister att tolka statistiken och använda sig av den på rätt sätt	3. Möjligt för kvalificerade statistikanvändare att använda den officiella statistikens datamaterial på egen hand	4. Tillgängligheten till den officiella statistikens datamaterial inskränks inte mer än nödvändigt för att säkra statistiksekretessen	5. Möjligt att komma åt den officiella statistiken genom basjänster till låga priser i kombination med valfria tilläggs-tjänster
<b>DOK:</b> Kunskapsbaserad dokumentation	x	x	x		
<b>TEK:</b> Teknik och presentation	x	x	x	x	x
<b>EKO:</b> Ekonomi: finansiering och prissättning	x		x		x
<b>SEK:</b> Hantering av statistiksekretessen			x	x	
<b>LAG:</b> Lagar och regler	x		x	x	
<b>ORG:</b> Organisation	x		x	x	

**Figur 23. Matris för att ange åtgärder för att bättre tillgodose vanligt förekommande önskemål och klagomål beträffande den officiella statistikens tillgänglighet.**

## 4.1 Lätt att söka och komma åt all officiell statistik, som är relevant för ett visst problem

### 4.1.1 Fullständiga dokumentationer och kvalitetsdeklarationer

DOK.1 *All officiell statistik och tillhörande observationsregister dokumenteras och kvalitetsdeklareras fullständigt enligt fastlagda regler och riktlinjer*

Såsom redogjordes för i avsnitt 1.3 finns redan ett enhetligt och tillräckligt omfattande beskrivningssystem för officiell statistik, det s.k. SCBDOK-systemet. Beskrivningssystemet är konkretiserat i form av mallar och handledningar för

- översiktliga produktbeskrivningar
- detaljerade observationsregisterdokumentationer
- översiktliga och detaljerade kvalitetsdeklarationer

Det gäller nu att konsekvent införa beskrivningssystemet i all officiell statistik, och att sedan fortlöpande aktualisera dokumentationerna och kvalitetsdeklarationerna, samt att vid behov förändra och vidareutveckla själva beskrivningssystemet, mallarna och handledningarna.

De tre nämnda grunddokumentationerna är sinsemellan systematiskt relaterade till varandra och kan i sin tur läggas till grund för mera skraddarsydda, användar- och problemanpassad information om den officiella statistiken och dess datamaterial. Transformationerna mellan olika dokumentationsvarianter bör så långt som möjligt åstadkommas med hjälp av datorstödda rutiner, så att de manuella arbetsinsatserna i sammanhanget minimeras. Jfr figur 5 i avsnitt 1.3.4. Se även åtgärdsförslag DOK.2 i avsnitt 4.2.1.



#### 4.1.2 Tekniska åtgärder för att förbättra sökbarhet och åtkomstmöjligheter

*TEK.1a All officiell statistik tillgängliggörs på ett standardiserat sätt genom Sveriges statistiska databaser och Svenska statistiknätet*

Officiell statistik kan naturligtvis publiceras och tillgängliggöras på många olika sätt, men ett av dem bör alltid vara genom Sveriges statistiska databaser och Svenska statistiknätet. Denna publicering bör vara den primära. Genom att all officiell statistik lagras och beskrivs på ett enhetligt sätt i Sveriges statistiska databaser, säkerställer man att statistik användarna på enklast möjliga sätt kan hitta och kombinera officiell statistik från olika statistikansvariga myndigheter. Svenska statistiknätet ger å andra sidan de olika statistikansvariga myndigheterna samtidigt goda möjligheter att profilera publiceringen av den officiella statistik de svarar för på olika sätt, t.ex. genom att länka den till annan information, även icke-statistisk, som är relevant för deras användare.

*TEK.1b Alla slutliga observationsregister från den officiella statistiken tillgängliggörs på ett standardiserat sätt genom en särskild registerfunktion inom Sveriges statistiska databaser*

Den nuvarande centrala databasverksamheten, Sveriges statistiska databaser, har också ansvaret för att tillgängliggöra 17 särskilt utpekade observationsregister, som räknas upp i Förordningen om vissa personregister för officiell statistik.<sup>40</sup> Enligt förordningen får nämnda register utgöra grund för SCB:s databasservice, som definieras i 2-3§§ i instruktionen (1988:137) för SCB, och de får sambearbetas med varandra.

SCB har sålunda i uppdrag att tillgängliggöra de 17 utpekade observationsregistren för statistikbearbetningar inom ramen för den centrala databasverksamheten. Detta uppdrag är fortfarande långt ifrån slutfört. Viss dokumentation tillgängliggörs för närvarande för några få av

<sup>40</sup> Förordningen (SFS 1995:1060, senaste ändring SFS 1998:1330) om vissa personregister för officiell statistik reglerar hur personregister får bearbetas och samköras för framställning av officiell statistik. I en bilaga till förordningen räknas de berörda personregistren upp per statistikansvarig myndighet. Ändamål, innehåll och regler för sambearbetning anges register för register. I punkt 1.15 specificeras de 17 registren som ingår i SCB:s databasservice. Det stadgas där, att "register för databasservice får sambearbetas med annat sådant register".

registren via databasverksamheten. Inget av registren är ännu vare sig direkt eller indirekt tillgängligt för statistikanvändarna via databasverksamheten, utan den som vill göra sådana bearbetningar får, liksom tidigare, vända sig till respektive statistikansvarig.

Det här framlagda förslaget innebär två saker:

1. Ansträngningarna intensifieras, att via den centrala databasverksamheten tillgängliggöra de 17 utpekade registren för statistikbearbetningar enligt användarnas önskemål. Detta kräver att alla idag tillgängliga tekniska möjligheter och programvaror utnyttjas på bästa sätt, t.ex. som beskrivits i avsnitt 3.4.5.
2. Verksamheten utvidgas successivt till att omfatta alla s.k. slutliga observationsregister, som ligger till grund för officiell statistik.

Den nuvarande verksamheten inom Sveriges statistiska databaser är fokuserad på makrodata och metadata, d.v.s. på aggregerade statistiska datamaterial och de beskrivningar som hör till dem. Tillgängliggörandet av observationsregistrens mikrodata innebär en verksamhet av något annorlunda karaktär, och det kan därför vara motiverat att markera denna verksamhet genom en särskild registerfunktion inom ramen för databasverksamheten. Se vidare åtgärdsförslag ORG.4.

*TEK.1c Det statistiska systemets gemensamma infrastruktur för att tillgängliggöra den officiella statistiken och dess datamaterial vidareutvecklas med hjälp av den snabbt föränderliga informationstekniken*

Här kommer det att finnas mycket att göra under överskådlig tid. Att inte göra något är inte att stå stilla, utan att falla tillbaka. Några av de viktigaste åtgärderna är följande:

- (i) *Den tekniska infrastrukturen för att tillgängliggöra den officiella statistiken baseras på standardkomponenter och standardiserade gränssnitt*  
Detta är det bästa sättet att kombinera hög flexibilitet och låg komplexitet, och det gör det möjligt att vid varje tidpunkt ha ett väl fungerande system, samtidigt som nya tekniska framsteg och utvecklingsmöjligheter snabbt kan tas till vara.
- (ii) *Den snabba utvecklingen på Internet-området utnyttjas även i fortsättningen för att göra den tekniska infrastrukturen ännu mera kraftfull och användarvänlig*  
Återigen gäller det att dra maximal nytta av den allmänna utvecklingen på Internetområdet utan att behöva bedriva ett alltför

omfattande och dyrbart utvecklingsarbete för den officiella statistiken på egen hand.

- (iii) *Nya metoder och verktyg för effektiv hantering av data- och metadatabaser utnyttjas: datalager, OLAP-verktyg, objektorienterade databaser m.m.*

De nämnda teknikerna är under stark utveckling på den kommersiella marknaden, och de är högst relevanta för möjligheterna att på effektivare, flexibla och mera användarvänliga sätt tillgängliggöra den officiella statistiken.

- (iv) *Verktyg för associativa sökningar, t.ex. tesaurusar och nya typer av sökmotorer, anskaffas och anpassas för den officiella statistikens behov*

Med associativa sökningar avses bl.a. sökningar där användaren kanske använder sig av andra termer och begrepp än de som är gängse i den officiella statistiken, t.ex. "kylskåp" i stället för "vitvaror". Med hjälp av tesaurusar och andra avancerade sökverktyg, gärna självlärande, kan sådana sökningar genomföras på ett effektivt och användarvänligt sätt.

- (v) *Skapa många olika ingångar till den officiella statistiken, så att problembaserade sökningar underlättas*

Samma statistik kan mycket väl vara relevant för många olika problem. Sökvägarna bör därför inte vara strikt hierarkiska, utan de måste kunna korsa och överlappa varandra på olika sätt.

- (vi) *Utnyttja nya tekniker för att presentera och levandegöra den officiella statistiken, t.ex. länkar, animationer, guider och expertsystem*

Även de gamla etablerade presentationsteknikerna, tabeller, grafik och texter, behöver standardiseras, moderniseras och läras ut med hjälp av datorstödd pedagogik.

### 4.1.3 Marginalkostnadsbaserad prissättning

- EKO.1 Den av Riksrevisionsverket (RRV) föreslagna huvudprincipen för prissättning av informationstjänster<sup>41</sup> bör tillämpas konsekvent vid tillgängliggörande av officiell statistik och informationstjänster i samband därmed

Om modern, Internet-baserad teknik används, såsom redan sker i bl.a. Sveriges statistiska databaser, innebär RRV:s huvudprincip, att all offi-

<sup>41</sup> "Principer för prissättning av informationstjänster", regeringsuppdrag, RRV 1995:64.

ciell statistik kan tillgängliggöras med hjälp av ett antal bastjänster till mycket låg marginalkostnad. Vidarebearbetningar och mera avancerade tjänster i anslutning till bastjänsterna betraktas som tilläggstjänster, vilka användaren antingen kan utföra själv eller köpa på en konkurrensutsatt marknad. Den fasta infrastrukturen, i första hand själva utvecklingen av informationssystemet och bastjänsterna, utgör fasta kostnader, d.v.s. de är oberoende av användningsvolym, och anslagsfinansieras. De rörliga kostnaderna för drift, uttag, vidarebearbetningar m.m. bekostas av användarna genom avgifter av något slag.

Dessa frågor har diskuterats utförligt i avsnitt 3.3 av denna uppsats.

#### 4.1.4 Begreppet ”officiell statistik” preciseras, och enhetliga publiceringsregler föreskrivs

*LAG.1a Datamaterial på olika aggregeringsnivåer, som ligger till grund för officiell statistik, skall tillgängliggöras enligt samma regler som gäller för den officiella statistiken och dess observationsregister*

Tydliggörandet förhindrar bl.a. att statistikansvariga mer eller mindre självsvåldigt bort definierar datamaterial från den officiella statistiken, som *de facto* är tillgängliga eller möjliga att tillgängliggöra till låg marginalkostnad.

Traditionellt är det bara de statistiska slutprodukterna på relativt hög aggregeringsnivå, som har räknats som officiell statistik, och som har betraktats som anslagsfinansierade. Mera finfördelade redovisningar och specialbearbetningar av bakomliggande observationsregister har varit relativt arbetskrävande och därmed dyrbara att utföra. Idag är situationen helt annorlunda. Det är fullt tekniskt möjligt att tillhandahålla även finfördelad statistik och specialbearbetningar av mikrodata enligt användarens specifikationer till mycket låga marginalkostnader. Den officiella statistiken bör därför anses omfatta såväl aggregerad statistik (makrodata) på olika aggregeringsnivåer som bakomliggande observationsregister (mikrodata) med tillhörande, utförliga dokumentationer (metadata).

*LAG.1b All officiell statistik skall publiceras genom Sveriges statistiska databaser och Svenska statistiknätet*

Detta innebär en skärpning jämfört med nuläget, där den statistikansvarige fritt får välja mellan traditionella publiceringsformer och elektroniska media. Den föreslagna publiceringsregeln innebär, att det

blir obligatoriskt att publicera officiell statistik genom den gemensamma infrastrukturen för Sveriges statistiska system: Sveriges statistiska databaser och Svenska statistiknätet. Statistikansvarig kan självklart därutöver välja att även publicera på andra sätt, t.ex. genom elektroniska publikationer på webben eller traditionellt tryckta publikationer.

#### 4.1.5 Skärpt tillsynsverksamhet

*ORG.1 Det centrala tillsynsorganet för den officiella statistiken får en stärkt ställning och bevakar*

- (i) att skärpta förfaranden för dokumentation och kvalitetsdeklaration av officiell statistik följs av de statistikansvariga
- (ii) att den officiella statistiken och bakomliggande observationsregister tillgängliggörs genom den centrala databasverksamheten, och att tillgängligheten till observationsregistren inte inskränks mer än nödvändigt för att skydda statistiksekretessen
- (iii) att enhetliga prissättningsprinciper, baserade på RRV:s förslag, tillämpas på all officiell statistik och alla tjänster i anslutning till den officiella statistiken, och att statistikansvariga inte, t.ex. genom olämpligt utformade informationstjänster eller prisstrukturer, *de facto* monopoliserar tillgången till den officiella statistikens data-material

### 4.2 Lätt även för icke-specialister att tolka officiell statistik och använda sig av den på rätt sätt

#### 4.2.1 Skräddarsydd information om statistiken för olika användarkategorier

*DOK.2 Officiell statistik förses med problemorienterad dokumentation, som är särskilt utformad för olika användares behov*

Den obligatoriska, grundläggande dokumentationen av officiell statistik består av

- produktbeskrivningar
- observationsregisterdokumentationer
- kvalitetsdeklarationer

Sammantaget bör de grundläggande dokumentationerna innehålla all information som kan behövas för att tolka och använda officiell statistik. Men för många statistik användare kan det vara alltför svårt och mödosamt att relatera den grundläggande dokumentationen till en aktuell användningssituation. Om inte annat saknar användaren ofta den tid och motivation, som skulle behövas. Det gäller därför för den statistikansvarige att kunna presentera olika varianter av den grundläggande dokumentationer på sätt som passar respektive användare. För att inte detta arbete skall bli alltför resurskrävande och kräva nya dokumentationsinsatser av ungefär samma omfattning som den grundläggande dokumentationen, måste man försöka hitta systematiska metoder att omforma till grundläggande dokumentationen till nya ändamål.

Som framhållits tidigare (se förslag DOK.1) är de tre nämnda grunddokumentationerna sinsemellan systematiskt relaterade till varandra, och de kan i sin tur läggas till grund för den mera skraddar-sydd, användar- och problemanpassade information om den officiella statistiken och dess datamaterial, som vi är ute efter här. Transformationerna mellan olika dokumentationsvarianter bör så långt som möjligt åstadkommas med hjälp av datorstödda rutiner, så att de manuella arbetsinsatserna i sammanhanget minimeras. Jfr figur 5 i avsnitt 1.3.4.

#### 4.2.2 Datorstödd framställning av användar- och problemanpassad dokumentation

*TEK.2 Metoder och tekniska system utvecklas, som stöder en systematisk och helst automatisk omformning av den grundläggande dokumentationen av officiell statistik till mera användar- och problemorienterade dokumentationsvarianter*

Detta förslag innebär att man utvecklar det erforderliga tekniska stödet för föregående förslag, DOK.2.

### 4.3 Möjligt för kvalificerade statistikanvändare att använda den officiella statistikens datamaterial på egen hand

En kvalificerad statistikanvändare, som vill använda officiell statistik, inklusive bakomliggande observationsregister, är mest beroende av följande saker:

- att *all* relevant information om de data och processer, som har legat till grund för statistiken, är så väl dokumenterad, att användaren på egen hand, utan att behöva konsultera (de kanske numera döda eller avgångna) producenterna av statistiken, kan tolka (och omtolka) statistiken, och alla de förutsättningar den bygger på, på ett ansvarsfullt och självständigt sätt
- att datamaterialen rent tekniskt är tillgängliga på ett sådant sätt att användaren lätt och till låg kostnad kan komma åt dem för egna bearbetningar eller åtminstone för bearbetningar, som (inom ramen för vad statistiksekretessen medger) så långt som möjligt utförs enligt användarens önskemål, vad gäller interaktivitet, flexibilitet, funktionalitet och kostnader

Åtminstone vissa typer av kvalificerade användare, t.ex. forskare, är däremot inte lika beroende av att dokumentationen presenteras på ett speciellt användarvänligt sätt. Det är viktigare att den är korrekt och fullständig. Forskaren har också förhållandevis gott om egen tid men ont om pengar för att köpa tjänster av andra. Bl.a. därför vill forskaren ofta göra så mycket som möjligt på egen hand, enligt självbetjäningsprincipen.

Andra typer av användare, t.ex. professionella utredare, har mera ont om tid och mera gott om pengar och kan därför vara beredda att köpa tjänster. Men även dessa användare vill i så fall göra detta av eget fritt val, och helst på en konkurrensutsatt marknad, så att de får valuta för pengarna. Det tidigare föreslagna upplägget med bastjänster och tilläggstjänster för tillgängliggörande av officiell statistik passar därför bra även för denna typ av användare, de kvalificerade statistikanvändarna.

Följande förslag (i avsnitt 4.3.1 – 4.3.6) är avsedda att öka tillgängligheten till den officiella statistiken för kvalificerade statistikanvändare.

#### 4.3.1 Fullständig dokumentation av dokumentationsregister, så att de kan återanvändas utan förstahandskunskap

*DOK.3 Alla observationsregister, som ligger till grund för officiell statistik, är så fullständigt dokumenterade, att en kvalificerad användare kan använda sig av dem på egen hand även efter lång tid, när inga experter med förstahandskunskap finns att fråga*

#### 4.3.2 Standardiserade tekniska lösningar för att göra datamaterialen tillgängliga

*TEK.3 Infrastrukturen för att tillgängliggöra den officiella statistikens datamaterial baseras på standardiserade tekniska lösningar, som gör det lätt för en kvalificerad användare att genomföra bearbetningar och analyser på egen hand och enligt egna önskemål, inom ramen för statistiksekretessen*

#### 4.3.3 Bastillgänglighet till lågpris för ”gör-det-självaren”, prisvärda tilläggstjänster till marginalkostnad

*EKO.3 Tjänsterna kring den officiella statistikens datamaterial utformas så, att en grundläggande tillgänglighet säkerställs med hjälp av billiga bastjänster, varefter användaren kan välja mellan självbetjäning eller avgiftsbelagda tilläggstjänster; all prissättning sker enligt RRV:s marginalkostnadsprincip<sup>42</sup>*

#### 4.3.4 Skydd för statistiksekretessen som inskränker tillgängligheten minimalt

*SEK.3 Skyddet för statistiksekretessen utformas så, att statistik-användarnas tillgång till datamaterialen inskränks minimalt*

<sup>42</sup> "Principer för prissättning av informationstjänster", regeringsuppdrag, RRV 1995:64.



#### 4.3.5 Tillgänglighetskraven på officiell statistik föreskrivs gälla även observationsregistren

*LAG.3 Begreppet "officiellt statistik" definieras så, att reglerna beträffande den officiella statistikens tillgänglighet blir tillämpliga även för observationsregister, som ligger till grund för officiell statistik*

#### 4.3.6 Offensiv bevakning av tillgängligheten

Statistiksekretessen nödvändiggör vissa inskränkningar i den officiella statistikens (och speciellt de bakomliggande observationsregistrens) tillgänglighet. Den statistikansvarige har förstahandsansvaret för avvägningen mellan skydds krav och tillgänglighetskrav. Tillsynsorganet övervakar att avvägningen är rimlig.

*ORG.3a Tillsynsorganet övervakar, att statistikansvariga arbetar konstruktivt med att göra den officiella statistikens observationsregister så tillgängliga som möjligt inom ramen för statistiksekretessen*

Det gäller att säkerställa att statistikansvariga på ett konstruktivt och resurssnålt sätt använder sig av alla till buds stående metoder och tekniska lösningar.

*ORG.3b Tillsynsorganet övervakar, att statistikansvariga inte inskränker tillgängligheten mer än nödvändigt och inte etablerar kunskaps- eller datatillgångsmonopol*

Statistikansvariga kan frestas att *de facto* etablera monopol av nämnda slag, t.ex. genom att underlåta att dokumentera ordentligt, eller genom att komponera datatjänster, som är orationellt utformade eller mera omfattande än vad användaren egentligen har behov av, och som därför blir onödigt dyra. Den statistikansvarig myndigheten kan därigenom tillförsäkra sig uppdragsverksamhet, som inte skulle komma till stånd, om den kvalificerade statistikanvändaren kunde göra mera av jobbet själv.

## 4.4 Tillgängligheten till den officiella statistikens datamaterial inskränks inte mer än nödvändigt för att säkra statistiksekretessen

Problematiken behandlades relativt utförligt i avsnitt 3.4. Se särskilt avsnitt 3.4.5. Här skall vi bara sammanfatta de viktigaste förslagen till åtgärder.

### 4.4.1 Kombinera legala, administrativa och tekniska åtgärder

Huvudpoängen är här, att just välkomponerad kombination av legala, administrativa och tekniska åtgärder i praktiken ger det bästa skyddet för statistiksekretessen. I teorin är det däremot praktiskt taget omöjligt att under alla omständigheter garantera hundra procentig statistiksekretess, oavsett hur bra lagar man har, och vilka raffinerade metoder man inför.

*SEK.4a Skyddet för statistiksekretessen skall säkerställas med en kombination av legala, administrativa och tekniska åtgärder, som inskränker tillgängligheten till den officiella statistiken och dess datamaterial så lite som möjligt*

Man kan också uttrycka detta krav på följande sätt:

*SEK.4b Den statistikansvarige skall utnyttja alla tekniska, legala och administrativa möjligheter, som står till buds, för att göra den officiella statistiken och dess datamaterial så tillgängliga som möjligt för användarna inom ramen för statistiksekretessen*

Bägge formuleringarna ställer hårdare krav på statistikansvariga, än vad som gäller idag, att aktivt och konstruktivt verka för praktiskt väl fungerande och säkra lösningar på problemet att tillgängliggöra den officiella statistikens datamaterial på de sätt, som användarna önskar sig, samtidigt som statistiksekretessen upprätthålls. Den statistikansvarige får inte frestas att upprätthålla statistiksekretessen ”med marginal”, om detta går ut över tillgängligheten, trots att brott mot statistiksekretessen helt säkert är förenade med större olägenheter för den statistikansvarige än brott mot tillgänglighetskraven.

#### 4.4.2 Minska tolkningsutrymmet och inför tillgänglighetsprincip

*LAG.4a Statistiksekretessen förtydligas i lagen, så att den statistikansvariges tolkningsutrymme minskar*

*LAG.4b Lagreglerna kring statistiksekretessen kompletteras med en lagstadgad tillgänglighetsprincip för effektiv hantering av statistiksekretessen*

Innebörden av dessa båda förslag har redan utvecklats och exemplifierats i avsnitt 3.4 liksom det följande förslaget, som avser den tekniska hanteringen av statistiksekretessen.

#### 4.4.3 Bästa kända tekniska lösningar för att maximera tillgängligheten inom given statistiksekretess

*TEK.4 Bästa kända metoder, tekniker och programvaror utnyttjas för att säkerställa skyddet av statistiksekretessen, samtidigt som tillgängligheten till statistikmaterialen inskränks så lite som möjligt*

Följande metoder och tekniker bör utnyttjas (se avsnitt 3.4.5 för utförligare resonemang):

- (i) Forskare "svärs in" vid forskararbetsplatser hos den statistikansvarige*
- (ii) Skapa elektroniskt tillgängliga forskararbetsplatser med interaktiv fjärrstyrning av statistikbearbetningar av känsliga data via brandvägg*
- (iii) Avidentifiera och tunna ut känsliga data, innan de görs fritt tillgängliga t.ex. som s.k. public files, som används mycket i USA*

Vad som här sägs om "forskare" bör gälla även andra typer av kvalificerade och seriösa statistikanvändare, t.ex. utredare i näringsliv och förvaltning.

#### 4.4.4 Registerfunktion inom databasverksamheten

*ORG.4 En registerfunktion inrättas inom den centrala databasverksamheten för bearbetningar och samkörningar av identifierade mikrodata för statistiska ändamål*

### 4.5 Möjligt att komma åt den officiella statistiken genom bastjänster till låga priser i kombination med valfria tilläggstjänster

Alla kategorier av statistikanvändare framhåller värdet av att de själva, i enlighet med sina egna preferenser, kan välja eller välja bort tjänster i anslutning till den officiella statistiken och bakomliggande observationsregister. De vill inte känna sig tvungna att använda tjänster som den statistikansvarige erbjuder, särskilt inte om dessa tjänster betingar ett högt pris, eller om användaren känner på sig att tjänsterna inte är prisvärda utan skulle kunna göras bättre av någon annan.

Det skulle därför i hög grad bidra till att användarna blir nöjda med officiella statistikens tillgänglighet, om infrastrukturen för tillgängliggörandet i ännu högre grad än för närvarande byggdes upp och prisattes på ett sådant sätt, att användaren kan se ett smörgåsbord av tjänster framför sig, bestående av

- billiga bastjänster, som säkerställer en miniminivå av tillgänglighet till den officiella statistiken
- självbetjäningstjänster, som fristående eller i kombination med billiga bastjänster, gör det möjligt för varje medborgare, som är beredd att lägga ner en hel del eget arbete, att i princip komma åt all information som finns i den officiella statistiken och bakomliggande datamaterial
- tilläggstjänster, som gör det möjligt för den statistikanvändare, som så önskar, att komplettera bastjänsterna med förädlande, mervärdeskapande tilläggstjänster, som tillhandahålls av den statistikansvarige eller av någon annan aktör på marknaden för statistiktjänster

I det följande (avsnitt 4.5.1 – 4.5.2) beskrivs ett antal åtgärder behöver vidtas.

#### 4.5.1      Infrastruktur som möjliggör flexibla kombinationer av billiga bastjänster och prisvärda tilläggstjänster

*TEK.5   Infrastrukturen för att tillgängliggöra den officiella statistiken och dess datamaterial byggs upp med hjälp av standardkomponenter och standardiserade gränssnitt, så att billiga bastjänster och dyrare tilläggstjänster lätt kan fogas in i strukturen och väljas eller väljas bort av statistikanvändarna*

Denna infrastruktur i kombination med marginalkostnadsbaserad prissättning skapar förutsättningar för en konkurrensutsatt marknad för tilläggstjänster.

#### 4.5.2      Marginalkostnadsbaserad prissättning

*EKO.5   EKO.5 Bastjänsterna och tilläggstjänsterna prissätts enligt RRV:s förslag till marginalkostnadsprissättning av informationstjänster*













## Referenser

Denell, Chris & Lundholm, Gunilla & Nyström, Per & Sundgren, Bo: ”*Produktbeskrivningsmall 1997 för officiell statistik – Med anvisningar*”, SCB 1997, reviderad version 1998.

Elvers, Eva & Rosén, Bengt: ”*Quality Concept for Official Statistics*”, R&D Report 1997:2, SCB.

Magalhães, Rodrigo: ”*DSIS Impact Project – Final Report*” (draft), London 1998.

Nordbotten Svein: ”*A Statistical File System*”, Statistisk Tidskrift, No. 2, Stockholm 1966 sid 99-109.

Nygren, Rolf & Larsson, Jan & Åkerman, Sune: ”*Samhällsdokumentation inför framtiden – En bok om samhällsforskningens framtida materialtillgång*”, Delegationen för långsiktsoveträdd forskning, LiberFörlag, Stockholm 1982.

Rapaport, Edmund & Sundgren, Bo: ”*Output Protection in Statistical Data Bases*”, Conference of the International Statistical Institute (ISI), Warszawa, 1975. Finns också tillgänglig som rapport från SCB.

Riksrevisionsverket (RRV): ”*Principer för prissättning av informationstjänster*”, regeringsuppdrag, RRV 1995:64.

Rosén, Bengt & Sundgren, Bo: ”*Dokumentation för återanvändning av mikromaterial från SCB:s undersökningar*”, Stockholm 1991. Finns även i engelsk översättning.

Samuelson, Per: ”*Statistikrätt – Den rättsliga regleringen av den statliga statistiken*”, SCB och CE Fritzes AB 1994.

SFS: ”*Lag om den officiella statistiken*”, SFS 1992:889.

SFS: ”*Förordning om den officiella statistiken*”, SFS 1992:1668, ändrad i SFS 1994:1108, omtryck.

SFS: ”*Förordning om vissa personregister för officiell statistik*”, SFS 1995:1060, senast ändrad i SFS 1998:1330.

SOU: ”Effektivare statistikstyrning. Den statliga statistikens styrning och samordning.”, SOU 1992:48.

SOU: ”Ändrad ansvarsfördelning för den statliga statistiken”, Betänkande av Genomförandekommittén, SOU 1994:1.

Statistiska centralbyrån (SCB): ”Kvalitetsbegrepp och riktlinjer för kvalitetsdeklaration av officiell statistik”, SCB, Meddelanden i samordningsfrågor (MIS) 1994:3.

Statistiska centralbyrån (SCB): ”Den framtida databasverksamheten vid SCB – Slutrapport”, SCB 1994.

Statistiska centralbyrån (SCB): ”Slutrapport från databasprojekten”, SCB 1997.

Statistiska centralbyrån (SCB): ”Föreskrifter och allmänna råd om offentliggörande av officiell statistik”, SCB, Meddelanden i samordningsfrågor (MIS) 1995:5.

Sundgren, Bo: ”Att dokumentera statistiska undersökningar – Användarhandbok för SCBDOK version 2.0”, SCB 1996.

Sundgren, Bo: ”Sweden's Statistical Databases: an infrastructure for flexible dissemination of statistics”, Report to the UN/ECE Conference of European Statisticians, June 1997.

Winberg, Christer & Åkerman, Sune: ”Forskningens framtida datatillgång”, Samarbetskommittén för långsiktigt motiverad forskning, Stockholm 1976.

# OFFISIELL STATISTIKKPRODUKSJON

## VISJONER OG STRATEGI<sup>1</sup>

Av Svein Nordbotten

<sup>1</sup> Dette notat er skrevet etter oppdrag fra Gruppen for å evaluere erfaringer fra det svenske systemet for offisiell statistikk fra 1.juli 1994. Forfatteren har hatt ulike tilknytninger til statistikk og statistikkarbeid i innpå et halv hundrede år, blant annet i Sverige. De visjoner og strategier som legges frem bygger derfor på synspunkter fra mange statistikere. Forfatteren har imidlertid det hele og fulle ansvar for eventuelle feiltolkninger av andres ideer.

## Innehåll

1. Innledning
  2. Om visjoner, mål og strategier
  3. Scenario for statistikkmarkedet omkring år 2010
  4. Strategi for utvikling av statistikkproduksjon
  5. Sammenfatning
- 
- A. Henvisninger
  - B. Begreper og termer

# 1. Innledning

For å kunne velge mål og og planlegge for fremtidig offisiell statistikkproduksjon er det nødvendig å utvikle sannsynlige visjoner for fremtiden i form av scenarier. Betegnelsen *statistikkproduksjon* blir i dette notat brukt i vid betydning og dekker all virksomhet fra planlegging til markedsføring av statistisk informasjon. Med en offisielle statistikkprodusent sikter vi her til en produsent som har ansvar for en del i det offisielle statistikksystemet.

Notatet gir neste avsnitt en kort omtale av tidligere nasjonal og internasjonal innsats for å skape mål og strategier for fremtidig statistikkproduksjon. I tredje avsnitt presenteres et scenario av informasjonssamfunnet slik vi tror det kan være omkring år 2010 basert på den teknologiske og samfunnsmessige utvikling i de siste 10 årene. Deretter drøftes hvilken rolle og mål det vil være ønskelig å gi den offisielle statistikkproduksjonen i et slikt informasjonssamfunn. I fjerde avsnitt understrekes det at forutsetningen om at effektiv oppnåelse av målene krever en strategi og noen steg i en slik strategi skisseres. Avslutningsvis oppsummeres de viktigste synspunkter og konklusjoner som er fremført i notatet.

I to vedlegg følger en litteraturhenvisning og et avsnitt om begreper og terminologi nyttet i notatet.

## 2. Om visjoner, mål og strategier

### 2.1 Hvorfor det er nødvendig å ha visjoner, mål og strategier

Mange har en motvilje mot å tenke langsiktig. Arbeid med fremtidsplaner omtales som bortkastet fordi de ofte bygger på sviktende forutsetninger. Skeptikerne hevder at å planlegge fremtiden i verste fall kan føre til ineffektivitet og i beste fall til sløsing av ressurser fordi planene vil være foreldet før de er fullførte. Det finnes mange eksempler på at slik holdning til fremtiden kan være vel berettiget blant annet fordi arbeidet ikke tar hensyn til informasjon som kommer under planleggingen og heller ikke omfatter nødvendig løpende vedlikehold og oppdatering av planene.

Vi vil argumentere for at det er nødvendig å generere visjoner og mål for å kan få et sikte for målrettet utvikling av statistikkproduksjonen. Statistikkproduksjon i seg selv er en tidkrevende prosess, og verdien av statistisk informasjon avhenger ofte av at den beskriver en historisk utvikling. Skal statistikkproduksjonen bli vellykket, må derfor statistikkprodusenten helst både kunne forutse etterspørselen på fremtidige tidspunkter, sette mål for å møte etterspørselen og utarbeide strategier og planer for hvordan disse mål skal oppnås. Samtidig må mål og planer oppdateres underveis når det foreligger ny informasjon.

### 2.2 Svensk statistikkproduksjon

Offisiell svensk statistikkproduksjon har alltid vært rettet og styrt etter fremtidige mål. Premissene for målene ble tradisjonelt satt av noen få store statistikkbrukere, i første rekke de statlige styresmakter, gjennom ressurstildeling, og av forpliktende internasjonalt samarbeid også sterkt påvirket av de samme store brukere.

Oppfatningene om hvordan en kunne oppnå en effektiv statistikkproduksjon har variert opp gjennom tidene. På 1700 og 1800 tallene ble produksjonen styrt gjennom tabellkommisjoner med oppgave å se til at berettigede brukerinteresser ble forsvarlig tilgodesett. Etterhvert som behovet for statistisk underlag i den statlige virksomhet tiltok i bredde, ble det innført en sterk *sentralisering* av statistikkproduksjonen i Statistiska Centralbyrån (SCB).



I de siste tiår har også andre brukere kommet inn med større tyngde og mer nyanserte synspunkter på statistikkproduksjonen har kommet frem. I den siste svenske statistikkreformen av 1. juli 1994, var det således 3 synspunkter som gjorde seg gjeldende [Förordning (1992:1668) om den officiella statistikken]. *Statistikk-produksjonen* skulle ta hensyn til de ulike brukerinteresser som skulle få større innflytelse på utforming av produktene, *markedsføringen* skulle tilrettelegges slik at flest mulig skulle kunne dra fordeler av statistikkproduktene og *produksjonsprosessen* skulle gjøre bruk av tilgjengelig teknologi for å bli mest mulig *effektiv*. Disse mål ga seg utslag i en *desentralisering* av den offisielle statistikkproduksjonen til et tjue-tall ansvarlige produsenter. Det koordinerende ansvar for hele den offisielle statistikk ble fortsatt forankret i SCB. I en publikasjon med tittelen *Fremtiden redan idag* viser SCB også sitt sterke engasjement i brukernes fremtidige problemer [Statistiska Centralbyrån 1996b].

### 2.3 Internasjonal statistikkproduksjon

I internasjonalt samarbeid er det blitt lagt premisser for hvordan de nasjonale statistikk-systemene burde utformes, og hvordan internasjonal statistikk skulle ivaretas. Disse premisser har konkret ført til utvikling av statistiske begrepssystemer og standarder som har hatt og har betydelig innflytelse på hvordan den nasjonale statistikkproduksjonen blir utformet. Behovene for internasjonal statistisk informasjon har vært ivaretatt av forskjellige internasjonale statistiske organer i nært samarbeid med de nasjonale statistikkprodusenter.

For å få oversikt over og muligheter for å studere ulike former for organisering av statistikkproduksjon, ble det i 1970 årene etter initiativ av UN Statistical Commission gjennomført omfattende undersøkelser av nasjonale statistikkkontorers oppbygning. Det ble også ført diskusjoner og utformet synspunkter av mer generell karakter om hvordan statistikkprodusentene oppfattet sin fremtidig rolle. Det kan være både interessant og lærerikt å studere disse arbeider [Duncan 1978, Kennessey 1994].

I de siste årene er verdifulle synspunkter lagt frem av to komiteer oppnevnte av UN Statistical Commission, [United Nations 1992] og [United Nations 1998]. I vår sammenheng er det viktig å merke seg anbefalinger om at den *internasjonale koordineringen* av utvikling på de begrepsmessige og metodiske områder må styrkes.

### 3. Scenario for år 2010

#### 3.1 Beskrivelse av den informatiske domene

##### Scenario

Med et scenario mener vi en visjonær beskrivelse av en fremtidig situasjon sammensatt av to komponenter. En komponent består av forhold som ligger utenfor vårt påvirkningsområde. Den andre komponenten søker vi å gi en utforming innen rammen av de føringer som den første forventes å gi. For den første komponenten kan vi bare støtte oss til forutsetninger om at visse utviklingstrender vil fortsette, for den andre vil vi søke å skissere en handlingsstrategi som kan realisere den form vi ønsker å gi den andre komponenten.

I den sammenheng som opptar oss i dette notat, vil vi således på den ene side ha en komponent som beskriver samfunnsmessige og teknologiske forhold som de offisielle statistikkprodusente bare i liten grad og indirekte kan påvirke. På den annen side kan vi innenfor de begrensninger og med de muligheter som denne tekniske, samfunnsmessige ramme gir, forsøke å stille opp ønskelige mål for den offisielle statistikkproduksjon og en strategi for hvordan målene skal kunne realisere.

Det scenario for informasjonssamfunnet som ligger til grunn for våre drøftinger i dette avsnitt, er illustrert i Figur 1.

**(Figur 1: Offisiell statistikk som del av et informasjonssystem)**

##### Horisont

Før vi kan gi oss i kast med å skissere et fremtidig scenario hvor den offisielle statistikken inngår, må vi ta standpunkt til hvor langt frem i tid det kan være hensiktsmessig å legge horisonten. Et perspektiv med en tidsnær horisont vil åpenbart kunne ha en høyere grad av realisme enn et annet perspektiv med tidsfjern horisont. På den andre side har vi allerede påpekt at statistikkproduksjon er tidkrevende og at etterspurt statistisk informasjon ofte vil være tidsserier som krever en viss lengde for at de

skal gi kunnskap om utvikling og ha interesse og verdi for brukerne. Det siste argument taler for at vi bør legge horisonten så lang frem i tiden at den gir rom for en realistisk gjennomføring.

I det følgende har vi valgt et 10–12 års fremtidsperspektiv med begrunnelse at stort lengere er det vanskelig å ha noen nyttig oppfatning av informasjonsteknologiske betingelser i samfunnet, mens et kortere perspektiv har en tvilsom nytte fordi større endringer neppe kan gjennomføres innen en stram tidsramme. Et annet argument for det valgte perspektiv er at ofte går det 8–10 år fra nye teknologiske ideer gjøres kjent til de er tatt i bruk og innarbeidet i større omfang. Innen vår horisont regner for eksempel romforskerne med å ha utviklet teknologi for å sette det første menneske på Mars, teknologi som uten tvil også vil sette preg på det informatiske samfunn.

## Teknologi

Ved utgangen av det 20. århundrede doubles *IT verktøyene* kraft pr. krone over få år.

Datamaskiner med en kraft som var reservert de aller største brukere for 20 år siden, finnes idag som personlige datamaskiner på arbeidsplasser, i skoler og hjem.

Parallelt har det i det siste ti-år skjedd en rask utvikling av telekommunikasjons-mulighetene. Mest bemerkelsesverdig er utviklingen av *Internet* som for 30 år siden ble initiert for militære oppgaver og romforskning, og så sent som for 10 år siden i stor utstrekning var reservert for den akademiske verden. Idag er brukermassen dominert først og fremst av brukere fra de kommersielle og private sektorer. I organisasjoner og interessegrupper vokser det frem skjermede *intranet* og mellom samarbeidene intranet etableres *extranet*. Markedsledere innen kommunikasjonsteknologi forventer at allerede om få år vil en femtedel av jordens befolkning ha adgang til Internet, i vår tidshorisont vil kanskje hvert fjerde menneske kunne kommunisere globalt og ha adgang til kunnskapssamlinger over hele verden.

Nye og raskere *telekanaler* er under utvikling og vil om få år tillate overføring av levende bilder og lyd uten særlige begrensninger. Med fortsatt utvikling av mobilnett, vil kravet om kabeltilknytting til kommunikasjonsnettet falle bort. Teknologien blir stadig mindre *energikrevende* og fjerner bruksbegrensning til steder med kontinuerlig energitilførsel. I løpet av få år vil det være mulig å lade opp en datamaskin i løpet av natten for hele neste dags bruk.

Over kort, lokale distanser vil data formidles trådløst i *mobilnett* som en naturlig utvidelse av den telefonteknologi vi kjenner fra slutten av de

20. århundrede. Den regionale trafikk vil formidles ved *høyhastighetskabelnett* mens global langdistansetrafikk vil formidles over *satellitnett*. Det regnes med at tallet på kommersielle kommunikasjonssatellitter vil 5 dobles allerede i de neste 5 år. I løpet av de neste 10-12 år vil utviklingen føre til bærbare datamaskiner får en kapasitet som kan sammenlignes med dagens tunge vertsmaskiner. De vil være selvforsynt med energi for minst en dag, ha trådløs tilnytting til nettverk, være integrert med mobiltelefon, og formidle både tekst, lyd og video.

I horisonten kan vi se for oss et teknologisk scenario hvor svært mange kan kommunisere og oppnå kunnskap av forskjellig slag uavhengig av tid og sted. Informasjonsteknologisk vil verden være uten grenser.

## Informasjonssamfunnet

Den teknologiske utvikling vil påvirke samfunnet på mange måter. Det har lenge vært fremhevet at den representerte kunnskap øker betydelig raskere enn det har vært mulig å gjøre bruk av kunnskapen. Noen hevder at den kunnskap som blir representert i et eller annet fysisk medium fordobles hvert tiår, andre mener det går raskere. Å gjøre den stadig raskere voksende kunnskap *mer* tilgjengelig for *flere* med sikte på å skape bedre samfunn, vil derfor være et av de viktigste utviklingsmål i tiden fremover.

I horisonten kan vi se industrielt utviklede samfunn hvor alle vil ha mulighet for tilgang til enorme kunnskapsmengder både i eget og andre samfunn. Disse samfunn vil ha andre utdannings-, arbeids- og fritidsmønstre enn de som er vanlige idag. Det vil i mange tilfeller ikke være nødvendig å være tilstede fysisk i skole, på arbeidsplass eller på fritidsarena for å lære, bidra i næringsvirksomhet eller la seg underholde.

*Publisering* på nettet vil tilta. Ved utgangen fra det 20. århundrede er veksten av nyhetsstoff på nettet raskt økende og det er sannsynlig at avisenes nettutgaver vil få voksende betydning som kilde for mer utdypende nyhetsstoff fordi disse vil tillate brukerne å bla tilbake til tidligere reportasjer. De fleste bøker vil etterhvert bli tilgjengelige i nettutgaver enten for direkte avlesning eller utskrift av personlige kopier. Film vil kunne hentes ned og vises privat når og hvor som helst og overta videokassettenes rolle.

Mengden av tilgjengelig informasjon vil bli mangeartet og stor. Den vil leveres i abonnement fra *informasjonsmeglere* i samsvar med personlige/ organisasjonsmessige *informasjonsprofiler* eller skreddersydd etter mål for spesielle behov. Ved hjelp av sterke

*søkesystemer*, vil brukerne få den informasjon de har behov for når behovet oppstår. Både klientutstyr og vertsmaskiner vil ha sterke innebyggede *redigeringsverktøy* som integrerer og tilrettelegger informasjon av ulike slag fra flere kilder til den form som søkerens kontekst og personlige ønsker krever.

Lett tilgang til aktuell og relevant informasjon vil stimulere innbyggerne til sterkere samfunnsengasjement. De politiske prosesser vil gå raskere og mer i takt med velgernes oppfatninger gjennom interaktive prosesser mellom politiske representanter og deres velgere.

Næringslivet vil gjennom tilgjengelig informasjon få større innsikt i sin markedsituasjon og bli mer tilpassingsdyktig. Når kjøperen ikke har behov for å se varen eller være tilstede for å nyttiggjøre seg tjenesten, vil de fleste økonomiske transaksjoner bli gjennomført over nettet.

Interaktiv informasjonsutveksling vil prege samfunnsstrukturen. Løpende samarbeidkontakt over ubegrensede distanser vil være vanlig ved hjelp av multimedia som tekst, audio og video. Samarbeid i virtuelle grupper og arbeid utenfor oppdragsgivernes fysiske lokaler vil bli presset frem av effektivitetshensyn, ressursbegrensninger og av miljøhensyn. Kontakt blant medlemmer i ulike interessegrupper vil utvikles uavhengig av avstand.

Undervisning på lavere trinn vil fortsatt foregå i gruppe-/klasserom i nærmiljøet så lenge læringen krever mellommenneskelig kontakt, men den vil bli supplert av økende støtte fra undervisningsstoff og –verktøy på nettet. En betydelig del av den høyere undervisning og studier vil foregå som interaktiv fjernundervisning over nettet og tillate personlige tilpassede utdanningsstier og hastighet i kunnskapsoppbyggingen. Mulighetene for videre- og etterutdanning vil øke og bidra til at arbeidstakerne kan være bedre oppdaterte på sine områder. Ideen om livslang læring vil realiseres. Generelt vil utdanningsstoff være tilgjengelig for å hentes ned over nettet. En økende del av underholdning vil flyttes fra kino, sportsarena og teater til hjemmet.

De fleste aktiviteter vil legge igjen elektroniske spor og skape personvernsproblemer. De spor som registreres vil bidra til økt risiko for misbruk, men vil på den annen side også kunne medvirke til bekjempelse av kriminalitet. Omfattende og vel overveiede regelverk for å regulere organisasjon, tilgjengelighet og formidling av informasjon vil derfor bli en av fremtidens mest krevende politiske oppgaver.

Tilgang til informasjon vil bli overveldende i brukervennlige *informasjonsvarehus* for standard informasjon som ved hjelp av sterke søke- og editeringsverktøy vil lete frem ønsket informasjon fra mange kilder på forespørsel. Informasjonsproduktene vil få alternative presentasjoner som vil spenne fra tradisjonell tekst, grafikk inklusive 3D, lyd og video. Brukerne vil enten få den presentasjonsform som synes

mest hensiktsmessig for deres bruk eller vil kunne velge presentasjonsform. På linje med andre økonomisk varer og tjenester vil informasjonen få en markedsverdi bestemt av tilbud og etterspørsel.

## 3.2 Etterspørsel og bruk av statistiske produkter

Figur 2 illustrerer hvordan vi kan tenke oss den offisielle statistikkproduksjonen i horisonten knyttet opp til lovverk, datakilder og brukere av produktene, og hvordan statistikkprodusentene systematisk får feedback fra brukere og markedet som påvirker det statistiske begrepssystem, det regulerende lovverk og produksjonsplaner. Den offisielle statistikkproduksjon vil i likhet med alle andre informasjonsprosesser, bli en økonomisk aktivitet med omfang bestemt av sin inntjeningsevne og driftskostnader.

### Offisiell statistikk

For statistikkprodusenten vil tekniske og samfunnsmessige utvikling frem til den overskuelige horisont ha flere konsekvenser. Det vil bli færre utpregede brukere av statistikk, de vil bli brukere av informasjon hvor statistikk vil være en av flere komponenter. Brukerne vil ha liten interesse for hvem de enkelte komponenter i et budskap kommer fra så lenge kvaliteten er tilfredstillende. På samme måte som andre produkter og tjenester, vil brukerne assosiere informasjonsprodusentene med produktens kvalitet.

Offisiell statistikk er i dagens svenske lovgivning knyttet til bestemte slag av statistisk informasjon. I fremtiden vil offisiell statistikk kunne sammenlignes med et varemerke som garanterer statistisk kvalitet. Offisiell statistikk vil særlig være fokusert på samfunnsbestemte oppgaver fastlagt i det generelle regelverket for informasjonsformidling og finansiert gjennom offentlige bevilgninger. Disse oppgavene vil omfatte de offentlige myndigheters oppgaver og fri informasjonsformidling om samfunnets tilstand og utvikling.

Dette vil gi de offisielle statistikkprodusentene eksklusive rettigheter til å samle inn, bearbeide og formidle bestemte typer av informasjon som krever behandling av konfidensielle opplysninger. Begrunnelsen for å gi statistikkprodusentene eksklusive rettigheter til å samle inn konfidensielle data fra administrative kilder og direkte, vil være at informasjon om samfunnsutviklingen ofte forutsetter behandling av konfidensielle data. Denne type av statistisk informasjon vil få økende betydning i et

samfunn hvor trivsel for alle vil være en sentral målsetting. Det vil bli alminnelig politisk enighet om at samling av disse oppgaver i et offisielt statistikkssystem underlagt streng offentlig kontroll vil innebære minst risiko for misbruk.

Fordi det vil være naturlig å knytte de offisielle statistikkprodukter til annen samfunnsorientert statistikk, vil de offisielle statistikkprodusenter også tilby andre statistikkprodukter og -tjenester i konkurranse med andre produsenter.

## Statistikkbrukere

Blant de som vil etterspørre informasjon hvor offisiell statistikk inngår kan vi sonde mellom brukergrupper med følgende forventede orientering:

*Nasjonale myndigheter:* Statistikk om nasjonale og internasjonale forhold.

*Regionale myndigheter:* Stedfestede sosio-økonomisk statistikk.

*Næringsliv:* Nasjonal og internasjonal stedfestede statistikk.

*Interessegrupper:* Detaljert statistikk om befolknings-/næringsgrupper.

*Nyhetsmedia:* Statistikk om endringer.

*Undervisning:* Oppdatert statistikk om nasjonale og internasjonale samfunn.

*Forskning:* Anonymiserte mikrodata

*Publikum:* Orienterende statistikk om samfunnet tilstand og utvikling.

Det er rimelig å forvente at hovedtyngden av brukerne i fremtiden ikke vil henvende seg direkte til den offisielle statistikkprodusent, men henvende seg til informasjons-varehus for å få mer vanlige typer av informasjon eller til informasjon-smeglere for å få spesiell informasjon. Det vil i første ledd være disse som vil statistikkprodusentenes vesentligste kunder.

## Statistikkprodukter

Etterhvert som offisiell statistikk blir mer etterspurt i flere grupper, vil etterspørselen trekke i følgende hovedretninger:

*Detaljerte faggruppering:* Avanserte analyseverktøy for å følge med i utviklingen og støtte opp om beslutningsprosessene vil stille større krav til statistikk om mer detaljerte faglige grupperinger.

*Stedfestet informasjon:* Regionale myndigheter og næringsliv vil bli mer interessert i å få fastlagt hvor ressurser finnes og utvikling foregår. Etterspørsel av informasjon om spesielle stedfestede områder vil bli viktig, men vil også skape problemer.

*Internasjonal kompatibilitet:* Med økende mobilitet og informasjonsutveksling vil avhengigheten mellom og interessen for andre land tilta. Markedene vil i økende grad bli internasjonale. Behovet for å være vel orientert vil øke og føre til voksende krav til og etterspørsel etter komparativ internasjonal statistikk.

*Informasjon om utvikling:* En stor del av den tradisjonelle statistikk har vært knyttet til periodiske undersøkelser. Fremtidens statistikk vil i langt større utstrekning være orientert mot å belyse utvikling, endring i og mellom ulike grupper.

Som all annen informasjon, vil også statistikkproduktene presenteres på flere ulike måter som tekst, tabeller, grafer av forskjellige slag, tale og video.

## Statistikkmarked

Det vil knytte seg tap til ikke å være orientert eller mangle kunnskap i fremtidens informasjonssamfunn. Offisiell statistikk vil derfor få økonomisk verdi og brukerne vil etterspørre statistiske informasjonsprodukter for å unngå de tap som knytter seg til uvitenhet.

Produktene vil bedømmes etter sin kvalitetsdeklarasjon og den tiltro produsentene opparbeider som statistikkleverandører. Statistikkprodusenten må vokte seg for å deklare bedre kvalitet enn produktene holder fordi markedet på sikt vil lære hvilke produsenter som leverer produkter i samsvar med sine deklarasjoner.

## Organisasjon av statistikktilbudet

Allerede i midten av det 19. århundrede gjorde statistikerne sine produkter sammenlignbare gjennom bruk av de statistiske standarder. Fra internasjonale organisasjoner har arbeidet med å gjennomføre og



tilby internasjonal statistikk etterhvert fått store dimensjoner i form av regionale og globale statistiske organisasjoner.

Behovet for å se statistisk informasjon i en videre kontekst krever at de enkelte deler kan integreres med hverandre og med annen informasjon. I hvert nasjonalt miljø vil det derfor være et sentralt koordinerende statistikkorgan med hovedansvar for å utvikle og formidle statistiske begrepssystemer og standarder i internasjonalt samarbeid med hverandre. I en verden uten informatiske grenser sterkt påvirket av IT vil det være voksende behov for felles internasjonal standardisering av begreper og prosesser for å kunne få sammenlignbare produkter.

Den konkrete fremstilling av de offisielle statistiske produkter vil være distribuert til de organer som har de beste faglige og metodiske forutsetninger for å produsere de enkelte deler av den offisielle statistikken. Hvor produksjonen fysisk foregår vil være av underordnet betydning. Tilbudet av produktene vil omfatte mer metainformasjon for å tilfredstille kravene om brukervennlighet og innsikt, og bli distribuert gjennom forskjellige kanaler. Mange av produktene vil samles og tilbys fra nettverk av store nasjonale og internasjonale informasjonsvarehus og -meglere. Det vil være spesielt utviklede verktøy for å finne og integrere statistikk fra ulike kilder og tilrettelegge denne for den faktiske etterspørsel. For å dekke globale informasjonsbehov, vil såvel nasjonale statistikkprodusenter som nasjonale/regionale informasjonsvarehus og -meglere være knyttet sammen i globale leverandørnettverk.

(Figur 2: Den fremtidige offisielle statistikkproduksjon og dens omgivelser)

### 3.3 Produksjon av offisielle statistiske produkter

#### Spesialisering

Den statistiske informasjon vil komme fra mange kilder. Etterhvert som statistisk informasjon får en markedsverdi vil kommersielle statistikkprodusenter etablere seg og konkurrere med de offisielle produsenter om å levere statistisk informasjon [Franchet 1994]. En viktig del av den statistiske informasjon som samfunnet vil ha behov for vil imidlertid kreve adgang til mikrodata fra administrative registre, og gjennomføring av spesiell datainnhenting med hjemmel i oppgaveplikt. Disse rettigheter supplert med strenge plikter om behandling av data, vil være fastlagt ved lov og delegert til det offisielle statistikkssystem.

Produksjon av slik statistikk vil kreve en innsikt som svært ofte vil være knyttet til utøvelse av andre administrative myndighetsoppgaver. Det er rimelig å anta at produksjonen av offisiell statistikk vil ligge tilrette for betydelig distribusjon til og spesialisering i samarbeidende statistiske produsenter knyttet til flere offentlige myndigheter. Disse produsenter vil samlet utgjøre det offisielle statistiske system.

På den annen side vil lovhjemlet innhenting av data utelukkende for statistiske formål og samkjøring for statistiske formål av mikrodata fra flere administrative registre som ansees som konfidensielle, foregå i hos en sentral statistisk produsent med disse oppgaver som særlige ansvar. Det er naturlig at produsent også har ansvaret for utvikling, vedlikehold og bruk av statistiske databaser bygget på slike mikrodata fra flere kilder med tilhørende formidling.

## Konkurransen

I vår horisont vil endel av det som nå faller inn under offisiell statistikk bli produsert utenfor det offisielle statistikkssystemet. Det vil være andre produsenter som finner det regningssvarende å innhente egne data og viderefordre data fra de offisielle produsentene og utføre disse oppgaver mer effektivt enn det offisielle statistikkssystemet.

Når den statistiske informasjon får en klar markedsverdi, må vi derfor forvente et marked med konkurranse mellom flere produsenter som utnytter spesiell kompetanse og sikrer seg markedsledelse på flere felter.

De offisielle statistikkprodusentene vil på sin side som særlige fortrinns ha:

- databaser med historiske statistikk utviklet fra administrative registre,
- adgang til ad hoc å utnytte administrative mikrodata innsamlet av offentlige,
- rett til å innkreve informasjon fra publikum for statistiske formål,
- oppgavegiverens tillit på grunn av lovbestemt plikt til konfidensiell behandling,
- statistisk opplysningsplikt til publikum offentlig finansiert.

## Kontinuerlig registrering

En betydelig del av den offisielle statistikkproduksjonen vil være basert på mikrodata samlet inn for andre formål enn statistiske av offentlige myndigheter. Som følge av IT-utviklingen, vil denne innsamling i økende grad bli hendelsesorientert, det vil si registrering vil foregå når de

aktuelle hendelser inntreffer, ikke retrospektivt som registrering av hvilke hendelser som har inntruffet under en gitt periode eller som periodisk statusregistrering.

Kontinuerlig registrering av administrative data vil gi grunnlag for kvalitetsmessig bedre statistikk fordi grunnmaterialet kan oppdateres frem til den konkrete etterspørsel inntreffer.

### Avtalt dataoverføring

Selv om en hoveddel av den offisielle statistikkproduksjonen vil produseres fra registreringer for andre formål, vil det alltid være forhold som krever spesielle statistiske observasjoner og hvor de offisielle statistikkprodusenter om nødvendig gjør bruk av sin lovgitte rettighet til å hente inn opplysninger direkte fra de statistiske enheter i de grupper som er i fokus.

En del av innsamlingen vil kunne foregå automatisk etter avtale. Det vil inngås avtaler direkte med oppgavegiverne om å legge inn standard rutiner i deres IT-systemer som automatisk overvåker og registrerer de transaksjoner som statistikkprodusentene har behov for og videresender disse til statistikkprodusentene uten annen medvirkning fra oppgavegiverne.

### Standardisering

Når data fra mange kilder integreres, er det særdeles viktig at komponentene er logisk avstemt så langt som mulig. Det gjelder både når statistiske data settes sammen, og når de settes sammen med data av annet slag.

Statistikkprodusentens aktive medvirkning i det internasjonalt samarbeid om utvikling av standardiserte begrepssystemer som reflekterer den virkelige verden på en hensiktsmessig måte, vil sikre globale sammenligninger. Det vil være viktig at bruk av begrepsstandardene følges opp nasjonalt. Dette vil være et kontinuerlig ansvar for det offisielle statistikksystemet.

## Metadata

Metadata, begreps-, prosess- og andre bruksbeskrivelser av statistiske data, vil være like viktige som numeriske statistiske data for å gi statistisk informasjon. Tidlig på 1990-tallet ble det allminnelig forståelse for at systematisk registrering og tilrettelegging av metadata var like nødvendig som utarbeiding av de statistiske tall.

I et fremtidsperspektiv vil vi se et voksende formalisert statistisk metadata-system som bearbeides like intensivt som den numeriske part og gir grunnlaget for fleksibel og effektiv integrering av mikro-, makro- og metadata til et etterspurte statistisk informasjonsprodukt.

## Kvalitetsdeklarasjon

Et siste, men ikke minst viktig, aspekt som vil karakterisere et statistisk produkt i et statistisk konkurransemarked er kvalitetsdeklarasjonen. Kvalitet i denne sammenheng er et vidt begrep som gir uttrykk for hvor godt statistikktilbudet tilfredstiller et brukerbehov. Det favner således om både produktets relevans og aktualitet såvel som dets nøyaktighet. Kvalitetsdeklarasjonen vil inngå som en naturlig og standard komponent i et produkts metainformasjon og gi et grunnlag for brukerne til valg mellom alternative informasjonstilbud.

## 3.4 Interaksjon mellom brukere og produsenter

### Forventede statistikkbehov og -formidling.

Markedskunnskap vil bli viktig for fremtidens offisielle statistikkprodusent. Den forventede konkurranse vil kreve at statistikkprodusenter må være lydhøre for brukernes behov for statistisk informasjon i nåtid og fremtid.

Statistikkprodusentene må både gå aktivt ut og hente inn informasjon om hvilke problemer brukerne forventer skal bli statistisk belyst i fremtiden, og selv gjøre prognoser om behovsutviklingen fordi statistikkproduksjonen er tidkrevende og svært ofte først gir resultater etter at en utvikling har vært observert over en periode.

Innhenting av behøvsdata fra markedet vil både kreve videre begrepsutvikling for å få beskrevet nye fenomener tilfredstillende og verktøy for å bearbeide de markedsdata som blir samlet inn.

I det fremtidige samfunn med sterkt utviklede globale kontakter, vil det komme klart frem at brukerne for mange formål vil ha behov for internasjonal statistikk som tillater sammenligninger mellom forhold i ulike land. En betydelig del av fremtidige behov vil synliggjøres som etterspørsel rettet mot internasjonal statistikk hvor den enkelte statistikkprodusent bare vil bli en av flere medvirkende leverandører som illustrert i Figur 3. I figuren viser de stiplede linjene hvordan mikrodata samles inn fra oppgavegiverne, mens de heltrukne linjer viser hvordan statistikkproduktene flyter ut til brukerne langs forskjellige veier. Doble linjer viser hvordan brukernes beslutninger tilslutt påvirker samfunnsutvikling.

(Figure 3: Den offisielle statistikkproduksjon i internasjonalt perspektiv)

## Markedsføring

For en statistikkproduksjon som må tilpasses etterspørselen i markedet, vil også markedsføring av statistikkproduktene kreve vesentlig mer oppmerksomhet av statistikkprodusentene enn hva som var nødvendig for en statistikkproduksjon i en tid hvor noen få store dominerende brukere på forhånd har bestemt og betalt for hva som skal produseres.

En viktig komponent i markedsføringen av offisielle statistikkprodukter vil være de muligheter som fremtidens elektroniske reklame vil by på. Med den mengde av informasjon som vil tilbys i et informasjonssamfunn, må den offisielle statistikkprodusent regne med å være en aktiv markedsfører, synliggjøre og fremheve både anvendelsesmuligheter og kvaliteten av sine produkter. I god tid i før nye statistikkprodukter vil foreligge og kunne leveres vil statistikkprodusentene måtte annonsere sine produkter.

## Tilbakemelding fra brukerne

Etterspørsel etter de offisielle statistikkprodukter vil avhenge av kvalitets-deklarasjonenes pålitelighet. Tilbakemelding fra brukerne vil være den viktigste kilde for statistikkprodusentene for å evaluere påliteligheten av egne kvalitetsprediksjoner og gjøre nødvendige tilpassninger.

Mens det i varemarkedet kan gis returrett når en vare ikke holder den kvalitet som er lovet, er kvalitetsbedømmelsen av informasjonsprodukter så komplisert at returrett neppe lar seg gjennomføre. Ofte vil kvalitetssvikt ikke komme tilsyne før den er lagt til grunn for beslutninger og det

vil være for sent å søke etter alternativ informasjon. Statistikkmarkedet i horisonten vil derfor være avhenge av en kontinuerlig tilbakemelding fra brukerne om deres erfaringer med statistikkproduktene for å bygge opp og beholde brukernes tillit. Går vi tilbake til Figur 2 er denne tilbakemelding illustrert ved stiplede linjer.

## 4. Mål og strategi for statistikkproduksjon

I dette avsnittet skal vi skissere noen av de tiltak som må gjennomføres for å møte de utfordringer som visjonene i de foregående avsnitt gir.

### 4.1 Statistikkmarkedet, struktur og konkurranse

#### Forberede statistikkproduksjonen på et konkurransemarked

De offisielle statistikkprodusenter må på kort sikt forberede seg på at produksjon av statistisk informasjon i fremtiden ikke vil bli forbeholdt de tradisjonelle produsenter. Som på andre felter som tradisjonelt har vært forbeholdt offentlige organer vil også statistikkproduksjon bli åpen for de som ønsker å investere i nødvendig kompetanse, tekniske ressurser og data. Den offisielle statistikkprodusent må forberede seg til å identifisere sine komparative fortrinn.

Kompetanse og tekniske ressurser vil kunne kjøpes av private statistikkprodusenter. Det offisielle statistikksystemet må identifisere til hvilke felter det har eksklusiv eller spesiell adgang til grunndata, på hvilke felter det har spesielle rettigheter på grunn av konfidensialitet og på for hvilke statistikkprodukter det har spesiell markedsføringsplikt.

På felter hvor de offisielle statistikkprodusentene stiller på linje med kommersielle konkurrenter, kan det også bli spørsmål om statistikkprodusentene har spesielle kompetansefortrinn som gjør dem konkurransedyktig både når det gjelder innsamling, bearbeiding og formidling, eller om de med fordel kan overlate eller sette bort deler av produksjonen til produsenter utenfor det offisielle statistikksystemet.

#### Bygge ut et offisielt statistikktilbud

På sikt må statistikkprodusentene være innstilt på at mange sluttbrukerne vil henvende seg til informasjonsvarehus eller -meglere som videreselger informasjon av alle slag og i alle kombinasjoner. Statistikkprodusentene må derfor arbeide for å bygge ut et offisielt statistikktilbud og få avtaler med de virksomheter som etablerer seg for å selge og formidle

informasjon. Bare spesielle sluttbrukere vil ha fordeler av å få produktene levert direkte fra de offisielle statistikkprodusentene.

## 4.2 Lov om offisiell statistikk

### Revidere loven om offisiell statistikkproduksjon

Gjeldende lov er *Lag (1992:889) om den offisiella statistiken*. Loven er supplert med *Förordning (1992:1668) om den offisiella statistiken*. Dette regelverk definerer offisiell statistikk ved en liste som oppdateres etter behov. For bedre å få understreket hva offisiell statistikk er og hvorfor ansvaret for denne må gis til det offisielle statistikksystemet, bør loven revideres slik at den definerer offisiell statistikkproduksjon som nødvendige virksomhet for å

- produsere statistisk samfunnsinformasjon for offentlig undervisning, folkeopplysning og utøvelse av offentlige myndigheters oppgaver,
- koble konfidensielle data innsamlet for andre formål av offentlige myndigheter med sikte på produksjon av statistisk informasjon,
- innhente oppgavepliktige data for statistiske formål,
- produsere statistisk informasjon fra databaser bygget på konfidensielle mikrodata.

Ellers synes det rimelig at ansvaret for den offisielle produksjonen kan være fordelt på flere produsenter avhengig av faglig kompetanse, men at en sentral statistikkprodusent gis ansvar for å

- utvikle statistiske standard begrepssystemer, herunder registerstrukturer.
- samarbeide med tilsvarende utenlandske organer om felles standarder,
- overvåke bruk av vedtatte begrepssystemer.

Det sentrale organ bør også ha ansvar for å samle de offisielle statistikkprodusenter som faggruppe og være gruppens felles talerør.

Statistikkprodusentene må i et lengere tidsperspektiv være forberedt på at lovgivningen må oppdateres og tilpasses de nye betingelser som en informatisk samfunnsutvikling vil kreve.



## 4.3 Kontinuerlig oppdatering av statistikkproduksjon

### Utrede hvilke statistikkbehov som administrative data kan dekke

I de senere år har ideen om en arkiv- eller registerbasert statistikkproduksjon fått ny næring [Danmark Statistik 1994, Statistiska Centralbyrån 1996a og Statistiska Centralbyrån 1997]. Drivkraften bak denne trend har vært den teknologiske utvikling og målsettingen om en omfattende og effektiv offisiell statistikkproduksjon med høy kvalitet. I stadig flere sammenhenger ser vi at spørsmålet fremmes om hvordan vi best kan nyttiggjøre oss de omfattende mengder av mikrodata som offentlig myndigheters administrative registreringer representerer.

Et nært mål for den offisielle statistikkprodusent vil være å identifisere hvilke administrative data som kan utnyttes statistisk, hvordan de kan harmoniseres og hva som kan gjøres for å skape et økonomisk grunnlag for å produsere offisiell registerstatistisk fra disse datakilder.

### Videreutvikle metoder for registerstatistikk

Det offisielle statistikksystemet må videreutvikle og foredle metodene for effektiv utnytting av administrative datakilder. Med eksklusive rettigheter til å koble slike sammen for statistiske formål, vil den offisielle statistikkprodusent ha et fortrin som ikke andre statistikkprodusenter vil ha. De enorme mengder av administrative data representerer vil være det aller viktigste grunnlag for fremtidig offisiell statistikkproduksjon. Utnyttelse krever imidlertid både nye modeller, metoder og teknikk for statistikkproduksjonen.

### Videreutvikle metoder for ad hoc undersøkelser

Med utgangspunkt i de statistiske bakgrunnsdata som administrative datakilder vil gi, vil de offisielle statistikkprodusenter få et forbedret utgangspunkt for å gjennomføre spesielle *ad hoc* undersøkelser med økt effektivitet i datainnhenting og bearbeiding. Teknologisk utvikling vil også gi nye muligheter for raskere innsamling og oppfølging. Dette vil sette nye krav til surveymetoder med tyngere vekt på utnytting av tilgjengelig bakgrunnsinformasjon om de statistiske enheter.

## 4.4 Informasjon om markedet og dets aktører

### Øke innsikten i statistikkbehovene

Mens produsentene historisk har fått sin finansiering gjennom noen få, store brukere som har dominert utformingen av produktene, vil det fremtidige marked for offisiell statistikk være preget av et voksende antall sluttbrukere med ulike profiler. For å finansiere og planlegge fremtidig virksomhet, må statistikkprodusentene bruke en større del av sine ressurser på få innsikt i de statistikkbehov forskjellige brukergrupper har og vil få, hvilke krav til kvalitet de vil knytte til etterspørselen etter produktene, hvilken pris de vil betale for produktene, og hvordan produktene best kan markedsføres.

Blant de mange spørsmål som i denne forbindelse må utforskes, vil også være hvordan brukerne reagerer på og prefererer ulike former for presentasjon av statistisk informasjon. For å fylle sin oppgave som formidler av samfunnsviden til publikum, vil det være nødvendig for statistikkprodusentene å ha kunnskap om hvilke presentasjonsformer ulike statistiske informasjoner bør gis for at de best mulig skal forstås av publikum.

Det må derfor være et mål å bygge opp rutiner for innhenting og bearbeiding av brukernes statistiske informasjonsbehov, markedsatferd og etterspørsel. Disse data vil være en naturlig utvidelse i metadata systemet knyttet til de produkter atferden er rettet mot.

### Planlegge bedre tilbud

På lengere sikt må statistikkprodusenten i sin planlegging og virksomhets-administrasjon ta konsekvensene av den endring som vil skje i markedet. Det vil blant annet være tale om å nyttiggjøre seg den markedsinformasjon som er samlet inn med sikte på å kunne møte markedsetterspørselen etter offisiell statistikk på beste måte ved planlegging av kvalitetsmessig relevante tilbud av offisielle statistikk-produkter.

## 4.5 Metadata system

### Videreutvikle metadatasystemet

Opp til ganske nylig var det tallene i tabellene som fikk størst oppmerksomhet. For få år tilbake initierte SCB innføring av et metadatasystem som skal gi produsent og brukerne bedre kjennskap til de offisielle statistikkprodukter [Sundgren 1993]. Et metadatasystem har som mål å gi bruks- og produksjonsrelevante opplysninger om de numeriske størrelser som inngår i statistikkproduktet. Oppbygningen av et dokumentasjonssystem hvor data om definisjoner, og prosesser som leder til de numeriske resultater inngår, er allerede vel igang. [Statistiska Centralbyråen 1995].

Tilpassing til et informasjonsmarked vil kreve at metadatasystemet utvides med resultater fra både effektivitetsmålinger i produksjonen og kvalitetsmålinger av de ferdige produkter. Metadata og numeriske data må kunne leveres integrert slik at de gir fullverdig statistisk informasjon. Det må derfor utvikles integrasjonsverktøy som tilpasser de to typer data til hverandre og i en form som etterspørres av brukeren.

### Utvikle kvalitetsindikatorer

Utvidelser i metadatasystemet i form av objektive kvalitetsindikatorer må innarbeides i rutine for den offisielle statistikkproduksjonen slik at brukerne kan få informasjon om den usikkerhet som knytter seg til bruk av de forskjellige statistikkprodukter og at å planleggerne av statistikkproduktene kan få muligheter for å vurdere kvalitet mot produksjonskostnad. Systematisk innføring av kvalitetsindikatorer vil være en kostbar, men nødvendig delprosess i statistikkproduksjonen som hittil har vært forsømt, og må tas opp som en viktig oppgave.

## 4.6 Harmonisering av de ulike statistikkprodukter

### Videre harmonisering

Statistikkprodusentene har tradisjonelt oppfattet mange av sine produkter som separate informasjonsprodukter. Det har vanligvis vært overlatt til brukerne selv å tilpasse de ulike statistikkprodukter til hverandre og til andre informasjonsprodukter for å få den informasjon som skulle dekke deres behov. Den økende bruk av ulike nasjonale og internasjonale statistikkprodukter som forventes, vil kreve at produktene er så langt som mulig logisk konsistente.

Det vil være behov for fortsatt harmonisering mellom de ulike statistikkprodukter, både ved innføring av nye begrepsstandarder og ved tilpassinger av produkter fra eksisterende begrepssystemer. Metadata systemet som vil være et naturlig hjelpemiddel for å identifisere behov for og styre harmoniseringen av statistikkproduktene.

### Harmoniseringsverktøy

På sikt må automatiserte harmoniseringsverktøy bli tilbudt i tilknytning til generelle søkemotorer. Faren ved slike verktøy kan bli at de ikke utnytter tilgjengelig metadata og foretar ineffektive tilpasninger av produktene. Ved utvikling og bruk av egne harmoniseringsverktøy som utnytter tilgjengelig metadata, særlig kvalitetsdata, om mikro- og makrodata, vil statistikkprodusentene være rustet til å tilby de offisielle statistikkproduktene mest mulig konsistente for ulike anvendelser.

## 4.7 Markedsføring

Skal den offisielle statistikkproduksjonen komme i økonomisk balanse, vil det være viktig for statistikkprodusentene å profilere seg i markedet som eksklusive leverandører av offisielle statistikkprodukter som bygger på adkomst til datakilder som er utilgjengelige for andre. Andre statistikkprodusenter som søker å tilby konkurrerende informasjon vil derfor ikke kunne gi produkter med samme kvalitet. De offisielle statistikkprodusentene må derfor utvikle markedsføringsprogram som fremhever disse fordelene ved de offisielle statistikkproduktene for brukerne.

I et informatisk samfunn basert på høyteknologi vil bruken av informasjon bli langt mer bevist og statistikkprodusenten må kontinuerlig være forberedt på nye produksjons- og markedsføringsløsninger. Fordi det forventes at brukerne i økende grad vil rette sin etterspørsel etter informasjon som løser eller bidrar til å løse konkrete problemer og hvor det offisielle statistikkproduktet bare inngår som en av flere komponenter, må statistikkprodusenten innstille seg på at den direkte kundekontakten blir overtatt av systemer som samler inn informasjon fra mange kilder, integrerer og komponerer det etterspurte produkt. Kontakt med informasjonsvarehus, -meglere og andre som viderebearbeider produktene blir derfor desto viktigere.

## 4.8 Kunnskapsoppbygning

Utviklingen inn i et informatisk samfunn hvor de offisielle statistikkprodukter får en markedsverdi bestemt av produktenes kvaliteter og etterspørsel av produktene, vil kreve at produsentene må utvide sin metodiske og tekniske kompetanse på flere felter. Etter- og videreutdanning av personalet må bli enda viktigere aktiviteter hos statistikkprodusentene.

I tillegg til de mer tradisjonelle felter, må utvikling av produsentenes kompetanse i tillegg til de rene statistiske emner også omfatte større satsing på disipliner som psykologi, administrasjons- og organisasjonsteori, IT-teknologisk kompetanse og på kommunikasjonsteknologi med sikte på nye retninger innen telekommunikasjon.

## 4.9 Systemutvikling og IT utbygging

### Statistiske databasesystemer og produksjonsprosesser

De statistiske datasystemer har gjennomgått en sterk utvikling i den senere tid, ikke minst på grunn av etablering av de statistiske databasesystemer som fortsatt vil bli ryggraden i enhver offisiell statistikkproduksjon. Den teknologiske trend viser at tids- og kapasitetsbegrensninger som hittil har virket bremsende på anvendelse av disse hjelpemidler vil reduseres. Arbeid med videreutvikling av de statistiske databaser må derfor fortsatt være en primær oppgave for statistikkprodusenten. Spesielt vil integrasjon av metadata og makrodata skape nye utfordringer for hensiktsmessig konstruksjon av databasenes struktur.

Den teknologiske behandlingsskapasitet vil tillate at mer avanserte bearbeidings-algoritmer som gjør mer intensiv bruk av tilgjengelige data, tas i bruk. Med de store mengder av tilgjengelig mikrodata vil utvikling av programvare som kan utnytte store mengder av bakgrunnsdata være en systemteknisk utfordring.

## Teknologi

På lengere sikt vil de nye kommunikasjonsverktøy som sette sitt preg på datainnhenting og -formidling, mens bedre grafikk, animasjon og 3-D teknologi vil åpne for tilbud av nye produkter til brukerne. Teknologiu utviklingen vil gi statistikkprodusentene store utfordringer med hensyn til videreutvikling av den egenkompetanse som skal til for å nyttiggjøre seg slik teknologi i videreutvikling av produksjonssystemene.

### 4.10 Organisasjon av den offisielle statistikkproduksjon

#### Organisasjon

Den svenske offisielle statistikkproduksjonen fikk ny organisasjon fra 1.juli 1994 som fordelte ansvaret for produksjon av statistikkproduktene til flere samarbeidende produsenter. SCB ble imidlertid gitt ansvaret for å koordinere statistikkproduksjonen, for det internasjonale samarbeidet og for utbygging og vedlikehold av de sentrale databaser med sikte på å kunne betjene brukere fra disse. Organisasjonen tillater også at de enkelte ansvarlige produsenter med visse begrensninger kan sette bort sine arbeidsoppgaver helt eller delvis til andre.

Med den utvikling vi kan tenke oss i den nærmeste fremtid er det lite behov for store endringer i denne arbeidsfordeling. En stadig voksende del av statistikkarbeidet vil trolig av andre grunner foregå som fjernarbeid. Hvor den tunge tallbehandlingen foregår har forlenget vært et uinteressant spørsmål.

Det vil bli viktig å delta i, videreutvikle og styrke standardiserings- og koordineringsarbeidet i FN og i regionale samarbeidsorganer for at statistikkproduktene skal bli konsistente og sammenlignbare [United Nations 1998]. Med den faglige internasjonale posisjon SCB har, bør det bli et særlig bidrag fra svensk side å ta opp de spørsmål som integrasjon av meta- og makrodata vil kreve.

Etablering av internasjonale forskningssentra for utvikling og evaluering av metoder relevante for den offisielle statistikken vil være et annet felt hvor vi må håpe på samling av de nasjonale utviklingskrefter i et felles utviklingsprogram [Kennessey 1994].

Innen nettbasert formidling av statistisk, må det forberedes organisasjonsmessige endringer som vil tillate bedre flyt i informasjonsformidlingen samtidig som ansvars- og rettighetsforhold blir ivaretatt. Mens Internet hittil har vært oppfattet som fri ressurs for global formidling av informasjon, må statistikkprodusenten som andre, forberede seg på at også tjenestene over Internet er en ressurs som ikke er kostnadsfri.

## 5. Sammenfattende synspunkter

### Tekniske premisser

I de neste 10-12 år vil *IT* verktøy bli mer effektive pr.krone og bli tatt i bruk i nye anvendelser. Parallelt ser vi at markedsledere innen *telekommunikasjon* forbereder en sterkere integrasjon av de to teknologier. I slutten av denne perioden vil myndigheter, næringsliv og publikum få adgang til og formidle informasjon uavhengig av sted og distanse.

Informasjon vil bli et *økonomisk gode* med et marked hvor informasjon og informasjonstjenester omsettes mellom *produsenter*, *informasjonsvarehus*, *informasjonsmeglere* og *brukere*.

### Mål for den offisielle statistikken

Det offisielle statistikkproduktet vil bli assosiert med *økonomisk verdi* og være en av flere informasjonskomponenter i mange anvendelser. Verdien vil i høy grad avhenge av en troverdig beskrivelse av produktenes sammensetning og *kvalitet*. Det må være et primært mål for det offisielle statistikksystemet å utmerke seg som en *pålitelig informasjonsleverandør*.

Når statistikkproduktene får en markedsverdi vil den offisielle statistikkprodusenten også kunne få *konkurrenter*. Det er ikke noe mål i seg selv å ha en offisiell statistikkprodusent dersom andre kan utføre produksjonen like godt eller bedre til samme pris. I et moderne samfunn vil det alltid være legitim etterspørsel etter statistisk informasjon, f.eks. om sosiale forhold, som må bygge på *konfidensielle* mikrodata. Det å fremskaffe slik informasjon på forsvarlig måte vil berettige *eksklusive* rettigheter for det offisielle statistikksystemet og må være et av det offisielle statistikksystemets viktigste mål.

(Figure 4: Strategi for fremtidig utvikling av statistikkproduksjonen)

### Strategi for å arbeide mot den ønskede tilpassing

For å forsvare sin plass i det fremtidige informasjonssamfunnet, må den offisielle statistikkprodusenten ha en handlingsstrategi. Figur 4 illustrerer noen av de oppgaver som må overveies for en slik strategi.



## A. Henvisningar

Danmark Statistik (1994): PERSONSTATISTIK I DANMARK – ET REGISTERBASERT STATISTIKSYSTEM. Danmark Statistik. København.

de Vries, W. (1998): GETTING BETTER ORGANIZED – SOME CHALLENGES FOR OFFICIAL STATISTICS IN NEAR FUTURE. ISI Annual Report on International Statistics. Voorburg.

Duncan, J. W (1978): STATISTICAL SERVICE IN TEN YEARS' TIME. A Seminar of the United Nations Economic Commission of Europe. Pergamon Press. Oxford.

FÖRORDNING (1992:1668) OM DEN OFFICIELLA STATISTIKEN.

Franchet, Y. (1994): A PRODUCER'S PERSPECTIVE ON FUTURE DATA NEEDS. In Kennesey, Z. (ed.): The Future of Statistics. International Statistical Institute. Voorburg.

Kennesey, Z. (1994): THE FUTURE OF STATISTICS- AN INTERNATIONAL PERSPECTIVE. International Statistical Institute. Voorburg.

LAG (1992:889) OM DEN OFFICIELLA STATISTIKEN.

Statistiska Centralbyrån (1995): DOKUMENTATIONSSYSTEMET SCBDOK MED DATORSTÖDET PCDOK. Användarhandbo. Stockholm.

Statistiska Centralbyrån (1996a): REGISTER STATISTICS, RESPONDENT COSTS, AND SURVEY PROCESS IMPROVEMENT. Proceedings of the 2<sup>nd</sup> International Conference on methodological Issues in Official Statistics. Stockholm.

Statistiska Centralbyrån (1996b): FRAMTIDEN REDAN IDAG.  
Stockholm

Statistiska Centralbyrån (1997): VERKSAMHETSSEMINARIUM  
KRING REGISTERSTATISTIKFRÅGOR 8.9 APRIL 1997.  
Stockholm.

Sundgren, B (1993). GUIDELINES ON THE DESIGN AND  
IMPLEMENTATION OF STATISTICAL MEATAINFORMATION  
SYSTEMS. Statistics Sweden. Stockholm.

United Nations (1992): REPORT OF THE REVIEW GROUP ON  
STRENGTHENING INTERNATIONAL STATISTICAL  
COOPERATION. N.Y.

United Nation (1998): REPORT OF THE AD HOC GROUP  
ESTABLISHED BY THE STATISTICAL COMMISSION. N.Y.

## B. Begreper og termer

I dette vedlegg vil vi gi en oversikt over endel av de begreper som er brukt i notatet.

*Data:* Data en beskrivelse av et fenomen som formidles/oppbevares ved hjelp av et fysisk medium med sikte på å gi mottaker informasjon om fenomenet.

*Harmonisering:* Tilrettelegge betingelsene for at to statistikkprodukter kan nyttes sammen som en logisk enhet. Harmonisering oppnås gjennom arbeid med begrepssystemer og standarder og ved tilpasninger i operasjonelle definisjoner og rutiner.

*Horisont:* I våre drøftinger vil horisont gi uttrykk for hvor langt inn i fremtiden vi vil forsøke å ha visjoner om et system.

*Informasjonsmegler:* En organisasjon som på bestilling og mot vederlag tilbyr å hente frem og sette sammen tilgjengelig informasjon, herunder statistikk, fra en eller flere informasjonskilder.

*Informasjonsvarehus:* En organisasjon som selger et bredt spekter av informasjons-produkter fra et fysisk eller virtuelt lager.

*Integrering:* Sette sammen to eller flere statistikkprodukter til et nytt produkt. Vellykket integrering betinger at produktene er harmonisert.

*Internet:* Elektronisk informasjonsnett med internasjonal utbredelse. Det gjør bruk av vanlige telekommunikasjonsnett, men gjør bruk av spesielle protokoller for global formidling av informasjon.

*Kvalitet:* Et uttrykk for hvor godt et statistikkprodukt tilfredstiller brukerne.

*Makrodata:* Makrodata er statistikkprodukter som beskriver egenskaper ved systemer eller grupper av objekter.

*Metadata:* Metadata representerer beskrivelser av andre mikro- og makrodata, og av de prosesser som frembringer data og karakteriserer bruk av data. De representerer bl.a. definisjoner av mikrodata for individuelle objekter, hvordan mikrodata observeres, av definisjoner for makrodata og prosessene som transformerer mikrodata til makrodata.

*Mikrodata:* Mikrodata er representasjoner av observerte egenskaper ved individuelle objekter i de systemer og grupper som observeres. Det skilles mellom identifiserbare og anonymiserte mikrodata.

*Offisiell statistikk:* Offisiell statistikk er den samfunnsinformerende statistikk som helt eller delvis utarbeides med hjemmel i et lovverk som gir den offisielle statistikkprodusenten eksklusive rettigheter og pålegger ham bestemte plikter.

*Offisielt statistikksystem:* Samlingen av alle statistikkprodusenter med ansvar for å utarbeide offisiell statistikk.

*Statistikkbruker:* En person eller organisasjon som gjør bruk av statistikk. I mange tilfeller kan en statistikkprodusent også selv være statistikkbruker.

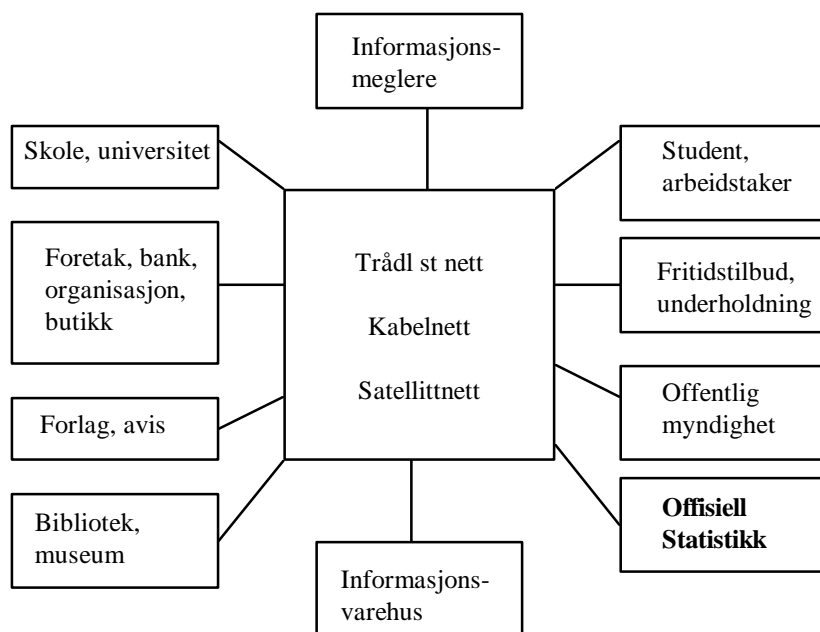
*Statistikketerspørsel:* Statistikketerspørsel er brukernes betingede forespørsel etter statistiske produkter. Eterspørselen vil vanligvis være betinget av det enkelte statistikkprodukts kvalitet og det vederlag brukeren må gi for produktet/tjenesten.

*Statistikkmarked:* Statistikkmarked er det reelle eller virtuelle møtested for statistikkprodusenter og -brukere hvor avtaler om økonomiske formidling av statistikkprodukter inngås.

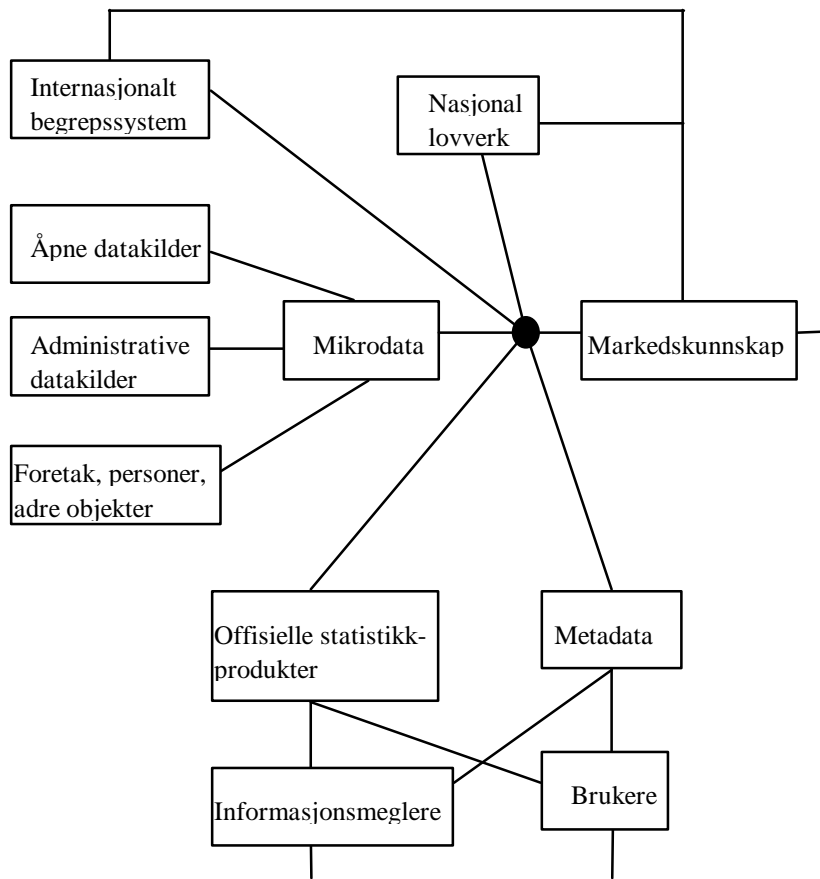
*Statistikkprodusent:* En statistikkprodusent er den organisasjon som er ansvarlig for å planlegge statistikkproduksjon, samle inn, bearbeide og ta vare på mikrodata, tilby makrodata, og utvikle metadata med sikte på å dekke etterspørsel etter statistikk.

*Statistikkproduksjon:* Statistikkproduksjonen er all aktivitet som utføres av en statistikkprodusent med mål å tilby brukerne statistikkprodukter. Statistikkproduksjon omfatter her således både planlegging, fremstilling og markedsføring av statistiske produkter.

*Statistikkprodukt:* Statistikkprodukt er et resultat av en sammensetning av statistiske makro- og metadata med sikte på å dekke et statistisk informasjonsbehov. Et statistikkprodukt kan ha bearbeidingsgrad som varierer fra anonymiserte mikrodata til avanserte makrodata med tilhørende metadata.



Figur 1: Offisiell statistikk som del av et informasjonssystem



Figur 2. Den femdige offstette statistikkproduksjon og dens omgivelser

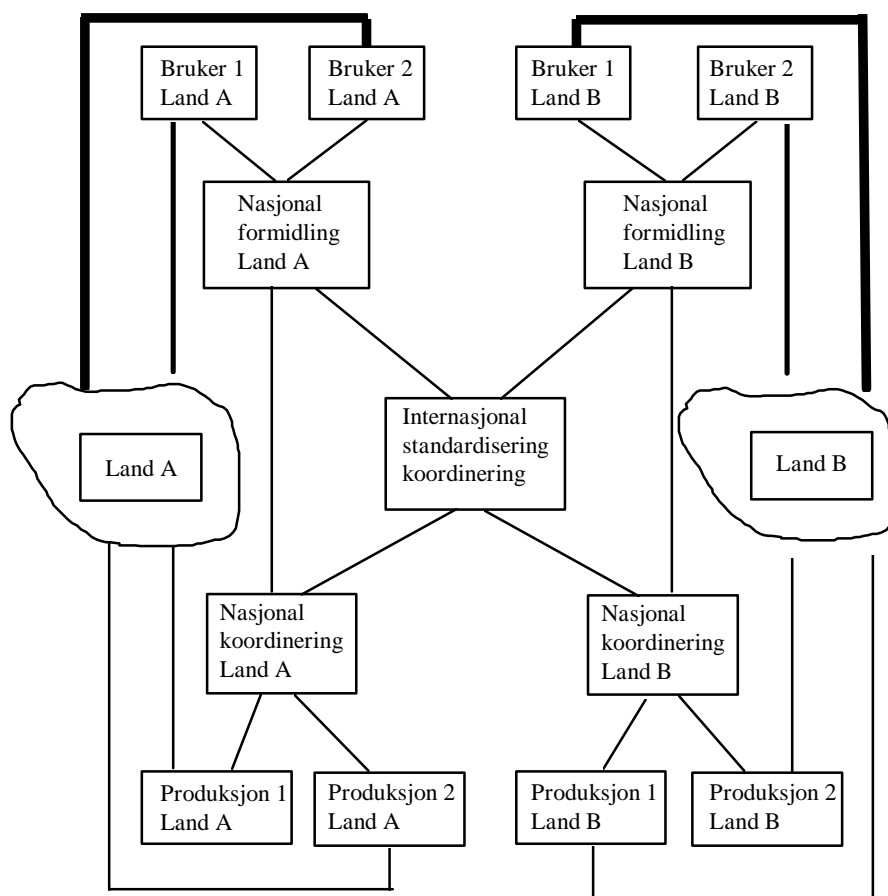


Figure 3: Den offisielle statistikkproduksjon i internasjonalt perspektiv



## STRATEGI FOR DEN OFFISIELLE STATSTIKK

Utvikle et konkurransedyktig offisielt statistikktilbud

Få presise mandat i lov for den offisielle statistikkproduksjon

Fortsette utvikling av registerstatistikk og beredskap for ad hoc undersøkelser

Søke innsikt i statistikkmarkedet og om dets aktører

Fortsette utvikling av metadata systemet

Harmonisere statistikkproduktene

Aktivere markedsføring av de offisielle statistikkproduktene

Utvikle metode- og teknologikompetanse på nye områder

Fortsette utbygging av systemer for innsamling, databaser og nettleveranser

Tilpasse statistikkorganisasjonen til nasjonale og internasjonale nettverk

Delta aktivt i utvikling av internasjonal begrepssystemer

**Figure 4: Strategi for fremtidig utvikling av statistikkproduksjonen**

## Idépromemoria från Statskontoret

Uppdragsenheten

1999-02-05

# 1 Nuläge

SCB:s uppgift är att ansvara för *förvaltningsuppgifter* såsom samordning och bevakning av den officiella statistiken, grundläggande metodutveckling, nationell statistksamordning, nomenklaturarbete, internationellt statistiksamarbete, information och statistikservice samt vissa centrala databaser.

SCB och ett 25-tal sektorsmyndigheter, har som *statistikansvariga myndigheter (SAM)* i uppgift att inom sina respektive statistikområden ansvara för att beställa statistik samt genom programråd eller liknande ge andra användare inflytande på statistikens utformning m.m. SCB ansvar för officiell statistik som griper över flera sektorer eller som saknar en entydig huvudanvändare eller där det påtagligt gagnar förtroendet för statistiken (ekonomisk, fördelnings- och välfärdsstatistik). SCB ansvar även för långsiktiga prognoser/analyser om befolkning, utbildning och arbetsmarknad. De ca 25 sektorsmyndigheterna ansvarar inom sina fakområden för sektorbunden statistik och har frihet att utan lagstadgade krav på konkurrensutsatt anbudsupphandling välja mellan SCB, andra fristående statistikproducenter eller produktion i egenregi.

Ovanstående verksamheter finansieras i regel genom SCB:s anslag och anslag vid de berörda sektormyndigheterna. Övrig statistik och statistiska tjänster finansieras av beställaren.

SCB:s tredje uppgift är att producera statistik på uppdrag av statistikansvariga myndigheter m.fl. Huvuddelen av den officiella statistiken (SOS) produceras även efter reformen av SCB på uppdrag av statistikansvariga myndigheter.

## 2 Utgångspunkter för övervägandena

### 2.1 Statistikverksamhetens mål

Som grund för politiska beslut, samhällsplanering, opinionsbildning, forskning etc. har statistik en betydelsefull samhällsfunktion. Mot denna bakgrund har statistiken en rad viktiga mål att leva upp till. Statistikens mål kan i övergripande ordalag sägas vara:

#### *Trovärdighet*

- Objektiv och förtroendeingivande statistik
- Ökad närhet mellan produktions- och analyskompetens

#### *Möjlighet att påverka*

- Ökad möjlighet för användarna att påverka statistiken
- Bättre dialog mellan statistikansvariga och användare
- Användaranpassad kvalitet (dvs. kvalitetsarbetet bör utgå från användarnas behov och önskemål)

#### *Effektivitet*

- Effektivare statistik
- Ökad konkurrens mellan olika producenter (i syfte att effektivisera produktionen)

#### *Tillgänglighet*

- Ökad tillgänglighet till den offentliga statistiken
- Ökat användande av statistik

## Trovärdighet – ett avgörande mål med flera dimensioner

Trovärdighet för statistikens och SCB:s **objektivitet** och **kvalitet** är ett mycket centralt mål som dels påverkar, dels är beroende av möjligheten till användarinflytande, tillgänglighet och effektivitet. Statistiken och SCB:s agerande måste framstå som trovärdig inför:

- Användare och statistikbeställare
- Konkurrenter
- Allmänhet

Att de som använder statistiken har möjlighet och förmåga att påverka utformning och utveckling är sannolikt av avgörande betydelse för vilket förtroende statistikproduktionen åtnjuter. Hur stort inflytandet skall vara är sannolikt en avvägningsfråga från fall till fall. Generellt bör dock inte glappet mellan producent och användare vara allt för stort.

Möjligheten att påverka och känna förtroende bör även inverka på graden av användning. Olika användares möjlighet att påverka statistiken och en mera användaranpassad produktion bör utgöra ett viktigt incitament i strävandet efter att öka samhällets statistikanvändning.

Förtroendet för statistik är även beroende av tillgängligheten. Med tanke på den genomslagskraft och opinionsbildande funktion statistik har i samhälldebatten är det, inte minst ur ett medborgarperspektiv, viktigt att det finns trovärdig och objektiv statistik lätt tillgänglig. Att enbart ta del av olika aktörers statistiska analyser utan egen tillgång till objektiv data riskerar att grundlägga ett demokratiskt glapp mellan de som har grundinformation och de som enbart tar del av bearbetad information.

Vad gäller effektivitet är det för det första, ur förtroende och legitimitetssynpunkt, viktigt med ett så effektivt resursutnyttjande som möjligt. För det andra är ambitionen att via konkurrens öka effektiviteten beroende av att SCB agerar på ett konkurrensmässigt förtroendeingivande och korrekt sätt gentemot sina konkurrenter och beställare.

## Målkonflikter

De ovan presenterade målen är alla viktiga och har, som framhållits, ett samröre med varandra. Samtidigt är de troligen inte sinsemellan helt förenliga. Det finns flera tänkbara målkonflikter att beakta.

Beroende på vem som använder statistik och i vilket syfte den används kan ökad tillgänglighet och användning hota trovärdigheten och objektiviteten. Om allt för många använder och presenterar statistik på ett förvanskat och oseriöst sätt riskerar objektiviteten och tilltron till statistiskt material att undermineras.

Det kan även ligga en målkonflikt i avvägandet mellan objektivitet och användarpåverkan i den meningen att anpassning och differentiering kan hota kontinuitet och vissa grundläggande krav på likhet (enhetlighet, jämförbarhet m.m.).

Slutligen kan konkurrens och eventuella tendenser till allt för långt gången prispress inverka negativt på kvalitetsambitionerna.

Hur målen och konflikternas allvar värderas och rangordnas får avgörande betydelse för vilken organisationsmodell och struktur som väljs.

## 2.2 Statistik – ett offentligt åtagande?

### Motiv för offentliga åtaganden

De klassiska motiven för offentliga åtaganden är att tjänsten eller varan som skall produceras har karaktären av en kollektiv eller fri nyttighet, att det finns en risk för marknadsmisslyckanden och att det föreligger naturliga monopol.

Behovet av offentliga åtaganden brukar man försöka bedöma utifrån den samhällsnyttan ställd mot kostnaden för åtagandet samt en analys av samhällspolitiska konsekvenser av ett uteblivet offentligt åtagande.

Ett marknadsmisslyckande kan dels förstås som risken för att marknadens vinstbehov inverkar negativt på en för samhället nödvändig nytta i den meningen att nyttigheten inte får generell spridning och att prissättningen inte gör den universellt åtkomlig. Produktion och distribution behöver inte nödvändigtvis vara ett offentligt åtagande, men ett marknadsmisslyckande, kopplat till naturliga monopol, kan uppstå om det inte är företagsekonomiskt attraktivt för en privat aktör att producera nyttigheten. I detta läge får staten ett naturligt monopol.

Utifrån de gängse motiven och kriterierna för offentliga åtaganden finns mycket som talar för att statistik kan vara att betrakta som ett nödvändigt offentligt åtagande. Tillgänglighet, spridning och objektiv statistik utan inslag av sär- och vinstintressen är faktorer som skulle kunna motivera ett offentligt statistiskt åtagande för att beställa och finansiera statistik. Vidare är det troligt att det ekonomiska intresset för privata aktörer är begränsat när det gäller att upprätthålla de databaser och den information som forskare, Regeringskansliet och flera myndigheter efterfrågar för forskning, utvärdering och samhällsplanering. Den avgörande frågan är hur långt det offentliga åtagandet för statistik skall sträcka sig och hur ansvaret skall struktureras inom den offentliga sektorn. Vilka uppgifter som faller inom det offentliga och vad som kan avlövas det

offentliga får betydelse för hur statistikhanteringen och SCB skall organiseras.

## Officiell statistik (SOS)

Inom begreppet officiell statistik ryms flera variabler som skulle kunna motivera ett offentligt åtagande. Genomförandekommitténs betänkande från 1994 anger bl.a. följande för den officiella statistiken:

Officiell statistik är sådan statistik för samhällsplanering, forskning, allmän information och internationell rapportering som en myndighet framställer enligt föreskrifter som regeringen meddelar.<sup>1</sup>

Den officiella statistiken skall omfatta regelbundet producerad primärstatistik samt sådana statistiska sammanställningar och vidarebearbetningar som utgör väsentligt underlag för samhällsplanering, forskning, allmän information och internationell rapportering.

Däremot ska SOS inte innehålla debattinlägg, värderande kommentarer eller analyser.

Någon entydig definition av officiell statistik har vi inte funnit. SOS-definitionen varierar också bland statistikansvariga myndigheter, kopplat till finansiering, kvalitetskontroll, publiceringsform m.m.

Hur SOS definieras har här bedömts främst påverka ansvarsområdet för SCB:s SOS-relaterade uppgifter som tillsyn och samordning m.m. Vidare påverkas statistikansvariga myndigheternas antal och uppgifter samt vilken statistik som skall produceras med statliga anslag. Organisationen i övrigt har bedömts påverkas marginellt.

<sup>1</sup> Ändrad ansvarsfördelning för den statliga statistiken (SOU 1994:1).

## 2.3 Faktorer som i framtiden kan påverka statistiken och SCB:s roller och förutsättningar

En organisations utvecklings- och förändringsbenägenhet bestäms och begränsas av såväl externa som interna faktorer. Externt tryck i form av samhälleliga förändringar, opinionstryck och politiska intentioner påverkar givetvis en organisation. Tack vare SCB:s dominerande ställning är frågan om vilken utveckling och vilka förutsättningar statistiken och SCB kan komma att ställas inför i framtiden i mångt och mycket avhängigt verkets eget agerande och strategiska val. Tre andra faktorer av betydelse är statsmakternas beslut (mål och resurser för statistiken), EU:s krav samt de statistikansvariga myndigheternas agerande. Nedan presenteras några tänkbara utvecklingsscenarier som kan få betydelse för statistiken och SCB:s organisation i framtiden.

### Ökad konkurrensutsättning

Konkurrensen är idag i praktiken tämligen begränsad. Relationen mellan SCB och statistikansvariga myndigheter kan i princip karaktäriseras som ett sektorsinternt köp-/säljssystem, dvs. offentliga medel mellan offentliga aktörer. Detta kan ha befrämjat effektiviteten inom den officiella statistiken. Någon egentlig konkurrens föreligger emellertid inte ännu, med undantag för ett fåtal statistikområden som varit föremål för en konkurrensutsatt upphandling där alternativa statistikproducenter lagt anbud.

Marknadsförutsättningarna är viktiga att beakta. Vilka producenter finns/kommer att finnas och vilka tjänster? Vilka hinder för konkurrens finns? Blir det endast fråga om olika mindre nischföretag som vänder sig till en viss sektor/kundkrets m.m. eller kommer det att finnas till SCB jämbördiga och breda konkurrenter? Vilken valfrihet kommer/bör statistikansvariga myndigheter ha att välja mellan olika producenter? Kommer det att uppstå marknadsuppdelning och prissamverkan?

SCB:s nuvarande dominans som producent och SCB:s samtidiga ägande av stora centrala grundläggande statistiska databaser som mycken statistik produceras ur, gör konkurrerande producenter mycket beroende av SCB samtidigt som det finns en risk att SCB genom sin ställning/position får insyn i konkurrenternas affärer (anbud etc.). SCB:s prissättning gentemot dessa konkurrenter kan även påverka konkurrensförhållandena.

När det gäller de statistiska grunddatabaserna är det inte av ekonomiska skäl möjligt för alla konkurrenter att bygga upp egna. Det är inte heller



lämpligt med flera stora nya personregister ur statsmakternas synpunkt med hänsyn till skyddet för samhällets sårbarhet och säkerhet samt individens integritet.

SCB erbjuder idag en rad tjänster som inte faller inom ramen för den officiell statistiken och som därmed inte är ett självklart offentligt åtagande. Exempel på sådana tjänster är intervjuundersökningar, regionala konjunkturbarometrar, servicemätningar, kundstudier och marknads- och branschprofiler. Genom sin breda verksamhet och dominerande ställning har SCB här ett stort eget utrymme att påverka framtida konkurrensförhållanden och välja hur man vill utforma sitt tjänstesortiment. Idag är konkurrensen tämligen begränsad vad gäller stora delar av statistikproduktionen. Genom att omstrukturera och konkurrensutsätta delar av SCB:s verksamhet finns det troligen möjligheter att öka antalet aktörer och därmed konkurrensen. Framförallt är frågan om hur SCB:s databaser skall organiseras avgörande i detta sammanhang.

## Analyskompetens

Idag producerar SCB i huvudsak basstatistik och utför relativt lite eget analysarbete. I detta kan det uppstå såväl trovärdighets- som konkurrensproblem. Bilden av SCB:s statistik- och analyskompetens är inte entydig. Vid intervjuer av företrädare för statistikansvariga myndigheter har det framkommit att SCB i vissa situationer ansetts ha sämre möjligheter och kompetens för kvalitetskontroll ("känsla för siffrorna"), analyser och annan vidareförädling. Å andra sidan marknadsför SCB sådana tjänster.

SCB behöver kompetens inkl. spetskompetens för att kunna förstå, påverka och granska statistiken. Som läget är idag riskerar SCB att bli gisslan i den meningen att SCB:s namn används som ett kvalitetssignum utan att man egentligen har någon möjlighet att påverka det resultat som presenteras. Om SCB:s datamaterial används i tveksamma analyser och presentationer kan detta på sikt underminera SCB:s trovärdighet.

Om SCB väljer att utveckla sin analysverksamhet kan detta även ha viktiga implikationer för konkurrenssituationen. Genom att göra mera analysarbete borde SCB kunna bredda sitt tjänsteutbud till områden/marknader där det finns flera tänkbara aktörer och en mera reell konkurrens. Å andra sidan kan man då riskera att inkräkta på de statistikansvariga myndigheternas ansvar.

## Teknikutveckling

Den tekniska utvecklingen i form av ökade möjligheter till Internet- och databasanvändande borde göra statistik mera lättillgängligt och lätt-spridd vilket kan leda till stora nyttor i samhället om statistik används i högre grad i olika beslut. Frågan är vad detta innebär för SCB:s nuvarande struktur? Det avgörande problemet ligger i hur ansvaret, finansieringen och tillgängligheten för databaserna skall ordnas. Ökad tillgänglighet till databaserna måste naturligtvis vägas mot integritets- och finansieringsproblemen. Baserna måste finansieras men det får för tillgänglighetens skull inte bli för dyrt att köpa information eller innebära försämrat integritetsskydd m.m.. Även trovärdighetsaspekten måste beaktas. Ökad tillgänglighet och oseriös användning kan underminera både statistikens och SCB:s trovärdighet. Detta gäller även de sammanhang när statistik används i opinionsbildande syfte.

## Krav på skärpt tillsyn

SCB har ansvar för samordning och att tillse att de mål och krav som riktas mot statistiken uppfylls. Tillsynsfunktionen är idag emellertid såväl otydlig som tämligen svag eftersom tillsynens mandat, uppgifter och befogenheter inte är tillräckligt preciserade. Sannolikt kan det uppstå behov och resas krav på en tydligare tillsynsroll inom statistiken, något som får konsekvenser för SCB:s organisering.

## 2.4 Förvaltningspolitiska implikationer?

De uppgiftskombinationer och situationer där dubbla eller motstridiga roller uppmärksammas är till synes mycket vanligt förekommande i offentliga utredningar och i riksdagstryck – i princip i alla de sammanhang där det finns kombinationer av verksamheter med skilda syften eller fokus.

## Rollkonflikter inom myndighetsutövningen

Det handlar bl.a. om rollkonflikten mellan rättsskipning och myndighetsutövning, inklusive varianten mellan överklagnings- och tillsynsärenden. En snarlik uppmärksammas rollkonflikt är den mellan tillståndsgivning och tillsyn.

Vidare förekommer rollkonflikter mellan service- och kontrollfunktioner, mellan rådgivning/information och tillsyn/kontroll samt mellan normerande och revisionell verksamhet. Dylika kombinationer mellan verksamheter med värnande/främjande syfte och med mera polisiär inriktning har lett till rollkonflikter som bl.a. ansetts påverka förutsättningarna för sanktioner. Riksdagens revisorer<sup>2</sup> har också betonat att förutsättningarna för en ändamålenlig tillsyn särskilt måste beaktas i samband med omorganisationer och ändrade ansvarsförhållanden.

### Rollkonflikter mellan myndighetsutövning och annan verksamhet

Rollkonflikter finns mellan myndighetsutövning och t.ex. uppdrags- och affärsverksamhet. Det gäller verksamhetsområden där klienten i ena stunden får tjänster utförda på kommersiella grunder och i andra sammanhang möter samma organisation som beslutande myndighet. Myndighetsutövningen är omgärdad av bestämmelser som skall garantera rättssäkerhet, obundenhet och medborgarnas förtroende, men hur en myndighet uppträder vid sidan av myndighetsutövningen har stor betydelse för detta förtroende.

Staten kan också som ägare ställa ekonomiska krav som är svåra att förtroendefullt förena med statens roller som reglerare, kontrollant, främjare av konkurrens och huvudansvarig för konsumenterna inom området.

### Rollkonflikter vid upphandling

Konflikter har ofta uppmärksammats i samband med konkurrensutsatt upphandling och val mellan egenregi och extern entreprenad, t.ex. när kommuner och landsting både skall beställa från sig själv och från privata vårdgivare. Även staten via bolagsägande förekommer i liknande rollkonflikter som ytterst ansvarig för både köpare och säljare.

<sup>2</sup> 1994/95:RR9

## Rollkonflikter i brett sammansatta verksamheter

Rollkonflikter förekommer också i samband med att olika nivåer och intressen är inblandade eller att gränserna dem emellan är oklara, t.ex. mellan central och regional nivå, mellan statlig och kommunal nivå eller när kommunala företrädare eller andra samhällsintressen finns representerade i statliga myndigheters ledning. En bred representativ sammansättning ger god förankring i samhället, ger fördelar i form av möjligheter till intresseavvägningar och samförståndslösningar inom myndigheten. Men samtidigt skapas en osäkerhet om vem man representerar. Intressekonflikterna ställs också på sin spets när det kan bli fråga om myndighetsingripande mot det intresse som man representerar. Partsammansättning kan innebära att myndighetsledningen blir konfliktfylld och ansvarsutkrävandet oklart.

Rollkonflikter är också vanliga i "breda" verksamheter, av typ länsstyrelse, kommun, medborgarkontor etc. Kommuner eller t.o.m. samma kommunala nämnd kan uppträda både som myndighet, kommersiell aktör och sakägare/intressent, t.ex. ifråga om fastigheter. I fallet friskolor uppträder kommuner både som finansär och utövare. Medborgarkontoren löser förvaltningsuppgifter för olika myndigheter genom en och samma person.

Rollkonflikter har även uppmärksammats för myndigheter ansvariga både för genomförande/verkställande och uppföljning/utvärdering av densamma, t.ex. att både utforma och utvärdera planer. RRV var tidigare både initiativtagare, verkställare, granskare och utvärderare. Ytterst ifrågasätts myndigheternas möjligheter att utvärdera/ompröva sin egen verksamhet.

## Andra faktorer än rollkonflikter kan väga tyngre

Förutom rollkonflikter har ofta andra samtidigt nackdelar uppmärksammats, nackdelar som delvis är en följd av rollkonflikterna, bl.a. risker för diskriminerande särbehandling, konkurrenssnedvridningar, kors-subsventionering, försämrat integritets och sekretesskydd, hotad rättssäkerhet m.m. Det kan också vara svårt att särskilja om det är rollkonflikter eller andra skäl som legat bakom beslut att renodla myndigheter. Myndighets- och förvaltningsuppgifter har även skilts från serviceinriktade verksamheter av skäl som att de senare kunnat ha en annan finansierings- och verksamhetsform samt huvudman (avgiftsfinansiering, bolagsform resp. privat). Vanligt förekommande mål bakom dylika reformer är också:

- att stärka den politiska styrningen. En tilltagande komplexitet hade gjort offentlig sektor politiskt svårstyrd och svåröverskådlig,
- att förbättra de offentliga finanserna. Ekonomiska problem och budgetunderskott hade försvårat möjligheterna att finansiera en offentlig tjänstesektor utanför statens traditionella åtaganden,
- att förnya och effektivisera den offentliga verksamheten genom att ompröva och renodla verksamheten till statens s.k. kärnfunktioner, i första hand kollektiva nyttigheter och myndighetsutövning med därtill knuten verksamhet,
- att skapa konkurrens, konkurrensutsätta, eftersträva styrning genom en marknadslogik och att göra ansvaret för målformulering, finansiering och utförande tydligare. Inom myndigheterna skulle beställarrollen separeras från producentrollen och entreprenadupphandlingar prövas som alternativ till produktion i egen regi. Konkurrensutsatt affärsverksamhet borde inte bedrivas i myndighetsform. Om verksamheten över huvud taget bedrivs i statlig regi borde bolagsformen övervägas. Renodlingsförslag förekommer även i syfte att minska en myndighets dominerande ställning.

## Renodling – statsmakternas huvudlinje

Statsmakterna har bl.a. i förvaltningspolitiska propositionen (1997/98:136) angivit inriktningen att värna om kärnverksamheterna och att fortsätta att renodla den statliga verksamheten. Som en naturlig del i detta arbete ingår fortsatta omstruktureringar och omprövningar av det offentliga åtagandet. Strukturförändringar behövs för att statsförvaltningen skall kunna utvecklas och effektiviseras. Vissa tjänster i de statliga myndigheterna, som idag utförs i egen regi, bör kunna effektiviseras betydligt om de upphandlas i konkurrens.

Konkurrensutsatt verksamhet bör inte bedrivas i myndighetsform, om det inte finns särskilda skäl. För konkurrensutsatt verksamhet bör bolagsformen övervägas. Om konkurrensen bedöms som otillräcklig eller om behovet av offentlig insyn och kontroll är särskilt stort, är myndighetsformen lämpligast. Bedriver myndigheten affärsverksamhet, bör affärsverksamheten särskiljas. Detta kan ske genom en tydlig organisatorisk uppdelning inom myndigheten.

## För- och nackdelarna är av olika art och tyngd

Nackdelarna med rollkonflikter är ett principiellt problem, som praktiskt främst tar sig uttryck genom otydlighet, ifrågasättande och trovärdighetsproblem i omvärlden. Rollkonflikter är vidare ett problem vars signifikans ofta ifrågasatts med hänvisning att det snabbt lösts internt inom organisationen.

Fördelarna med sammansatta/blandade roller är å andra sidan mera konkreta. Stordriftsfördelar, samordnings- eller effektivitetsvinster och synergieffekter genom att resurser, kompetenser och roller/uppgifter hålls samlade på ett enda ställe. De underlättar helhetssyn, ökar genomslagskraft, kraftsamling samt möjligheter till löpande kunskaps- och erfarenhetsutbyte. Bl.a. har affärsverksamhet ansetts kunna ge ökad förnyelse- och förändringskompetens. Breda kompetenser ger ärenden en bra och allsidig belysning. Avvägningsfrågor och ibland oundvikliga intressekonflikter löses på ett smidigt administrativt sätt inom en myndighet istället för att frågorna lyfts till och löses av regeringen. Renodling för att bl.a. minska rollkonflikter har också ansetts leda till försämrad överblick och fördyringar.

Tidpunkten för bedömningen av rollkonflikter kan också vara avgörande. Det gällde t.ex. tidigare i fråga om RRV, där en samlad kompetens ansågs vara värdefull för reformen att förändra styrningen och uppföljningen av statsförvaltningen. Men när reformen i huvudsak genomförts ansågs inte längre behovet av en samlad kompetens vara lika stort, varvid ESV bröts ur RRV, bl.a. för att stärka RRV:s oberoende ställning och integritet m.m. Tidigare hade kostnads- och effektivitetsskäl ansetts väga tyngre än rollkonflikter mellan systemstöd, normering och revision.

## Varierande utfall i praktiken av renodlingsfrågan

I enskilda fall och tillämpningar har det gjorts skilda bedömningar och/eller funnits skilda meningar som lett till kompromisser och olika lösningar, bl.a. när det gäller val av organisatorisk lösning och verksamhetsform (myndighet, BU-modeller, resultatenheter, bolag) samt huvudmannaskap (offentligt/privat).

I vissa fall har man funnit att principiella skäl till invändningar finns, men å andra sidan gjort bedömningen att några avgörande olägenheter inte uppstått i praktiken. I andra fall har man gått längre och framhållit att man bör vara uppmärksam på rollkonflikten, samt noga följa upp utvecklingen och vid behov/praktiska problem ta initiativ till åtgärder. Men det förekommer också fall där man strävat efter att begränsa riskerna för intressekonflikter genom preciserade mål och uppdrags-

beskrivningar, lagreglering, avtalskonstruktioner, begränsningar i registeråtkomst, minskade ekonomiska beroenden m.m.

Organisatoriska och strukturella åtgärder förekommer även för att minska rollkonflikterna. Antingen ändras myndighetens organisationsstruktur i syfte att på ett tydligt sätt hålla isär rollerna inom en och samma organisation. Genom att ställa stora krav på att de olika rollerna kan utföras oberoende och med stor integritet i förhållande till varandra tillåts inte samma personal eller organisatoriska enhet vara ansvarig för båda rollerna eller ens utbyta information med varandra. I vissa fall har t.o.m. särskilda krav på bemanning/kompetens ställts. Slutligen finns fall där man gått så långt att man lagt olika roller i skilda organisationer eller ytterst på olika huvudmän. En mellanvariant är de fall när myndigheter skapats "inom myndigheten", t.ex. inspektionsverksamheter på trafikområdet som formellt sett är självständiga myndigheter.

### 3 Bedömning av tänkbara rollkonflikter

De SCB-roller som har kunnat identifieras har grupperats nedan efter typ av uppgift. Observera att grupperingen efter uppgiftstyp går att ifrågasätta.

#### *Myndighetsuppgifter*

- helhets-/systemansvarig för officiell statistik (SOS)
- tillsynsorgan och kvalitetskontrollant för SOS
- normgivare, standardiserare m.m. för SOS
- samordnare, koordinatör för SOS nationellt och internationellt
- ägare av statistiska grunddatabaser inom SOS, ett informationskapital av stor betydelse för produktion av SOS och för förädling
- bevakare av integritetsskydd, företagshemligheter etc. i SOS

#### *Uppgifter som statistikansvarig myndighet (SAM)*

- beställansvarig finansiär (och upphandlare) av främst sektorsövergripande statistik i egenskap av statistikansvarig myndighet

#### *Stabsuppgifter, såsom en gemensam SOS-resurs för Regeringskansli och statistikansvariga myndigheter*

- expert för SOS/statistikutvecklare
- metodutvecklare för SOS
- statistik- och databasservice åt statistikansvariga myndigheter

#### *Serviceuppgifter (f.n. normalt avgiftsfinansierade)*

- statistikproducent för egen och för andras räkning (SAM m.fl.) såsom underleverantör av statistik på entreprenad
- rådgivare, tjänstekonsult, systemleverantör, partner



- vidareförädlare, dvs. analys av samt avancerad och/eller skräddarsydd tillämpning av statistik informatör, marknadsförare, försäljare, återförsäljare (grossist/ detaljist) och distributör av statistik (statistik- och databasservice)

SCB har att hantera flera olika roller och har flera mål att leva upp till. Det finns ett genomgående samband mellan mål och roller. Det är i sammanhanget särskilt intressant att se på vilka rollkonflikter som är av sådan art att de försvårar förverkligandet av de mål som inledningsvis angavs för statistiken.

Trovärdighet är ett viktigt mål i sig, men som vi inledningsvis framhöll är trovärdigheten dels avgörande för, dels beroende av målen om påverkansmöjligheter, effektivitet och tillgänglighet. Mot denna bakgrund kommer vi att hävda att trovärdighet är det viktigt, gemensamt och övergripande mål.

### 3.1 Helhetsansvar kontra egenintresse

Vad gäller SCB:s roller ligger den övergripande konflikten i förekomsten av ett helhetsansvar som riskerar att kollidera med ett egenintresse. SCB har ett statistiskt helhetsansvar som rymms inom ett offentligt statistiskt åtagande och innefattar bl.a. samordning och tillsyn. Samtidigt har man en säljare- och producentroll. Denna övergripande konflikt kan brytas ned i flera mera avgränsade problem.

#### Stabsuppgifter kontra uppdragsverksamhet

SCB har genom sina förvaltningsuppgifter till viss del rollen av ett statistiskt stabsorgan. SCB har exempelvis till uppgift att arbeta med metodutveckling, produktionsteknisk utveckling, utveckla och ansvara för standarder på databaser och sprida statistik.

Främst ur finansieringssynpunkt är frågan hur mycket av dessa stabsuppgifter som bör friläggas respektive inrymmas i den uppdragsfinansierade verksamheten och den beställarutförarrelation SCB har gentemot statistikansvariga myndigheter? Stabsuppgifterna skall i princip vara anslagsfinansierade medan tjänsteproduktionen skall vara kopplad till uppdragsfinansieringen från statistikansvariga myndigheter och andra beställare. Otydligheten i hur dessa roller skall avgränsas och vad som skall tillhandahållas fritt och inte ger kanske inte upphov till någon allvarligare rollkonflikt. Det finns dock en fara för att otydliga finansieringsprinciper på sikt kan vara till skada för SCB:s förtroende. Till följd

av budgetrestriktioner kan t.ex. SCB misstänkas för att ta betalt för uppgifter de egentligen är ålagda att tillhandahålla fritt.

## Tillsynsrollen

Tillsynsrollen är ett exempel på hur helhetsansvar kan kollidera med egenintresse. Ansvar för att tillse kvalitet och måluppfyllelse kan, bl.a. genom att direkt eller indirekt underkänna annan produktion än den egna, användas för att gynna egen produktion. Detta riskerar givetvis att påverka trovärdigheten främst gentemot konkurrenter.

Tillsynsrollen påverkar även trovärdigheten gentemot användare och allmänhet. Användarna och beställarna av statistik tillses av sin leverantör och affärsrättsliga motpart. Allmänhetens förtroende för statistiken är i sin tur beroende av att målsättningarna på ett effektivt sätt tillses. I kravet på en jämbördig affärsrelation med beställare och kraven från allmänheten på tillsyn, kontroll och effektivt resursutnyttjande kan SCB hamna i en rollkonflikt.

Ett fungerande SOS-system kan behöva en tillsyn som kan hävda sig gentemot statistikmyndigheter och statistikproducenter, dvs. som har tillräckliga resurser, kompetens/erfarenhet/expertis samt befogenheter att begära uppgifter, få insyn m.m.

Det är svårt att värdera allvaret i de potentiella konflikter som ligger i tillsynsrollen. Där föreligger dock en, åtminstone principiellt, inte oviktig risk för motsättningar. Problemet kan till stor del ligga i att tillsynsfunktionen idag är mycket otydlig och innan tillsynsrollen förtydligats kommer denna potentiella konflikt troligen att bestå.

Utöver att vara oklar tycks det som om tillsynsrollen idag är tämligen svag. Om kraven på tillsyn skulle stärkas kan den beskrivna konflikten troligen bli mera tydlig och svårhanterlig. I ett läge där det t.ex. skulle åligga SCB:s att utdela tydliga sanktioner skulle detta bli mycket svårt att förena med rollen som affärspartner.

## Myndighetsutövning kontra försäljning/produktion

Det ovan beskrivna problemet utgör ett tydligt exempel på hur myndighetsutövning kan kollidera med rollen som producerande och säljande affärspartner.

Ytterligare ett problem ligger i SCB:s troligen mycket omfattande inflytande över statistikens utformning. Trots att statistikansvaret decentraliserats till statistikansvariga myndigheter får SCB genom att vara både försäljare, producent samt tillsyns- och samordningsansvarig en

mycket styrande roll över statistiken. I synnerhet om beställaren har liten statistisk kunskap/kompetens är det lätt att SCB blir mycket styrande. Detta riskerar givetvis att underminerar målet om användarpåverkan och kan i en förlängning skada SCB och statistikens förtroende gentemot beställare och användare.

Problemet med helhetsansvar kontra egenintresse och myndighetsutövning kontra försäljning/produktion märks även när det gäller rollkonflikter kopplade till konkurrens.

## Konkurrensfördelar

SCB:s stora breda verksamhet inklusive det statistiska helhetsansvaret ger SCB en mycket dominerande ställning. Detta vore inte något större problem om SCB är ensam möjlig producent eller om det handlade om en klart uttalad produktion som skall ske i statlig regi. Med en beställare-utförarerelation där statistikansvarig myndighet har möjlighet att vända sig till andra leverantörer och det faktum att SCB även producerar andra tjänster än basstatistik och officiell statistik blir gränserna mera flytande. Ställt mot den, inte minst från SCB, uttalade ambition om en effektivitetsstimulerande konkurrenssituation medför SCB:s nuvarande ställning problem. Det finns en inte obetydlig risk att SCB:s dominerande ställning ger upphov till betydande konkurrensfördelar. Konkurrensfördelarna är tydligast i SCB:s informationsövertag vid sökandet efter uppdrag och offertgivning. Genom att konkurrerande aktörer i många fall är tvingade att vända sig till SCB för statistisk grundinformation får SCB dels kännedom om uppdrag, dels får man möjligheten att anpassa offerterna gentemot konkurrenten. Att det kan vara svårt att förstå SCB:s prissättningsprinciper medverkar även till att försämra förutsättningarna för konkurrens (och statistikbeställning).

Genom att ha i princip ensamrätt på produktion av officiell basstatistik får SCB även ett försprång gentemot konkurrenter när det gäller kompletterande tjänster.

I uppdragsverksamheten ingår också konsultationer i statistikfrågor och specialbearbetningar av olika slag. Den vanligaste tjänsten är att kunderna vill komplettera den anslagsproducerade statistiken med andra tabelluttag som bättre tillfredsställer deras specifika informationsbehov.<sup>3</sup>

Det är svårt att värdera hur allvarlig den beskrivna rollkonflikten är. Ett problem kan ligga i att en databasansvarig myndighet säljer information till konkurrenter och samtidigt via offentlig finansiering ges konkurrens-

<sup>3</sup> Statistiska centralbyrån, Årsredovisning 1997. s. 20.

fördelar i bl.a. prissättning. Detta riskerar att underminera förtroendet gentemot både konkurrenter och beställare. Vidare kan ambitionen om att öka effektiviteten genom ökad konkurrens gå förlorad.

## 4 Organisatoriska lösningsalternativ

### 4.1 Modifierat nollalternativ

Alternativet har till syfte att genom tydligare organisatorisk uppdelning inom SCB minska risker för rollkonflikter och andra konsekvenser till följd av att ansvaret för sektorsövergripande statistik, förvaltningsuppgifter och uppdragsbaserad statistikproduktion ligger inom en och samma myndighet (SCB).

Denna lösning innebär i stort sett ett antal förtydliganden inom ramen för nuvarande organisation. Dessa förtydliganden skulle kunna åstadkommas utan några större strukturella ingrepp i nuvarande organisation. Den övergripande tanken är att tydligt hålla isär olika funktioner så att de kan verka oberoende av varandra.

Som tidigare påpekats är tillsynsfunktionen inom SCB idag såväl svag som otydlig vad gäller ansvar, mandat, befogenheter och position inom SCB. För att undvika risken för rollkonflikter och bristande internt och externt förtroende bör tillsynsfunktionen förtydligas och ges en mera fristående och neutral roll.

Ett andra förtydligande som bör framhållas är behovet av en klarare gränsdragning mellan SCB:s roll som statistikansvarig myndighet och rollen som producent. Syftet med detta är ökad tydlighet i kontakter mellan SCB och andra statistikmyndigheter.

Ett tredje förtydligande krävs i form av tydligare ekonomiska spelregler och en bättre gränsdragning när det gäller vilka tjänster som SCB skall producera fritt respektive ta betalt för. Vidare är det av yttersta vikt med en tydligare och mera genomskinlig prissättning från SCB:s sida. Till detta kan även fogas behovet av en mera öppen och konkurrensneutral åtkomst till statistiska grunddatabaser.

Slutligen vill vi framhålla vikten av att SCB verkar för att bevara och vidmakthålla sin statistiska kompetens. För att SCB inte skall förlora "känslan för siffrorna" och de bakomliggande samhällsfunktionerna och effekterna kan det krävas att analysrollen förstärks.

## 4.2 Maximal decentralisering – utvidgning av reformen

Alternativet innebär ett ökat decentraliserat beslutsfattande om statistik där fler av statistikens huvudanvändare – främst de myndigheter där statistiken kan ha/har en operativ roll i verksamheten – ges en tydligare beställarroll som statistikansvarig myndighet. Den ändrade rollfördelningen mellan SCB och befintliga eller nya statistikansvariga myndigheter innebär också att en större del av statistiken kommer att finansieras direkt av statistikansvariga sektorsmyndigheterna istället för via SCB:s anslag. Ansvar för tydliga sektorsövergripande statistikbehov hålls även fortsättningsvis samman, men lyfts ur SCB till befintliga eller för detta ändamål särskilt tillskapade statistikansvariga myndigheter. Av det 40–50 procentiga statistikansvar som fick ligga kvar på SCB efter reformen kvarstår mycket lite inom SCB.

Minskas SCB:s roll som statistikansvarig myndighet genom decentralisering, blir övriga roller som SOS-ansvarig och producent tydligare samtidigt som riskerna för rollkonflikter minskar. Till nackdelarna hör en svårare samordning och minskade synergier. Om enskilda statistikansvariga myndigheter utvecklas i riktning mot ”mini-SCB” med produktion i egen regi kan nya rollkonflikter uppstå.

## 4.3 Tredelning i SOS-myndighet, statistikansvarig myndighet och produktionsansvarig myndighet

I detta alternativ delas nuvarande SCB antingen upp i tre helt fristående myndigheter eller tre från varandra mycket tydligt organisatoriskt och finansiellt avgränsade enheter.

En myndighet/enhet får uteslutande stabsuppgifter såsom samordningsansvar, metodutveckling och tillsyns- och databasansvar. En andra myndighet/enhet får ansvaret för den statistik SCB idag är statistikansvarig myndighet för. Denna enhet utför ingen egen produktion utan fungerar som en fristående beställare och administratör av den statistik man är ålagd att ansvara för. En tredje myndighet/enhet har slutligen till uppgift att fungera som en ren produktionsenhet som skall förse de statistikansvariga myndigheterna och andra beställare med statistiska tjänster i konkurrens med andra producenter.

## 4.4 Maximal konkurrensutsättning – tudelning i myndighets- och produktionsdel

Den bärande tanken med detta alternativ är att SCB gör sådant som inte kan eller bör produceras av privata aktörer. Förutom att det utifrån ett samhällsintresse kan vara motiverat att SCB producerar och har ansvar för att det finns neutral och objektiv statistik så finns det ytterligare ett klassiskt motiv för ett offentligt åtagande. Det är sannolikt att de databaser och den statistik SCB producerar är så dyr att det inte är företagsekonomiskt attraktivt för privata aktörer att bygga upp en konkurrenskraftig verksamhet. Det uppstår med andra ord ett naturligt monopol.

Det presenterade alternativet kan kallas för en maximal renodling av SCB:s offentliga åtagandet. Allt som inte kan skötas av andra än den offentliga sektorn behålles. Där det inte finns några motiv för ett offentligt åtagande bör verksamheten öppnas upp för privata aktörer och konkurrensutsättning.

Ett sätt att organisatoriskt och finansiellt genomföra en renodling av det offentliga åtagandet är att öronmärka viss statistikproduktion i statlig regi. Ifråga om sådan statistik kan statistikansvariga myndigheter vända sig till SCB. Poängen är att i frånvaron av reell konkurrens tydliggöra och formalisera ett regelrätt beställare – utförarsystem myndigheter emellan.

För all annan statistik och övriga tjänster skall statistikansvariga myndigheter vända sig till en reellt konkurrensutsatt marknad. På denna marknad kan i sin tur SCB via ett fristående produktionsbolag medges möjlighet till likvärdig konkurrens. Renodlingen av SCB:s offentliga åtagande kombineras alltså i detta fall med inrättandet av ett fristående produktionsbolag vid sidan av SCB, ett bolag som är helt avskilt från den anslagsfinansierade produktion som ligger inom ramen för det offentliga åtagandet. Med en sådan lösning är det även viktigt att frilägga SCB:s databaser och göra det möjligt för alla aktörer, inklusive den bolagiserade delen av SCB, att på lika villkor köpa information. Därmed finns även möjligheter att finansiera databaserna samtidigt som SCB får en mindre dominerande ställning än idag.

SCB blir med denna delning en tudelad organisation där den ena delen har stabsuppgifter, ansvar för databaser och producerar anslagsfinansierad officiell basstatistik. Den andra delen utgörs av ett frikopplat produktionsbolag som på samma villkor som andra aktörer får köpa information av den databasansvariga delen av SCB.

Det bör observeras att myndigheter inte kan bilda eller äga bolag. Bolagsbildningar måste alltid ske med statsmaktsbeslut. Till skillnad

från myndigheter, som alltid är en del av den juridiska personen staten, är bolag alltid självständiga juridiska personer vilket påverkar regeringens förutsättningar att styra. Istället gäller styrning i bolagsordningen, genom ägardirektiv eller genom avtal mellan bolaget och annan part.<sup>4</sup>

#### 4.5 ”Återställare” – renodlad ursprungsmodell

Ytterligare ett alternativ är att centralisera det statistikansvar som ligger inom ramen för det offentliga åtagandet och återföra de nuvarande statistikansvariga myndigheternas ansvar till SCB. Detta skulle t.ex. kunna motiveras med att externa krav och budgetstyrning via öronmärkta statistikanslag minskat utrymmet för användarinflytande.

Detta alternativ skulle i stort sett innebära en återgång till läget före statistikreformen. För att garantera en statistik av god kvalitet bör en recentralisering av statistikansvaret kompletteras med att SCB även utför mera av analys- och förädlingsarbete.

<sup>4</sup> Statskontoret 1996:25 En förvaltningsrättslig översikt, Statskontoret 1997:25 Statens samverkan.



## 5 Alternativens för- och nackdelar

Ett sätt att sammanfatta denna idépromemoria är att göra följande samlade bedömning – att rollkonflikterna i nuläget är av sådan art och omfattning att någon form av åtgärd är motiverad. Redovisade organisationslösningar kan därvid ge uppslag kring möjliga handlingsvägar.

I princip alla redovisade organisationsalternativ skulle också troligen leda till mera renodlade roller än dagens organisationslösning (nollalternativet). Organisationsalternativen är dock i mångt och mycket renodlade ytterlighetslösningar, och det finns andra lösningsalternativ, bl.a. kombinationer i gråzonerna mellan olika alternativ. Det finns också risker med i stort sett varje alternativ. I en del fall kan nya rollkonflikter uppstå.

Men när det gäller bedömningen av vilka roller som är möjliga att förena i en och samma organisation utan risker för rollkonflikter, otydlighet eller minskat förtroende för SOS är det mycket av en bedömningsfråga – ju starkare garantier som eftersträvas, desto radikalare förändringar. Samtidigt kan en del av fördelarna med dagens organisation gå förlorade. Vinster i form av renodling av roller bör därför vägas mot fördelar med nuvarande organisation, bl.a. synergier, samordning, intern konflikthantering, tillgänglighet m.m.

Ytterst kan man tala om olika ”rollblock” i mer eller mindre lämpliga kombinationer, nämligen följande:

- *Ansvar för officiell statistik (SOS-rollen)*, kan främst kombineras med *rollen för ”gemensamma” stabs- eller SOS-resurser* (metodutveckling, databaser m.m.) för statistikansvariga myndigheter. Hur kostnadsfria eller avgiftsfinansierade dessa bör vara gentemot statistikansvariga myndigheter är dock främst en fråga för rådande budgetrestriktioner och regelverk för avgiftsfinansiering m.m.
- *Tillsynsrollen* kan eventuellt ingå i SOS-rollen, alternativt vara mera fristående. I övrigt är den svår att kombinera med övriga roller som statistikansvarig myndighet och som producent.
- *Rollen som statistikansvarig myndighet för övergripande statistikområden* är svår att kombinera med någon annan roll utan risker för otydlighet m.m. Frågan är också hur mycket statistikansvar som kan fördelas på andra myndigheter.

- *Rollen som kommersiell statistikproducent* är svår att kombinera med övriga roller, utan att riskera otydlighet m.m. i förhållande till statistikansvariga myndigheter och andra statistikproducenter. Rollen bör organiseras fristående och eventuellt delas upp på flera.

Ytterst bör också valet av organisationsalternativ utgå från frågan om renodlade roller enligt något av organisationsalternativen påtagligt skulle påverka statistikreformens eller statistikens mål. Olika mål kan dock tala för olika organisationsalternativ inklusive nollalternativet.

En slutsats är att ett tydligare uppdelat SCB eller ett mindre SCB troligen skulle gagna statistikmyndigheternas och konkurrenternas förtroende för statistiken. Hur denna uppdelning görs kan dock ha stor inverkan på om styrningen och samordningen av statistiken kan bli effektivare/enklare. Dagens lösning med ett inomoffentligt köp-/säljssystem kan ha gett lägre transaktionskostnader och ökat kostnadsmedvetande och därmed haft positiv inverkan på effektiviteten i den officiella statistiken. Detta skulle tala emot ett återställaralternativ.

En ökad avgiftsbeläggning skulle troligen påverka statistikens kostnads- och produktivetsutveckling positivt, men kan också riskera att negativt påverka statistikens tillgänglighet och användning i samhället. En uppdelning av SCB:s statistikproduktion i flera enheter skulle på samma sätt kunna stimulera kostnads- och produktivetsutvecklingen, givet att det inte finns några nämnvärda stordriftsfördelar att ta hänsyn till.

Organisationslösningarnas påverkan på mål som statistikens kvalitet (omfattning/innehåll, relevans, flexibilitet m.m.) samt påverkan på mål som integritetsskyddet och uppgiftslämnarnas börda är svåra att bedöma