

F1 5.3

TK

MUISTIO

TILASTOKESKUS

Tekijä
Jarmo Asikainen
Vt. yliaktuaari

Pvm
27.06.1972

No
12.

YHTENÄISTETTYJEN TILASTO-
AINEISTOJEN JÄRJESTELMÄ

Sisällysluettelo

	Sivu
1. Johdanto	1
2. Yhtenäistettyjen tilastoaineistojen järjestel- män periaatteet	2
2.1. Yleistä	2
2.2. Eräitä yhtenäistettyjen tilastoaineistojen järjestelmään liittyviä käsitteitä	2
2.3. Tilastoyksikköjen määrittäminen ja identifiointi	4
2.4. Tietojen keruu	7
2.5. Tietojen varastointi	10
3. Yhtenäistettyjen tilastoaineistojen järjestelmän tavoitteiden mittaamisesta ja määrittämisestä ...	13
4. Kustannukset vs. tilastojen laatu	15
5. Tilastolliselle tutkimukselle koituva etu	16
5.1. Yleisiä etuja	16
5.2. Erityisiä etuja	18
6. Eräitä näkökohtia	21
7. Lähdeluettelo	23

1. Johdanto

Seuraavassa hahmotellaan ns. yhtenäistettyjen tilastoaineistojen järjestelmäksi nimitettyä tilastojen tuotantometodia, jonka periaatteet Svein Nordbotten¹ ensimmäisenä esitti julkisuudessa Pohjoismaiden tilastokokouksessa v. 1960.

Kirjoitelma koskettelee aihetta muusta kuin automaattisen tietojenkäsittelyn näkökulmasta. Myös puhtaasti käytännölliset ja moraalista laa-
tua olevat pulmat jäävät tarkastelun ulkopuolelle. Tarvittavat määritelmät ja käsitteet esitetään varsinaisen tekstin yhteydessä. Määritellään tässä kuitenkin eräät jatkuvasti esille tulevat termit. Täten tilastovirasto on se elin, joka pääosin vastaa yhteiskunnan tilastollisen informaation tuottamisesta. - Perustieto on standardoimaton havaintoarvo, jota ei ole käsitelty. Jos sitä on käsitelty (esim. luokiteltu, koodattu tai muuten standardoitu), käytetään siitä yksinkertaisesti nimitystä - tieto, joka siis kelpaa panokseksi tilastotuotantoon. - Informaatio on tosiasioita koskevien tietojen yhdistelmä, joka ilmaisee jollekin jotakin. Osa tästä on luonteeltaan tilastollista. - Tilasto puolestaan tarkoittaa tässä muistiossa joukkoilmiöitä kuvaavaa numeroaineistoa². - Tilastohaara käsitetään siksi elimeksi, joka vastaa oman alansa tilastotuotannosta. - Tilastotuotanto määritellään vuorostaan tässä tutkielmassa siksi perustietojen keräämis- ja käsittelyaktiviteetiksi, jonka tuloksena syntyy tilastotietoa.

Muistio perustuu kirjoittajan tilastotieteen laudaturtyöhön, josta tässä kirjoitelmassa esitetään osa. Saadut huomautukset ja kritiikki on otettu huomioon mahdollisuuksien mukaan.

¹ Nordbotten käytti nimitystä arkistotilastojärjestelmä. Myöhemmin on myös puhuttu rekistereihin perustuvasta tilastojärjestelmästä, vaikka tällä termillä ei ole aina tarkoitettu samaa asiaa. Kts. Nordbotten, Svein: Elektronmaskinene og statistikkens framtidige utforming, Nordisk statistisk skriftserie 7.

² Englanninkielinen sana "statistics" tarkoittaa kansainvälisessä kielenkäytössä lisäksi sitä toimintaa, jolla tätä aineistoa tuotetaan (tässä tutkielmassa käytetty termi - tilastojen tuotanto) tai niitä metodeja, joita tässä toiminnassa sovelletaan (esim. yhtenäistettyjen tilastoaineistojen järjestelmä) sekä sitä tiedettä (tilastotiede), joka kehittää sovellettavia metodeja.

Kts. Laurila, E. H: Tilastollisen päätoimiston tilasto-, tutkimus- ja selvitystöiden ryhmittelyä, Muistio 1970:1 (Tilastollinen päätoimisto), s. 2.

2. Yhtenäistettyjen tilastoaineistojen järjestelmän periaatteet

2.1. Yleistä

Yhtenäistettyjen tilastoaineistojen järjestelmästä käytetään seuraavassa lyhennystä YTJ. - Svein Nordbotten määrittelee sen systeemiksi, jossa tietoja kerätään jatkuvasti ja riippumatta tavanomaisista tilastojen tuotanto- ja havainnointiajankohdista ja jossa kerättyjä tietoja muokataan aina, kun tarve vaatii³. YTJ:lle on edelleen ominaista Nordbottenin mukaan mahdollisuus tuottaa jälkikäteen (jopa kymmenien vuosien kuluttua) tilastotietoa halutussa muodossa sekä hallinnollisten tietovarantojen hyväksikäyttö. Lisäksi se kohentaa suuresti tilastojen tuotantovalmiutta; tietojen päällekkäiskeräys ja tiedonantajien vaivaaminen minimoituu. Se myös mahdollistaa sellaisten samaa yksikköä koskevien tietojen yhdistelyn, jotka on kerätty eri lähteistä ja/tai eri aikoina.

Järjestelmän tavoitteena on yhteiskunnan tilastollisen informaatiotarpeen optimaalinen tyydyttäminen. Käytännössä systeemin periaatteiden toteuttamisen on mahdollistanut tietojenkäsittelyvälineiden nopea kehittyminen. Toisaalta sen syntyyn ovat osaltaan vaikuttaneet nykyisen tilastotuotannon puutteellisuudet, jotka ikäänkuin ovat antaneet pontta nykyistä paremman menetelmän kehittämiseksi.

2.2. Eräitä yhtenäistettyjen tilastoaineistojen järjestelmään liittyviä käsitteitä⁴.

Tieto voidaan standardoituna esittää vektorina:

$$d = (i, k, t, x),$$

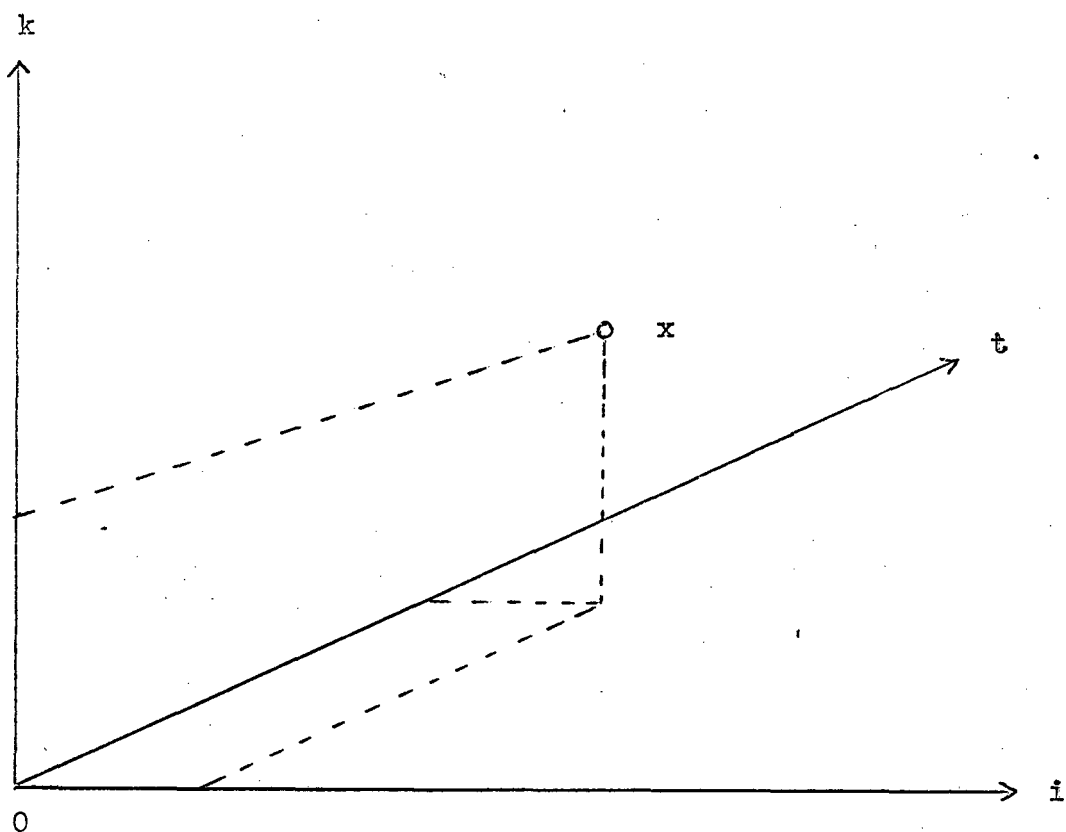
jossa i = sen tilastoyksikön identifiointi, johon tieto viittaa;

k = muuttujan nimi; x = muuttujan arvo ja t = ajankohta tai ajanjakso,

³ Nordbotten, Svein: A Statistical File System, Statistisk tidskrift 1966:2, s. 60.

⁴ Käsitteet on pääasiassa lainattu Nordbottenilta, joka näiden ja eräiden muiden avulla on rakentanut hyödykkeiden tuotantoterminologian käsitteillä operoivan mallin; kts.

johon tieto paikallistetaan ajan suhteen. Standardoitu tieto siis aina viittaa kuvattavaan tilastoyksikköön, joka on identifioitu. Muuttujan arvo x on löydettävissä (kun tieto on varastoitu myöhempää käyttöä varten) kuvion 1 esittämällä tavalla koordinaattiakselien i , k ja t avulla.



Kuvio 1.

Tietojen varastosta voidaan käyttää nimitystä tietopääoma, joka on jaettavissa-havainto- ja tulostieto- eli tilastopääomaan. Tietopääoman tuottavuus riippuu tietojen valmiusasteesta ja siitä, miten se tyydyttää YTJ:n tarpeet. Sen kasvattamista voidaan sanoa investoimiseksi, joka suoritetaan keräämällä tilastoyksiköjä koskevia tietoja

Nordbotten, Svein: Purposes, Problems and Ideas Related to Statistical File Systems, Invited Paper to the 36th Session of the International Statistical Institute in Sydney, 28.8.-8.9.1967; Nordbotten, Svein: On Statistical File Systems II, Statistisk tidskrift 1967:2.

tai lisäämällä valmiin tilaston määrää. Tilastopääoman avulla kyetään tuottamaan ainoastaan tilastollista informaatiota, kun taas havaintopääomasta voidaan tuottaa muutakin kuin tilastollista informaatiota.

2.3. Tilastoyksikköjen määrittelyminen ja identifiointi

Koska YTJ:ssä yksikköjä koskevia tietoja kerätään useista eri lähteistä ja/tai eri aikoina ja koska niiden muutoksia on kyettävä seuraamaan ajan ja ominaisuuksien suhteen, kukin tilastoyksikkö on kyettävä yksilöimään ja tunnistamaan (= identifioimaan) yksiselitteisesti. Tämä on eräs YTJ:n keskeisimpiä teknisiä edellytyksiä.

Tämä toimenpide ei ole ongelmaton, koska identifioinnille, jotta se toimisi tarkoituksenmukaisesti, asetetaan runsaasti vaatimuksia. Yleisperiaate on se, että on identifioitava sellaiset yksiköt, joista saatava tieto antaa mahdollisimman suuren hyödyn YTJ:lle. Vaikeutena on tällöin mm. se, että eri tilastotietojen käyttäjäpiirit haluavat tietoa erilaisista yksiköistä. Esimerkiksi taloustieteen tutkijat haluavat etupäässä informaatiota toimintayksiköistä, kun taas hallinnollisia viranomaisia kiinnostavat lähinnä juridisin perustein määritellyt yksiköt. Käytännön rajoituksena on toisaalta se, minkälaisista yksiköistä tietoja voidaan kerätä⁵. Identifioitavan yksikön valintaa rajoittaa myös se seikka, että yksikköä on kyettävä seuraamaan syntymästä häviämiseen saakka. Koska tämä aikaväli saattaa olla vuosikymmeniä, jona aikana yksikön monet tunnusmerkit voivat muuttua niin paljon, ettei sitä voida enää häviämishetkellä tunnistaa sillä syntymähetkellä olleen tunnusmerkistön mukaan samaksi yksiköksi, olisi oltava selvät säännöt siitä, milloin yksikkö katsotaan toiseksi siinä tapahtuneiden muutosten perusteella.

⁵Tilastoyksikköjen rajaamisen ongelmaa on lähemmin käsitelty artikkeleissa: Luther, Georg: The Acting Units in Statistics on the Sector of Enterprises, Reprint from: Kokonaistaloudellisia ongelmia, Vol. XXV; ja Strengell, Göran: Enheterna i person-, byggnads- och bostadsregistren, Rapport från nordiska seminariet i registerteknik, Hässelby slott, 18-22 mars, 1968.

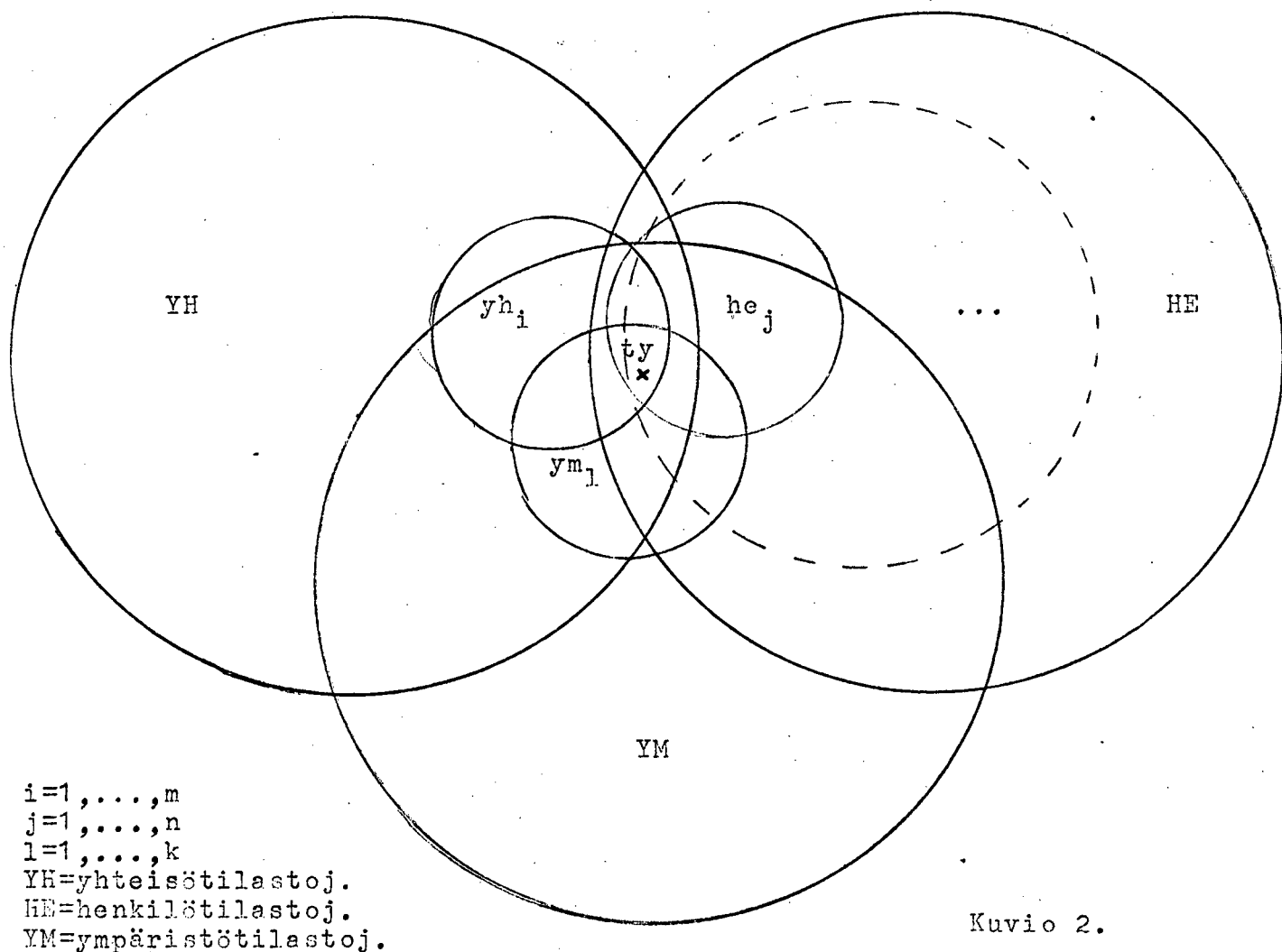
Jotta tilastotietojen käyttäjien hyvinkin erilaiset tarpeet tulisivat mahdollisimman suuressa määrin tyydytetyiksi, olisi tilastoyksiköksi valittava pienin mahdollinen fyysinen yksikkö hierarkisen järjestelmän yksityiskohtaisimmalta tasolta. Jos hierarkiseen jaotteluun ei voida mennä, niin sellainen yksikkö olisi valittava, joka takaisi suurimman kombinointimahdollisuuden. Yksiköitä rajattaessa on kuitenkin tärkeää, miten niitä koskevat tiedot voidaan pitää kohtuullisin kustannuksin ajan tasalla. Tämä seikka saattaa vaikuttaa siihen, että YTJ:ään ei voida ottaa mukaan kaikkia niitä yksiköitä, jotka tilastollisen tutkimuksen kannalta olisivat suotavia.

Seuraavassa sanotaan äskeisten periaatteiden mukaan tilastoyksikölle annettavaa yksilöivää ja tunnistavaa merkkiä identifiointitunnukseksi. Tällekin voidaan asettaa erinäisiä vaatimuksia. Hyvältä tunnuksesta pitäisi voida vaatia ainakin seuraavaa: 1) kullakin saman asteen yksiköllä (hierarkisessa systeemissä) on yksi ja vain yksi tunnus, 2) sama tunnus kuuluu yhdelle ja vain yhdelle saman asteen yksikölle ja 3) yksikön lakattua olemasta sen tunnusta ei anneta toiselle yksikölle. Näistä kolmesta vaatimuksesta voitaisiin käyttää nimitystä "identifioinnin minimivaatimukset".

Identifiointitunnuksella olisi lisäksi suotavaa olla seuraavat ominaisuudet. Sen tulisi olla mahdollisimman lyhyt, jotta i-akseli (kts. kuvio 1, s. 3) veisi mahdollisimman vähän tilaa tietoja varastoitaessa. Lisäksi sen tulisi olla neutraali, ts. se ei saisi sisältää informaatiota, ei ainakaan silloin, kun on olemassa informaatioasisällön muuttumisen mahdollisuus.

On tarkoituksenmukaista YTJ:n kannalta, että samasta yksiköstä samaan päätilastojärjestelmään kuuluvana käytetään samaa identifiointitunnusta kaikissa tilastohaaroissa ja YTJ:n osana olevissa hallinnollisissa viranomaisissa. Lisäksi edellytyksenä mahdollisimman korkeatasoisille, yhtäläisin laatuperustein tuotettaville tilastoille on yhteisten luokitusten, standardien ja määritelmien käyttö.

Joskus yksikkö ei näennäisesti täytä identifioinnin minimivaatimuk-
sia, koska se näyttää saavan useita tunnuksia. Tarkastellaan kuviota
2 (kuvio ei ole realistinen, vaan sen osoittama systeemi on voimak-
kaasti pelkistetty havainnollisuuden vuoksi). Siitä käy ilmi, että
mainittujen kolmen päätilastojärjestelmän kuvausalueet menevät osit-



tain päällekkäin. Kukin niistä sisältää alitilastojärjestelmiä ja
nämä puolestaan alitilastojärjestelmiä jne. kunnes päästään yksittäi-
seen tilastohaaraan ja lopulta yksittäiseen tilastoyksikköön (= ty).
Koska kuvien mukaan on olemassa ainakin yksi $ty_s \in he_j \subset \dots$ ja
 $ty_s \in ym_l \subset \dots$ ja $ty_s \in yh_j \subset \dots$ yht'aikaa, on ainakin näennäisesti

olemassa se mahdollisuus, että ainakin yksi tilastoyksikkö saa useamman kuin yhden identifiointitunnuksen. Näin eräässä mielessä onkin, mutta tällöin ko. fyysinen yksikkö on kussakin päätilastojärjestelmässä eri tilastoyksikkö⁶.

Kuviota tarkastelemalla voidaan vielä havaita, että kullakin yksiköllä täytyy olla viitetunnus, joka ilmaisee, mihin tilastojärjestelmään tai ylemmän tason yksikköön kukin yksikkö kuuluu. Hierarkisen järjestelmän ollessa kyseessä kunkin yksikön yhteydessä tarvitaan vain yksi viitetunnus, koska tarvittaessa kukin yksikkö voidaan löytää epäsuorasti toisten välityksellä.

Jos tässä kappaleessa hahmotellut vaatimukset ja ehdot on täytetty, silloin ei periaatteessa ole esteitä tietojen yhdistelylle tilastohaarojen tai -järjestelmien kesken tai tietojen siirtämiselle hallinnollisilta viranomaisilta tilastovirastolle.

Lennart Fastbom esittää eräässä SSDS:ää käsittelevässä artikkelissaan⁷ muuttujien aggregointimahdollisuutta ajan, paikan ja sekundäärisen yksikön suhteen. Tämä lienee eräs SSDS:n tehokkaaksi rakentamisen ehto. Jos näin on, niin tämä ei ole ongelma, jos äsken mainitut ehdot on tyydytetty. Tästä johtaen voitaisiin sanoa, että YTJ:n synty on edellytys tarkoituksenmukaisen SSDS:n luomiselle.

2.4. Tietojen keruu

Kuten aikaisemmin on mainittu, tietopääoman lisäämistä voidaan kutsua investoinniksi. Perustietojen keruu muodostaa siitä ehkä tärkeimmän

⁶ Esim. ihmistä kuvataan henkilötilastoissa pääasiassa luonnollisena henkilönä, kun taas elinkeinotilastoissa häntä pidetään elinkeinonharjoittajana juridisena henkilönä. Periaatteessa voitaisiin käyttää kustakin fyysisestä yksiköstä vain yhden päätilastojärjestelmän tunnusta. Käytännössä tällaisesta menettelystä aiheutuisi kuitenkin melkoisesti hankaluuksia.

⁷ Fastbom, Lennart: Some Aspects on the Construction of a System of Social and Demographic Statistics (SSDS), Statistisk tidskrift 1969:1, s. 12-15.

vaiheen. Toinen keino kasvattaa tietopääomaa on lisätä tulostietopääoman suuruutta muodostamalla perustiedoista tai valmiista tilastosta uutta, johdettua tilastoa. Tässä tyydytään tarkastelemaan investointia tietopääomaan siltä osin, kuin se ei koske tulostieto- eli tilastopääomaa.

Jotta tilastoyksikköjen vaivaaminen minimoituisi ja esim. päällekkäiskeräys häviäisi, suoritetaan YTJ:ssä keruu integroidusti siten, että kun kysytty tieto tulee YTJ:n johonkin osaan, se on kaikkien tilastohaarojen käytettävissä. Tämä merkitsee sitä, että kukin yksikkö antaa halutun tiedon vain kerran YTJ:ään. Tietojen yhdistely suoritetaan sitten identifiointi- ja viitetunnusten avulla⁸. Kuitenkin käytännössä lainsäädännöllisesti kehitettävä tietosuoja voi rajoittaa tietojen yhdistely- ja siirtomahdollisuuksia. Täten kaikkia haluttuja tietoja ei saada ehkä YTJ:n piiriin.

Jotta yllämainittu toiminta mahdollistuisi, kappaleen 2.3. identifiointia koskevien vaatimusten on oltava täytettyjä. Koska kuitenkin joidenkin tilastoyksikköjen väliset rajanvedot ovat jonkin verran harkinnanvaraisia, olisi myös seuraavien viiden ehdon oltava tyydytettyjä⁹:

- 1) saman asteen yksiköt on aina määriteltävä samalla tavalla,
- 2) kukin yksikkö viedään aina samaan luokkaan tilastohaarojen yhteisissä luokituksissa,
- 3) kaikki saman asteen yksiköt otetaan järjestelmässä huomioon,
- 4) saman asteen yksiköt peittävät yhdessä kyseisen kuvausalueen,
- 5) minkään toiminnan osa tai funktio ei sisälly yht'ainakaan kahteen saman asteen yksikköön.

⁸ Se mahdollisuus, että tilastovirasto itse on kerännyt jonkin tiedon voimatta tietosuojan vuoksi siirtää sitä hallinnollisten viranomaisien käyttöön ja jos nämä joutuvat sitten kysymään ko. tietoa itseään varten, ei muuta tilastoyksikköjen vaivaamisen minimoinnin periaatetta, koska tällöin ko. viranomaisia ei katsota YTJ:n osaksi.

⁹ Tämän tutkielman kirjoittaja on jonkin verran yleistänyt G. Lutherin esittämiä periaatteita; kts. Luther, Georg: Yritysrekisterin kehittämisen ja toiminnan suuntaviivoja, Tilastollinen päätoimisto 8.4.1969 (moniste).

Tunnusomainen YTJ:n piirre on myös tietojen kerääminen jatkuvana tietovirtana aina, kun tämä on mahdollista, jolloin resurssien käyttöaste pidetään mahdollisimman tasaisena. Tähän sisältyy myös tietoaistien jatkuva muokkaaminen. On kuitenkin huomattava, että kaikkia status-tietoja, so. yksikköjen tilatietoja, ei voida kerätä jatkuvana virtana. Toisaalta, jos kaikki halutut tapahtumat voidaan kerätä yhtäjaksoisesti, saadaan myös tilatiedot selville, sillä rajoittavathan yksikön tilaa molemmin puolin tapahtumat. Näin ollen ainakin teoriassa jatkuva keruutapa on YTJ:ssä vallitseva, kuten mainittiin.

Jatkuvan keruun seurauksena tilastojen tuotantoaika ilmeisesti lyhenisi ja tuotantovalmius kohentuisi vaikkakin joskus laadun kustannuksella. Laatu saattaisi heiketä mm. jatkuvaan keruuseen liittyvän, eräässä mielessä otantaan perustuvan tuotantosysteemin vuoksi (kts. alempana). Vastakkaiseen suuntaan vaikuttaa kuitenkin se seikka, että jatkuvassa keräyksessä voidaan havaita sellaisia tapahtumia, jotka muutoin kätkeytyisivät harvemmin kerättäviin tietoihin, kuten on asianlaita käytetäessä aggregaattitietoja. Joka tapauksessa etujen punninnassa on asetettava vastakkain toisaalta ekstrapolointi suurine epävarmuustekijöineen yhä nopeutuvassa yhteiskunnan kehityksessä ja luotettavampi tilastotieto, mutta usein moniin käyttötarkoituksiin vanhemmeena ja toisaalta taas tuore, joskin ehkä joskus hieman epävarma tilasto. Kuitenkin luotettava tilasto saataisiin julkaistuksi vähintään yhtä nopeasti kuin nykyisinkin.

Eräs Svein Nordbottenin ehdottama jatkuvan keruun tavallaan otantaan perustuva muoto olisi seuraavan sisältöinen¹⁰. Jaetaan kuvattava perusjoukko n :ään osajoukkoon, joista kukin ilmoittaisi tietonsa eri $n:n$ periodin kuluttua lukumäärältään n :ltä edelliseltä ajanjaksolta. Siten kunkin $n:n$ nen periodin lopussa olisi YTJ:ssä käytettävissä otosestimointia varten mitä tahansa aikaisempaa periodia varten. Samalla täydelliset tiedot olisivat käytettävissä kustakin sellaisesta ajanjaksosta, joka olisi vanhempi kuin edellinen n :ään osaan jaettu periodi.

¹⁰Tämän kirjoittaja on muuttanut yleiseksi Nordbottenin yhtä tapausta varten ehdottaman esimerkin; kts. Nordbotten, Svein: A Statistical File System, Statistisk tidskrift 1966:2, s. 101.

Tietojen keruu, siirto ja yhdistely on se YTJ:n lohko, jossa lienee vielä melkoisesti tutkimisen tarvetta. Systeemanalyysi ja verkkomenetelmät saattaisivat olla käyttökelpoisia välineitä. Ongelmana olisi tällöin mm. se, miten tiedot kerätään, siirretään ja yhdistellään erilaisia tarpeita varten tehokkaimmin ts. miten "tietojen kulkema matka" saataisiin minimoitua; edelleen, mitkä olisivat ne tietovarannot (rekisterit), joihin kaikki jonkin tilastojärjestelmän keskeiset tiedot varastoitaisiin ja mihin vain osa. Lisäksi olisi selvitettävä koko YTJ:ää ajatellen, mitä tietoja yleensä on syytä kerätä; onko mahdollista kenties rakentaa jonkinlainen palautemekanismi (feed-back), joka ilmoittaisi kulloisenkin YTJ:n ulkopuolelta tulevan tilastotiedontarpeen muutoksen. - Toisaalta lienee hyvin vähän jos ollenkaan tutkittu sitä, minkälaista yhteiskunnan tilastollisen informaation tulisi olla, jotta se tyydyttäisi riittävässä määrin nykyisen näkyvän ja piilokäytön.

2.5. Tietojen varastointi

Eräs vaihe investoinnissa tietopääomaan on kerättyjen tietojen varastointi (= arkistointi), joka YTJ:ssä suoritetaan viemällä tiedot rekistereihin. Tavallaan rekisteri on keino siirtyä kuviossa 1 (s. 3) i-akselia pitkin oikeaan paikkaan¹¹. Tarpeen vaatiessa muutkin akselit voivat tulla kysymykseen ainakin etsityn tiedon likimääräisen paikan määrittämisessä.

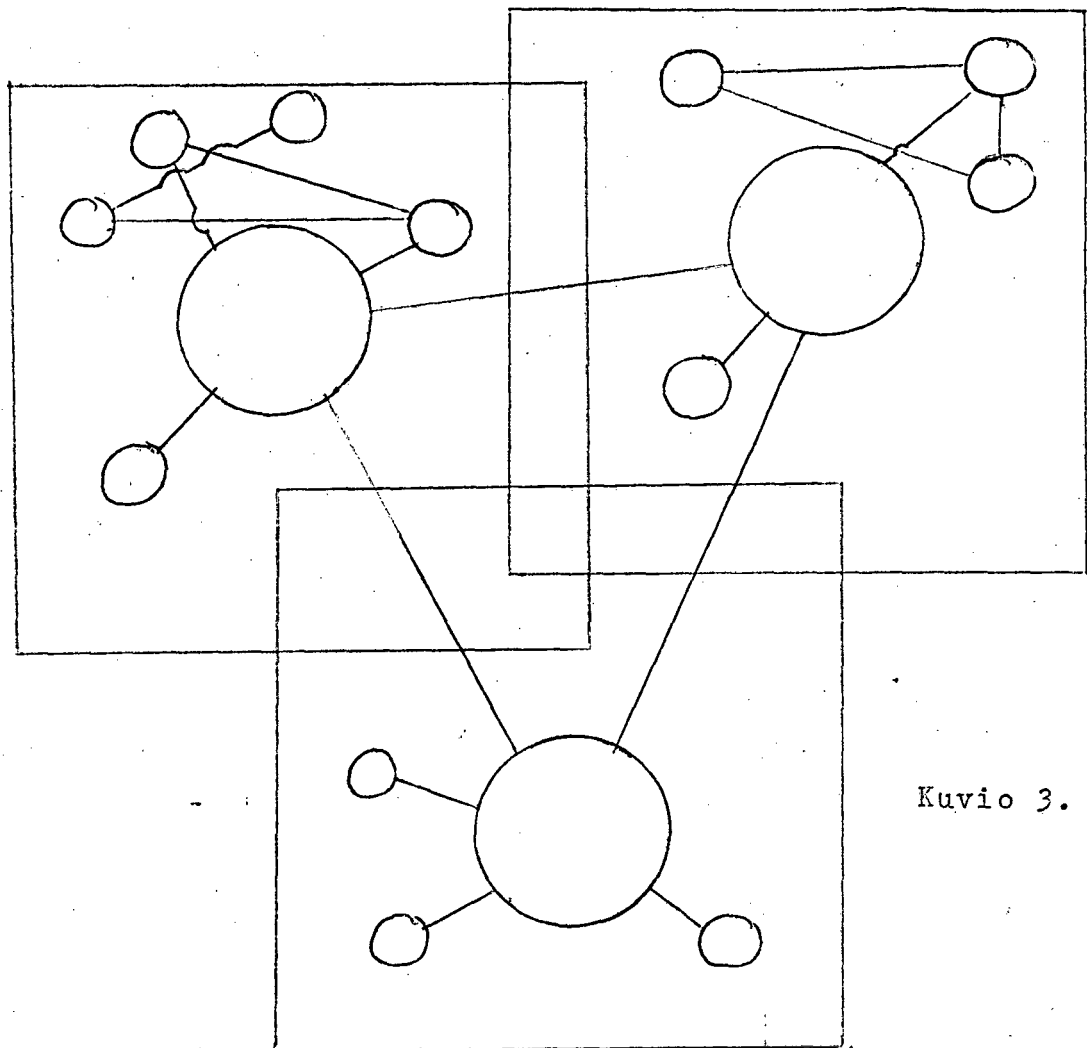
Ongelmanasettelu voidaan periaatteessa suorittaa samoin kuin yleensäkin minkä tahansa tiedon varastoinnissa (esim. kirjastot). Esimerkiksi voitaisiin asettaa seuraavat ei ehkä toisensa täydellisesti pois sulkevat ceteris paribus-vaatimukset¹²: 1) varastoidaan tiedot pienimpään mahdolliseen tilaan, 2) minimoidaan varastoitava tietomäärä, 3) minimoidaan tietojen siirto- ja yhdistelytarve. - Koska YTJ:n eräs tarkoitus on lyhentää tilastojen tuotantoaikaa, olisi myös tarkoituksenmukai-

¹¹ Käytännössä rekisteri on luettelo, joka voi olla millä tahansa tietovälineellä.

¹² Arvas, Christer: Databeskrivning och filorganisation, Rapport från nordiska seminarier i registerteknik, Hässelby slott 18-22 mars 1968.

nen vaihtoehto minimoida tietojen vapauttaminen arkistoista ajan ja kustannusten suhteen. - Koska kuitenkin monet vaatimukset olisivat ainakin osittain toistensa kanssa kilpailevia ja koska lisäksi olisi optimoitava tietty tavoitefunktio rajoituksena joukko niukkoja resursseja (kts. myös kpl 3), ongelman teoreettinen ratkaisu lienee parhaiten löydettävissä erilaisilla matemaattisilla ohjelmointiteknikoilla (esim. dynaamisen ohjelmoinnin avulla).

Tarkastellaan kuviota 3, joka kaaviomaisesti esittää erästä mahdollista rekisterien keskinäistä suhdetta.



Kuvio 3.

Olkoon siinä kukin suorakaiteen rajoittama alue tilastohaarojen muodostama tarkoituksenmukainen kokonaisuus, esim. tilastollinen pääkuvausjärjestelmä. Tällöin kussakin pääjärjestelmässä voisi olla yksi ns. keskusrekisteri ja pienet ympyrät ns. alirekistereitä, joiden suhteen voitaisiin noudattaa hierarkista jakoperiaatetta aina pienempiin rekistereihin (ei näkyvissä kuviossa). Näiden kuvausalue voisi olla sama kuin jonkun tilastohaaran tai sen osan. Rekistereitä yhdistävät viivat kuvaavat yhteyksiä, jotka ovat mahdollisia identifiointi- ja linkkitunnusten avulla. Se miten YTJ:ssä nämä yhteydet järjestetään, voisi kuulua osana edellä mainittuun optimointitehtävään.

Svein Nordbotten esitti jo v. 1960¹³, että yleensä rekisteri jakautuisi kolmeen YTJ:n tavoitteiden kannalta tarkoituksenmukaiseen osarekisteriin: varsinaiseen identifiointitunnusrekisteriin, historia-rekisteriin ja aktiivirekisteriin. Näistä ensimmäinen sisältäisi mainitun identifioinnin lisäksi yksikön maantieteellisen paikantamisen (so. postiosoite, karttakoordinaatit); toinen koostuisi tilastoyksiköjä koskevista historiallisista muutostiedoista (tai vaihtoehtoisesti kaikista aikaisemmista tiloista; historiarekisterin merkityksestä keskustellaan lähemmin kappaleessa 5.2.); viimeisin sisältäisi yksiköjä koskevat tuoreimmat tiedot. - Keskusrekisterien tehtävänä voisi olla mm. edellisessä kappaleessa mainittujen "Lutherin viiden ehdon" täyttäminen.

Joskus YTJ:n käyttöönoton jälkeen tulevaisuudessa saattaisi tulla harkittavaksi kuvion 1 (s. 3) 0-pisteen siirto pitkin aika-akselia, so. vanhentuneiden tietojen poisto rekistereistä, koska käytännössä ajan kuluessa tilastojen tuotantoperusteet muuttuvat yhteiskunnan kehittymisen myötä siinä määrin, että toisistaan ajallisesti kaukana olevien tietojen vertaaminen vaikeutuu ellei suorastaan tule mahdottomaksi. Olkoot esimerkiksi t , $t-1$, ..., $t-n$ (kuviossa 1 $t-n = 0$ -piste) sellaisia ajankohtia, jolloin on otettu käyttöön uusi luokitus. On ilmeistä, että käytettäessä minkä tahansa vierekkäisen ajankohdan $t-i$ ja $t-i-1$

¹³Nordbotten, Svein: ma. Elektronmaskinene ... s. 137.

luokituksia kartoitettaessa yksikköjä (vastaa funktioteorian mapping-termiä) jostain luokituksen $t-i$ luokasta luokituksen $t-i-1$ vastaavaan luokkaan, voidaan käytännössä saada riittävän suuri tarkkuus monia tarkoituksia varten, ellei sitten luokitusten rakenne ole perinpohjin muuttunut. Lähes mahdoton on keskinäinen vertailu kuitenkin silloin, kun käytetään luokituksia t ja $t-n$, johtuen mm. siitä, että samasta yksiköstä on tuskin ajankohtana $t-n$ käytettävissä samoja tietoja kuin ajankohtana t . Tämä on sitä ilmeisempää, mitä suurempi n on.

Tietojen siirto, yhdistely ja varastointi ovat käytännössä ATK-tekniikkaa. Tämän nopea kehittyminen onkin tehnyt - kuten mainittiin - YTJ:n periaatteiden soveltamisen käytännössä mahdolliseksi. Tulevaisuuden tekniikat näyttävät vain lisäävän näitä mahdollisuuksia¹⁴.

3. Yhtenäistettyjen tilastoaineistojen järjestelmän tavoitteiden mittaamisesta ja määrittämisestä

Jo aiemmin on todettu, että YTJ:n tarkoituksena on tyydyttää yhteiskunnan tilastollista informaatiotarvetta mahdollisimman hyvin. Tämä tavoitteenasettelu sisältää ongelman, koska yhteiskunnan tilastollisen informaation tarvetta ei ole juuri ollenkaan tutkittu. Ehdotuksia on tosin tehty, mutta nämä eivät ole päässeet toteutumisasteelle¹⁵. Owe Salomonsson¹⁶ on sitä mieltä, että informaatiosysteemin arvo mitataan sen tuottaman relevantin informaation avulla. Tätä voidaan pitää lähtökohtana, mutta mittaamisongelmaan Salomonsson ei puutu. Sisältäähän käsite "relevantti" subjektiivisen varauksen, ts. YTJ:n tuottama etu on mitattavissa tuotetun tiedon käyttäjän oman arvostuksen mukaisesti. Nykyisin menetelmin hyötyä ei kyetä vertaamaan ihmisten kesken yksiselitteisesti, koska nollapiste on ensin määriteltävä ja sitten löydettävä sopiva yhteismitta. Nollapisteen paikasta voidaan

¹⁴Biblioteksarbete och automatisk databehandling, Kommittén för ADB i forskningsbibliotek, Stockholm 1969, s. 24

¹⁵Dalenius, T.E: Official Statistics and their Uses, Review of the International Statistical Institute, Vol. 36, No 2, s. 121-140; sekä Valtion tilastotoimen kehittämisojelma, Liite 1, s. 14-16.

¹⁶Salomonsson, Owe: Offentliga data, Ekonomi och Samhälle, Stockholm 1971, s. 34.

sopia, mutta tilastojen ollessa kyseessä ei yleisin arvon mitta -
raha - voi toimia tilastojen tuottamisessa resurssien kohdentajana.

Edellistä taustaa vasten esitetään Svein Nordbottenin hyvin yksinker-
tainen malli tilastollisen informaatiopolitiikan ohjeeksi¹⁷: Olkoon
I = tilastollisen informaation tarjonta per aikayksikkö, S = tilasto-
eli tulostietopääoma, M = ilmaisee tilastollisen informaation laadun
ja S:n hyväksikäytön asteen, joten

$$(1) \quad I = I (M, S).$$

Olkoon D = tietopääoma, V ja U = ilmaisevat D:n ja S:n hyväksikäytön
asteen, t = aika, joten tilastopääoman muutos voidaan esittää funktio-
naalisessa muodossa seuraavasti:

$$(2) \quad dS/dt = S (V, U, S, D).$$

Olkoon lisäksi dD/dt = investointi tietopääomaan,
C = kustannukset per aikayksikkö, joten:

$$(3) \quad C = C (I, dS/dt, dD/dt, S, D).$$

Nyt voidaan YTJ:n tarkoituksena sanoa olevan löytää ne aikafunktiot
 dD/dt , V, U ja M, jotka tyydyttävät yhtälöt (1) - (3), kun S ja D
ovat annettuja, ja jotka maksimoivat funktion:

$$(4) \quad W = W \left(\begin{bmatrix} I \\ 0 \end{bmatrix}^T, \begin{bmatrix} C \\ 0 \end{bmatrix}^T \right),$$

jossa siis on T periodia ja W voisi ilmaista yhteiskunnalle YTJ:stä
koituvan hyödyn¹⁸, joka siis on maksimoitava.

¹⁷Nordbotten, Svein: ma. On Statistical ... s. 115-117.

¹⁸Nordbotten ei itse yhdistänyt tähän malliinsa hyötykäsitettä. -
Mallia voitaisiin laajentaa ottamalla mukaan lisää muuttujia ja
yhtälöitä sekä rakentamalla yhtälöitä, joiden muuttujavektorien
kertoimina olisivat aktiviteettivektorit ja joilla olisi erilai-
sia resurssirajoituksia. Näin päästäisiin kappaleessa 2.5. mai-
nittuun matemaattiseen ohjelmointitehtävään.

Nykyinen tilastojen tuotantojärjestelmä ei tunne termejä S ja D siten, kuin ne on mallissa esitetty, vaikkakin ne esiintyvät kertaluonteisesti kunkin tietojen keruun jälkeen. YTJ:ssä puolestaan S ja D ovat riippuvaisia myös aikaisempien periodien kehityksestä, joten voidaan sanoa, että tietopääoman arvo kasvaa ajan myötä, ainakin tiettyyn rajaan asti.

4. Kustannukset vs. tilastojen laatu

Tilastojen laadun voidaan sanoa olevan sitä parempi, mitä objektiivisemmän kuvan ne antavat kuvaamistaan ilmiöistä. Siirryttäessä YTJ:n käyttöön on odotettavissa yleensä laatutason kohoaminen, mutta on vaikea sanoa, kuinka paljon. Kustannukset puolestaan voitaneen mitata rahamitalla. Kuitenkin on hankala mitata tietojen keruun tilastoyksiköille aiheuttamia kustannuksia. Sen sijaan tilastoviraston keru-, muokkaus- ja varastointikustannukset ovat helpommin ilmaistavissa.

Vaikka suoran keruun kustannukset päällekkäiskeruun häviämisen vuoksi laskisivat, vaikuttaa kuitenkin jatkuva keruutapa päinvastaiseen suuntaan. Lisäksi laajojen hallinnollisten aineistojen käsittely, tilastoyksikköjen täsmentynyt identifiointi, rekisterien ajantasallapito ja entistä korkealaatuisempien resurssien tarve kohottaisivat tuotantokustannuksia nykyiseen verrattuna. Olisi odotettavissa, että kokonaiskustannukset nousisivat, vaikkakin on ilmeistä, että keskimääräiset kustannukset laskisivat.

On ilmeistä, että kerättävien perustietojen tarkkuutta (ja samalla tuotettavan tilaston laatua) voidaan lisätä lähes rajattomasti käyttämällä entistä enemmän ja entistä korkeatasoisempia resursseja yksittäisen perustiedon hankintaan. Tällöin on odotettavissa, että kustannusten kohotessa saavutetaan raja, jolloin perustietojen tarkkuuden kasvusta saatu rajahyöty (tilastojen laadulla mitattuna) on yhtä suuri kuin rajakustannukset, josta eteenpäin resurssien määrää ja ta-soa ei enää kannata lisätä.

Koko yhteiskunnan kustannuksia ajatellen olisi kustannus-hyötyanalyysi käyttökelpoinen menetelmä. Tämän käyttö vaatisi kuitenkin tämän ja edellisen kappaleen avoimien kohtien ratkaisua. Eikä itse metodikaan ole aivan ongelmaton.

Edellisessä kappaleessa esitetyn Nordbottenin mallin yhtälö (3) sisältää sekä liiketaloudellisen että kansantaloudellisen tarkasteltavan. Kysymys on vain siitä, mitä merkityksiä mallin muuttujille annetaan. Mallia voidaan laajentaa ja jokainen muuttuja ilmaista muiden äskeisessä mallissa mainitsemattomien muuttujien funktiona.

5. Tilastolliselle tutkimukselle koituva etu

5.1. Yleisiä etuja

Jo aikaisemmin on useassa yhteydessä mainittu YTJ:n käyttöönottamisesta seuraavia etuisuuksia. Tässä tarkastellaan vielä eräitä tärkeimpiä.

On oletettavaa, että on sellaisia tietoja, joita tiedonantajat eivät mielellään ilmoita tilastovirastolle, mutta jotka hallinnolliset viranomaiset saavat käyttöönsä lähes kaikkien yksikköjen osalta. Näin siksi, että ko. viranomaisilla on mahdollisuus käyttää positiivista tai negatiivista palkitsemista, kun taas yleensä tilastovirastolla on käytettävissä negatiivinen palkitseminen eikä sekään täydellisesti. Näin olisi odotettavissa vastausprosentin kasvu eräiden tietojen osalta ja sitä tietä tilaston laadun kohentuminen. Brita Larson¹⁹ mainitsee laatutasoa parantavana tekijänä sen, että hallinnollisilta viranomaisilta voidaan perinteisesti odottaa suurta tarkkuutta tietojen käsittelyssä. Usein kuitenkin käsitteet poikkeavat siitä, mikä tilastotoimen tarkoituksiin olisi soveliainta. - Lisäksi YTJ:ssä perustietojen looginen kontrollimahdollisuus kasvaa keskusrekisteritasolla, koska kerätyt tiedot voidaan asettaa vertailtaviksi vertikaalisesti

¹⁹Larson, Brita: Identifikation av statistisk information, Del. 1, Statistisk tidskrift 1969:2, s. 85.

ajan suhteen ja horisontaalisesti eri tilastohaarojen kesken²⁰. Samaan suuntaan vaikuttaa myös mahdollisuus rekisteröidä tiedot halutulla tarkkuudella ajan suhteen, jolloin ne voidaan entistä paremmin viedä oikeaan periodiin kuuluviksi.

Kuitenkin on eräitä laatuun negatiivisesti vaikuttavia puolia, mutta ei ole uskottavaa, että niiden vaikutus kumoaisi aikaisemmin mainittujen seikkojen aikaansaaman positiivisen lopputuloksen. Eräs on se, että yksikköjako saattaa olla erilainen hallinnollisella kuin tilastollisella sektorilla. Tämän merkitys ei liene niin suuri kuin ehkä yleensä luullaan. Ensiksikin: tähänkin saakka tilastovirasto on toiminut yhteistyössä hallinnollisen sektorin kanssa ja tällöin on usein käytetty hallinnollisille viranomaisille sopivaa yksikköjakoa. Toiseksi: julkinen sektori lienee tärkein tilastojen käyttäjä. Sen etujen mukaista on usein käyttää hallinnollista yksikköjakoa. Kolmanneksi: yksiköt voitaneen eräissä tapauksissa rajata uudelleen tilastovirastolle sopivammiksi, jos on tarpeen. - Haittapuolena väitetään myös olevan sen, että tilastovirasto ei voi kontrolloida kaikkia perustietoja. Tämäkään ei liene merkitykseltään kovin suuri. Ensiksikin: tilastovirasto ei nykyisenkään järjestelmän vallitessa voi tarkistaa täydellisesti kaikkia niitä tietoja, jotka nykyisinkin saadaan hallinnolliselta sektorilta. Toiseksi: YTJ:ssä edellytetään nimenomaan yhteisten standardien ja luokitus-ten käyttöä koko YTJ:n piirissä. Kolmanneksi: on myös hallinnollisen sektorin etujen mukaista kontrolloida sille tulevat tiedot mahdollisimman hyvin.

Toinen tilaston käyttökelpoisuuden mitta on sen laajuus. Tässäkin suhteessa YTJ on parempi kuin nykyinen järjestelmä. On nimittäin runsaasti sellaisia tietoja, jotka ovat hyvin vaikeasti kerättävissä eri lähteistä. YTJ:ssä siirto- ja yhdistelymahdollisuuksien vuoksi nykyisin hajallaan olevat tiedot on saatavissa yhdestä paikasta. Reaalista laajentumista merkitsee mahdollisuus, että YTJ:n piiriin saadaan sellaisia tietoja, jotka aikaisemmin eivät ole sisältyneet panoksina tilastototuotantoon esimerkiksi kustannussyistä. Kaiken lisäksi YTJ:ään voi-

²⁰ Ohlsson, Ingvar: EDB-systemets betydelse för statistikproduktionen och statistisk analys, Nordisk statistisk skriftserie 12, s. 27.

daan joustavasti liittää uusia syntyviä tiedostoja. Göran Svanfelt²¹ korostaa mahdollisuutta kytkeä YTJ:ään myös yksityisiä tiedostoja, tilastollisia tutkimuksia ym. informaatiota.

Kolmas tilaston käyttökelpoisuuden mitta on tilaston sisällys, so. min-kälaisia tauluja ym. joistakin ilmiöistä laaditaan. On ilmeistä, että YTJ:ää sovellettaessa julkaistaan tilastotietoja samoin kuin nykyisin-kin, mutta lisäksi on mahdollisuus tuottaa erikoistilastoja käyttäjien haluamalla tavalla, kunhan vain salassapitosäännöksiä ei rikota.

Neljänneksi mitaksi voidaan mainita tilaston jatkuvuus ajan ja paikan suhteen. YTJ tyydyttää tämänkin vaatimuksen. Pikemminkin tilastojen vertailukelpoisuutta parantaa yhteisten standardien, luokitusten ja määritelmien käyttöönotto. Kuitenkin niin nykyisessä järjestelmässä kuin YTJ:ssäkin yhteiskunnan kehittyminen pakottaa näiden muuttami-seen, mutta tähän on YTJ:stä täysin riippumatonta.

Kaiken kaikkiaan voitaneen sanoa, että YTJ vaikuttaa positiiviseen suuntaan ns. yleisten etujen kannalta katsottuna.

5.2. Erityisiä etuja

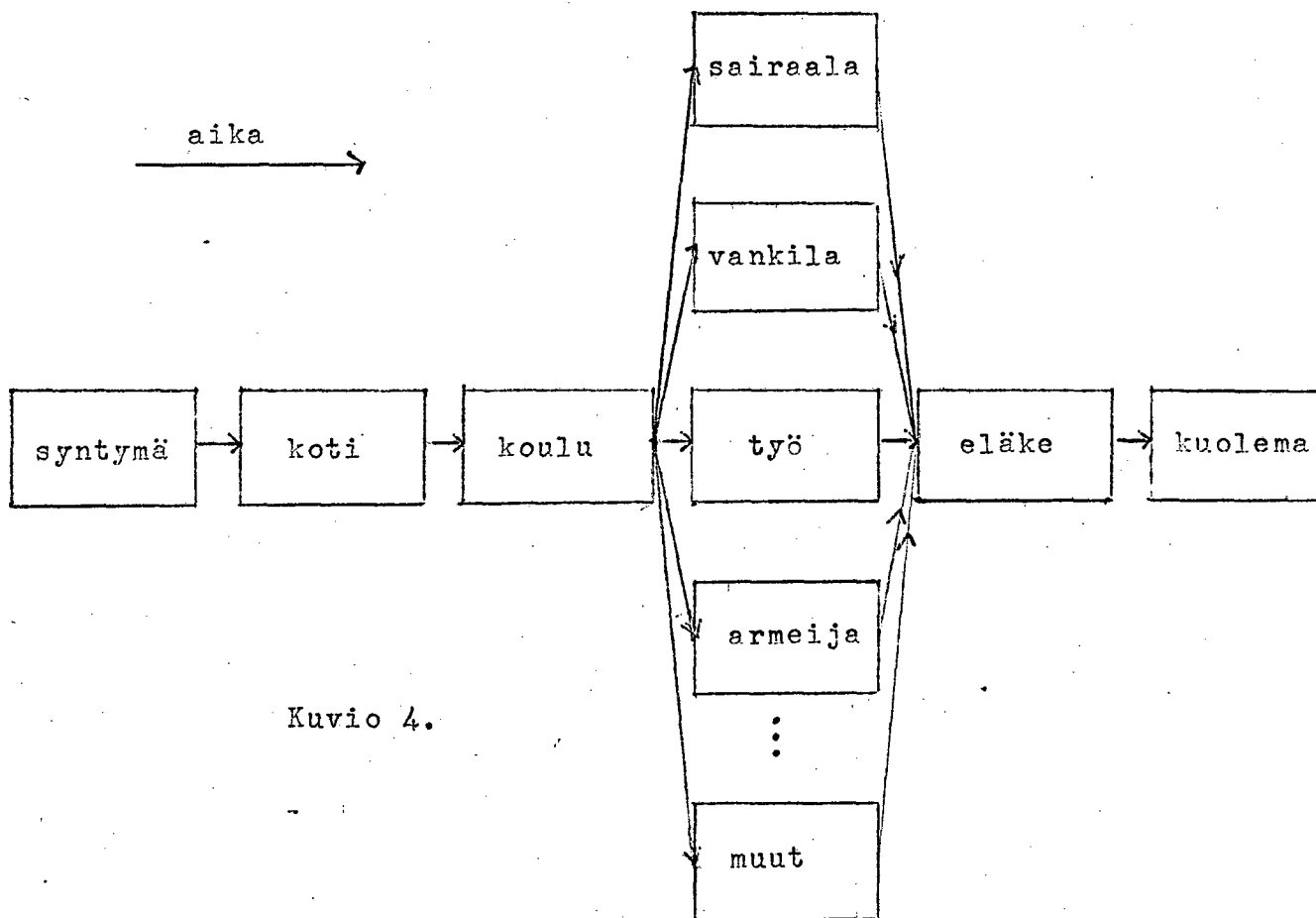
Suurimmalle osalle edellisessä kappaleessa mainittuja kriteerejä on yh-teistä se, että tilastojen käyttökelpoisuutta ilmaistuna niiden avulla voidaan parantaa nykyisenkin tuotantojärjestelmän puitteissa. Tässä puolestaan tarkastellaan niitä etuisuuksia, joita ei liene mahdollista nykyisin menetelminsaavuttaa ilman kohtuuttomia ponnistuksia.

Aiemmin on sivuttu tietojen identifiointia. Laajemmassa merkityksessä käsitettynä voidaan puhua identifioinnista ajan, paikan ym. tekijöiden suhteen. Tehokas keino liittää tiedot paikkaan on sisällyttää niihin myös karttakoordinaatit. Tämä lienee samalla yksinkertainen tapa saada osa ympäristötilastollisen kuvausjärjestelmän aluetta integroitua YTJ:ään.

²¹ Svanfelt, Göran: Korreferat til foredrag af Johan-Kristian Tønder: utdanningsstatistikk i relasjon til eit socio-demografisk system, Rapport fra Nordisk Symposium vedrørende et integreret befolknings- og socialstatistisk system, Vedbaek 8. - 10. april 1970.

Täten tilastolliselle tutkimukselle avautuu Owe Salomonssonin²² mukaan aivan uusia mahdollisuuksia: saadaanhan siten selville sellaiset muut-
tajat kuin havaintojen keskinäinen etäisyys, havaintojen keskinäinen
suunta ja havaintojen sijaintien rajoittama alue. Erityisesti hyötyvinä
tutkimusaloina olisivat tällöin aluetutkimus ja sosiaalitieteet. - Kai-
ken lisäksi tilastotietoja olisi teoriassa mahdollisuus saada kaikenlai-
silta halutuilta alueilta. Näin siksi, että karttakoordinaattien määrit-
tämä piste on pienin mahdollinen alueyksikkö. Pisteiden yhdistelymahdol-
lisuudethan ovat rajattomat.

Historiarekisterien sisällöllä tulisi olemaan aivan erityinen merkitys



Kuvio 4.

²² Salomonsson, Owe: Koordinater i register och dataarkiv, Statistisk tidskrift 1968:4, s. 313-321.

tutkimustyölle. Tarkastellaan kuviota 4²³, joka havainnollistaa esimerkiksi tilastoyksikön (tässä fyysisen henkilön) aktiviteettia yhteiskunnassa. Vastaavat tiedot olisivat kustakin yksiköstä. Piirrokselta voidaan nähdä, että YTJ:ssä pitkän ajan tutkimukset, esim. kohorttianalyysit, tulevat mahdollisiksi. Tilastoyksikköjen aktiviteettia kyetään seuraamaan yksikköjen syntymästä häviämiseen saakka. Tämä luo suotuisat edellytykset kausaalisten yhteyksien selvittämiseksi yksikkötasolla varsinkin, kun ympäristön (niin yhteiskunnallisen kuin maantieteellisenkin) vaikutukset saadaan tarkasteluun mukaan. Lisäksi, jos lokero "muut" sisältäisi vaikkapa yksikön jatkokoulutusta koskevia tietoja, olisi käytettävissä sitä koskeva koulutukseen viittaava kumulatiivinen tietosarja. - Kun kustakin yksiköstä saadaan kuviossa osoitetun tapaisia tietoja, on Klas Wallbergin²⁴ mukaan mahdollista saada esimerkiksi empiirisiä siirtymätodennäköisyyksiä stokastisten prosessien avulla suoritettavaa tutkimusta varten.

Edellisen lisäksi YTJ tarjoaa tietoja käyttäytymisyksiköistä, mikä seikka on omiaan lisäämään mallinrakentajan mahdollisuuksia. YTJ mahdollistaa vielä tietojen saamisen joukkojen sisäisestä hajonnasta eikä yksinomaan niiden välisestä. Aggregaattitietojen avulla tähän ei päästä.

Rekistereitä voidaan käyttää otostutkimusten pohjana. On myös mahdollisuus käyttää keskusrekistereitä silloin, kun perusjoukko ulottuu laajemmalle kuin erikoisrekisterien tietoalue.

Tässä kappaleessa on esimerkin vuoksi käsitelty lähinnä henkilötilastoja apunaan käytettävälle tutkimukselle tulevia etuisuuksia ja onkin ilmeistä, että tämä ala hyötyy eniten YTJ:n käyttöönottamisesta. Mutta vastaavanlainen hyöty tulee myös muun yhteiskunnallisen tutkimuksen osaksi.

²³ Kuvio lainattu: Hjerpe, Reino - Niitamo, Olavi: Some Aspects of Developing the System of Social Indicators in Finland, International Association for Research in Income and Wealth, Twelfth General Conference, Ronneby Brunn, Sweden, Aug. 30 - Sept. 4, 1971 (esitelmä).

²⁴ Näistä asioista on keskusteltu vilkkaasti SSDS:ää käsittelevässä kirjallisuudessa ja kansainvälisissä kokouksissa. Kts. esim. Wallberg, Klas: Kravet på statistik som underlag för ett socio-demografiskt system, Statistisk tidskrift 1971:3.

6. Eräitä näkökohtia

Tässä tutkielmassa kirjoittaja on esittänyt mielipiteitä lähinnä ajatellen Suomen ja muiden Pohjoismaiden oloja, mutta osa niistä lienee kuitenkin universaalisia. On tarkasteltu paitsi itse menetelmää niin myös sen soveltamisen tekniikkaa tilastotuotannon näkökulmasta, jotta nähtäisiin mihin YTJ todella perustuu, sillä menetelmään on periaatteessa hyvin yksinkertainen. Lisäksi on pohdittu YTJ:n tilastojen käyttäjille suomia hyötypuolia. - On mainittu, että YTJ palvelee yhteiskunnan tilastollisen informaation tuotantoa. Mutta on myös ajateltavissa, että menetelmä olisi sovellettavissa muunkinlaisen informaation tuottamiseen. Kysymykseen tulisivat sellaiset tapaukset, joissa esiintyvät jatkuvasti samat tilastoyksiköt ja joita koskevaa tietoa joudutaan hankkimaan eri aikoina ja/tai eri lähteistä. Kysymyksen ei edes tarvitse olla tilastojen tuottamisesta. Mahdollinen käyttöalue voisi yritystasolla olla asiakastiedostojen hoito. Myös monet elinkeinoelämän keskusjärjestöt voisivat ottaa menetelmän käyttöönsä ja mikä vielä edullisempaa molemmiin puolin, ne voisivat integroitua valtakunnan YTJ:ään.

On myös ajateltavissa, että YTJ:n periaatteita voitaisiin käyttää integroidusti kansainvälisellä tasolla. Täten voitaisiin paremmin saada selvyyttä esim. sellaisista yrityksistä, joiden toiminnan voidaan sanoa olevan niin kansainvälistä, että kokonaiskuvan saaminen niistä muutoin olisi kovin työlästä. Lisäksi henkilötilastojen puolella voitaisiin paremmin tutkia esim. maiden välistä muuttoliikettä. Tämä ei liene teknisesti mahdotonta toteuttaa, koska yhteisten standardien, luokitusten ja määritelmien käyttö on yleistymässä YK:n ja sen alajärjestöjen sekä muiden kansainvälisten elinten toimesta. On kuitenkin ilmeistä, että tämänlaatuinen yhteistoiminta voitaisiin eri maissa vallitsevien institutionaalisten erojen vuoksi ajatella toteutettavaksi ainakin aluksi vain Pohjoismaiden kesken, vaikka tälläkin tasolla on vielä runsaasti käsiteanalyysin tarvetta. Mittavuu- tensa vuoksi olisi käytännöllisintä saada ensin menetelmä sovellutusasteelle kussakin maassa erikseen, mutta kuitenkin yhteisten periaatteiden mukaisesti toimien.

Järjestelmään liittyy vielä useita ongelmia. Ratkaistavina ovat mm. miten tietojen joutuminen asiattomien käsiin kyetään estämään; siirretäänkö tietoja vain hallinnollisilta viranomaisilta tilastovirastolle muttei päinvastoin; kuinka paljon yhteiskunnan on syytä uhrata resursseja siihen, että se saa itsestään tarkoituksenmukaista tietoa; kuinka tiedotustoiminta järjestetään; kuinka erilaiset tekniset yksityiskohdat ratkaistaan jne. Menetelmään liittyvät perusongelmat on jo ratkaistu. Näyttääkin siltä, että lähiaikoina Suomessakin on edullista ottaa YTJ käyttöön, sillä niin ilmeisiä ovat sen edut - huolimatta vaikeasta hyödyn mittaamisongelmasta ja tilaston tuotantokustannusten huomattavasta kohoamisesta. Osittain YTJ:ää menetelmänä sovelletaankin jo joidenkin tilastohaarojen puitteissa, muttei niissäkään vielä täydellisesti.

7. Lähdeluettelo

- Arvas, Christer: Databeskrivning och filorganisation, Rapport från nordiska seminariet i registerteknik, Hässelby slott 18-22 mars 1968.
- Biblioteksarbete och automatisk databehandling, Kommittén för ADB i forskningsbibliotek, Stockholm 1969.
- Bjerke, Kjeld: An Integrated Social and Demographic Statistical System, Statistisk tidskrift 1970:3.
- Dalenius, T.E: Official Statistics and their Uses, Review of the International Statistical Institute, Vol. 36, No 2.
- Fastbom, Lennart: Some Aspects on the Construction of a System of Social and Demographic Statistics (SSDS), Statistisk tidskrift 1969:1.
- Hjerppe, Reino -
Niitamo, Olavi: Aspects of Developing the System of Social Indicators in Finland, International Association for Research in Income and Wealth, Twelfth General Conference, Ronneby Brunn, Sweden, Aug. 30 - Sept. 4, 1971 (esitelmä).
- Larson, Brita: Identifikation av statistisk information, Del. 1, Statistisk tidskrift 1969:2.
- Laurila, E.H: Tilastollisen päätoimiston tilasto-, tutkimus- ja selvitystöiden ryhmittelyä, Muistio 1970:1.
- Luther, Georg: The Acting Units in Statistics on the Sector of Enterprises, Reprint from: Kokonaistaloudellisia ongelmia, Vol. XXV.
- " Yritysrekisterin kehittämisen ja toiminnan suuntaviivoja, Tilastollinen päätoimisto 8.4.1969, (moniste).
- " Rekisterijärjestelmämme tulevasta kehityssuunnasta, Tilastokeskus 15.12.1971, (moniste).
- Nordbotten, Svein: Elektronmaskinene og statistikkens framtidige utforming, Nordisk Statistisk skriftserie 7.
- " A Statistical File System, Statistisk tidskrift 1966:2.
- " On Statistical File Systems II, Statistisk tidskrift 1967:2.

- Nordbotten, Svein: Purposes, Problems and Ideas Related to Statistical File Systems, Invited Paper to the 36th Session of the International Statistical Institute in Sydney, 28.8. - 8.9.1967.
- Ohlsson, Ingvar: EDB-systemets betydelse för statistikproduktionen och statistisk analys, Nordisk statistisk skriftserie 12.
- Salomonsson, Owe: Koordinater i register och dataarkiv, Statistisk tidskrift 1968:4.
- " Offentliga data, Ekonomi och Samhälle, Stockholm 1971.
- Svanfelt, Göran: Korreferat til foredrag af Johan-Kristian Tønder: utdanningsstatistikk i relasjon til eit socio-demografisk system, Rapport fra Nordisk Symposium vedrørende et integreret befolknings- og socialstatistisk system, Vedbaek, 8. - 10. april 1970.
- Strengell, Göran: Enheterna i person-, byggnads- och bostadsregistren, Rapport från nordiska seminariet i registerteknik, Hässelby slott 18-22 mars 1968.
- Wallberg, Klas: Kravet på statistik som underlag för ett socio-demografiskt system, Statistisk tidskrift 1971:3.