

Tilastokeskus

Tekijä	Päiväys	N:o
Jouko Kaartinen Kari Molnar Riitta Wallenius	29.2.1980	58

TILASTOKESKUKSEN PITKÄN TÄHTÄYKSEN SUUNNITELMA

Erillisselvitys:

KÄSITTEIDEN JA LUOKITUKSEN YHTENÄISTÄMISEN PTS-SELVITYS

SISÄLLYS

- Alkulause
 - 1. Johdanto
 - 2. 1970-luvulla kirjattuja yhtenäistämisperiaatteita ja tavoitteita
 - 3. Nykytilan selvitys
 - 3.1. Yleistä
 - 3.2. Luokitusstandardit
 - 3.3. Johtopäätöksiä
 - 4. Yhtenäistämistarpeisiin vaikuttavia tekijöitä
 - 4.1. Päällekkäisen tiedonkeruun ja tiedonantotaakan vähentäminen
 - 4.2. Atk-tekniikan lisääntyvä käyttö
 - 4.3. Lisääntyvä tiedontarve
 - 4.4. Tietojen vaihdon kansainvälistyminen
 - 5. Yhtenäistämistoimenpiteistä
 - 6. Yhtenäistämisen voimavarat
 - 7. Yhtenäistämisen strategiat
-
- Liite 1 Yhtenäistämismenetelmistä
 - Liite 2 Esimerkki informaation luokittelussa ja koodauksessa käytettävistä peruskäsitteistä

ALKULAUSE

1. Koko tilastotointa koskevan PT-selvityksen toiminta-ajatukseksi on paljolti se, että
Tilastokeskuksen rooli muuntuu 1980-luvulla perustietoja tuottavasta virastosta myös aktiivista tilastopalvelua tuottavaksi virastoksi.

Yhtenäisen tavoiteohjelman muodostamisen pohjana ovat seuraavat PT-erillisselvitykset.

- a) Tuotantomentelmien kehittäminen ja tietokantapolitiikka.
- b) Tilastotietojen välittäminen käyttäjille.
- c) Viraston henkilöstön kehittäminen.
- d) Tilastotoimen asema 1980-luvulla.
- e) Käsitteiden ja määritelmien yhtenäistäminen.
- f) Yhteiskunnan tietoa-aineistojen hyväksikäytön editäminen.

On selvää, että nämä erillisselvitykset eivät ole toisiaan pois-sulkevia, vaan paljolti "päällekkäisiäkin": ne pyrkivät samaan em. "P-T-tavoitteeseen" painottaen eri näkökulmaa.

On jo intuitiivisesti selvää oheisen PT-erillisselvityksen kiinteät kytkennät edellä lueteltuihin "a"-, "b"- ja "f"-selvityksiin.

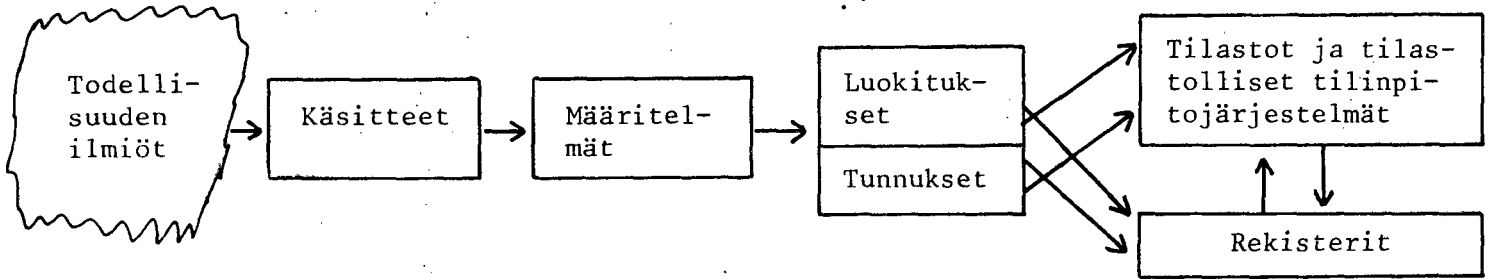
Edellä esitettyä koko tilastotoimen PT-selvityksen päätavoitetta, toiminta-ajatusta on täsmennettävä erilaisilla sitä luonnehtivilla erityiskriteereillä. Saatammekin asettaa tilastotoimen kehittämisen erityiskriteereiksi mm. seuraavat "laatuvaatimukset":

- oikea-aikaisuus
- luotettavuus
- käyttöönsaattamismuoto
- jalostusaste
- tekninen joustavuus
- alueellisuus
- perustiedon yhdenmukaisuus erilaisiksi kokonaisuuksiksi.

Tällöin PT-selvityksessä on mahdollista antaa noille kriteereille toinen paino kuin esimerkiksi KT- ja LT-suunnitelmissa. Oheisen erillisselvityksen erityispaino on ilmeisesti viimeisellä edellä luetelluista laatuvaatimuksista. Mutta samalla se auttaa esimerkiksi kahden edellisen kriteerin toteuttamista.

2. Oheisen erityisselvityksen toiminta-ajatusta saatetaan - sen erityisen keskeisyyden vuoksi - tarkastella varsin laajasti. Jopa niin laajasti, että tarkastelukulma näyttää ehkä abstraktilta: Ihminen pyrkii erilaisin käsittein järjestämään havaintonsa olemisesta - todellisuudesta - niin että tajuaa sen rakennetta ja liikelakeja: Hän pyrkii käsittämään todellisuutta, käsitelläkseen sitä ("mukautumaan todellisuuteen mukauttamalla sitä"):

Todellisuutta hahmottavan käsitejärjestelmän taustana on tällöin jonkinlainen "teoria" tuon todellisuuden olemuksesta. Tilastotoimen tehtävää tässä todellisuuden jäsentämisessä, käsitteellistämässä voidaan konkretisoida ns. yhtenäistettyjen tietoa-aineistojen toimintaidean avulla. Tämän erityisselvityksen suhdetta yhtenäistettyjen tietoa-aineistojen toimintaideaan voidaan kuvata seuraavan intuitioon vetoavan kuvion avulla.



Tästä kuviosta ainakin vihjeellisesti havaitsee, mikä yhteys tällä PT-selvityksellä on tilastotoimen kokonaisuuteen sinänsä, tunnusjärjestelmien kautta tuotantomenetelmien standardisointiin, rekistereiden kautta yhteiskunnan tietoaaineistojen hyväksikäyttöön & PT-selvitykseen tms.: Kytkentöjä em. tuotantomenetelmien ongelmiin on tarkasteltu lähemmin käytettyjä termejä koskevassa esityksessä.

Kytkentä yhteiskunnan tietoaaineistojen hyväksikäyttöön (em. "f"-selvitykseen) tapahtuu mm. siten, että:

- käsitteitä, määritelmiä tms. koskevat suositukset antavat varsinaiselle tilastotoimelle - paljolti ns. välittömälle tilastointia palvelevalle tiedonkeruulle - lähtökohtana olevat yhtenäistävät standardit
- myös muiden tietoaaineistojen tilastolliselle hyödyntämiselle ne luovat erilaisen välillisiin menetelmiin perustuvan yhtenäistämisen- ja yhteiskäyttöpohjan.

Muita tietoaaineistoja ovat toisaalta ns. hallinnolliset tietoaaineistot, toisaalta yksityisen sektorin tilastotyyppiset tietoaaineistot.

Hallinnollisista tietoaaineistoista muodostavat pääosan rekisterityypiset tiedostot.

Tilastointia palvelevat käsitejärjestelmät perustuvat

- Toisaalta aivan arkikielen käyttämään todellisuuden käsitteellistämiseen (tuoli, pöytä, auto, kalastus, mies, nainen, lääkäri jne.)
- Toisaalta ao. käsitejärjestelmän omiin erityisratkaisuihin, joilla
 - täsmennetään eri ilmiöiden kuulumisen peittävästi, toisena poissulkevasti johonkin luokkaan (ts. etteivät esimerkiksi miltei kokonaan puusta tehdyt laivat yhtäaikaan esiinny puunjalostusteollisuudessa ja metalliteollisuudessa - laivanrakennustoimialassa)
 - luodaan erityiskäsitteitä, joilla tiettyjen varantojen (reaali pääomakanta, rahoitusomaisuus, koulutustaso, terveystaso jne.) rakenteet ja niiden muutokset, virrat (tuotanto, koulutus, vienti, opetus, terveydenhoito) muodostavat nämä ilmiöt johdonmukaisesti ja mielekkäästi toisiinsa kytkeytyviä mielekkäitä kokonaisuuksia.

Tämän tyyppiseen tilastotoimen suorittamaan todellisuuden käsitteellistämiseen liittyvät niihin pohjautuvien maailmankuvien standardisointiin liittyvät vaarat:

Loogisen, aukottomaan tietosysteemiin pyrkivän käsitekehikon kehittämiseen ja soveltamiseen liittyvä vaara ja vastuu siitä on suuri. Tieto olevaisesta "pakotetaan" luokituksiin, jotka ovat soveltumattomia ja harhaanjohtavia selvitetessä todellisuutta. Pyrittäessä yksinkertaisuuteen saattaa hintana olla siis harhainen yksipuolisuus perusaineistona kerättävän (kerätyn ja kerättävissä olevan) tiedon olennainen supistuminen ja vääristyminen. Yksinkertaisuuden kautta havainnollisuuteen pyrkivällä järjestelmällä jo on sinänsä pakkopaidan luonne: Aineiston keruu- ja hyödyntämismahdollisuuksia saatetaan supistaa liikaa. Tehokkuuden hintana on mahdollisten maailmankuvien vähentäminen. Mitä suurempia kokonaisuuksia yhtenäistetty tietojärjestelmä "sitoo" sitä suurempi on tuo yksipuolinen "totalitarismin" vaara.

3. Vaikka myöhemmin määritellään tässä selvityksessä käytetyt käsitteet, on syytä lyhyesti kosketella itse LUOKITUS käsitteen keskeistä sijaa tilastotoimessa:

Saatamme sopia, että yleensä informaation näkökulmasta meitä ympäröivä olevainen koostuu

- havaitsemistamme yksittäisistä "asioista",
a l k i o i s t a
- näiden alkioiden eräiden ominaisuuksien perusteella muodostetuista l u o k i s t a
- alkioiden tai luokkien s u h t e i s t a muihin alkioihin tai luokkiin.

Ongelmanamme on tällöin edellä esitetyn perusteella tarkastella tilaston, ts. hyvin keskeisen joukkoilmiötä koskevan kuvauksen muodossa ilmenevän tietojärjestelmän suhdetta syy-analyysiin, tms. tilastollisen perustiedon jatkokäsittelyyn. Tällöin saatamme sanoa, että tilastollinen tietojärjestelmä- ja nimenomaan luokitusten soveltaminen perustuu

- noita alkioita koskeviin määritelmiin ("talonrakennus on ...").
- ns. johdettuihin määritelmiin eli jakautumalauseisiin (talonrakennustoiminta = asuinrakennustoiminta + liikerakennustoiminta).
- em. suhteita koskeviin väittämiin, siis väittämiin "mistä mikin johtuu".

Alkioita koskevissa määritelmissä - siis tavallaan perusmääritelmissä - joudumme "ymmärrettävästi" kuvailemaan, mitä on henkilö (taloudenpitäjä, kulutusyksikkö jne.), yhteisö (yritys, järjestö jne.), fyysisen ympäristön alkiot (rakennus, silta, suo, joki jne.), ts. ne elementit joiden tiloja tilastojärjestelmissä tarkastellaan.

Johdetuissa määritelmässä eli jakautumalauseissa esitämme miten tietty kokonaisuus jakautuu - aukottomasti, poissulkevasti ja tarkoituksenmukaisesti osiin ts. miten muodostamme edellä mainitut alkioiden luokat. Kun jäsenämme kokonaisuudet relevantteihin osiin - tai kääntäen: kokoamme (aggregoimme) alkioista tiettyjä luokkia, kokonaisuuksia, emme varsinaisesti vielä väitä mitään vaikutussuhteita.

Sen sijaan väitteet "mistä mikin johtuu" lähtevät riippuvuussuhteista, joiden oletetaan tavalla tai toisella vallitsevan jakautumalauseiden jäsenten välillä (tai em. toisen jaon mukaan: alkioiden tai luokkien suhteina toisiinsa alkioihin tai luokkiin). Nämä suhteet voivat olla teknisiä (panos-tuotoskertoimet), insitutionaalisia (veroprosentti) tai psyykkisiä käyttäytymissuhteita (kulutusalttius).

Tässä tarkastelussa tilastotoimi on pääpainotteisesti jakautumalauseiden järjestelmä. Tämä havainto koskee sekä yksittäistä tilastoa että sellaista pitkälle kehitettyä tilastollista tietojärjestelmää kuten kansantalouden tilinpito. Esimerkiksi viime mainitun neljän perustilin rakenne saatetaan esittää seuraavien jakautumaalueiden avulla:

1. ¹⁾ $B + U = C + I + E$
2. $C + S + Y^S = B - D + Y^F$
3. $I - D - L = S$
4. $E + Y^F = U + Y^S - L$

Tämän tyyppisen tilastojärjestelmän - k ä s i t e - j a l u o k i - t u s k e h i k o n - pohjana on kokonainen sarja alenevan kerta-luokan luokituksia. Esimerkiksi:

Kulutus = yksityinen kulutus + julkinen kulutus

Yksityinen kulutus = Elintarvikkeet + ...

Elintarvikkeet = Juomat + ...

.

.

Kunne päädytään alkiota koskevaan perusmääritelmään.

Hyödyke on ...

1)

Symbolien selitykset:

B = bruttokansantuote, U = tuopti, C = kulutus, I = investoinnit, E = vienti, S = säästäminen, Y^S = tulonsiirrot ulkomaille (netto), D = kiinteän pääoman kuluminen, Y^F = tuotannontekijätulot ulkomailta (netto), L = lainanotto ulkomailta (netto).

Toisaalta tuo kansantalouden tilinpito¹⁾ kytkeytyy tietö- ym. "linkein" sekä "henkilö- ja sosiaali-tilastojen kokonaiskehikkoon"²⁾ että "ympäristötilastojen kokonaiskehikkoon"³⁾. Tällöin nuo kolme perusjärjestelmää yhdessä muodostavat tavallaan korkeamman kerta- luokan tilastollisen kokonaiskuvausjärjestelmän.

Viitaten edellä esitettyihin luokituskehikkoihin ja niitä yhdistä- viin linkkijärjestelmiin⁴⁾ saatetaankin hyvin yleistäen sanoa, että

- tilastotoimen pyrkiessä omalta osaltaan yhtenäistämään eko-, tekno- ja sosioekonomista perusinformaatiota, ni- menomaan "luokitustoimi" muodostaa tuon perustehtävän pääasiallisimman välineen.

O. E. Niitamo

-
- 1) SNA = System of National Accounts (ks. A System of National Accounts, Stud. in Methods Series F. N:o 2, Rev. 3. Departement of Economic and Social Affairs, Statistical Office of the United Nations, New York 1968).
 - 2) FSDS = Framework of Social and Demographic Statistics (ks. "Concepts, Definitions, and Classifications of a System of Demographic and Social Statistics. Conf. Eur. Stats./WG. 34/9):
 - 3) FES = Framework of Environmental Statistics (ks. Aarno Laihonon. The Framework of an Information System of Environmental Statistics. Tilastokeskuksen tutkimuksia n:o 23/1973).
 - 4) Konkreettimmin näitä "linkkaus"-ongelmia tarkastellaan mm. Yhteiskunnan tietoa-aineistojen yhteistämisen PT-selvityksessä - erityisesti rekistereiden hyödyntämistä koskevassa osassa.

1. Johdanto¹⁾

Tilastotoimen tehtävänä on tuottaa tilastotietoja niistä yhteiskunnallisista ilmiöistä, joita yhteiskunnassa esiintyvät, tavalla tai toisella määritellyt ja priorisoidut, käyttötarpeet edellyttävät. Tätä tilastotietoa tulee tuottaa mahdollisimman tehokkaasti. Jos ajatellaan tilastotoimintaa kokonaisuutena, merkitsee näin esitetyn tehtävän suorittaminen melkoista optimointiongelmää.

Kuvattavien ilmiöiden valintaa rajaavat erilaiset tiedonsaanti- ja luottamuksellisuusseikat sekä tietenkin se, että tilastollinen kuvaus on massailmiöiden selvittämistä - se on tietoa "pelkistävä" menetelmä.

Tilastoinformaation tuottaminen kohdeilmiöistä tapahtuu näitä ilmiöitä vastaavien käsitteiden avulla. Jotta samaa ilmiötä kuvaavat tilastotoimen eri haaroilla tuotetut tiedot olisivat yhteisesti hyödynnettävissä, on kohdeilmiö oltava tunnistettavissa samaksi; ilmiö on määriteltävä tämän vaatimuksen tyydyttävällä tavalla.

Jotta tietojen yhteisesti hyödynnettävyys toteutuisi, on sekä kansallisella että kansainvälisellä tasolla yhä enemmän kiinnitetty huomiota tilastotoimessa käytettävien käsitteiden vakioimiseen, standardien aikaansaamiseen. Näiden standardien, tilastotoimen "työkalujen" laatiminen ja käyttöönsaattaminen edustaa samanlaista tehokkuuden ja taloudellisuuden tavoittelua kuin mittastandardien ja rationalisoinnin soveltaminen materiaalisessa tuotannossa.

Yhtenäistämässä joudutaan hyväksymään se, että samalla ilmiöllä on monta nimeä (on olemassa synonyymeja), ja että tiettyä kohdealuetta on usein tarve luokitella erilaisin kriteerein aina tiedon käyttötarkoituksen mukaan (esim. monet tavaraluokitukset).

Tietojen yhdistely- ja vertailumahdollisuus on kuitenkin otettava huomioon. Tällaisissa tapauksissa on tärkeää, että käytettävissä on ohjeistoja käsitteiden ja luokitusten muodostamiseksi sekä käsitteiden että luokitusten välisiä avaimia.

Käsitteiden ja luokitusten yhtenäistämässä päämääränä on tehdä mahdolliseksi tietojen kesto- ja monikäyttö. Yhtenäistämisen tavoite on siten mitä suurimmassa määrin konkreettinen, tilastotoimen "omimpaan" aiheeseen ja sen tehostamiseen liittyvä pyrkimys, joka on ymmärrettävä sellaisena esim. suhteessa informaatiotekniikkojen kehittymiseen toisaalta ja yhteiskuntailmiötä koskevan tutkimuksen tarpeisiin toisaalta. Tietojen kesto- ja monikäyttö tarkoittaa sitä, että

1)

Esitämme suuret kiitokset toimistopäällikkö Kaarlo Siltarille, jonka rakentavat ajatukset yhtenäistamisestä ovat niin paljon edesauttaneet tämän selvityksen laatimista.

1. Ilmiöitä voidaan vertailla
 - ajan suhteen (ilmiö on identifioitavissa samaksi eri ajankohtina, jolloin voidaan laatia aikasarjoja)
 - paikan suhteen (ilmiöitä voidaan vertailla esimerkiksi kansainvälisesti)
 - taloudellisen prosessin eri vaiheissa (saadaan virta-analyysimahdollisuus, kun ilmiötä voidaan vertailla eri vaiheissa: tuonti, tuotanto, kuljetus, kauppa, kulutus, vienti)
2. Ilmiöitä voidaan yhdistellä erilaisiin kokonaisuuksiin, niistä voidaan rakentaa erilaisia hierarkioita.
3. Samoja tietoja voidaan käyttää eri järjestelmissä.

Korostettakoon, että käsitteiden ja luokitusten yhtenäistämisyrittämisellä ei tarkoiteta yhteiskuntailmiöiden tilastollisen kuvauksen kangistamista vain tiettyjen vakioitujen käsitesäännösten ja luokitusten avulla tapahtuvaksi. On varmasti olemassa erikoiskäsitteitä, joita meillä ei ole oikeuskaan standardisoida. Tilastotoimessa on kuitenkin löydettävissä yli eri tilastonosien käsitteitä ja luokituksia, jotka on vakioitavissa; näiden standardointiin on pyrittävä tilastojen sisällön täsmällistämiseksi ja tilastotuotannon tehokkuuden lisäämiseksi.

Tämä selvitys lähtee olettamuksesta että ensisijainen edellytys yhtenäistämiseksi on luokitus- ja käsittestandardien aikaansaaminen sellaisista ilmiöistä, jotka toistuvasti esiintyvät tilastotoimen kuvauskohtina ja joiden merkityssisältö kuvattavien rakenteiden kannalta pysyy ajassa "riittävän" vakiona. Yleensä standardin laatiminen vaatii yhteistyötä, sillä standardi on sen tarvitsijoiden yhteiseen sopimukseen perustuva menettelytapa tms. ohje. Vähänkin suuremman luokituksen aikaansaaminen on todella suurta valmistelua ja yhteistyötä vaativaa. Sellaisen laadinnan ja ylläpidon aiheuttamia kustannuksia ei ainakaan Suomessa ole katsottu vielä riittävän perustelluiksi kaikissa niissä tapauksissa, joissa se olisi aiheellista.

Standardien laatiminen ei kuitenkaan yksin ratkaise yhtenäistämistä. Vähintään yhtä voimakkaasti on tehtävä työtä standardien käytön edistämiseksi.

2. 1970-luvulla kirjattuja yhtenäistämisperiaatteita ja tavoitteita

Vuonna 1971 säädettiin Tilastokeskusta koskevalla asetuksella suunnitteluosaston erääksi tehtäväksi käsitteiden, määritelmien ja luokitusten yhtenäistäminen. Vuonna 1972 asetti Tilastokeskuksen ylijohtaja työryhmän laatimaan valtion tilastotoimen kehittämissuunnitelmaan kytkeytyvä suunnitelma kyseisen tehtävän suorittamiseksi. Työryhmältä ilmestyi kaksi väliraporttia 8.9.1972 ja 31.5.1973. Työryhmä ei saanut toimeksiantoaan päätökseen. Työryhmän siihen asti saavuttamia tuloksia voidaan pitää PTS-selvitysten luonteisina, vaikka tehtäväksiannossa ei tätä nimenomaan korostettukaan. Raporteissa esitetyt kannanotot yhtenäistämisen periaatteista sekä yhtenäistämistyön suuntaviivoista, organisoinnista ja resursseista ovat suurimmalta osaltaan yhä ajankohtaisia ja joitakin niistä voidaan tiivistetysti toistaa tässäkin:

Yhtenäistämishankkeen on nojaututtava

- suhteellisen yksimielisesti hyväksytyihin loogisiin periaatteisiin, joita käytetään määrittely- ja luokitus-työn pohjana.
- tämän ajatusrakenteen voimakkaaseen propagointiin tilastotoimen piirissä, mihin on liitettava yhtenäistämiseen tähtäävä koulutustoiminta.
- huomion kiinnittämiseen standardimääritelmien ja luokitusten systemaattiseen kehittämistyöhön käytettävissä olevien resurssien puitteissa. Tähän kehittämistyöhön on saatava organisoiduksi mukaan eri tilastojen vastuu- ja laadintahenkilökuntaa yhteistyössä suunnitteluosaston koordinoitukysymyksiä hoitavien henkilöiden kanssa. Tämä sen vuoksi, että päätoimisen "luokitushenkilökunnan" lisäys ei näköpiirissä olevien mahdollisuuksien puitteissa mitenkään vastaa ratkaisua vaativien ongelmien laajuutta ja lukumäärää. Edellä viitattu yhtenäisten periaatteiden propagointi tulee juuri tästä syystä tärkeäksi
- nykyisin käytettävien käsitteiden, määritelmien ja luokitusten kartoittamiseen edellä viitatuista periaatteista sekä valmiista tai suunnitellusta standardeista lähtien. Tämä merkitsee sitä, että kaikkien nyt sovellettavien käsitteiden, määritelmien ja luokitusten kartoitus ilman johdonmukaisesti asetettua tavoitetta ei ole resurssien puolesta mahdollista eikä tarkoituksenmukaista. Tällainen kaikki tilastot peittävä kartoittaminen liittyy oleellisesti sellaiseen tilastoaineistojen dokumentointiin, jonka kehittämisen tulee tapahtua erityisesti atk-sovellutusten yhteydessä.

Yhtenäistämiseen liittyviä suuntaviivoja ja toimenpiteitä on esitetty myös "Valtion tilastotoimen kehittämisojelma (VATKO)" asiakirjoissa. Niissä yhtenäistämiseksi asetetut tavoitteet ovat edelleen ajankohtaisia.

Tilastojen tuotannossa käytettävien luokittelujen yhdenmukaistaminen on tilastojärjestelmään olennaisesti ja pysyvästi kuuluva perustehtävä (VATKO 1973-1977).

Eri valtion viranomaisten laatimien tilastojen luokittelujen yhtenäistämiseksi on

- a) rekistereissä ja muissa tietoaaineistoissa käytettävä samoista yksiköistä samoja tunnuskoodeja ja
- b) luotava tilastolliset keskusrekisterit, joissa perusyksiköt on lueteltu ja merkitty yleisillä standardoituilla tunnistamiskoodeilla ja kuhunkin tilastoyksikköön on tilastollisessa keskusrekisterissä merkittävä tarpeellinen määrä Tilastokeskuksen vahvistamia luokittelutietoja (VATKO 1973-1977).

Tilastokeskuksen ajankohtaisena tehtävänä luokittelujen alalla on yhtenäistää omassa tilastojen tuotannossa käytettävät luokittelut ja nimikkeistöt käyttäen tarpeellisessa määrin hyväksi kansainvälisesti sovittuja suosituksia ja standardeja (VATKO 1973-1977).

Sekä oman toiminnan rationalisoimiseksi että koko tilastotoimen edistämiseksi on tarpeen, että Tilastokeskus laatii ja julkaisee luokittelukäsikirjoja ja muita vastaavia selvityksiä käytettävistä luokitteluista ja niiden perusteista (VATKO 1973-1977).

Yhtenäisten käsitteiden ja määritelmien sekä luokitusten kehittäminen tapahtuu pääpainoisesti sellaisia standardimääritelmiä ja -luokituksia laatimalla ja julkaisemalla, jotka mieluummin monella eri tilaston alalla käytettynä ovat erityisen keskeisiä tilastotoimelle (VATKO 1974-1978).

3. Nykytilan selvitys

3.1. Yleistä

Valtion tilastotoimessa käytettävien luokitusten ja käsitteiden sovellutuskenttä on melko hajanainen. Valtionhallinnossa on yli 30 organisatorisesti riippumatonta tilastoja tuottavaa viranomaista. Lisäksi tilastotuotantoa harjoitetaan myös paikallishallinnossa, jossa muutamalla suurimmista kunnista on omat tilastotoimistonsa; näihin ei Tilastokeskuksen koordinoiva ote lain ja asetuksen mukaan ulotu. Tilastotuotannon organisatorinen hajanaisuus on merkinnyt tilastojen päällekkäisyyttä ja epäyhtenäisyyttä.

Epäyhtenäisyyttä on paitsi tilastojen tuottajien, myös tietojen antajien välillä sekä antajien ja tuottajien välillä. Tilastotoimen kannalta yhtenäistämisen pitäisi ulottua ensinnäkin tilastojen tuottajiin Tilastokeskuksessa ja muualla valtionhallinnossa sekä paikallishallinnossa ja toiseksi sekä välittömiin että välillisiin tiedonantajiin. Minkään tilastotoimen ulkopuolisen tietojärjestelmän ei voida olettaa tuottavan tietoa ensisijaisesti vain tilastointia varten. Esimerkiksi välittömän tiedonkeruun piiriin kuuluvan yrityksen laskentatoimen käsitteistön tehtävänä on palvella yrityksen päätöksentekoa: siinä käytettävät käsitteet ja luokitukset perustuvat kirjanpitolakiin ja hallintoviranomaisten, kuten veroviranomaisten vaatimukseen - sen sijaan taloustilastoissa käytetyt käsitteet ovat yleensä peräisin kansantaloustieteestä ja kansainvälisistä tilastointisuosituksista. Välillisen tiedonkeruun piiriin luuluvien hallinnollisten tietoaaineistojen käsitteet ja luokitukset ovat niiden erityissäännösten mukaisia, joita tietoaaineisto koskee. Ne ovat tarpeen kyseisen viranomaisen tehtävien suorittamisessa.

3.2. Luokitusstandardit

Keskeisin tehtävä tähänkin asti yhtenäistämrintamalla on ollut pääasiallisimman yhtenäistämisvälineistön, luokitustandardien, tekeminen ja - joskin huomattavasti vähäisemmin - niiden huoltaminen.

Tilastokeskus on vahvistanut seuraavat kuusi luokitusta valtion tilastotoimessa sovellettaviksi standardeiksi:

- Koulutusluokitus (1971)
- Toimialaluokitus (1972)
- Institutionaalisten sektoreiden luokitus (1975)
- Rahoitusvaateiden luokitus (1975)
- Julkisyhteisöjen tehtäväluokitus (1979) ja
- Yksityisten voittoatavoittelemattomien yhteisöjen tehtäväluokitus (1979)

Lisäksi voidaan pitää tilastotoimen standardiluokituksina myös seuraavaa kolmea, muiden virastojen toimesta käyttöön otettua ja vakiintuneen käytön saamaa, luokitusta:

- Suomen tullitariffinimikkeistö (1958)
- Pohjoismainen ammattiluokitus (1963) ja
- Tautiluokitus

Yllämainittujen lisäksi on alueluokituksista olemassa suositus, jonka mukaan tilastotuotannossa olisi ensisijaisesti käytettävä kunta-, lääni- ja seutukaava-aluejakoja tai ainakin pyrittävä keräämään tiedot niin, että niiden tulostaminen olisi mahdollista joillakin yllämainituista tavoista.

Luokituksia, jotka julkaistaan Käsikirjoja-sarjassa vuoden 1980 aikana, on kymmenen:

- Väestötilastollinen valtio- ja maa-alueuokitus
- Pääasiallisen toiminnan luokitus
- Pääasiallisen toimeentulolähteen luokitus
- Kotitaloustavaroiden ja -palvelusten luokitus
- Ammattiluokitus
- Rakennus- ja toimitilaluokitus
- Sosio-ekonominen asema luokitus
- Ammattiasemaluokitus
- Kotitaloustyyppiluokitus
- Perhetyyppi luokitus

Tämänhetkisistä standardeista tullitariffinimikkeistöllä ja toimialaluokituksella on erittäin laaja käyttö, joka ulottuu tilastotoimen ulkopuolelle. Myös koulutusluokitusta käytetään paljon, vaikka muitakin luokitustarpeita esiintyy koulutuksen suhteen.

70-luvun alkupuoliskolla standardointi keskittyi kansantalouden tilinpidon luokituksiin käynnissä olleen tilinpidon uudistuksen johdosta. Tämä luokitustyö on jäänyt jossakin määrin kesken.

Lähestyvä vuoden 1980 väestö- ja asuntolaskenta (VAL-80) herätti 70-luvun jälkipuoliskolla kiinnostusta laskennassa käytettäviin keskeisiin tilastoja yhtenäistäviin luokituksiin. Näihin kohdistuva tarkistus- ja standardoimistyö on nyt siis valmistumassa.

VAL-80:een liittyen valmistui v. 1979 laaja kartoitus tilastokeskuksessa käytettävistä tulokäsitteistä ja -luokituksista. Osittain näihin liittyen on julkisen talouden luokituslautakunnan jaostossa käynnissä työ julkisen hallinnon toiminnan ja talouden suunnittelussa ja seurannassa käytettävien tulo- ja menolajiluokitusten yhtenäistämiseksi.

3.3. Johtopäätöksiä

1970-luvun alussa yhtenäistämiseksi asetetut tavoitteet eivät ole toteutuneet niin, että voitaisiin puhua systemaattisesta tai edes hyvin organisoitusta kehittämistyöstä. Koordinointi on edelleen puutteellista - erityisesti tämä pätee Tilastokeskuksen ulkopuolisiin tilastoviranomaisiin. Ylipäänsäkin tilastotoimen kehittämistyö on ollut "sisään päin kääntynyttä", Tilastokeskuksen omien tilastojen yhtenäistämiseen keskittyvää. Lisäksi tietojen vertailtavuuteen ei ole kiinnitetty siinä määrin huomiota kuin verrattuna esimerkiksi tilastojen sisällön kehittämiseen. Syynä yhtenäistämistyön heikkoon tilaan ovat osaksi suunnitteluosaston heikot resurssit, jotka yhtenäistämispaineiden lisääntymisestä huolimatta ovat pysyneet samoina, mutta osaksi myös suunnittelutyön painopisteiden tähän astiset valinnat.

Standardien laatiminen on tapahtunut suurimmaksi osaksi ryhmätyönä niin, että ryhmien jäsenet ovat pääasiassa olleet Tilastokeskuksesta. Muiden käyttäjien mielipiteet on pyritty saamaan lausunnotmenettelyllä. Ryhmätyö onkin suositeltavin työskentelymuoto standardoinnissa, joskin ulkopuolisten osuutta voisi nykyisestä lisätä.

Vaikka standardien laatimismenettelyä voidaan pitää tyydyttävänä on koko yhtenäistämistyö silti vaivallisesti organisoitu. Osastojen väliset tehtävät ovat muutamista yrityksistä huolimatta jääneet täsmentämättä. Ajankohtaisin ongelma on luokitusstandardien ylläpito, joka hoituu "kivuttomasti" vain koulutusluokituksen osalta. Tullitariffinimikkeistöä pidetään luonnollisesti jatkuvasti ajan tasalla, samoin kuin tullitariffia tiivistä seuraavaa teollisuuden tavaranimikkeistöä.

Positiivisena tekijänä valtionhallinnon puolella on julkisen talouden luokituslautakunnan perustaminen. Sen jäsenten ja näiden edustamien taustaorganisaatioiden asiantuntemusta voidaan käyttää hyväksi sellaista tilastollisten luokitusten laatimisessa ja käyttöönsaattamisessa, joilla on merkitystä julkisen hallinnon toiminnan ja talouden suunnittelussa ja seurannassa. Yhtenäistämistä edistävää toimenpide on myös valtion hallinnon standardi - (VHS) menettelyn aloittaminen.

Tämän hetkisellemme keskustelulle luokitusten, käsitteiden ja tunnusjärjestelmien yhtenäistämistä on tunnusomaista sen epämääräisyys. Puhutaan yhtenäistämisen tarpeellisuudesta tai yhtenäisten tietoa-aineistojen rakentamisen mahdollisuudesta. Harvoin kuitenkin täsmennetään mitä epäyhtenäiset aineistot ovat tai mihin käsitteisiin ja luokituksiin epäyhtenäisyys kohdistuu. Tämä epätietoisuus yhtenäistämistarpeen todellisesta tilanteesta vaikeuttaa osaltaan myös käytettävissä olevien yhtenäistämisvälineiden ohjausta oikeisiin kohteisiin. Yhtenäistämistarpeiden täsmentämiseen valtion tilastotoimessa olisikin kiinnitettävä huomiota kartoittamalla ne aineistot, joihin pääasiallisimmat tietojen yhtenäistämispaineet kohdistuvat.

Eräänä haikkana yhtenäistämisen tiellä voidaan pitää myös epätietoisuutta siitä, mitä kaikkea tilastotietoa yhteiskunnassa tuotetaan. Tämän poistamisen edellytys olisi mm. ns. muuttujaluettelon aikaansaaminen ja tehokas tiedottaminen sen olemassaolosta.

Yhtenäistämisen kannalta tärkeää on myös että Tilastokeskus menee eri hallintoalojen ja virastojen tietoa-aineistojen kehittämistyöhön mukaan jo lainvalmisteluvaiheessa, ts. silloin kun asianomaisen toiminnan tai tehtävän sisältö, käsitteet, määritteet ja luokitukset lyödään lukkoon.

4. Yhtenäistämistarpeisiin vaikuttavia tekijöitä

4.1. Pälletekkäisen tiedonkeruun ja tiedonantotaakan vähentäminen

Tiedonantotaakan vähentäminen nykyisen laajuudessa tiedon tuotannossa edellyttää siirtymistä välilliseen tiedonkeruuseen. Siten pälletekkäisen tiedonkeruun ja tiedonantotaakan vähentäminen ovat suuresti riippuvaisia käsitteiden, luokitusten ja tunnusjärjestelmien yhtenäisyydestä.

Hallinnollisten tietoaaineistojen käyttö tilastotuonnossa kasvaa ja samalla myös halu hyödyntää niiden tietoja entistä monipuolisemmin. Tiedostojen uudistamisen ja uusien perustamisen myötä lisääntyy perustajien tarve saada aineistoihinsa valmiita luokituksia. Yksityiskohtaisten luokitusten lisäksi hallinnollisiin tietoaaineistoihin halutaan myös suuren aggregaattitason yleisluokituksia, joiden kehittämiseen ei ole tähän mennessä kiinnitetty juuri huomiota. Esimerkki tällaisesta suuresta aggregaattitasosta on toimialaluokituksen l-numerontason kolmeen luokkaan yhdistäminen alkutuotantoon, jalostukseen ja palveluksiin.

4.2. Atk-tekniikan lisääntyvä käyttö

Suuren paineen yhtenäistämisyrittämiselle aiheuttaa atk-tekniikan lisääntyvä käyttö niin tietojen keruussa, taltioinnissa, käsittelyssä kuin käyttöönsaattamisessakin. Käsitteiden, luokitusten ja tunnusjärjestelmien on oltava yhteensopivia, jotta tietojen yhteiskäyttö atk:lla mahdollistuisi. On hyvin todennäköistä, että tulevaisuudessa välitönkin tietojenkeruu tapahtuu pääasiassa atk:n avulla. Käsien suoritettava lomaketäyttö vähenee ja "henkilökohtaisuus" tietojen muokkaamisessa lomakkeille häviää.

Konekielisten linkkijärjestelmien kehittämiseen on kiinnitettävä huomiota. Tällaiset linkkijärjestelmät sisältäisivät luokitusten välisiä avaimia ja rekistereiden välisiä avaimia, jotka mahdollistavat tietojen siirron rekisteristä toiseen.

Päätekäytön lisääntymisen myötä tulee automaattinen merkintä yleistymään tietojen taltioinnissa. Tämä asettaa omat vaatimuksensa useille luokituksille ja varsinkin niiden nimikkeistöille. Automaattisessa merkinnässä voidaan käyttää mm. seuraavia nimikkeistöjä: alue-, ammatti-, toimiala-, tutkinto- ja tavaranimikkeistö. Automaattisella merkintäjärjestelmällä saavutetaan:

- suurempi luotettavuus
- suurempi merkinnän nopeus
- mahdollisuus virheiden havaitsemiseen ja korjaamiseen
- henkilöajan säästö

Yleisenä periaatteena luokitusten hyväksikäytölle tietoaaineistojen yhtenäistetyn hyväksikäytön toteutumisessa olisikin hyväksyttävä se periaate, että mahdollisimman paljon käytetään hyväksi automaattisen merkinnän suomia mahdollisuuksia silloin, kun luokitukset ja niiden tulkintaohjeet tähän soveltuvat. Tällaisten luokitusten aluetta pyritään laajentamaan ja luokitustoimen varsinainen "manuaali- tai aivotyö" pyritään keskittämään laajempien luokitusjärjestelmien ja toisaalta myös ulkopuolelle suuntautuvan koordinoitujen kehittämiseen.

1) Automaattisen merkintäjärjestelmän kehittämissuunnitelma, Muistio 12.11.1979, Tilastokeskus/tietojenkäsittelyosasto.

4.3. Lisääntyvä tiedontarve

Päätöksentekijöiden ja yleisön nälkä hyviä tilastotietoja kohtaan tulee luultavasti kasvamaan. Samoin kasvaa kriittisempi suhtautuminen tilastotietoihin. Aihekohtaisia suuntauksia on vaikea ennakoida. Aina on tietysti mukana myös täysin ennakoimattomien uusien kuvauskohteiden mahdollisuus. Sosiaalisuunnittelun ja tarkkailun systematisoituessa ja kasvaessa voitaneen sosiaalitulostatot ennakoida kasvusektoriksi. Taloudellisista muutoksista johtuvat sosiaaliset seuraukset vaativat analysointia ja tilastovirastoja johdetaan mittausalueille, jotka tällä hetkellä koetaan vaikeiksi. Tässä mielessä sosiaalitulatojen kehittäminen on haaste tilastotoimelle. Toisaalta, näistä dynaamisista seikoista huolimatta, eivät sosiaalityilat 80-luvulla pystyne ohittamaan talustilatoja absoluuttisessa laajuudessa.

Aluetason tiedontarve lisääntyy, erityisesti hallinnon desentralisoimisen myötä. Desentralisoimisen myötä kasvaa myös aluetason tietotuotanto. Tiedontuotannon mahdollinen hajatuminen nykyistä laajempiin mittoihin asettaa jatkuvia haasteita yhtenäistämislle. Keskusviranomaisen silmissä nämä hankkeet kehittyvät "villisti" kun ne ovat sen säätelyn ulkopuolella, ja totta onkin että koordinointi esim. kuntien itsenäisen tilastotuotannon suhteen voi tapahtua vain vapaaehtoisuuden ja kuntien oman edun pohjalta. Nyt on kuitenkin syntymässä rinnakkaisia tietojärjestelmiä, joita ei voi yhteisesti hyödyntää. Tämä kehitys lisää tietojen päällekkäistä keruuta ja käsittelyä yhteiskunnassa.

Lakisäateistä säätelymahdollisuutta ei voi kuitenkaan pitää ehdottomana rajana, jonka yhdellä puolella on koordinoitumahdollisuuksia ja toisella ei, käytännössä vaikuttaminen ei tule tehokkaaksi (vain) laki- ja asetuspykälää vahvistamalla. Jos rekisterinpitäjä kokee TK:n säännöt ja ohjeet itselleen sopimattomiksi tai hankaliksi, sillä on oma ministeriönsä, joka kykenee luomaan vastapykälää. Yhtenäistämisen pääongelmana käsitteiden ja luokitusten luomisen ohella tulee olemaan niiden markkinointi siten, että hallintovirastot eivät vastusta niille tarjottujen luokitusten ym. vastaanottoa vaan tulevat hakemaan TK:lta ratkaisuja omiin ongelmiinsa. Tämä näkökohta liittyy olennaisesti Tilastokeskuksen imagon kehittämiseen mutta sen takana tulee olla myös määrätietoinen valmiuksien kehittäminen siihen haasteeseen vastaamiseksi, jonka tällainen tavoite virastolle asettaa.

4.4. Tietojen vaihdon kansainvälistyminen

Tietojen vaihdon kansainvälistymisessä on kyse kahdesta näkökulmasta: 1) tietojen vaihto ja 2) tietojen vertailtavuus. Jälkimmäinen aiheuttaa yhtenäistämistarvetta. Tämäkin tietojen vaihto tapahtuu yhä useammin konekielisessä muodossa ja jo tietyllä tavalla yhtenäistettynä. Näiden tietojen hyväksikäyttö edellyttää, että kotimaiset tiedot ovat vertailtavissa kansainvälisiin tietoihin. Hyviä esimerkkejä kansainvälisistä yhtenäistetyistä tietopankeista ovat YK:n kansainvälisen kaupan tietopankki, energiatilastotietopankki, kansantalouden tilinpidon tietopankki, teollisuuden tuotantotilatojen tietopankki jne., jotka kaikki ovat konekielisessä muodossa jäsenvaltioiden saatavilla.

5. Yhtenäistämistoimenpiteistä

Uudet tilastot ja tilastojärjestelmät, kuten julkisyhteisöjen talous- ja toimintatilastojärjestelmä (JULTTI) tai energiatilastojärjestelmä, lisäävät standardoinnin tarvetta. Vaikka tilastojärjestelmien kehittäminen ja laajentaminen jo sinänsä merkitsee tiettyjen tilastojen yhtenäistämistä, tarvitaan myös eri tilastojärjestelmien tai niiden ja yksittäisten tilastojen välille kytkentöjä. On siis lisättävä toimenpiteitä kaikkien tilastojen yhteisesti muodostaman kokonaisjärjestelmän koossapitämiseksi. Toisaalta on varmistettava, että perustettavat tai uudistettavat tilastot ovat sopu-soinnussa muiden tilastojen kanssa, ja toisaalta varmistettava, että uusien tilastojen muille tilastoille asettamat vaatimukset leviävät kokonaissyteemiin.

1980-luvun haasteena Tilastokeskukselle on kehittää tiedonhallintajärjestelmä, joka tekee mahdolliseksi eri tilastoissa ja rekistereissä olevan tietoa-aineiston yhdistelyn ja hyödyntämisen (esimerkkinä rekisteripohjainen väestölaskenta). Rekistereiden tehokas hyväksikäyttö edellyttää, että niissä olevat yksiköt pystytään helposti tunnistamaan, ja että yksikkötiedot voidaan aggregoida niiden erilaisten ominaisuuksien perusteella. Yksikkötunnusjärjestelmien olemassaolo on ehdoton edellytys hallinnollisten aineistojen käytölle tilastotuotannossa. Rekistereiden monilukuisuuden vuoksi ei ole mahdollista, että Tilastokeskus pitäisi yllä näiden tunnusjärjestelmiin liittyvien yksiköiden luokitustietoja. Koska Tilastokeskus kuitenkin tulee käyttämään rekisterien tietoja, on sen tietojensa luotettavuuden varmistamiseksi saatava takeet rekisteritietojen oikeellisuudesta. Se että rekisteriviranomaisten on saatava luokitustiedot rekistereihinsä mahdollisimman yhdenmukaisesti ja oikein, asettaa suuria vaatimuksia luokituksille ja niiden oikealle tulkinnalle. On epärealistista ajatella että luokkien nimet ja määritelmät saataisiin sellaisiksi, että ne eivät jättäisi sijaa väärille tulkinnoille. Tämän vuoksi tulevat yhä tärkeämmiksi luokituksiin liittyvät nimikkeistöt, joilla helpotetaan yksiköiden jakamista oikeisiin luokkiin.

Suomi on ollut halukas kansainvälisten luokitussuosittelujen käyttäjä. Kansainvälisen vertailukelpoisuuden lisääminen tai ainakin säilyttäminen - kansainvälisen tilastoyhteisön osana pysyminen - on eräs Tilastokeskuksen tavoitteista. Sen ensisijaisena tehtävänä on silti tyydyttää kansalliset tilastotarpeet. Tasapainottelu kansallisten ja kansainvälisten vaatimusten välillä tulee merkitsemään sitä, että Suomi paitsi työstää kansainvälisiä suosituksia entistä enemmän omia olojaan vastaaviksi, myös tuottaa ns. harmonisoituja luokituksia¹⁾ ja käsitteitä sekä avaimia kansainvälisiin standardeihin

1) Kun luokitukset ovat harmonisoituja keskenään voidaan niiden välille laatia kaksisuuntainen avain. Keskenään harmonisoiduissa luokituksissa on siis samat perusrakennuspalat, jotka kuitenkin on ryhmitelty eri tavoin, eri kriteerien mukaan.

On alettava laatia 3-5 vuoden pituisia ohjelmia, jotka sisältävät mitä standardeja laaditaan ja uusitaan, mitä tunnusjärjestelmiä ryhdytään yhtenäistämään ja mitkä rekisterit ja tilastot kaipaavat minkäkinlaista yhtenäistämistä. Lisäksi on muistettava varautua ennakoimattomien yhtenäistämisiongelmiin käsittelyyn. Kun koko kansallista tilastotointa ei voida odottaa saatavan yhtenäistetyksi, on toimenpiteiden priorisointi oleellinen osa ohjelmatyötä. Ohjelmasta tulisi käydä ilmi seuraavat seikat:

- luettelo kaikista olemassaolevista tilastostandardeista ja niihin odotettavissa olevista sisällöllisistä tai teknisistä muutoksista
- luettelo tärkeimmistä tilasto- ja rekisterihankkeista: mitä yhtenäistämishankkeita ja -mahdollisuuksia ne luovat
- luettelo keskeisistä kansainvälisistä luokitushankkeista
- luettelo uusista standarditarpeista

6. Yhtenäistämisen voimavarat

Käsitteiden ja luokitusten yhtenäistämistyö on ensinnäkin ongelmien ja puutteiden kartoittamista ja painopisteiden valintaa. Toiseksi se on standardien tuottamista, ylläpitoa, uusimista sekä poikkeuksista sopimista ja harmonisointia. Kolmanneksi se on yhtenäistämisen organisointia ja voimavarojen kohdentamista sekä olemassaolevien standardien käyttöönoton edistämistä.

Edellä mainitussa 31.5.1973 päivätyssä KÄMÄ-työryhmän väliraportissa päädyttiin seuraavaan yhä ajankohtaiseen näkemykseen yhtenäistämistyön organisoinnista:

Työryhmän näkemyksen mukaan standardien vahvistamiskuntoon saattaminen ja ylläpito siihen liittyvine seuranta- ja yhteydenpitotehtävineen edellyttäisi tullakseen tarkoituksenmukaisesti hoidetuksi moninkertaisesti nykyisen määrän erityisesti näihin tehtäviin keskittyneitä henkilöitä. Tämä kannanotto merkitsee sitä, että jatko-toimenpiteet yhtenäistämiskysymyksessä Kämän kannalta jäävät viraston johdon määriteltäviksi, ilman nimenomaista esitystä yhtenäistämistyön organoinnista. Työryhmä edellyttää kuitenkin, että tällainen organisointiesitys on mahdollisimman pikaisesti saatava aikaan. Tällöin on korostettava sitä seikkaa, että yhtenäistämisen organisointimallina ei suinkaan välttämättä tarvitse eikä saakaan olla koko mainitunkaltaisen henkilöstön sijoittaminen suunnitteluosastoon. Päinvastoin standardointiin perustuva yhtenäistäminen edellyttää tilastokeskuksen eri yksikköjen kiinteää yhteistyötä ja koko tuotantoketjun kehittämistä sekä standardointia. Tällöin organisointia koskevan ehdotuksen perusajatuksena pitäisi olla se, että kuhunkin "riittävän suureksi" määriteltyyn viraston toimintayksikköön määrättäisiin henkilö suunnitteluosaston ao. tehtäviä hoitavan yksikön kanssa yhteistyössä hoitamaan luokitus- ja muita standardointitehtäviä.

Standardointiin nykyisin käytettävä työpanos on edelleen riittämätön. Suunnitteluosaston tehtävinä on lähinnä kartoittaa yhtenäistämistarpeet ja toimia yhtenäistämistyön koordinoijana. Sen tehtävänä on myös seurata tiiviisti standardien laatimista sekä välittää tiedot niihin kohdistuvista vaatimuksista tai suosituksista, kuten hallinnollisista luokituksista tai kansainvälisistä ehdotuksista, standardeja laativille elimille. Standardien muokkaaminen tulee sen sijaan hoitaa työryhmissä, joissa on edustajat tärkeimmistä standardia käyttävistä tilastoista tai rekistereistä.

Standardien laatimisessa on siis edullista turvautua sen pääasiallisimpiin käyttäjiin. Näin maksimoidaan työhön käytettävä asiantuntemus ja standardin sopivuus eri tilastoihin sekä helpotetaan sen käyttöönottoa.

7. Yhtenäistämisen strategiat

Standardien laatimisesta ei ole hyötyä, ellei niitä ryhdytä käyttämään mahdollisimman monessa käyttäjäksi soveltuvassa tilastossa tai rekisterissä. Siksi resursseja tarvitaan paitsi standardien laatimiseen, myös käyttäjien saavuttamiseen. Toistaiseksi on kiinnitetty liian vähän huomiota standardien käytön edistämiseen, ja siten tämä työ vaatisi erityistä paneutumista.

Standardien postittaminen ei riitä strategiaksi. Kun toistaiseksi ei ole kunnolla kiinnitetty huomiota yhtenäistämisvälineiden markkinointiin, on hankalaa laatia ohjelmaa tämän seikan kehittämiseksi. Parannusehdotuksina voidaan kuitenkin mainita seuraavat:

- käyttäjien osallistuminen standardien laadintaan: Käyttäjät kokevat standardin relevanttina, jos heidän näkemyksensä on alusta alkaen otettu huomioon laadintatyössä. Käyttöönotto tulee myös nopeammaksi ja helpommaksi.
- standardin laatiminen ja markkinoiminen sellaiselle alueelle, jolle joka tapauksessa ollaan laatimassa luokitusta. Toisin sanoen standardointityö tulee kohdistaa keskeisten tarpeiden mukaan, jolloin standardin olemassaolo ratkaisee luokituksia tarvitsevien viranomaisten ongelmat.
- standardointityössä keskittyminen tärkeimpiin tiedostoihin ja tärkeimpiin tiedontuottajiin. Tällöin marginaalisempien tiedostojen ja tiedontuottajien ei ole varaa olla käyttämättä standardeja.
- luokitusten markkinointi erilaisten projektien yhteydessä
- potentiaalisten käyttäjien etsiminen ja aktiivinen tiedottaminen. Etenkin tässä työssä tarvitaan koko Tilastokeskuksen, eikä vain yhtenäistämiseen erikoistuneen henkilökunnan aktiivista panosta.

LIITE 1

YHTENÄISTÄMISMENETELMISTÄ

Yhtenäistämiskeskustelua on hämärtänyt epätietoisuus siitä mitä eri yhtenäistämismenetelmillä tarkoitetaan, ja millä tavoin ne poikkeavat toisistaan.

Seuraava esimerkki valaiskoon neljää eri yhtenäistämismenetelmää, jotka rakentuvat kahdesta dikotomiasta: 1) yksikötason yhtenäistäminen vs. summatason yhtenäistäminen ja 2) horisontaalinen yhtenäistäminen vs. vertikaalinen yhtenäistäminen.

	Horisontaalinen	Vertikaalinen
Yksikkö- taso	Henkilö:010110123A Tulot: a mk Henkilö:010110123A Osoite: nnn Asunto: 3H + k Henkilö:010110123A Kieli: suomi jne.	Henkilö:020220234B Osoite: mmm Tulot: b mk Henkilö:030330345C Osoite: mmm Tulot: c mk Kotitalous xyz Osoite: mmm Tulot: b + c mk
Summa- taso	Toimiala 3 Tuotantoindeksi: x Toimiala 3 Työvoimaindeksi: y Toimiala 3 Tuottavuusindeksi: $100 \frac{x}{y}$	Tavara: 60.01 1000 Tuonti: a mk Vienti: b mk Tavara: 60.01 1000 tuotanto: c mk Tavara: 60.01 1000 kotimaan käyttö: a + c - b mk

Yksikötaso/horisontaalinen esittää kolmessa eri rekisterissä olevaa tietoa ja sitä miten ne voidaan yksikötunnuksen avulla yhdistää - edellyttäen, että yksiköt on määritelty samalla tavalla. Henkilö-yksikön kohdallahan tämä ei ole ongelma, mutta esim. toimipaikkaa koskevien tietojen yhdistäminen eri tietojärjestelmistä edellyttää paitsi yhteistä tunnusta myös yksiköiden rajaamista samalla tavalla.

Yksikötaso/vertikaalinen esittää miten kasi henkilöä, jotka asuvat samassa osoitteessa mmm muodostavat kotitalouden xyz. Rekisterin henkilökohtaisista tulotiedoista voidaan konstruoida kotitalouden tulot tarvitsematta järjestää niistä erillistä kyselyä. Samoin jos yksi yritys koostuu kahdesta toimipaikasta a ja b, voidaan näitä koskevista tiedoista rakentaa yritystä koskevat tiedot. (Tämä koskee tietenkin vain tietoja, joita voi kerätä toimipaikkatasolla. Esim. tasetietoja ei voi rakentaa näin).

Summataso/horisontaalinen edustaa erästä kaikkein tavallisinta tilastojen käyttötapaa. Kahdesta luokitellusta tiedosta laskeaan kolmas tieto; esimerkki tapauksessa jopa sellainen, jota ei millään muulla tavalla voi hankkia. Tässä on tärkeintä, että toimialajako on sama molemmissa tilastoissa.

Summataso/vertikaalinen osoittaa miten käytetään hyväksi (kuten yksikkötaso/vertikaalisessa) määritelmällistä yhteyttä Kotimaan käyttö = Tuonti + Tuotanto - Vienti. Näin voidaan hyödyntää ulkomaankauppatilastoa ja teollisuustilastoa uuden tiedon tuottamiseen ja säästyään käyttäjäpohjaiselta tilastolta.

Yksikkötason yhtenäistämässä tarvitaan yksiköiden identifiointia ja niiden yhdenmukaista määrittelyä. Luokituksista on turha muurehtia niin kauan kuin tunnukset ja yksiköiden rajausta on yhtenäistämättä. Ajatellaan esim. verotustietojen käyttöä yritystilastoissa. Tietojen yhdistäminen edellyttää samoja tunnuksia (tai avainta, mikä on käytännössä sama asia) ja samoja yksikkömääritelmiä. Yksikkötason yhtenäistäminen on rekisteritason yhtenäistämistä. Mainittakoon kuitenkin, että, esimerkiksi henkilötilastoja ja yhteisötilastoja (tai rekistereitä) ei voida linkata yksikkötasolla, koska yksiköt ovat erilaisia. Niitä voidaan kuitenkin yhdistellä jonkun yhteisen luokituksen avulla. Tällainen luokitus-tieto voi olla esimerkiksi toimiala.

Tällöin on siis kysymys summatason yhtenäistämässä, jossa tarvitaan samoja luokituksia ja luokkien yhtenäisiä määritelmiä. Yksikkökysymys ei ole siinä mielessä tärkeä, että onko joku tuotantoindeksi saatu yritystä, toimipaikkaa vaiko tuotetta koskevien tietojen pohjalta. Tärkeätä on, että luokittelu on tapahtunut oikein. Samoin esimerkiksi ulkomaankauppatiedoissa on tärkeätä että tavarat on luokiteltu oikein, ei niinkään onko ne kerätty maahantuojien ja viejien ilmoituksista vaiko tullausilmoituksista.

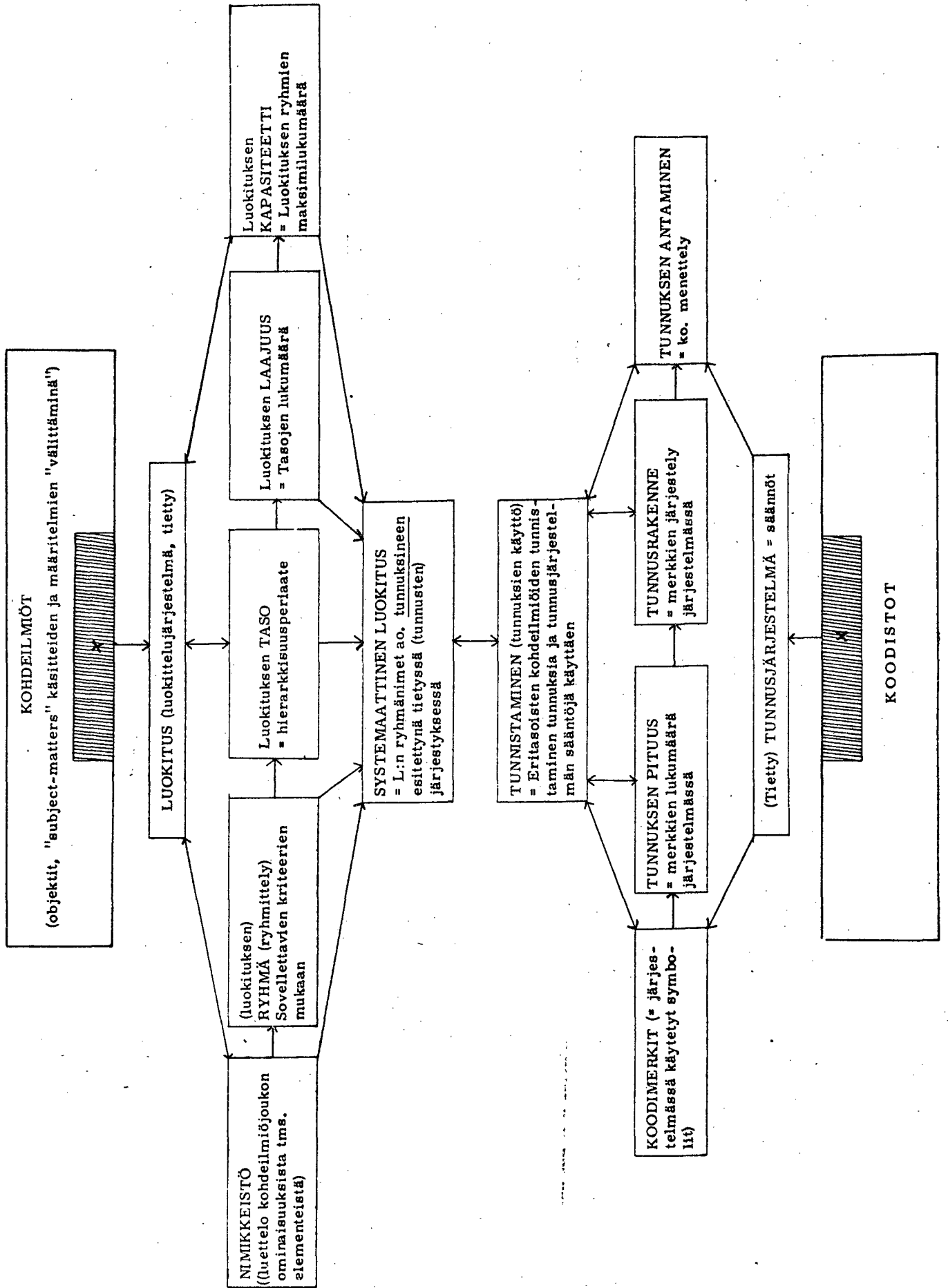
LIITE 2

CMEA-CCIAT (SEV-maiden rakenteilla oleva tuoteluokittelujärjestelmä)

KAAVIO INFORMAATION LUOKITTELUSSA JA KODAUKSESSA KÄYTETTÄVIEN PERUSKÄSITTEIDEN MÄÄRITELMISTÄ (Sisältyy YKn dokumenttiin ST/ESA/STAT/78, The Harmonization of Statistical Classifications)

Systemaattinen luokitus (Classifier)	Järjestetty esitys luokituksen ryhmänimistä asianmukaisine tunnuksineen
Luokitus (Classification)	Järjestelmä jakaa tietty luokiteltavien objektien joukko (esim. tuotteet) erillisiin ryhmiin (alajoukkoihin), joihin kuuluvat alkiot ovat keskenään tietyssä määrin homogeenisia
Nimikkeistö (Nomenclature)	Luettelo sen objektijoukon käsitteistä, ominaisuuksista, ominaispiirteistä tai muista osatekijöistä, jota varten luokitus järjestelmä on laadittu
Luokituksen ryhmä-, ittelu (Classification group(-ing))	Kategoria (alajoukko) johon kuuluu osa luokiteltavista olioista ja jonka muodostamisessa (eli näiden olioiden sijoittamisessa) on sovellettu yhtä tai useampaa luokitusjärjestelmässä käytettävistä kriteereistä
Luokittelutaso	Luokitus on analoginen portaittaisen pyramidin kanssa. Kukin porras vastaa tasoa. Korkein (aggregoiduin) taso käsittää vähiten ryhmiä. Jokainen sitä alempi taso saadaan edellisen korkeamman tason alajaotteluna sopivaksi katsotun jaotteluperusteen mukaan
Luokituksen (-järjestelmän) laajuus (Depth of the classification system)	Luokitusjärjestelmään kuuluvien luokittelutasojen lukumäärä
Luokituksen (-järjestelmän) kapasiteetti (Capacity of the classification system)	Tietyn luokitusjärjestelmän luokittelutasojen suurin mahdollinen lukumäärä
Tunnusjärjestelmä (koodausj.) (Codingsystem)	Joukko merkkien (symbolien) järjestelmää koskevia sääntöjä ja merkkien käytössä tietojen esittämis-, siirto-, käsittely- ja tallennustarkoituksiin sovellettavat menettelytavat
Tunnusmerkit (koodimerkit) (Coding alphabet)	Tunnusjärjestelmässä käytetyt merkit

Tunnistaminen (tunnuksen käyttö) (Identification by code)	Objektin (tason, ryhmän, luokan, alaluokan, tuotteen jne.) identifiointi merkin tai merkkien yhdistelmän avulla tietyn tunnusjärjestelmän sääntöjä noudattaen
Tunnuspituus (Code length)	Tietyn järjestelmän tunnuksiin kuuluvien merkkien lukumäärä
Tunnusrakenne (Code structure)	Merkkien järjestely tietyn tunnusjärjestelmän tunnuksissa
Tunnuksen antaminen (Coding)	Menettely tunnuksen antamiseksi tietylle objektille tai objektiryhmälle



KOHDEILMIÖT

(objektit, "subject-matters" käsitteiden ja määritelmien "välittäminä")

LUOKITUS (luokittelujärjestelmä, tietty)

NIMIKEISTÖ
(luettelo kohdeilmiöjoukon ominaisuuksista tms. elementeistä)

RYHMÄ (ryhmittely)
Sovellettavien kriteerien mukaan

TASO
= hierarkkisuusperiaate

LAAJUUS
= Tasojen lukumäärä

KAPASITEETTI
= Luokituksen ryhmien maksimilukumäärä

SYSTEMAATTINEN LUOKITUS
= L:n ryhmänimet ao. tunnuksineen esitettyinä tiettyssä (tunnusten) järjestyksessä

TUNNISTAMINEN (tunnuksien käyttö)
= Eritasoisten kohdeilmiöiden tunnistaminen tunnuksia ja tunnusjärjestelmän sääntöjä käyttäen

KOODIMERKIT (= järjestelmässä käytetyt symbolit)

TUNNUSRAKENNE
= merkkien lukumäärä järjestelmässä

TUNNUSRAKENNE
= merkkien järjestely järjestelmässä

TUNNUSJÄRJESTELMÄ = säännöt

TUNNUSJÄRJESTELMÄ = säännöt

KOODISTOT