

Statistikcentralens databokslut 2017, rapport över behandlingen av personuppgifter

26.6.2018

Innehåll

1	Sammandrag	4
2	Statistikcentralens strategi	5
	2.1.1 Bild: Strategiska målsättningar.....	5
	2.1.2 Bild: Kritiska framgångsfaktorer	6
3	Statistikcentralens roll i Finlands statistikväsen.....	7
4	Lagstiftning, dataskydd och datasäkerhet	8
	4.1 Lagstiftning som styr Statistikcentralens verksamhet.....	8
	4.1.1 Lagen om Statistikcentralen.....	8
	4.1.2 Statistiklagstiftning	8
	4.1.3 Personuppgifter och personuppgiftslagen.....	8
	4.1.4 Lagen om offentlighet i myndigheternas verksamhet.....	9
	4.2 Inverkan av ändringen av dataskyddslagstiftningen på Statistikcentralens verksamhet	9
	4.3 Datasäkerhet.....	9
	4.3.1 Hantering av datasäkerhet.....	9
	4.3.2 Hantering av behörighet.....	9
	4.3.3 Teknisk datasäkerhet.....	10
	4.3.4 Incidentrapportering	10
	4.3.5 Kontinuitet och riskhantering.....	10
	4.3.6 Bild: Riskhanteringsprocess.....	10
	4.4 Dataskydd för statistik och tabeller	11
	4.5 Material på enhetsnivå och tabellmaterial	11
	4.6 Statistiska dataskyddsmetoder	11
	4.7 Pseudonymisering dvs. ersättning av personuppgifter med tillfälliga koder	12
	4.8 Intern revision	13
	4.9 Att säkerställa de anställdas kunnande.....	14
5	Kvaliteten på statistik, yrkesetik och riktlinjer (CoP)	15
	5.1 Kvaliteten på statistik	15
	5.2 Yrkesetik och riktlinjer	15
	5.2.1 Bild: Riktlinjerna för europeisk statistik.....	16
	5.3 Peer review	16
6	Datainsamling.....	17
	6.1 Centralisering av datainsamling	17
	6.2 Allmänna principer för datainsamling	17
	6.2.1 Informering av uppgiftslämnarna.....	17
	6.2.2 Minimering av uppgiftslämnarbördan	18
	6.2.3 Datainsamlingsmetoder	18

6.3	Användning av material och deras aktualitet.....	19
6.3.1	De största leverantörerna av material	20
6.4	Statistikcentralens datainsamlingar riktade till personer och hushåll	20
7	Statistikproduktion.....	21
7.1	Statistik som produceras.....	21
7.2	Livslängd för personuppgifter	21
7.2.1	Bild: datainsamlingens datalager	22
7.3	Processmodell för statistikproduktionen.....	23
7.3.1	Bild: processmodell för statistikproduktionen	23
7.4	Statistikdatalager.....	23
7.5	Nuvarande persondatalager.....	24
7.6	Nuvarande lönedatalager.....	25
7.7	Vägtrafikolyckor	26
7.7.1	Bild: Datasystem	27
8	Distribution av statistisk information	28
8.1	Statistikcentralens principer för offentliggörande av statistik.....	28
8.2	Öppna data	28
9	Utlämnande av uppgifter	28
9.1	Utlämnande av uppgifter för statistiska ändamål	28
9.2	Utlämnande av uppgifter för forskningsändamål.....	29
9.2.1	Behandlingen av individmaterial kräver tillstånd av Statistikcentralen	29
9.2.2	Tabell: Antalet beviljade användningstillstånd och utlämning av material efter produktgrupper år 2017.....	30
9.2.3	Sisu-mikrosimuleringsmodell	30
9.2.4	Distansanvändningssystemet	30
9.3	Företags- och arbetsställeregistret.....	31
9.4	Register över uppgiftslämnande företag	32
10	Behandlingen av personuppgifter i form av uppdrag	33
10.1	Kundfinansierade intervju- och datainsamlingstjänster	33
10.1.1	Tabell: Kundfinansierade datainsamlingar	33
10.2	Kundspecifika avtal om informationstjänst	34
10.3	Statistikföringstjänster	34
10.4	Nationell distansanvändningstjänst.....	34
11	Kund- och intressegruppsuppgifter	35

12	Ärendehantering och arkivering	36
12.1	Ärendehantering och analogiskt dokument-material	36
12.2	Arkivering av statistikmaterial.....	36
13	Utvecklingsåtgärder för år 2018	37

1 Sammandrag

Detta är Statistikcentralens första databokslut. I databokslutet beskrivs den totala processen för produktionen av Statistikcentralens statistik, insamling av uppgifter och undersökningar som görs via uppdrag, informationstjänst om det statistiska grundregistret samt utlämnande av uppgifter för forskningsändamål för att läsaren ska kunna få en bild av utgångspunkterna för behandlingen av personuppgifter och därtill hörande verksamhetsätt samt omfattningen av behandlingen av uppgifter vid Statistikcentralen. Uppgifterna i rapporten gäller år 2017.

Statistikcentralen producerar nästan 160 olika statistikgrenar, av vilka största delen hör till helheten av Finlands officiella statistik. Merparten av allt statistiskt datamaterial fås ur den offentliga förvaltningens register eller andra administrativa källor. Statistikcentralen samlar också in uppgifter med enkäter och intervjuer om det inte är möjligt att erhålla de uppgifter som behövs på något annat sätt. Statistikcentralen frågar direkt privatpersoner eller hushåll efter uppgifter i omkring tio direkta datainsamlingar per år. Uppgifterna från de direkta individ- eller hushållsbaserade datainsamlingarna utnyttjar Statistikcentralen i tolv olika statistikgrenar.

Dessutom behandlar Statistikcentralen personuppgifter vid skapande av undersökningsmaterial, vid olika uppdrag samt vid skötsel av kund- och intressegruppsrelationer.

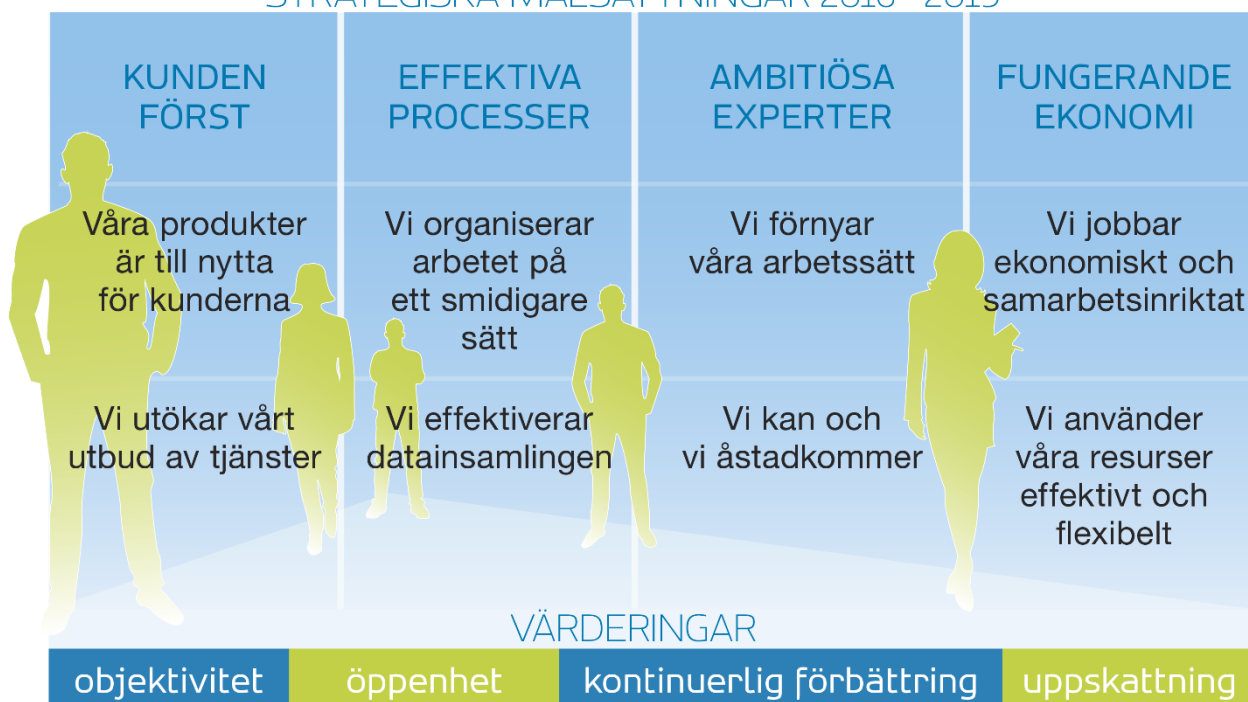
Nyckelord: databokslut, dataskydd, dataflöden, materialtyper

2 Statistikcentralens strategi

2.1.1 Bild: Strategiska målsättningar

VISION FAKTA INFÖR FRAMTIDEN – TILLFÖRLITLIG STATISTIK TILL NYTTA FÖR KUNDERNA

STRATEGISKA MÅLSÄTTNINGAR 2016–2019



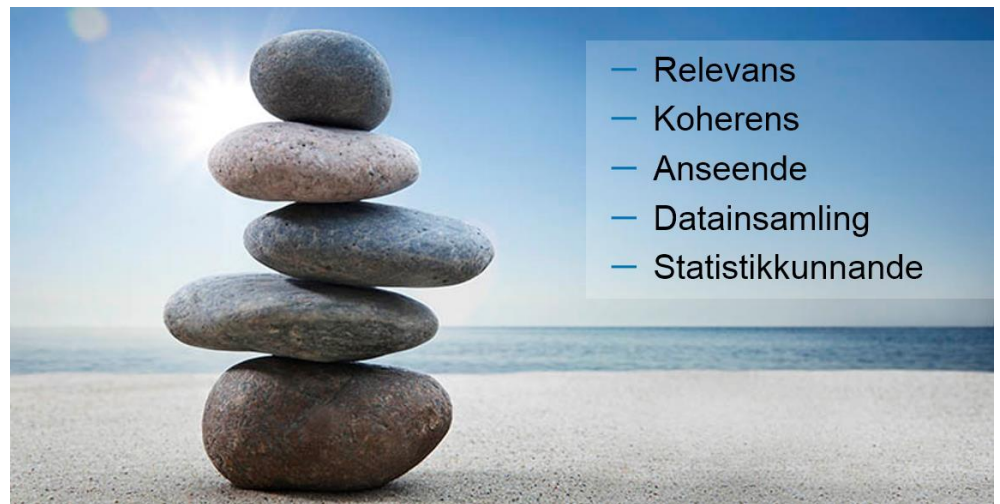
Utgångspunkten för Statistikcentralens verksamhetsidé är de uppgifter som definierats i lagstiftningen: *Statistikcentralen stöder ett demokratiskt och informationsbaserat beslutsfattande samt forskning genom att producera tillförlitlig statistik, utredningar och material som beskriver samhällsförhållanden. Statistikcentralen utvecklar statens statistikväsen i samarbete med andra statliga myndigheter.*

Under strategiperioden 2016–2019 har Statistikcentralen som mål att sätta kunderna i centrum för verksamheten. Statistikcentralens vision är *Fakta inför framtiden – tillförlitlig statistik till nytta för kunderna*. De tre centrala temana i strategin är kunderna, digitalisering och verksamhetssätt. Enligt strategin är målsättningen i det dagliga arbetet att generera nytta för kunderna. Till kunderna hör utöver uppgifts användarna också uppgiftslämnarna, dvs. medborgare, företag och sammanslutningar.

Uppgiftslämnarnas och -användarnas syn på sekretess kan vara vitt skilda. Vissa grupper av uppgifts användare önskar eventuellt få mycket detaljerad information. Ur uppgiftslämnarnas synvinkel är det nödvändigt att inte enskilda uppgiftslämnarnas uppgifter röjs. Statistikcentralen gör kontinuerligt en uppskattning av individens dataskydd och uppgifts användarnas behov av detaljerad information.

I Statistikcentralens verksamhet har man identifierat fem kritiska framgångsfaktorer. Av dem hör anseende och de risker som hotar bevarandet av anseendet nära ihop med kvaliteten på och tillförlitligheten av statistikproduktionen och -produkterna samt med dataskyddet och datasäkerheten.

2.1.2 Bild: Kritiska framgångsfaktorer



3 Statistikcentralens roll i Finlands statistikväsen

Statens statistikväsen är ett nätverk av myndigheter inom statistikbranschen, som producerar Finlands officiella statistik (FOS), Europeiska statistiksystemets (EES) statistik samt annan statistik. Till dess uppgifter hör att producera sådan statistik som beskriver samhällsförhållanden och deras utveckling och som man rent allmänt behöver i samhället. Till statens statistikväsen hör 14 ämbetsverk och inrättningar, varav fyra är statistikmyndigheter, åtta är andra myndigheter som framställer statistik och två är andra inrättningar som framställer statistik.¹

Statistikcentralen har i statistiklagen definierats som allmän myndighet inom statistikväsendet. Som allmän myndighet är det Statistikcentralens uppgift att i samarbete med andra statistikproducenter sörja för att de statistikföringsuppgifter som grundar sig på nationella och internationella ålägganden utförs.

Statistikcentralens uppgift vid den nationella koordineringen är att främja och stödja produktionen av högklassig officiell statistik. Delegationen för Finlands officiella statistik (FOS-delegationen) grundades år 2002 för att effektivisera Statistikcentralens koordinerande roll och intensifiera samarbetet mellan myndigheter som producerar statistik.² Till delegationen hör alla myndigheter och inrättningar som framställer officiell statistik. Dess uppgift är att främja riktlinjerna för utvecklingen av statistikväsendet och målsättningarna för koordineringen av det europeiska statistiksystemets nationella nivå.

Delegationen som tillsattes för åren 2016–2017 hade till uppgift att regelbundet behandla frågor om publikationen av statistik, kvaliteten och utvecklingen av statistikproduktionen främja koordineringen av det nationella statistikväsendet och samarbetet i enlighet med målsättningarna för det europeiska statistiksystemet, främja den officiella statistikens synlighet och kännedomen om den i Finland och internationellt, utarbeta riktlinjerna för utveckling av Finlands officiella statistik och följa upp verkställandet av dem samt att följa upp förändringar i den nationella statistiklagen och EU:s statistiklag och i samarbete anpassa statistikväsendet efter dem.

¹ http://tilastokeskus.fi/org/tilastotoimi/index_sv.html

² **Virhe. Hyperlinkin viittaus ei kelpaa.**

4 Lagstiftning, dataskydd och datasäkerhet

4.1 Lagstiftning som styr Statistikcentralens verksamhet

4.1.1 Lagen om Statistikcentralen

Enligt lagen om Statistikcentralen har Statistikcentralen till uppgift att framställa statistik och utredningar som gäller samhällsförhållanden samt att ha hand om den allmänna utvecklingen av statens statistikväsen i samarbete med andra statliga myndigheter. Statistikcentralen är en allmän myndighet inom statens statistikväsen.

4.1.2 Statistiklagstiftning

EU:s statistikförordningar är förpliktande för Finlands statistikväsen. EU:s befogenheter vid produktionen av europeisk statistik grundar sig på artikel 338 i fördraget om Europeiska unionens funktionssätt (FEUF). EU:s statistikförordningar är oftast rådets och parlamentets förordningar som kompletteras med kommissionens genomförandeakt. Den mest centrala av förordningarna är Europaparlamentets och rådets förordning (EG) nr 223/2009, dvs. EU:s statistiklag, som ger de yttre ramarna för framställningen av statistik i Europeiska unionen. De statistikspecifika förordningarna reglerar harmoniseringen av EU-ländernas statistik, såsom klassificeringar och andra statistikstandarder, och uppgifter som ska levereras till kommissionen och tidtabeller som ska följas.

Statistiklagen är en allmän lag som gäller hela statens statistikväsen. Dess syfte är att säkerställa tillgången till tillförlitlig information som rent allmänt behövs för beslutsfattandet och planeringen i samhället. Målet är också att uppfylla förpliktelser i anslutning till det internationella statistiksamarbetet, att förenhetliga och effektivisera statistikväsendet samt att tillämpa god statistiksed. Lagens syfte är att förtydliga procedurerna i förhållande till dem som lämnar och använder uppgifter samt att säkerställa rättigheterna för dem som lämnar uppgifter eller som uppgifterna gäller. Syftet med lagen är dessutom att främja användningen av uppgifter som insamlats för statistiska ändamål i vetenskaplig forskning.

Statistiklagen styr framställningen av Statistikcentralens och också andra statliga myndigheters statistik. I lagen föreskrivs om Statistikcentralens rätt att samla in uppgifter med stöd av uppgiftsskyldighet. Enligt lagen ska man i första hand utnyttja uppgifter som samlats in i något annat sammanhang. Största delen av uppgifterna kan fås ur olika register. Från uppgiftslämnarna samlas endast sådana nödvändiga uppgifter in som inte kan fås någon annanstans ifrån. Statliga myndigheter har en allmän uppgiftsskyldighet ifråga om de uppgifter de förfogar över. Företag, kommunala organisationer och icke-vinstsyftande organisationer är förpliktade att lämna uppgifter om de ärenden som särskilt fastställts i lagen. En enskild medborgare är inte skyldig att lämna uppgifter till Statistikagen, utan det är frivilligt att lämna uppgifter.

4.1.3 Personuppgifter och personuppgiftslagen

Med personuppgifter avses alla slags anteckningar som beskriver en fysisk person eller hans egenskaper eller levnadsförhållanden som kan hänföras till honom själv eller till hans familj eller någon som lever i gemensamt hushåll med honom. Till exempel namn och adress utgör personuppgifter liksom också exempelvis noggrant yrke, kommun och ålder. Med personuppgifter avses identifierbara uppgifter om alla fysiska personer också när dessa verkar som yrkesutövare.

För behandling av personuppgifter i statistikföring tillämpas utöver statistiklagen också personuppgiftslagen. Enligt personuppgiftslagen krävs det vid framställningen av statistik omsorgsfull planering av informationshanteringsprocessen innan personregistren inrättas. Behovet av uppgifter utreds noggrant och bara de uppgifter som är nödvändiga för framställningen av statistik begärs av uppgiftslämnarna. En liknande planeringsskyldighet krävs också enligt statistiklagen.

4.1.4 Lagen om offentlighet i myndigheternas verksamhet

Om sekretess beträffande uppgifter som insamlats för statistiska ändamål föreskrivs i lagen om offentlighet i myndigheternas verksamhet (621/1999, offentlighetslagen), enligt vilken statistiskt primärmaterial och dokument som lämnats till statistikmyndigheter för framställning av statistik är sekretessbelagda. Undantagna från sekretessen är offentliga uppgifter om statliga och kommunala myndigheters verksamhet och ekonomi samt vissa uppgifter om företag om vilka föreskrivs särskilt.

4.2 Inverkan av ändringen av dataskyddslagstiftningen på Statistikcentralens verksamhet

Efter ikraftträdandet av EU:s allmänna dataskyddsförordning tillsattes ett projekt vid Statistikcentralen vars uppgift var att utreda nuläget för behandlingen av personuppgifter vid Statistikcentralen. Projektet kartlade det statistikmaterial som hösten 2016 var i produktionsbruk, deras behandlings- och lagringsmiljö, personregistrens registerbeskrivningar, arkiveringssituationen, gällande avtal om köp av tjänster, avtal om utlämnande av uppgifter, olika auditeringar och granskningar som gjorts nyligen, anvisningar för dataskydd och datasäkerhet samt introduktioner och utbildning gällande personalens dataskydd och datasäkerhet.

I samband med utredningsarbetet för projektet konstaterades också att det vid Statistikcentralen pågår flera projekt som främjar och till en del också är en förutsättning för att de krav EU:s allmänna dataskyddsförordning ställer kan uppfyllas. Dessa projekt hänför sig till pseudonymisering av en personbeteckning, dvs. byte av beteckningar mot tillfälliga koder, utvecklingen av behörighetshantering, dokument- och ärendehantering samt verkställandet av loggpolitiken.

Efter kartläggningen av läget inleddes våren 2017 ett genomförandeprojekt vars uppgift är att bl.a. säkerställa att de registrerades rättigheter tillgodoses i enlighet med förordningen, dessutom att säkerställa att anvisningarna uppdateras, utbildning och introduktion förnyas samt att ansvarsskyldigheten säkerställs i enlighet med förordningen. Projektets tidsfrist går ut i juni 2018.

4.3 Datasäkerhet

4.3.1 Hantering av datasäkerhet

Statistikcentralens principer för datasäkerhet uppfyller lagstiftningen om ämbetsverkets verksamhet. Principerna för datasäkerhet har godkänts av ledningen och de anställda tillämpar dem i sitt arbete. Statistikcentralen utnyttjar i sin verksamhet också de anvisningar som utarbetats av ledningsgruppen för informations- och cybersäkerhet inom statsförvaltningen (VAHTI).

4.3.2 Hantering av behörighet

Vid Statistikcentralen baserar sig behörigheten på arbetsuppgifter. Vid beviljandet av behörigheter följs principen om minsta behörighet, dvs. rättigheter ges bara till den information som krävs för att utföra arbetsuppgifter.

4.3.3 Teknisk datasäkerhet

För det tekniska underhållet av Statistikcentralens system ansvarar Statens center för informations- och kommunikationsteknik Valtori tillsammans med Statistikcentralens ICT-enhet.

Statistikcentralens ICT-enhet underhåller och utvecklar ämbetsverkets egna system och applikationer avsedda för statistiska ändamål. För datasäkerheten för dessa system och applikationer sörjer man bl.a. genom regelbundna auditeringar. Auditeringarna utförs alltid av en utomstående aktör.

Statistikcentralens underleverantörer eller tjänsteleverantörer, vars personal rör sig i Statistikcentralens lokalteter eller vars anställda behandlar sekretessbelagda uppgifter på uppdrag av Statistikcentralen, undertecknar ett sekretessavtal. Förpliktelser med anknytning till behandlingen av personuppgifter, såsom undertecknande av sekretessavtal, finns också angivna i avtalen med personer som behandlar uppgifterna.

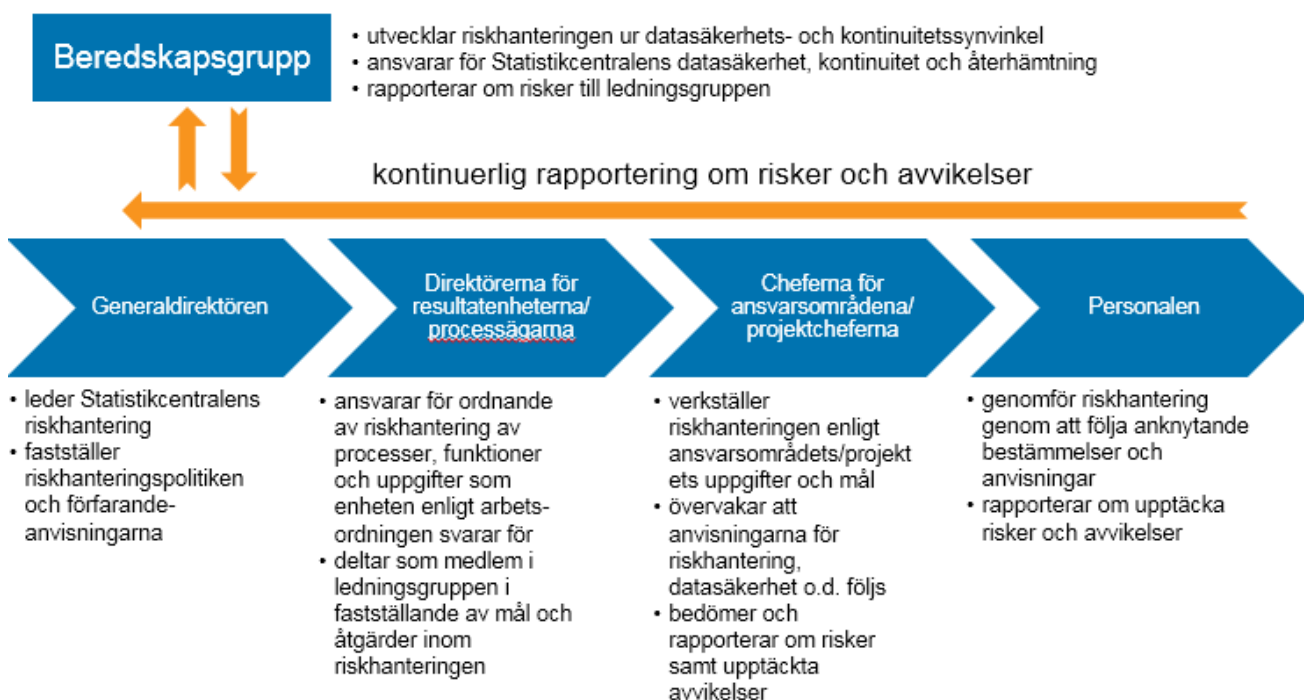
4.3.4 Incidentrapportering

Datasäkerhetsincidenter i Statistikcentralens system och verksamhetsmiljö rapporteras och anges i det centrala systemet för incidenthantering. För utvärdering av incidenter svarar ämbetsverkets beredskapsgrupp. Säkerhetsincidenterna behandlas av ämbetsverkets ledningsgrupp minst två gånger per år.

4.3.5 Kontinuitet och riskhantering

För kontinuiteten och riskhanteringen i verksamheten svarar ämbetsverkets beredskapsgrupp. Gruppen kartlägger risker som en del av riskhanteringsprocessen tillsammans med ledningen för ämbetsverket. Riskerna utvärderas och ansvariga aktörer utses. Riskerna i datasystemen bevakas i ämbetsverkets ICT-enhet

4.3.6 Bild: Riskhanteringsprocess



4.4 Dataskydd för statistik och tabeller

De uppgifter som Statistikcentralen samlar in i statistiksyfte är sekretessbelagda utöver vissa undantag som föreskrivs i statistiklagen. Målet med dataskyddet är att skydda uppgiftslämnarnas integritet samt affärs- och yrkeshemligheter.

Dataskyddet säkerställs med hjälp av lagar, bestämmelser, anvisningar, metoder och tillämpningen av dessa. Dataskyddet beaktas redan vid datainsamlingen och vid planeringen av insamlingen genom att sörja för att uppgifter som inte är nödvändiga för framställandet av statistik inte samlas in. Uppgifterna bearbetas med iakttagande av dataskyddsprinciper, dvs. så att man i alla skeden av bearbetningen hindrar att sekretessbelagda uppgifter röjs för utomstående.

Syftet med dataskyddsprinciperna är att upprätthålla ett förtroendefullt förhållande till uppgiftslämnarna. På så sätt säkerställer man tillgången till heltäckande och tillförlitliga uppgifter för framställningen av statistik. Principerna för dataskydd begränsar vilka uppgifter som kan samlas in, hur uppgifterna får bearbetas och i vilken form de kan publiceras. Dataskyddet inverkar sålunda på förfaringssätten som tillämpas på framställningen av statistik och på innehållet i statistiken som publiceras.

De uppgifter som Statistikcentralen publicerar eller lämnar ut är i en sådan form att den som uppgifterna hänför sig till eller sekretessbelagda uppgifter om denna inte röjs. Tabellmaterialen som ska publiceras skyddas genom användning av tillämpliga statistiska dataskyddsmetoder. Dataskyddsmetoderna kan påverka innehållet i materialet, vilket innebär att man fäster uppmärksamhet vid kvaliteten på uppgifterna, hur känsliga de är och för vilket ändamål materialet ska användas när man väljer metod. Sekretessbelagda uppgifter kan publiceras utan att de skyddats om den som uppgifterna hänför sig till har gett sitt skriftliga samtycke till det. Denna möjlighet har främst använts ifråga om storföretagens uppgifter, aldrig ifråga om personuppgifter.

4.5 Material på enhetsnivå och tabellmaterial

Material som innehåller uppgifter på enhetsnivå, t.ex. om personer eller företag, kallas material på enhetsnivå. Uppgifterna hänför sig direkt till en viss enhet och därför är det i allmänhet lätt att dra slutsatser beträffande den som uppgifterna hänför sig till om materialet är oskyddat, även om direkta identifikationsuppgifter har raderats.

När man hänvisar till aggregerat material som är i tabellformat talar man om tabellmaterial (tabeller).

4.6 Statistiska dataskyddsmetoder

För att skydda sekretessbelagda uppgifter finns det olika statistiska dataskyddsmetoder. Syftet med alla dataskyddsmetoder är att bearbeta materialet till ett format som gör röjandet av en enhet eller dess egenskaper nästan omöjligt. Uppgifterna betraktas inte som identifierbara, när identifieringen skulle kräva oskäligt med tid, kostnader eller arbetskraft. För tabellmaterial och material på enhetsnivå finns det egna dataskyddsmetoder som avviker från varandra.

Både tabellmaterial och material på enhetsnivå kan avvika från varandra till både innehåll, struktur, publiceringskoncept, användare och ändamål. Hur de olika dataskyddsmetoderna inverkar på materialet som ska skyddas kan i hög grad

avvika från varandra. På grund av detta är det svårt att ge detaljerade skyddsanvisningar som passar alla tabell- och enhetsmaterial. Dataskyddsmetoderna ska alltid väljas efter situation.

De viktigaste kriterierna vid val av skyddsmetod är uppnående av behövlig skyddsnivå och bevarande av materialets viktiga egenskaper. Ju bättre särdragen för varje material har beaktats desto bättre kan man skydda materialet och samtidigt bevara dess nyttighet. Ofta påverkas valet av metod allra mest av disponibla resurser, såsom tid och tillgängligt program som används för skyddet. Skyddsmetoden ska också vara transparent på så sätt att den som använder det skyddade materialet förstår huvuddragen i metoden och då kan beakta eventuella förändringar i materialet som dataskyddet orsakar.

Allmänna dataskyddsprinciper tillämpas på skyddspraxisen för tabellerna. Materialen på enhetsnivå omfattar i allmänhet mer information än tabellmaterialet och det finns därför en större röjanderisk än för tabellerna. Således planeras och genomförs skyddet av material på enhetsnivå alltid materialvis. Ju större volymer och ju känsligare information som ska behandlas desto striktare dataskyddsmetoder används.

4.7 Pseudonymisering dvs. ersättning av personuppgifter med tillfälliga koder

Statistikcentralen har tagit fram en tjänst som gör det möjligt att på ett enhetligt sätt pseudonymisera sådana uppgifter i statistikproduktionen som innehåller personuppgifter. Pseudonymisering av personuppgifter är ett sätt att hindra en direkt identifiering av personer. Pseudonymisering innebär att direkta identifieringsuppgifter raderas eller ersätts med andra identifierare i materialen. Raderade eller ersatta uppgifter lagras väl skyddade separat från annat material.

Alla personuppgifter som anländer till Statistikcentralen pseudonymiseras i fortsättningen genast i första skedet av mottagandet av materialet oberoende av om det är fråga om uppgifter från myndigheter eller privata aktörer eller uppgifter som Statistikcentralen samlar in direkt av personer eller företag.

Pseudonymiseringstjänsten är en del av den nya processen för mottagande av material, som införs stegvis fr.o.m. början av hösten 2018. Uppgifterna styrs till pseudonymiseringen i det första skedet av mottagandeprocessen. Vid pseudonymiseringen ersätts personbeteckningarna med en slumpmässig identifierare som är kopplad till en person. Identifieraren innehåller inte information om personens egenskaper, t.ex. kön eller födelsedatum, liksom är fallet med personbeteckningen. Utöver att personbeteckningen ersätts raderas personernas namnuppgifter och e-postadress från den behandlade filen. Efter pseudonymiseringen går materialet vidare till de följande skedena i mottagandeprocessen och därifrån till statistikproduktionens behandlingsskeden.

Mottagandeprocessen är automatisk och den styrs med hjälp av metadata. Behörigheten till uppgifter i det första skedet av mottagandeprocessen, där materialet ännu omfattar personernas direkta identifieringsuppgifter, är väl begränsad och grundar sig på utredning av eventuella felsituationer som sker i processen.

De personbetecknings- och namnuppgifter som ersätts och raderats i pseudonymiseringen sparas i ett separat skyddat datalager. Behörighet att

återställa uppgifter kan bara beviljas personer som behöver tillgång till personernas direkta identifieringsuppgifter för att utföra sina arbetsuppgifter. Sådana arbetsuppgifter är till exempel att kontakta uppgiftslämnare.

All behandling av uppgifter i pseudonymiseringstjänsten registreras i ett centralt loggsystem där det är möjligt att utreda transaktionernas gång och ta reda på orsaken till behandlingen.

4.8 Intern revision

Den interna revisionen är en oberoende och objektiv utvärderings-, säkrings- och konsultverksamhet, utformad för att stöda Statistikcentralen i dess arbete med att uppnå sina mål.

Den interna revisionens uppgift är att för ledningens räkning klarlägga ändamålsenligheten och tillräckligheten med den interna kontrollen och riskhanteringen samt utföra de revisionsuppgifter som generaldirektören fastställt. Föremål för granskning kan vara Statistikcentralens alla funktioner och resultatenheter.

Statistikcentralens interna revision utför årligen omkring 3–4 revisioner samt några utredningar. Under tidigare år har tyngdpunkten legat på granskningen av förvaltningen, men under de senaste åren har den flyttats över till Statistikcentralens kärnverksamhet – statistikproduktionen.

År 2015 granskades dokumentationen av statistikproduktionen. År 2016 granskades statistikproduktionens kvalitetssäkring, funktionaliteten i processerna för grundande och nedläggning av statistikgrenar, sekretessförbindelser och tidsenligheten i datainsamlingsbeslut. År 2017 låg tyngdpunkten i revisionsverksamheten på kraven i EU:s dataskyddsförordning. Föremål för granskningen var behandlingen av personuppgifter i fem statistikgrenar med personuppgifter, som valts med för revisionen. Dessutom gjordes en utredning om behandlingen av personuppgifter i intervjuarbetet, där man granskade insamlingen och behandlingen av material från insamlingar av uppgifter om hushåll och personer, som genomfördes i form av intervjuarbete.

Utöver sin egen interna kontroll köper Statistikcentralen experttjänster inom den interna revisionen som köptjänst. År 2016 var företagsstatistikens integrerade datasystem föremål för granskning i form av köptjänst. Tyngdpunkten i granskningen låg på administrativa kontroller och syftet var att uppskatta ändamålsenligheten och funktionaliteten i produktionssystemets kontroller och dataskyddsmekanismer. År 2017 utfördes som extern revision en granskning av dataskyddet för planerna för socialstatistikens integrerade datasystem under utveckling. I granskningen av systemet var syftet att utreda hur väl dataskyddet motsvarade kraven enligt dataskyddsförordningen. Det är meningen att behandlingen och publiceringen av Statistikcentralens material med personuppgifter ska överföras till det nya systemet i olika skeden fr.o.m. år 2018.

I alla granskningar som blev klara år 2017 har Statistikcentralens verksamhet, processer, datasystem och anvisningar utvärderats i förhållande till kraven i EU:s dataskyddsförordning. Likaså har man lyft fram utvecklingsrekommendationer. Verksamheten har förbättrats på basis av utvecklingsförslagen.

I slutet av år 2017 påbörjades en granskning med målet att säkerställa att personalens behörighet är à jour och att inte för omfattande rättigheter är i bruk.

4.9 Att säkerställa de anställdas kunnande

År 2015 införde Statistikcentralen sekretessförbindelser som följer riktlinjerna för europeisk statistik. Varje ny anställd undertecknar en sekretessförbindelse på sin första arbetsdag. I samband med detta får den nya arbetstagaren information om dataskyddet och sekretessen i uppgifterna som samlats in för statistikföring. Ett kort PM har upprättats om sekretesskyldigheten och den ges också till den nya arbetstagaren. Vid introduktionsutbildningen går man tillsammans med den nya arbetstagaren igenom dataskydds- och datasäkerhetsbestämmelser och vad de innebär i praktiken.

Introduktion och utbildning som gäller dataskyddet anordnas vid Statistikcentralen både inom enheten och som en allmän introduktion. Till den allmänna introduktionen hör ett utbildningsprogram för statistik-kunnande som planerats och genomförts vid Statistikcentralen. Programmet tar upp yrkesetiska principer och lagstiftningen inom statistikbranschen där dataskyddet har en central roll. Utbildning för hela organisationen har anordnats i samband med ändringarna av lagstiftningen

År 2017 anordnades utbildning om den nya dataskyddslagstiftningen. Den var riktad till olika personalgrupper. Dataskyddslagstiftningen var temat under workshopen om riskhantering som anordnades för den högsta ledningen, vid samevenemanget för chefer och på de olika resultatenheternas möten för hela personalen. Statistikcentralen har deltagit i samprojektet Vahti-Juhtas workshopar. Dessutom har anställda deltagit i olika evenemang och seminarier om den nya dataskyddslagstiftningen.

5 Kvaliteten på statistik, yrkesetik och riktlinjer (CoP)

5.1 Kvaliteten på statistik

I alla skeden av statistikprocessen iakttas kvalitetskriterier som överenskommit gemensamt. Det finns flera uppsättningar av kriterier och de gäller funktioner inom statistikprocessen eller statistikprodukter eller båda. De uppsättningar av kriterier som är i bruk vid Statistikcentralen är inbördes kompatibla.

Vid Statistikcentralen används principerna för det europeiska kvalitetspriset (EFQM-principer) som totalram för kvalitetsstyrning. Statistikbranschens egen ram för kvalitetsstyrning är riktlinjerna för europeisk statistik (CoP). Ramarna kompletterar varandra. Också kvalitetskriterierna för Finlands officiella statistik (FOS) är kompatibla med riktlinjerna för europeisk statistik. Om kvaliteten rapporteras till användarna med hjälp av kvalitetsbeskrivningar skilt för offentliggörandet av varje statistik: Kvalitetsbeskrivningarna är offentligt tillgängliga på statistikgrenarnas webbsidor.

Statistikcentralen följer halvårsvis på ämbetsverksnivå upp att kvaliteten realiserats. Både resultat och kvalitetshantering följs upp med hjälp av mätare. Utöver antalet statistikprodukter och produktivitet uppföljs bl.a. offentliggörandenas rättidighet, andelen felfria offentliggöranden, eftersläpning vid publicering och svarsandelar.

I Statistikcentralens strategi är kunderna, kundernas behov och växelverkan med kunderna i fokus och detta följs regelbundet upp genom undersökningar och enkäter. Det gäller såväl företagsbilden, kundernas syn på förändringar i användningen av information, webbtjänsterna som enskilda produkter och tjänster. Dessutom följer man upp synpunkter som inkommer i stat.fi-tjänstens responslåda, kursutvärderingar, gensvar från uppgiftslämnarna, nöjdheten bland kunder inom den avgiftsbelagda verksamheten och användningen av webbtjänster.

5.2 Yrkesetik och riktlinjer

Inom statistikbranschen betraktas yrkesetik som viktig. En förutsättning för statistikmyndighetens verksamhet är att de statistiketiska principerna följs. Grund för tillförlitlig statistik är de traditionella värdena inom statistikväsendet, dvs. tillförlitlighet, objektivitet, obundenhet och opartiskhet.

Arbetstagare inom statistikbranschen har förpliktelser gentemot samhälle, kunder och finansärer samt gentemot de individer och sammanslutningar som statistiken beskriver. Vid framställning av statistik ska man beakta statistikens användningsbehov, dataskyddet och uppgiftslämnarbördan för uppgiftslämnare.

Riktlinjerna för europeisk statistik (European Statistics Code of Practice, CoP) utgör standardrutiner som produktionen av europeisk statistik ska bygga på. Utgångspunkten för sammanställningen av riktlinjer var behovet att sammanställa en minimistandard med vilken man garanterar obundenhet, ärlighet och ansvarsfullhet. I riktlinjerna fastställs också rollerna för aktörerna i det europeiska statistiksystemet. Riktlinjernas ställning har senare fastställts i Europeiska parlamentets och rådets förordning om europeisk statistik och realiseringen av principerna uppföljs regelbundet.

Riktlinjerna gäller statistikmyndigheternas oberoende och ansvar samt kvaliteten på processer och uppgifter som publiceras. Principerna överensstämmer med de principer för officiell statistik som godkänts av FN:s statistikkommission och kompletterar dem. Principerna överensstämmer också med principerna för det europeiska kvalitetspriset (EFQM).

Riktlinjerna för europeisk statistik omfattar 15 principer. Realiseringen av principer utvärderas med hjälp av indikatorer:

5.2.1 Bild: Riktlinjerna för europeisk statistik

Institutionell miljö	Statistikprocesser	Producerade statistikuppgifter
1. Yrkesmässigt oberoende 2. Mandat för datainsamling 3. Tillräckliga resurser 4. Kvalitetsåtagande 5. Statistiksekretess (dataskydd i statistiken) 6. 6. Opartiskhet och objektivitet	7. God metodik 8. Lämpliga statistiska förfaranden 9. En rimlig uppgiftslämnararbörda 10. Kostnadseffektivitet	11. Relevans 12. Noggrannhet och tillförlitlighet 13. Tidsenlighet och rättidighet 14. Konsekvens och jämförbarhet 15. Tillgänglighet och tydlighet

5.3 Peer review

Realiseringen av riktlinjerna för europeisk statistik säkerställs genom peer review. I EU:s statistikväsen har man genomfört två peer review-omgångar, av vilka den första under åren 2006–2008. Utvärderingen genomfördes av utvärderingsteam med medlemmar från olika statistikbyråer. Åren 2013–2015 genomfördes en andra utvärdering med hjälp av externa utvärderare. Under den andra omgången utvärderades också EU:s statistikbyrå Eurostats verksamhet. Rapporterna har publicerats på Eurostats hemsida. De utvecklingsåtgärder som konstaterats i samband med en peer review följs upp årligen.

6 Datainsamling

6.1 Centralisering av datainsamling

År 2013 koncentrerade Statistikcentralen en stor del av datainsamlingen till en separat enhet genom en omorganisation. Genom att koncentrera har man försökt effektivisera användningen av administrativt material och andra massdata, automatisering av mottagande av material, omfattande utnyttjande av material, aktivt letande av nytt material samt kostnadsbesparingar. Genom ändringar av verksamhetssätt försöker man få ett intensivare samarbete med intressegrupperna och bättre riskhantering samt minskning av uppgiftslämnarbördan.

Ur kvalitetssynvinkel är täckningen och kontinuiteten i materialen väsentlig och ur riskhanteringssynvinkel är det en tillräcklig tidig tillgång till uppgifter om förändringar i innehållet eller täckningen. Därför deltar Statistikcentralen aktivt i planeringen av nya administrativa register såsom inkomstregistret.

6.2 Allmänna principer för datainsamling

Vid Statistikcentralen styrs insamlingen av uppgifter framför allt av statistiklagen och personuppgiftslagen fram till 25.5.2018 (se närmare kapitel 4).

Statistikcentralen kartlägger behoven av statistikföring och eventuella datakällor med omsorg innan statistikföringen inleds och samlar in uppgifter direkt från enskilda uppgiftslämnare bara om det är nödvändigt. Vid framställningen av statistik utnyttjas i första hand uppgifter som samlats in för andra ändamål såsom uppgifter som samlats in då myndigheter skött förvaltningsuppgifter. Merparten av de personuppgifter som Statistikcentralen fått och bearbetat för statistikföringen härstammar från administrativa material som uppkommer i andra myndigheters egen verksamhet. Uppgifterna samlas in direkt från enskilda uppgiftslämnare bara om uppgifter som samlas in för andra ändamål inte står till förfogande. Av uppgiftslämnarna efterfrågas bara uppgifter som nödvändiga för statistikproduktionen.

Det är frivilligt att lämna uppgifter för statistikföringen, om det inte i lagen föreskrivs om uppgiftsskyldighet. För enskilda medborgare har inte någon uppgiftsskyldighet föreskrivits. Däremot har uppgiftsskyldighet föreskrivits i statistiklagen (14 och 15 §) för företag och andra organisationer såsom leverantörer av administrativa material.

Statistikcentralen strävar efter ett bra samarbete med uppgiftslämnarna och deras representanter vid datainsamlingen. Statistikcentralen samråder i huvudsak med företrädare för uppgiftslämnarna i samarbetsgrupper som grundats för ändamålet. Direkta samråd hålls med de myndigheter som lämnar administrativa material samt med andra aktörer som levererar material.

Statistikcentralen kan på uppdrag samla in uppgifter också åt andra myndigheter och sammanslutningar i samband med sin egen datainsamling. Sådana uppdrag kan tas emot, om det inte är till förfång för Statistikcentralens egen verksamhet, dvs. då ändamålet med uppgifterna har begränsats att användas för statistikföring och forskning. Statistikcentralen kan också samla in uppgifter tillsammans med andra myndigheter förutsatt att kraven på integritetsskydd för de uppgifter som samlas in för statistikändamål uppfylls.

6.2.1 Informering av uppgiftslämnarna

Uppgiftslämnarna och deras företrädare underrättas i god tid om att datainsamlingen inleds och om de uppgifter som samlas in så att uppgiftslämnarna på lämpligt sätt kan förbereda sig på att lämna uppgifter.

I samband med datainsamlingen informeras uppgiftslämnarna om centrala frågor som har att göra med uppgiftslämnandet: ändamålet med uppgifterna, hur skyddet av uppgifterna har ordnats, centrala förfaringssätt för framställning av statistik och bearbetning av uppgifter, eventuella övriga datakällor vid framställning av statistik och uppgifter som samlas in från annat håll, förvaringstid för uppgifterna, uppgiftslämnarens rättigheter, om uppgiftsskyldighet föreligger eller om datainsamlingen är frivillig och övriga uppgifter som är viktiga för uppgiftslämnaren, t.ex. om uppgifterna får lämnas vidare. Praxis för informering av uppgiftslämnarna blir något mer detaljerad då den nya dataskyddsförordningen delvis fastställer informeringen på nytt.

Den information som Statistikcentralen erbjuder uppgiftslämnarna är skiktad. Det innebär att de väsentligaste sakerna för uppgiftslämnandet har angetts i materialet som skickats till uppgiftslämnaren. I materialet anges att mer detaljerad tilläggsinformation fås på Statistikcentralens internetsidor eller genom att kontakta personer som har hand om datainsamlingen.

6.2.2 Minimering av uppgiftslämnarbördan

Statistikcentralen strävar efter att minimera kostnaderna för insamlingen av uppgifter både för uppgiftslämnaren och för Statistikcentralen och överlag uppgiftslämnarbördan.

Statistikcentralen följer upp uppgiftslämnarbördan för företagen och försöker på olika sätt minimera den.

Detta sker med noggrann planering av datainsamlingen, genom att använda befintliga datakällor, utnyttja statistiska och andra metoder för att minska kostnaderna för och arbetet med lämnande av uppgifter alltid då det är möjligt med tanke på statistikens tillförlitlighet. Till exempel användning av s.k. urval när man väljer ut personer eller företag för datainsamlingen minskar uppgiftslämnarbördan i målpopulationen. Sannolikheten för att komma med i urvalet är mycket liten särskilt för enskilda medborgare.

Syftet är att underlätta lämnandet av uppgifter genom blankettplanering och utveckling av datainsamlingsredskap. Genom att testa blanketterna säkerställer man kvaliteten på datainsamlingen. Vid testningen får en grupp uppgiftslämnare testa en autentisk enkätsituation och påverka att blanketterna är förståeliga och datainsamlingsredskapen användbara.

6.2.3 Datainsamlingsmetoder

Vid planeringen av en datainsamling beaktas när man bedömer de allmänna principerna för datainsamlingen, vilken insamlingsmetod som är det bästa alternativet med tanke på statistikens eller undersökningens mål, kvalitet och kostnader samt målpopulation.

Nedan anges de mest centrala datainsamlingsmetoderna:

I Finland är de **administrativa materialen och registren** omfattande. Att använda dem som grund för statistiken minskar medborgarnas eller företagens uppgiftslämnarbörda. En utmaning är att det ibland är svårt att bearbeta uppgifter som samlats in för administrativt ändamål så att de kan användas för statistiken samt att förändringar i lagstiftning och förvaltningspraxis kan försvaga jämförbarheten av uppgifterna. Uppgifter ur administrativa material kompletteras med uppgifter som samlats in med hjälp av andra insamlingsmetoder.

I en **telefonintervju** är frågorna sådana att de är lätta att förstå i telefon. Svartspersonen kan själv välja en lämplig plats och tidpunkt för intervjun.

Metoden är vanligast då uppgifter som gäller personer och hushåll samlas in. Intervjuerna är relativt korta. Vid insamling från företag görs få telefonintervjuer.

Besöksintervjuer, dvs. intervjuer som görs ansikte mot ansikte, görs oftast i svarspersonens hem eller på annan på förhand överenskommen plats. Besöksintervjuerna underlättar insamlingen av uppgifter med utmanande innehåll. De tar en längre tid än telefonintervjuerna. Denna metod används huvudsakligen i datainsamlingar som är riktade till individer. Ett format som kan jämföras med en besöksintervju är s.k. direkt observation, i praktiken betyder det att priserna på hyllvaror insamlas i butiker.

I **postenkäter** skickas en pappersblankett till svarspersonerna för ifyllning. Metoden används både vid datainsamlingar från personer och från företag, men ökningen av datainsamlingar via internet har minskat användningen av den här metoden. Dagboksmetoden är en speciell form av postenkät, där svarspersonen för bok över det fenomen som undersöks, till exempel tidsanvändning eller konsumtion.

Att **svara via internet** är snabbare och förmånligare än de andra datainsamlingsmetoderna. Metoden möjliggör mångsidiga frågeställningar till olika grupper av svarspersoner. Företag svarar huvudsakligen bara med webblankett. Också i fråga om vissa datainsamlingar från personer kan man svara via internet.

Statistikcentralens mål är att utveckla och ta i bruk nya former av datainsamling. Sådana är bl.a. automatiserad dataöverföring direkt från uppgiftslämnarens datasystem och automatiserad webb-insamling (s.k. web scraping).

6.3 Användning av material och deras aktualitet

Användningen av administrativa material och register i statistikproduktionen minskar både kostnaderna för statistikföring och uppgiftslämnarnas svarsbörda, vilket är ett av de centrala målen i statistiklagen.

De snabbaste administrativa materialen och registren är tillgängliga några veckor eller ett par månader från den aktuella perioden (fordonsregistret, registret över arbetssökande, befolkningsdatasystemet och skattemyndigheternas betalningsbevakningsmaterial), medan det tar upp till ett år innan de långsammaste materialen är färdiga (t.ex. beskattningsdatabasen). Hur aktuella uppgifterna är påverkas av hur lång tid det tar innan händelsen registreras och hur ofta uppgifter uppdateras i registret. Befolkningsuppgifter uppdateras varje dag, inkomstuppgifter uppdateras för närvarande främst en gång per år, men när skattemyndigheternas inkomstregister tas i bruk fås uppgifter om löner, anställningar och förmåner dagligen. Statistikcentralen tar emot uppgifter från inkomstregistret fr.o.m. början av år 2020.

Aktualiteten påverkar också hur täckande uppgifterna är. T.ex. när betalningsbevakningsmaterial inkommer första gången till statistikmyndigheten är täckningen omkring 80 procent av företagets omsättning och 75 procent av det slutliga beloppet när det gäller lönesumman som ordinarie arbetsgivare betalar. Beträffande enskilda företag är det problematiskt att veta om det saknas en anmälan på grund av dröjsmål med uppgiftslämnningen eller nedläggning av företagets verksamhet. Därför efterfrågas uppgifterna om den aktuella månaden för användning av statistikväsendet återkommande under sex månader till dess uppgifterna är täckande. Detta svarar emellertid inte på utmaningarna på tidsenlighet, som man försöker minska med statistiska metoder.

6.3.1 De största leverantörerna av material

Med hjälp av uppgiftsskyldigheten samlar Statistikcentralen in en stor mängd uppgifter från olika myndigheter ur de register som skapas i deras administrativa verksamhet. Betydande leverantörer av material är bl.a. Skatteförvaltningen, Befolkningsregistercentralen, Naturresurscentralen, Pensionsskyddscentralen, Finansinspektionen, Posti, Folkpensionsanstalten, Keva, Justitieministeriet, Rättsregistercentralen, Utbildningsstyrelsen, Patent- och registerstyrelsen, Polisstyrelsen, Finlands Bank, Institutet för hälsa och välfärd, Tullen, Arbets- och näringsministeriet, Statskontoret och Kommunikationsverket.

6.4 Statistikcentralens datainsamlingar riktade till personer och hushåll

Statistikcentralen samlar in uppgifter direkt från privatpersoner och hushåll. Störst av enkäterna är arbetskraftsundersökningen, där datainsamlingen månatligen riktar sig till omkring 12 500 personer. Arbetskraftsundersökningens undersökningsveckor omfattar årets alla kalenderveckor. Samma uppgiftslämnare intervjuas sammanlagt fem gånger per telefon under ett och ett halvt år. Intervjun tar i genomsnitt 4–8 minuter.

I konsumentbarometern är ett slumpmässigt urval på 2 350 personer, som helt byts ut varje månad, föremål för undersökningen. Samma urval används också i insamlingen av uppgifter för undersökningen om finländarnas resor.

Uppgifter samlas årligen in i enkäten om reparationer av bostäder och egnahemshus, undersökningen om inkomster och levnadsförhållanden samt undersökningen om befolkningens användning av informations- och kommunikationsteknik. Dessutom görs undersökningar oregelbundet eller inte så ofta, bl.a. undersökningen om vuxenutbildning, undersökningen om tidsanvändning, konsumtionsundersökningen, undersökningen om arbetsförhållanden och fritidsundersökningen. Dessutom genomförs datainsamlingar av engångsnatur om samhällsligt viktiga och aktuella ämnen.

År 2017 vara omkring 250 000 privatpersoner eller hushåll föremål för enkäter i nio direkta datainsamlingar.

7 Statistikproduktion

7.1 Statistik som produceras

Statistikcentralen producerar för närvarande statistik över omkring 160 statistikgrenar. Statistiken publiceras på Statistikcentralens webbplats. Omkring 560 nya offentliggöranden görs årligen om statistik.

En årlig publiceringskalender ges ut i slutet av föregående år på internet. Den innehåller alla statistiska offentliggöranden som publiceras under året i kronologisk ordning.

Ett årsprogram i rapportform Tilastokeskuksen vuosiohjelma – tilastot, tuotteet ja palvelut” (ung. Statistikcentralens årsprogram – statistikgrenar, produkter och tjänster) publiceras varje vår på webbsidan: <http://www.stat.fi/org/tilastokeskus/vuosiohjelma.html>

- Detaljerade beskrivningar av statistikgrenarna finns på adressen: http://www.tilastokeskus.fi/meta/til/index_sv.html
- Mera information om Statistikcentralens datainsamlingar finns på adressen: http://www.stat.fi/keruu/index_sv.html
- Mera information om ändringar i statistiken publiceras i realtid: http://tilastokeskus.fi/til/muutokset_sv.html

7.2 Livslängd för personuppgifter

De uppgifter som samlas in i statistiksyften bearbetas i statistikproduktionsprocesser som planerats och fastställts på förhand (bild 7.3.1). Hur lång livslängden i ämbetsverket blir för uppgifter som samlas in i statistiksyfte fastställs i de insamlings specifika besluten om datainsamling.

Statistikgrenarna och deras produktionsprocesser klassificeras vanligen på basis av källmaterialet. Statistiktyper är:

- **urvalsstatistik**, där målpopulationens egenskaper bedöms med hjälp av statistiska urvals- och estimeringsmetoder och uppgifterna erhålls i allmänhet genom direkt datainsamling
- **registerstatistik** och andra statistikgrenar, som utnyttjar indirekt insamlat mikromaterial, där källmaterialets variabler och målpopulation har fastställts för annat än för statistikföring.
- **samlingsstatistik** som utnyttjar färdiga kalkyler eller indikatorer för att beskriva fenomen som ofta är besvärliga att mäta direkt
- **kombinationer av föregående**. Ett växande antal statistikgrenar grundar sig på en sammanslagning av olika typer av datakällor.

Statistikcentralen bearbetar personuppgifter i flera skeden av statistikprocessen och i olika användningssituationer. Statistikföring på basis av personuppgifter kräver alltid utnyttjande av Befolkningsregistercentralens uppgifter. Befolkningsuppgifter från Befolkningsregistercentralen används både för uttagning av urval vid individbaserade datainsamlingar baserade på urval och produktion av totalstatistik. I individbaserad totalstatistik bearbetas uppgifter om befolkningen i hela Finland. För individbaserade urvalsundersökningar tas för sin del ut befolkningsuppgifter en gång på ett halvår. När insamlingen startar utökas urvalet med uppdateringsuppgifter, som hämtas till Statistikcentralen från Befolkningsregistercentralen varje vecka.

När Statistikcentralen samlar in personuppgifter direkt av privatpersoner och hushåll behandlas personuppgifterna för personer och hushåll som valts med i urvalet också i samband med att man kontaktar personerna. För sändningen av brev och ordnande av intervjuer hämtas utöver intervjuobjektets adressuppgifter också i vissa fall telefonnummer eller e-postadresser ur offentliga datakällor.

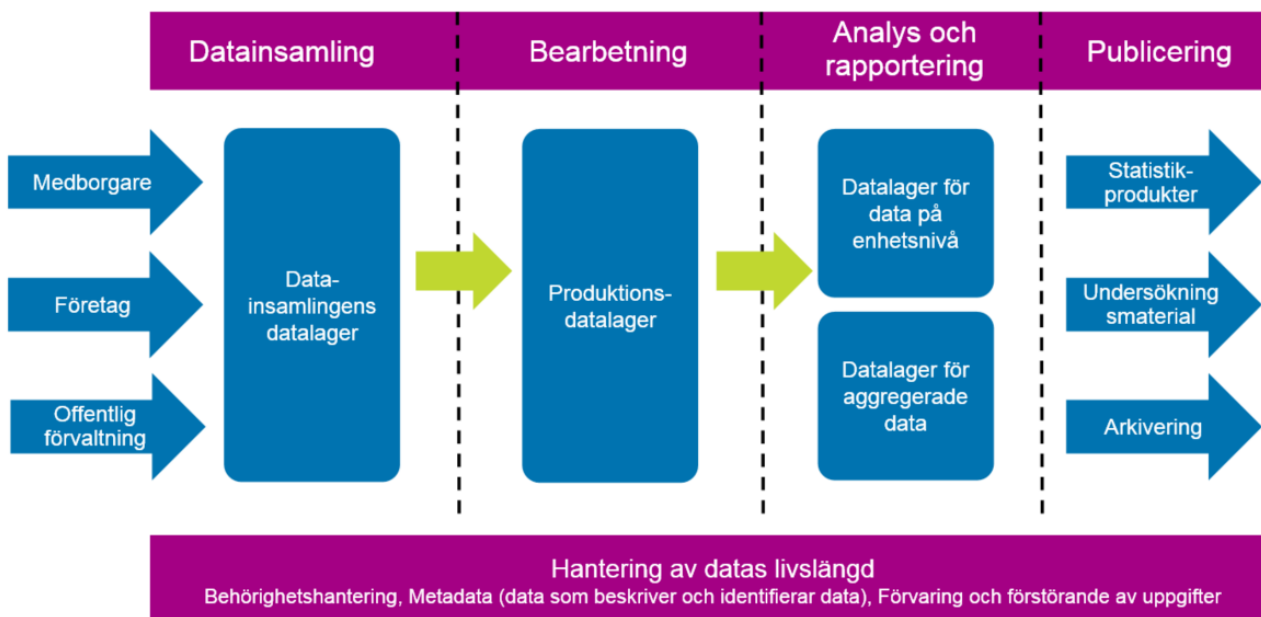
När det gäller företagsenkäter samlar Statistikcentralen in personuppgifter från företagen om deras anställda – i enkäter som vanligen har att göra med lön och anställning.

Datas livslängd vid Statistikcentralen börjar när Statistikcentralen får en annan myndighets eller annan dataproducents datamaterial till sitt förfogande. Livslängden för personuppgifter som samlas in direkt av medborgare eller företag börjar när uppgiftslämnaren besvarar Statistikcentralens enkät. Uppgifter som samlats in direkt av dataleverantörer och uppgiftslämnare lagras i insamlingsspecifika datalager där uppgiften är i samma form som Statistikcentralen fått den. Behandlingen av uppgifter sker på ett separat område för behandling av uppgifter där uppgifter som samlats in om enskilda personer eller organisationer ur olika källor slås samman och behandlas enligt kraven för produktion av statistiska uppgifter. Behandlingen sparas i s.k. produktionsdatalager. Produktionsdatalager och tidigare nämnda rådatalager bevaras olika länge beroende på kraven i statistikprocessen.

På basis av produktionsdatalagret skapas datalager på enhetsnivå (t.ex. person eller företag) som behövs för beskrivning av varje fenomen. De bevaras permanent. Att de bevaras permanent betyder att uppgifterna och beskrivningarna av uppgifterna (beskrivningarna av materialet och dess variabler) arkiveras så att det är möjligt att bearbeta dem på nytt för angivet ändamål också efter flera decennier.

Av enhetsdatalager skapas också statistikdatalager med aggregerade uppgifter, på basis av vilka man framställer olika statistikprodukter såsom offentliggöranden och tabeller. Uppgifter som behandlas i det här skedet är statistisk information som kan distribueras och ställas till samhällets förfogande som olika dataprojekter. Det är möjligt att läsa mer om dessa resultat av statistikföringen bl.a. på Statistikcentralens webbplats.

7.2.1 Bild: datainsamlingens datalager



7.3 Processmodell för statistikproduktionen

7.3.1 Bild: processmodell för statistikproduktionen

Hantering och kvalitetshantering av metadata							
1 Fastställande av informationsbehovet	2 Planering av statistikföringen	3 Byggande av produktions-systemet	4 Datainsamling	5 Bearbetning	6 Analys och rapport	7 Publicering	8 Utvärdering av processen
1.1 Kartlägg informationsbehov	2.1 Planera innehållet	3.1 Skapa datainsamlingsinstrument	4.1 Urval	5.1 Sammanslå källmaterial	6.1 Framställ uppgifter som offentliggörs	7.1 Uppdatera distributionsdata basen	8.1 Sammanställ uppföljningsuppgifter
1.2 Säkerställ informationsbehov	2.2 Beskriv variabler	3.2 Skapa bearbetnings- och analyslösningar	4.2 Förbered datainsamlingen	5.2 Klassificera och koda	6.2 Granska uppgifter som offentliggörs	7.2 Sammanställ publikationen	8.2 Utvärdera
1.3 Avtala om datainnehållet	2.3 Planera datainsamling	3.3 Genomför lösningar för distribution	4.3. Genomför datainsamlingen	5.3 Identifiera fel	6.3 Tolka och rapportera resultaten	7.3 Publicera	8.3 Avtala om åtgärder
1.4 Identifiera begrepp	2.4 Planera urval		4.4 Slutbearbeta datainsamlingen	5.4 Editera och imputera	6.4 Säkra dataskyddet	7.4 Främja informationsanvändningen	
1.5 Kartlägg datakällor	2.5 Planera metoder för bearbetning och analys			5.5 Bilda härledda uppgifter	6.5 Slutbearbeta uppgifter som offentliggörs	7.5 Stöd informationsanvändningen	
1.6 Förbered verkställighetsplan	2.6. Planera arbetsförloppet och verktyg			5.6 Beräkna vikter			
	3.4 Förbered arbetsförlopp			5.7 Estimera			
	3.5 Testa produktions-systemet			5.8 Slutbearbeta uppgifter			
	3.6 Testa statistikproduktionen						
	3.7 Slutbearbeta produktions-systemet						

Målet med processmodellen för den allmänna statistikproduktionen är att förenhetliga produktionsprocesserna för statistiken. Processmodellen för statistikproduktionen (Generic Statistic Business Process Model, GSBPM) är en beskrivning av produktionsprocessen för statistik (statistikföringsprocess) som FN:s ekonomiska kommission för Europa UNECE och de europeiska statistikbyråerna upprätthåller. På Statistikcentralen infördes modellen år 2015 och efter det har en strukturering enligt modellen använts bl.a. i nya arbetsinstruktioner, processbeskrivningar, gränssnitt för nya datasystem, vid tillämpning av strukturen för arbetskataloger och i andra tillämpliga objekt.

7.4 Statistikdatalager

Statistikproduktionen baserar sig på en kombinerad av **material** som beskriver fenomen inom nästan alla ämnesområden ur olika synvinklar. När mängden data

som ska sammanläggas ökar är det allt viktigare att uppgifterna kan administreras, att de är konsekventa och konsistenta.

De största materialhelheterna har sammanslagits till **datalager**. Genom att slå samman uppgifterna hindrar man att statistikuppgifterna är motsägelsefulla och att behandlingen av materialen blir överlappande. Huvudregeln är att en och samma uppgift sparas på ett ställe.

Statistikcentralens största statistikdatalager är:

1. Företagsdatalagret
2. Datalagret över registermaterial som beskriver befolkningen och levnadsförhållanden
3. Datalagret över inkomstfördelningen
4. Lönedatalagret
5. Datalagret över nationalräkenskaper och betalningsbalansen
6. Datalagret över kommunekonomi
7. Datalagret över miljö- och energistatistik
8. Datalagret över pris- och kostnadsindex
9. Datalagret över nybyggande.

7.5 Nuvarande persondatalager

Det nuvarande persondatalagret omfattar det registerbaserade totalmaterial med anknytning till personer som används i statistikföringen. De äldsta uppgifterna är från folkräkningen år 1950.

Datalagret innehåller bl.a. följande uppgifter:

- folkräkningsuppgifter från år 1970 (dessutom ett urval på 10 procent från folkräkningen år 1950)
- uppgifter om befolkningsstruktur, familjer och bostadshushåll
- uppgifter om byggnader och bostäder
- befolkningsförändringar: födda, döda, flyttningar osv.
- områdesuppgifter (delområden, tätorter, rutor och postnummer)
- examensregister, läroanstaltsregister, studerandematerial
- sysselsättningsuppgifter
- dödsorsaksuppgifter.

Uppgifterna i datalagret har producerats genom att kombinera, klassificera och koda om flera administrativa register. Källmaterialet behandlas i separata produktionsdatabaser och uppgifterna överförs till persondatalagret när de slutliga materialen blivit färdiga. I datalagret har personernas namn och personbeteckningar ersatts med personnummer (s.k. surrogat).

I datalagret består uppgifterna i huvudsak av tvärsnittsuppgifter om någon tidpunkt, i allmänhet årets slut. Databastabellerna omfattar dessutom perioduppgifter, t.ex. om anställningar, arbetslöshet och äktenskap.

Med stöd av persondatalagren produceras flera statistikgrenar bl.a. sysselsättningsstatistik, statistik över struktur- och förändringar i befolkningen samt statistik över bostäder och boendeförhållanden.

Största delen av klassificeringarna i datalagret baserar sig på internationella standarder fastställda med EU-direktiv (bl.a. näringsgrensindelningen och

utbildningsklassificeringen). Genom att använda klassificeringsrekommendationer främjar man jämförbarheten mellan statistiska uppgifter som producerats under olika tider och om olika områden.

Genom hela produktionsprocessen genereras metadata om enskilda statistikgrenar som ska dokumenteras och bevaras. Med metadata avses t.ex. definitioner, klassificeringar, standarder, urvalsram, tilläggsinformation, editeringar och variabeländringar. Material- och variabelbeskrivningar av tabeller i persondatalagret har sparats i metadatalagret.

En kvalitetsbeskrivning av var och en av de publicerade statistikgrenarna har gjorts där man mer detaljerat beskrivit undersökningsmetoderna i statistiken i fråga dvs. statistikens målpopulation, använda primärmaterial, urvalsdesign, datainsamlingsätt, de viktigaste klassificeringarna och estimeringsmetoderna. För en del av statistikgrenarna har man också utarbetat en separat detaljerad metodbeskrivning.

Vid Statistikcentralen används genomlysning av statistiken och självvärderingar som skräddarsytt för statistikproduktionen som centrala metoder för kvalitetsbedömning som beskriver den totala kvaliteten på statistiken.

Uppgifter från persondatalagret utnyttjas i ett flertal statistikgrenar som baserar sig på urval, utöver i individstatistik som publiceras på basis av totalmaterial, såsom i arbetskraftsundersökningen, inkomstfördelningsstatistiken, konsumtionsundersökningen och i statistiken över energiförbrukningen inom boendesektorn. Ur persondatalagret framställs dessutom undersökningsmaterial direkt åt forskare och andra kundspecifika specialutredningar.

7.6 Nuvarande lönedatalager

Nuvarande lönedatalager innehåller registerbaserade uppgifter för statistikföring och undersökningsändamål samt personuppgifter om löner och anställningar som samlats in med Statistikcentralens egna datainsamlingar. De äldsta uppgifterna är från 1970-talet och de nyaste uppgifterna är för närvarande från år 2017. Datalagrets uppgifter beskriver 1,6 miljoner löntagares inkomstuppgifter som upphöjda.

Datalagret innehåller:

- Uppgifter om den privata sektorns månads- och timlöner
- Uppgifter om kommunsektorns månads- och timlöner
- Statens löneuppgifter
- Lönestrukturstatistikens uppgifter

Datalagrets uppgifter har producerats genom att sammanslå, klassificera och koda om med hjälp av det statistiska ämnesområdets egna och Statistikcentralens allmänna klassificeringar samt de gemensamma datalagrens uppgifter bl.a. ur YTY-datalagret.

Behandlingen av källmaterial görs i en separat operativ databas och uppgifterna överförs till datalagret när datamaterialen blivit färdiga.

Uppgifterna i datalagret är tvärsnittsuppgifter från årets sista kvartal. Beträffande enskilda statistikprocesser inflyter metadata genom hela produktionsprocessen

bl.a. definitioner, klassificeringar, tilläggsinformation och editeringar. Lönedatalagrets materialbeskrivningar samt beskrivningar av informationsdatabaser har sparats i metadatalagret.

Ur lönedatalagret produceras följande statistik:

- statistik över månadslöner inom den privata sektorn
- statistik över timlöner inom den privata sektorn
- statistik över månads- och timlöner inom kommunsektorn (gemensamt)
- lönestatistik över den statliga sektorn
- Lönestrukturstatistik (EU-statistik)

En kvalitetsbeskrivning av var och en av de publicerade statistikgrenarna har gjorts där man mer detaljerat beskrivit undersökningsmetoderna i statistiken i fråga dvs. statistikens målpopulation, använda primärmaterial, urvalsdesign, datainsamlingssätt, de viktigaste klassificeringarna och estimeringsmetoderna.

Lönedatalagrets uppgifter utnyttjas förutom i löne- och lönestrukturstatistik också i produktionen av förtjänstnivåindexet. På basis av löneuppgifterna görs standardiserade informationstjänster och specialutredningar. Till de största aktörerna som använder sig av löneuppgifter hör arbetsmarknadsorganisationer, myndigheter, företag och media. Dessutom levereras lönestrukturuppgifter till Statistikcentralens forskartjänster, där enhetlig harmoniserad lönestrukturstatistik bildas på basis av uppgifterna för forskarna (tidsserie fr.o.m. år 1995)

7.7 Vägtrafikolyckor

Statistiken över vägtrafikolyckor innehåller information om vägtrafikolyckor med personskador och om personer som varit delaktiga i dem. Datainnehållet omfattar vägtrafikolyckor med personskador som kommit till polisens kännedom och som polisen meddelat Statistikcentralen. Uppgifterna samlas in om de vägtrafikolyckor som lett till personskador och som införts i polisens datasystem och meddelats Statistikcentralen. Polisens uppgifter kompletteras med Trafikverkets och Trafiksäkerhetsverkets uppgifter samt med uppgifter från Statistikcentralens statistik över dödsorsaker.

Materialet innehåller alla olyckor med dödlig utgång och en stor del av de olyckor som lett till personskador. Datamaterialet omfattar uppgifter med anknytning till personer som är delaktiga i olyckor, t.ex. väganvändargrupp, ålder, kön, medborgarskap, resultat av utandningsprov, användning av rusmedel och uppgifter om körkort. Dessutom omfattar materialet uppgifter om olyckan (allvaret i olyckan, olyckstyp, tidpunkt, antal delaktiga), uppgifter om olycksplatsen (landskap, kommun, plats på vägnätet, hastighetsbegränsning, korsning) och uppgifter och förhållanden vid olyckan (väglag, väder, ljusförhållanden).

7.7.1 Bild: Datasystem



Statistikcentralens datasystem är indelade i tre huvudgrupper. Statistikproduktionssystemen omfattar fyra olika huvudskeden i produktionen. Indelningen följer processmodellen för den allmänna statistikproduktionen (GSBPM).

De system som stöder statistikproduktionen har en inverkan genom hela statistikproduktionsprocessen. De är väldigt bundna till näringsgrenen och av systemen är största delen system som Statistikcentralen själv producerat.

8 Distribution av statistisk information

8.1 Statistikcentralens principer för offentliggörande av statistik

Tidpunkterna för publicering av statistik fastställs varje höst i samband med planeringen av verksamheten. Publiceringskalendern för följande år publiceras för användarna i december.

Statistikcentralens offentliggöranden av statistik, uppdateringar av statistikdatabaser samt meddelanden offentliggörs samtidigt till alla vardagar kl. 9.00 på internet. Uppgifterna är offentliga efter att de har uppdaterats på webbplatsen. Undantag är bl.a. statistik och meddelanden som offentliggörs under presskonferenser. Tidpunkten för offentliggörande av dem fastställs separat. Före publiceringen får de statistiska uppgifterna bara behandlas av personer som deltar i framställningen av statistiken ifråga.

- [Publiceringskalender](#)

8.2 Öppna data

Enligt internationella jämförelser producerar Finlands statistikväsen öppna data av hög kvalitet. [ODIN \(Open data inventory\)-mätaren](#), som lägger vikt särskilt vid material som är tillgängligt som öppna data på statistikbyråernas webbsidor i olika länder, placerar Finlands statistikväsen bland de bästa internationellt sett.

Kundens behov och förväntningar med anknytning till öppna data har ökat. Gensvar på öppna data har också erhållits via sammanslutningar som verkar inom öppna data. Enligt olika kundenkäter och respons finns det fortfarande behov av utveckling av Statistikcentralens öppna data, bl.a. av kvaliteten på metadata och datas användbarhet samt hur data kan hittas och hur synliga de är.

Statistikcentralen har satsat på att öka tillgången på öppna data (www.stat.fi/avoindata_sv). En kartläggning av kundernas behov av data ur landskaps- och vårdreformens synvinkel inleds. Kartläggningen är en del av en mer omfattande projekthelhet, där målet i framtiden är att öka utbudet av regionala uppgifter som öppna data.

Utöver att Statistikcentralen har öppnat upp data har Statistikcentralen deltagit i hackathon-evenemang som utnyttjar öppna data och anordnat kundutbildningar om öppna data.

9 Utlämnande av uppgifter

9.1 Utlämnande av uppgifter för statistiska ändamål

Om utlämnande av uppgifter som samlats in för statistiska ändamål och begränsningar med anknytning till utlämnandet föreskrivs i Europaparlamentets och rådets förordning om europeisk statistik och i statistiklagen. Sekretessbelagda uppgifter som samlats in för statistiska ändamål kan lämnas ut för statistikföring till en annan statistikmyndighet, en myndighet som framställer statistik för Europeiska unionens databehov och till Finlands Bank.

Avtal om utlämnande och byte av uppgifter har ingåtts med Naturresursinstitutet, Institutet för hälsa och välfärd, Ålands statistikbyrå ÅSUB och Finlands Bank.

Dessutom lämnar Statistikcentralen ut anonymiserade eller aggregerade uppgifter till Europeiska unionens statistikbyrå Eurostat som krävs enligt Europaparlamentets och rådets förordningar om statistikföring. Material i statistikform lämnas ut bl.a. till FN och OECD, vilket internationella avtal förpliktar till.

9.2 Utlämnande av uppgifter för forskningsändamål

Rätt att använda uppgifter som samlats in för statistiska ändamål kan beviljas, dvs. uppgifterna kan behandlas i övervakad fjärråtkomstmiljö, om det är möjligt att indirekt identifiera en enskild person i materialet under bearbetningens gång. Efter granskningsprocesserna är det inte längre möjligt att identifiera enskilda personer i de färdiga tabellerna. Det anonymiserade undersökningsmaterialet där inte ens indirekt identifiering är möjlig, kan lämnas ut till forskaren.

Vid Statistikcentralen verkar en statistiketisk nämnd som generaldirektören tillsatt. Dess uppgift är att ge utlåtanden om utlämnande av mikromaterial för forskningsändamål samt andra dataskyddsärenden och statistiketiska frågor som har att göra med t.ex. datainsamling, behandling eller publicering. Målet med den statistiketiska nämndens behandling är att säkerställa ett enhetligt förfarande i principiella frågor som gäller tillämpning av lagstiftningsprinciper och statistiketiska principer genom att föra samman statistisk och juridisk sakkunskap som stöd för beslutsfattandet.

9.2.1 Behandlingen av individmaterial kräver tillstånd av Statistikcentralen

Användning av Statistikcentralens material på enhetsnivå kräver tillstånd. Uppgifterna får användas bara av en person som har tillstånd till det och enbart för det ändamål som godkänts i beslutet. De personer som behandlar materialen undertecknar en sekretessförbindelse.

Med tanke på statistikföringsobjekten dvs. observationer i statistikmaterialet, är det viktigt att ha hand om dataskyddet. Att uppgifter om enskilda personer, hushåll, företag eller någon annan statistisk enhet röjs för utomstående ska hindras i uppgifter som publiceras på basis av material som Statistikcentralen överlåtit för forskningsändamål. Till villkoren för användningstillståndet hör att forskaren inte får försöka identifiera objekten i materialet. För användarna av materialet har man tagit fram anvisningar om dataskydd. För forskarna finns också en webbkurs om distansanvändningen, Tutkimusaineistot etäkäytössä (på finska).

Tillstånd beviljas för en bestämd tid. Statistikcentralen sänder en påminnelse till forskningsprojektet när tillståndet går ut. Material som lämnats ut till användaren ska återsändas eller förstöras inom den tid som fastställts i beslutet om användningstillstånd när projektet går ut. På samma sätt ska eventuella kopior av materialet förstöras och detta ska meddelas Statistikcentralen. Forskarnas inloggningskoder fungerar inte längre i distansanvändningssystemet efter att användningstiden gått ut. Det är möjligt att ansöka om en förlängning av användningstillståndets giltighet.

Som villkor för användningstillståndet kan man ställa att forskaren har tillgång till materialet bara i Statistikcentralens lokaler eller i fjärråtkomstsystemet (t.ex. totalmaterial). Av mottagaren av tillståndet kan man också kräva att han eller hon ska presentera utredningen eller undersökningen för Statistikcentralen för dataskyddsgranskning innan den publiceras. Slutprodukterna eller hänvisningarna till dessa ska också skickas till Statistikcentralen.

Statistikcentralen raderar personbeteckningarna från alla material som lämnas ut för forskningsändamål och ersätter dem med pseudokoder. I undantagsfall kan Statistikcentralen lämna ut uppgifter om ålder, kön, yrke, socioekonomisk ställning och utbildning också med identifieringsuppgifter, om forskaren har rätt att behandla personuppgifter. Med stöd av speciallagstiftning kan också

dödsorsaksuppgifter lämnas ut med identifieringsuppgifter när det är fråga om forskningsändamål.

Det är också möjligt att slå samman material med forskarens eget eller andra organisationers material. Då levereras material med identifierare (personbeteckning, namn) till Statistikcentralens forskartjänster där direkta identifierare raderas och ersätts med projektvisa pseudokoder innan materialet överförs till systemet för distansanvändning.

9.2.2 Tabell: Antalet beviljade användningstillstånd och utlämning av material efter produktgrupper år 2017

Produktgrupp	beviljade användnings-tillstånd	material-leverans
Dödsorsaksmaterial	82	84
Fortsättningstillstånd, utvidgning av tillstånd	141	
Mikrosimulering	21	14
Färdiga material/distansanvändning	49	40
Skräddarsydd/distansanvändning	84	76
Skräddarsydd/utlämning	29	19
Med identifierare	26	21
Övriga	5	6
Totalt	437	260

9.2.3 Sisu-mikrosimuleringsmodell

Statistikcentralen utvecklar och upprätthåller mikrosimuleringsmodellen SISU (Simuloitu Suomi) som modellerar Finlands personbeskattnings- och socialskyddssystem. SISU-modellen är ett beräkningsverktyg avsett för planering, uppföljning och evaluering av personbeskattning och socialskyddslagstiftning. Det registerbaserade omfattande mikromaterialet som används i SISU-modellen och som är en god representation av befolkningen (urval av omkring 800 000 personer) kan användas bara via distanssystemet.

År 2017 hade SISU-modellen 75 användare. Användarna kommer i regel från ministerierna, forskningsinstituten och intresseorganisationerna.

9.2.4 Distansanvändningssystemet

Statistikcentralen har ett system för distansanvändning av uppgifter (FIONA – Finnish Online Access). Via denna fjärråtkomsttjänst får forskaren tillgång till mikromaterial från sin egen arbetsstation i en övervakad och datasäker omgivning. Systemet använder stark identifiering. Forskaren loggar in i fjärråtkomstmiljön antingen till uppdragsservern eller mikrosimuleringsservern.

Att arbeta på fjärrskrivbordet motsvarar arbete på den egna datorn, men forskningsmaterialet förs över till användaren bara som skrivbordsvisning. I undersökningsmaterialet har enhetsnivåns direkta identifierare, såsom personbeteckningar eller FO-nummer ersatts med pseudokoder i systemet. Pseudokoden är en artificiell identifierare som ersätter den direkta

enhetsidentifieringen och med hjälp av vilken man kan följa upp enheten (eller individen) i materialet och länka enhetsuppgifterna mellan olika material. När forskaren vill föra ut filer och resultat från systemet, granskar Statistikcentralen dataskyddet före överföringen.

För det tekniska underhållet av distansanvändningssystemet ansvarar IT-centret för vetenskap CSC (som står under Undervisnings- och kulturministeriet) tillsammans med Valtori. För systemets datasäkerhet och för dataskyddet av undersökningsmaterial som förvaras på CSC:s server ansvarar Statistikcentralen och CSC enligt avtalet mellan organisationerna.

Forskarna har tillgång till analysprogram (SPSS, STATA osv.) och varje projekt har egna arbetsutrymmen för att förvara forskningsmaterial, analysresultat och koder. Vid mikrosimulering har forskaren tillgång till programmet SAS EG som är avsett för användning av SISU-mikrosimuleringsmodellen. Användningen av systemet är avgiftsbelagd och logguppgifter om användningen samlas in för fakturering samt för underhåll och övervakning av systemet.

Distansanvändningen förutsätter att forskarens egen organisation har upprättat ett avtal om öppnande och underhåll av en distansförbindelse samt att den har levererat en utredning om dataskydds- och datasäkerhetspraxis ifråga om distansförbindelsen. I avtalet anges villkoren för användningen samt organisationens administrativa och tekniska kontaktperson.

Antalet distansanvändare ökar hela tiden. Statistikcentralen utvecklar mätare tillsammans med CSC, med vilka användningen kan uppföljas. I slutet av år 2017 hade 71 organisationer ett avtal om distansanvändning med Statistikcentralen. Antalet projekt var totalt 163 och antalet användare 363.

9.3 Företags- och arbetsställeregistret

Statistikcentralen upprätthåller ett företags- och arbetsställeregister som innehåller alla företag, sammanslutningar och privata näringsidkare som är skyldiga att betala mervärdesskatt för sin affärsverksamhet eller som verkar som arbetsgivare. Registret omfattade i slutet av år 2017 omkring 735 000 fysiska personer av vilka omkring 287 000 personer fortfarande utövar verksamhet. En fysisk person bedriver företagsverksamhet under eget namn eller registrerat firmanamn. Till denna grupp hör yrkesutövare, de flesta lantbruksföretagare samt affärsidkare.

Registrets huvudsakliga användningssyfte är att fungera som ett statistiskt grundregister vars uppgifter används i datainsamlingar och vid framställning av statistik som källa för primäruppgifter, urvalsram samt som grund för registret över uppgiftslämnande företag. Dessutom används registret som informationstjänstregister och offentliga uppgifter enligt 18 § i statistiklagen levereras till kunderna från registret. Uppgifter i företagsregistret är bl.a. företagets adresser, näringsgrenar, omsättningens och personalens storleksklass, datum då företaget inlett sin verksamhet samt uppgift om importör/exportör. Uppgifter finns både på företag- och arbetsställenivå.

Företagsregistrets datakällor är ett flertal administrativa material samt egna enkäter som sänds direkt till företagen. De viktigaste materialen är Skatteförvaltningens kunddatabas och Skatteförvaltningens övriga datamaterial. Dessutom utnyttjas Patent- och registerstyrelsens och Skatteförvaltningens

gemensamma datasystem FODS, Tullens material över utrikeshandel, Patent- och registerstyrelsens handelsregister, Befolkningsregistercentralens register över byggnader och lägenheter, Posti Oy:s adressmaterial, Statskontorets och Kevas material över anställningsförhållanden.

I Statistikcentralens egna enkäter utreder man bl.a. företagens näringsgren, antalet anställda och uppgifter om arbetsställets struktur. Enkäter sänds årligen till merparten av företag med flera arbetsställen och till företag med ett arbetsställe och mer än 20 anställda. Till övriga företag sänds enkäterna mera sällan beroende på rotation. Enkäten sänds till nya företag rätt snart efter att verksamheten inletts.

9.4 Register över uppgiftslämnande företag

I registret över uppgiftslämnande företag samlas uppgifter om till vilka Statistikcentralens datainsamlingar det uppgiftslämnande företaget hör och tidpunkterna för datainsamlingarna. I registret över uppgiftslämnande företag visas alla direkta datainsamlingar från företag och alla datainsamlingar som riktar sig till ett enskilt företag. Uppgifter har samlats in till registret sedan år 2006.

Registret innehåller följande uppgifter om uppgiftslämnarna: FO-nummer, uppgiftslämnarbeteckning, uppgiftslämnarens namn, kontaktpersonens namn, svarsuppgift och svarssätt. Uppgifterna används bl.a. vid informeringen av uppgiftslämnare (datainsamlingstjänst för företag) och inom Statistikcentralen t.ex. vid utredning av vilka som är stora uppgiftslämnare och hur uppgiftslämnarbördan hänför sig till företagen.

Uppgifter lämnas ut från registret bara till uppgiftslämnaren själv via programmet Datainsamlingstjänst för företag, som finns på Statistikcentralens webbplats för datainsamling. Uppgifterna kan också lämnas ut till prokuristen i koncernens moderbolag om moderbolaget vill ha tillgång till sina dotterbolags datainsamlingsuppgifter. För användning av denna tjänst skickas användningsidentifikationerna separat till de uppgiftslämnande företagen.

10 Behandlingen av personuppgifter i form av uppdrag

10.1 Kundfinansierade intervju- och datainsamlingstjänster

Statistikcentralen erbjuder intervju- och datainsamlingstjänster som en avgiftsbelagd verksamhet. Externa kunder är t.ex. olika forskningsinstitut, ministerier, universitet och forskargrupper. Föremål för datainsamling kan beroende på insamlingen vara en person, ett hushåll, företagets verksamhetsställe eller en kommun. Uppgifter samlas in beroende av datainsamlingen via webben, per telefon, besöksintervju eller post, men oftast genom kombinationer av olika sätt.

Ämnen och storleksklasser för datainsamlingarna varierar mycket. Datainsamlingarna varierar t.ex. från webb-postenkäter riktade till tiotusentals svarspersoner till undersökningar genom besöksintervjuer med 3 000 personer och från datainsamlingar via webb-telefon med 5 000 personer till webbenkäter som omfattar alla Finlands kommuner och samkommuner.

Till genomförandet av enskilda undersökningar hör många arbetsskeden från planeringen av datainsamlingen till bearbetningen, leveransen och, beroende på avtal, också rapporteringen av materialet. Oberoende av ämnet för datainsamlingen hör planering av frågor tillsammans med kunden, produktion av enkätblanketter för olika insamlingsverktyg, planering och framställning av material för uppgiftslämnarna (följebrev och påminnelser, broschyrer och andra material samt datainsamlingssidornas innehåll), uppföljning och styrning av ackumuleringen av svar, putsning och grundanalysering av materialet samt framställning av en analysrapport. Analysrapporten, grundrapporten och materialet levereras till kunden när datainsamlingen upphör.

Arkiveringen av kundfinansierade undersökningar görs så att bara de personuppgifter som är nödvändiga för efterbehandlingen bevaras av allt material. Dessutom bevaras det slutliga medföljande materialet, blanketten och data.

På arkiveringen inverkar vad som står i avtalet med kunden. I vissa fall kan man komma överens om att bevara materialet eller annat i större omfattning eller längre t.ex. för nästa omgång av datainsamlingen eller för paneldesign.

10.1.1 Tabell: Kundfinansierade datainsamlingar

År	Antalet uppdrag
2010	23
2011	15
2012	11
2013	10
2014	7
2015	14
2016	11
2017	12

10.2 Kundenspecifika avtal om informationstjänst

Statistikcentralen utför specialutredningar som avgiftsbelagd serviceverksamhet utgående från kundens behov. I specialutredningen används de material som Statistikcentralen samlat in för statistikändamål t.ex. material gällande befolkning, sysselsättning, bostadshushåll, bostäder och byggnader. Ibland slås materialet samman med kundens material eller uppgifterna levereras på basis av en områdesindelning som kunden levererat. Uppgifterna rapporteras som en enskild tabell, ett tabellpaket, en statistisk utredning, grafiska figurer eller temakartor.

10.3 Statistikföringstjänster

Statistikcentralen producerar statistik på uppdrag också till andra organisationer. Statistikföringstjänsten kan omfatta hela processen för framställningen av statistik från datainsamling till distribution eller också kan den bara omfatta något deluppdrag av processen. Datainsamlingen kan då genomföras antingen genom intervjuer, som en elektronisk enkät eller genom att kombinera färdiga registermaterial. Ifråga om statistikföringstjänster tillämpas alltid samma statistiska arbetsätt och dataskyddspraxis som i all annan statistikproduktion.

Tjänsten planeras alltid i nära samarbete med kunden och ett detaljerat avtal ingår om genomförandet av tjänsten.

Som en del av statistikföringstjänsten kan man också publicera färdiga figurer, tabeller samt analystexter. Vid behov levereras statistikmaterialet till kunden också i icke-identifierbar form för forskningsändamål.

För närvarande utförs omkring 20 olika statistikföringstjänster vid Statistikcentralen. Kunder är bl.a. Skatteförvaltningen och Trafiksäkerhetsverket Trafi. Tillhandahållandet av tjänster har utvidgats ganska kraftigt under de senaste åren. Utvidgningen baserar sig i stor utsträckning på goda kundrelationer och en positiv image.

10.4 Nationell distansanvändningstjänst

Statistikcentralen har kommit överens med fyra myndigheter och 23 högskolor om att de använder FIONA-distansanvändningssystemet som en kanal för utlämning av forskningsmaterial. Då har registerförarens s.k. färdigmaterial levererats till Statistikcentralens forskartjänster, där materialet har försetts med pseudokod och överförs till FIONA. Statistikcentralens forskartjänster öppnar fjärråtkomsten till modulerna med färdigmaterial enligt de användartillstånd som registerföraren beviljat forskare som anges i tillstånden.

11 Kund- och intressegruppsuppgifter

Till de kunder och intressegrupper som använder Statistikcentralens slutprodukter hör t.ex. statens ämbetsverk och inrättningar, kommuner, företag och sammanslutningar, universitet och läroanstalter, forskare samt media. Kundrelationen bygger alltid på avtal, samtycke eller ställning och uppgifterna samlas i regel in från kunden eller kontaktpersonen själv. För kund- och intressegruppsuppgifterna har förvaringstider fastställts i planen för arkivbildningen. Förvaringstiderna varierar efter situation. Till exempel uppgifterna om en beställare av Statistikcentralens nyhetsbrev tas bort när kunden avslutar sin beställning. Uppgifter med anknytning till avtal förvaras för sin del tio år efter avtalets giltighetstid, varefter uppgifterna tas bort.

Hantering och behandlingen av uppgifter om kund- och intressegrupper vid Statistikcentralen har utvecklats under de senaste åren. I slutet av år 2016 infördes ett nytt kund- och intressegruppssystem, kund- och intressegruppsuppgifterna genomgick en grundläggande uppdatering och föråldrade uppgifter togs bort på behörigt sätt. Med hjälp av det nya systemet har det blivit enklare att kontrollera riktigheten av kund- och intressegruppsuppgifter och det är nu en systematisk del av kund- intressegruppsarbetet.

År 2017 integrerade man i systemet för kund- och intressegruppshantering nya verktyg för kund- och intressegruppskommunikation som stöd för e-postkommunikationen och kanalerna i sociala medier. På gång är också utvecklingen av Statistikcentralens statistiska offentliggöranden och beställningssystemet för olika nyhetstjänster bl.a. på så sätt att beställarnas möjlighet att hantera sina egna beställningar och uppgifter förbättras.

Tack vare de senaste årens utvecklingsarbete känner Statistikcentralen sina centralaste kunder och intressegrupper bättre än tidigare och samtidigt kan man säkerställa tidsenligheten och administrationen i kund- och intressegruppsuppgifterna i enlighet med EU:s dataskyddsförordning.

12 Ärendehantering och arkivering

12.1 Ärendehantering och analogiskt dokument-material

Statistikcentralens plan för bildande av arkiv av administrativa ärenden har uppdateras och fastställts när det gäller den allmänna administrationen, ekonomiadministrationen och personaladministrationen år 2012. År 2017 inleddes i ett projekt (AMS, TOS och processbeskrivningar) uppdatering av redan fastställda avsnitt till nivån för nuläget samt uppdatering av andra uppgiftsgrupper (statistikproduktion, informationshantering, kommunikation, marknadsföring och informationstjänst samt internationellt samarbete). En framställan om förvaringstid sändes till Riksarkivet våren 2018.

I oktober 2017 startade ett projekt för utveckling av ärende- och dokumenthantering vid Statistikcentralen (ADOHA). Syftet med projektet är att skapa klara processer och verksamhetssätt samt ta i bruk lättanvända elektroniska system inom ärende- och dokumenthanteringen. Projektet inom ärendehanteringen inleddes i slutet av år 2017 med ett projekt där man gör upp en datastyrningsplan för de väsentligaste delarna. I fortsättningsprojekten kommer man att skaffa och ta i bruk elektronisk ärendehantering.

Enligt statsrådets principbeslut av 21.6.2017 digitaliseras det analoga dokumentmaterialet som myndigheterna förfogar över fram till år 2030. Enligt beslutet arkiveras dokumentuppgiften bara i digitaliserad form, vilket innebär att den största delen av dokumenten i analog form förstörs efter digitaliseringen. Digitaliseringen sker i ett massdigitaliseringsprojekt där Riksarkivet är koordinator. Statistikcentralen kartlade år 2017 antalet manuella dokument som ämbetsverket förfogar över och kvaliteten på dem, som en del av Riksarkivets förutredningsprojekt.

12.2 Arkivering av statistikmaterial

En permanent arkivering av statistikmaterial säkerställer att information som ur samhälls synvinkel är värdefull bevaras och kan utnyttjas också i framtiden. Arkiveringsprocessen är smidigare tack vare arkiveringsanvisningarna för elektroniskt statistikmaterial och arkiveringsverktygen som förnyades år 2017.

Det är väsentligt att statistikmaterial som arkiveras beskrivs, med tanke på utnyttjandet av det, för att variablernas innehåll ska öppnas korrekt och så att man undviker feltolkningar. Kvaliteten på materialets och variablernas metadata innebär bl.a. att klassificeringarna kopplas till variablerna och att variablernas innehåll beskrivs täckande och lättförståeligt. Arkivpaketet innehåller egentligt material, beskrivningar av variabler och arkiveringsmetadata enligt normen SÄHKE2, ifall information om materialet innehåller uppgifter på enhetsnivå.

Varje statistikgren arkiverar en gång om året statistikmaterial som samlats in och bearbetats för publiceringen av statistiken och statistikmaterial på färdig observationsenhetsnivå. Om man producerar många statistikgrenar ur ett och samma omfattande material, arkiveras materialet bara en gång. Statistikavdelningen bestämmer själv när materialet är färdigt att arkiveras och anger måldatumet för arkiveringen i arkiveringsplanen.

Arkiveringsplanerna och hur de utfallit kan följas upp i realtid i rapporteringssystemet Portti. Utfallet av arkiveringen är år 2018 för första gången med som mätare i resultatenheternas resultatkort. Statistikcentralen skickade 2017 en framställan om förvaringstid till Riksarkivet för ett beslut om statistikmaterial som ska bevaras permanent i elektronisk form.

13 Utvecklingsåtgärder för år 2018

Utvecklingsarbetet fortsätter. År 2018 är Statistikcentralens utvecklingsobjekt med anknytning till dataskydd följande:

- Utveckling av helhetsarkitekturen
- Dataarkitekturen och logiska datalager
- Införande av pseudonymisering
- Utveckling av behörighetshantering
- Utveckling av ärende- och dokumenthantering
- Arkivering
- Verkställande av loggpolitiken och utveckling av loggövervakning
- Utvärderingar av dataskyddets effektivitet för den individbaserade statistikens del

Hur utvecklingsåtgärderna framskrider rapporteras i databokslutet för år 2018.