



KÖZPONTI
STATISZTIKAI
HIVATAL

Az EGYSÉGES STATISZTIKAI FOLYAMATMODELL (ESTFM)

v2.3

Budapest, 2018

Tartalom

Egységes Statisztikai Folyamatmodell	4
A statisztikai adatelőállítási folyamat folyamatszakaszai és részfolyamatai:	8
I. Igényspecifikáció	8
I.1. Igények meghatározása	9
I.2. Igények egyeztetése	9
I.3. Fogalmak meghatározása	9
I.4. Adatelőállítási folyamat outputjainak meghatározása	10
I.5. Adatforrások feltérképezése	10
I.6. Adatfelvételi javaslattev készítése	10
II. Tervezés	11
II.1. Adatelőállítási folyamat outputjainak részletes megtervezése	11
II.2. Fogalmak, mutatók, nomenklatúrák definiálása	11
II.3. Felvételi keret meghatározása	12
II.4. Adatfelvételi módszer és eszköz részletes megtervezése	12
II.5. Mintavételi terv elkészítése	13
II.6. Adatbegyűjtést követő folyamatok tervezése	13
II.7. Adatelőállítási folyamat részletes tervének előállítása	13
III. Fejlesztés, tesztelés	14
III.1. Adatfelvételi eszközök véglegesítése	14
III.2. Informatikai eszközök tervezése	14
III.3. Informatikai eszközök fejlesztése	15
III.4. Informatikai eszközök tesztelése	15
III.5. Próbafelvétel	15
III.6. Adatelőállítási folyamat informatikai eszközeinek véglegesítése	15
IV. Adatfelvétel megvalósítása	16
IV.1. Felvételi keret, adatszolgáltatói kör, minta kiválasztása	16
IV.2. Adatfelvétel-szervezés, oktatás	17
IV.3. Adatbegyűjtés, sürgetés, érkeztetés	17
IV.4. Adatbegyűjtés lezárása	17
V. Adatelőkészítés	18
V.1. Kódolás	18
V.2. Mikro- és mezovalidálás	18

V.3. Editálás	19
V.4. Adatelőkészítés lezárása	19
VI. Feldolgozás	19
VI.1. Adatintegrálás.....	19
VI.2. Pótlás	20
VI.3. Új statisztikai egységek és mutatók képzése	20
VI.4. Súlyok előállítása	20
VI.5. Aggregálás, becslés, hibaszámítás	20
VI.6. Feldolgozás lezárása	21
VII. Tájékoztatás előkészítése.....	21
VII.1. Adatelőállítási folyamat nyers outputjainak elkészítése	21
VII.2. Szezonálisan kiigazított adatok előállítása.....	22
VII.3. Makrovalidálás	22
VII.4. Felfedés elleni védelem.....	22
VII.5. Elemzés.....	22
VII.6. Adatelőállítási folyamat outputjainak véglegesítése	22
VIII. Tájékoztatás.....	23
VIII.1. Tájékoztatási termékek előállítása	23
VIII.2. Közzététel.....	24
VIII.3. Tájékoztatási termékek népszerűsítése	24
VIII.4. Felhasználók támogatása, információszolgálat.....	24
A statisztikai adatelőállítási folyamatot átfogó területek	25
Minőségmenedzsment (A.).....	25
Statisztikai adatmenedzsment (B.)	25
Metaadat-menedzsment (C.).....	26
Melléklet – Fogalmak	27

Egységes Statisztikai Folyamatmodell

Az ESTFM célja

Az Egységes Statisztikai Folyamatmodell (ESTFM) célja, hogy a Központi Statisztikai Hivatal munkája, folyamatai átláthatóbbá váljanak azáltal, hogy a KSH-ban lévő statisztikai adatelőállítási folyamatok leírása, bemutatása és értelmezése egységes módon történik. Az ESTFM a statisztikai adatfelvételek folyamatának modellje, mely alatt a KSH-ban elsődlegesen a statisztikai adatelőállítási folyamatot, valamint a statisztikai regiszterek kezelésével, felvételi keret előállításával kapcsolatos folyamatokat értjük. Az ESTFM alkalmas az egyéb statisztikai folyamatok leírására is, mint pl. osztályozások karbantartása, kérdőívek, adatállományok, publikációk előállítása, stb.

Az ESTFM a nemzetközi Generic Statistical Business Process Model-re (GSBPM v.5.0¹) épül, mely az UNECE² keretében kidolgozott teljes statisztikai adatelőállítási folyamatra vonatkozó folyamatmátrix. Az adatelőállítási folyamatnak nem képezik részét a funkcionális tevékenységek, mint például a hivatal működtetése, gazdálkodási, jogi, személyügyi vagy igazgatási területek, folyamatok. Részét képezik viszont olyan általános feladatok, mint a minőségmenedzsment, a statisztikai adatmenedzsment (beleértve az adatvédelmet) és a metaadat-menedzsment, amelyek a statisztikai adatelőállítási folyamat ún. átfogó területeit alkotják és az adatelőállítási folyamat egészére kiterjednek.

Az ESTFM felépítése

Az ESTFM modellje a statisztikai adatelőállítási folyamat szakaszait (folyamatszakaszok), a folyamatszakaszok alábontásaként meghatározott részfolyamatokat, valamint az átfogó területeket (minőségmenedzsment, statisztikai adatmenedzsment és metaadat-menedzsment), tartalmazza.

Az ESTFM-ben a **folyamatszakaszok** meghatározásakor az adott folyamatszakaszhoz kapcsolódó részfolyamatok kulcsfontosságú elemei kerültek kiemelésre a gyakorlati használhatóság elősegítése érdekében. Az ESTFM a statisztikai adatelőállítás folyamatára fókuszál, nem az azokat kiszolgáló meglévő informatikai alkalmazásokra, melyek kiszolgálhatják több folyamatszakasz több részfolyamatát is.

Az ESTFM-ben az egyes **részfolyamatok** leírásának felépítése a következő:

- az adott részfolyamat célja;
- az adott részfolyamathoz tartozó főbb tevékenységek meghatározása, beleértve a részfolyamat inputjait, illetve
- az adott részfolyamat lezárását követő tevékenységek és a létrehozott outputok leírása.

Az ESTFM-ben az **átfogó területek** tartalmi összefoglaló leírása tartalmazza a valamilyeni folyamatot jellemző általános leírást (pl. célját, főbb tartalmi elemeit). Ezen kívül az egyes folyamatszakaszok leírásában is szerepelnek az átfogó területek kulcsfontosságú tevékenységei.

¹ Lásd: [Generic Statistical Business Process Model GSBPM \(Version 5.0, December 2013\)](#).

² ENSZ Statisztikai részlege: United Nations Economic Commission for Europe <http://www.unece.org/info/ece-homepage.html>.

Az ESTFM használata

Akárcsak a GSBPM, az Egységes Statisztikai Folyamatmodell is egy **mátrix szerkezetű modell**, vagyis a folyamatszakaszok alatt szereplő részfolyamatok nem feltétlenül lineáris módon követik egymást a folyamat végrehajtása során. Egyes részfolyamatok bizonyos adatelőállítási folyamatokban meg sem jelennek, illetve egy adott folyamat során megismétlődhetnek, valamint a részfolyamatok ábrázolt sorrendje sem tükrözi feltétlenül a folyamat végrehajtási sorrendjét, irányát. Az ESTFM tehát egy nemlineáris modell, amelynek alkalmazása során nem feltétlenül kell minden részfolyamatot felhasználni. Az egyes részfolyamatok súlya is eltérő lehet a leírt folyamatban. Segítségével a valós gyakorlatot kell dokumentálni, ami esetenként kevesebb részfolyamatból áll, mint a teljes modell. Az ESTFM fontos célja, hogy a hivatalban fellelhető valamennyi statisztikai folyamatot le lehessen írni a modell meglévő elemeinek kombinációjával.

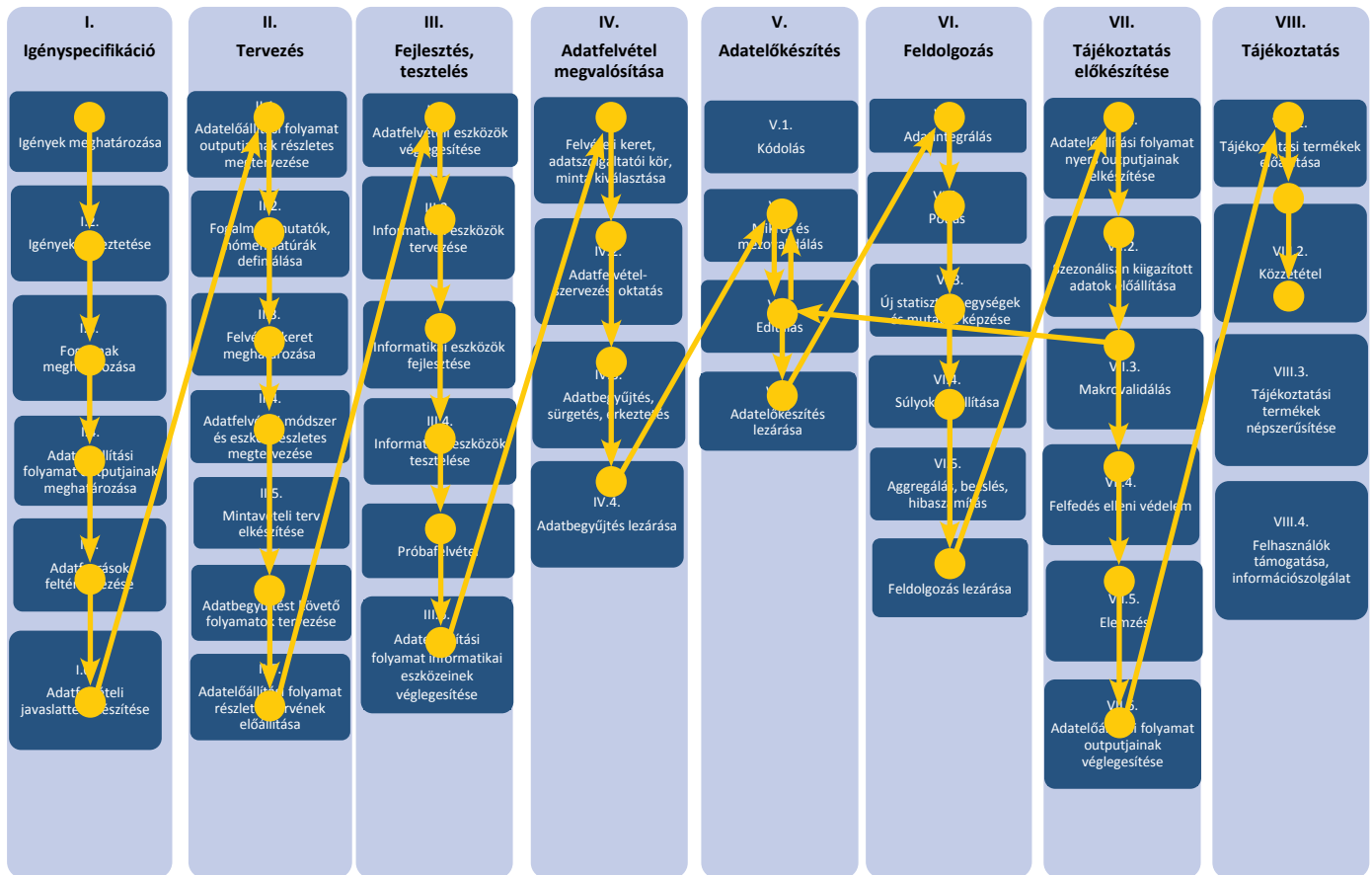
Az ESTFM gyakorlati felhasználási területei:

- módszertani standardizálás: a hivatali folyamatok áttekintése folyamatszakaszok vagy részfolyamatok mentén, a meglévő gyakorlatokon alapuló egységes módszertanok kialakítása, a meglévő megoldások standardizálása és az esetlegesen hiányzó elemekre módszertani megoldások kialakítása;
- folyamatdokumentáció: a folyamatok részletes, szakértőknek szóló dokumentálása az ESTFM mentén, melynek része lehet folyamatábra készítése, illetve a folyamatábrára épülő, egyes részfolyamatokban zajló tevékenységek dokumentálása is;
- folyamatmonitoring: a folyamatok monitorozása, a folyamatokhoz kötődő célértékek, valamint minőség mérése, értékelése, javaslatok a folyamatok javítására, esetleges beavatkozás a folyamatba;
- folyamatmenedzsment: rendszerszintű, szisztematikus folyamat monitoring, kitűzött célok teljesítése, hiányzó területeken célfejlesztések, akár projektek indítása;
- folyamatoptimalizálás: folyamatok állandó javítása (akár újraszervezése) a hatékonyság és az eredményesség érdekében.

Az ESTFM előnye:

- a folyamatszakaszok és részfolyamatok lehatárolásával a felelősségi körök beazonosíthatóak, a részterületek, résztevékenységek egymásra hatása egyértelműen megállapítható;
- az egyes adatelőállítási folyamatok egységes logika alapján kerülnek leírásra, ami növeli a hivatali folyamatok transzparenciáját és összehasonlíthatóságát. Ugyanakkor alapdokumentációként funkcionál a munkatársak fluktuációja esetén;
- a folyamatok tervezésekor meghatározott erőforrások könnyebben összesíthetőké válnak;
- az átfogó területek alapértelmezettek és egységes elvek mellett beépítettek az egyes statisztikai folyamatszakaszokba; a folyamatok leírása lehetővé teszi azok későbbi menedzselését is, vagyis a folyamat útjának pontos leírásával bármelyik szakasz minőségvizsgálatát el lehet végezni, illetve a folyamatok módosítása esetén az új részfolyamatok könnyen hozzáilleszthetők a meglévő leíráshoz.

A. Minőségmenedzment / B. Statisztikai adatmenedzment / C. Metaadat-menedzment



1. ábra: Példa az ESTFM alkalmazására egy teljes körű adatfelvételnél

Az ESTFM leírás köre, egysége

A statisztikai adatelőállítási folyamatok dokumentálásánál az ESTFM-ben leírt folyamatszakszokra, részfolyamatokra kell hivatkozni. Ilyen dokumentáció készülhet

- az egyes adatfelvételek (adatgyűjtések, adatátvételek) tervezéséről, végrehajtásáról, elsődleges feldolgozásáról és közléséről;
- hasonló logikai felépítésben, de magasabb aggregáltsági, pl. szakstatisztika szinten;
- a regiszterek, nyilvántartások tervezéséről, karbantartásáról, felvételi keretek előállításáról, a regiszterből való elsődleges tájékoztatásról;
- az általános folyamatok tervezéséről, működtetéséről, amelyekre az egyes adatfelvételek leírása hivatkozhat;
- az adatfelvételek másodlagos rendszeres és eseti feldolgozásáról, közléséről;
- egyéb adatelőállítási folyamatokról, például osztályozások karbantartásáról.

A. Minőségmenedzsment / B. Statisztikai adatmenedzsment / C. Metaadat-menedzsment

I.	II.	III.	IV.	V.	VI.	VII.	VIII.
Igényspecifikáció	Tervezés	Fejlesztés, tesztelés	Adatfelvétel megvalósítása	Adatelőkészítés	Feldolgozás	Tájékoztatás előkészítése	Tájékoztatás
I.1. Igények meghatározása	II.1. Adatelőállítási folyamat outputjainak részletes megtervezése	III.1. Adatfelvételi eszközök véglegesítése	IV.1. Felvételi keret, adatszolgáltatói kör, minta kiválasztása	V.1. Kódolás	VI.1. Adatintegrálás	VII.1. Adatelőállítási folyamat nyers outputjainak elkészítése	VIII.1. Tájékoztatási termékek előállítása
I.2. Igények egyeztetése	II.2. Fogalmak, mutatók, nomenklatúrák definiálása	III.2. Informatikai eszközök tervezése	IV.2. Adatfelvétel-szervezés, oktatás	V.2. Mikro- és mezovalidálás	VI.2. Pótlás	VII.2. Szezonálisan kigazdírt adatok előállítása	VIII.2. Közzététel
I.3. Fogalmak meghatározása	II.3. Felvételi keret meghatározása	III.3. Informatikai eszközök fejlesztése	IV.3. Adatbegyűjtés, sürgetés, érkeztetés	V.3. Editálás	VI.3. Új statisztikai egységek és mutatók képzése	VII.3. Makrovalidálás	VIII.3. Tájékoztatási termékek népszerűsítése
I.4. Adatelőállítási folyamat outputjainak meghatározása	II.4. Adatfelvételi módszer és eszköz részletes megtervezése	III.4. Informatikai eszközök tesztelése	IV.4. Adatbegyűjtés lezárása	V.4. Adatelőkészítés lezárása	VI.4. Súlyok előállítása	VII.4. Felfedés elleni védelem	VIII.4. Felhasználók támogatása, információszolgálat
I.5. Adatforrások feltérképezése	II.5. Mintavételi terv elkészítése	III.5. Próbafelvétel			VI.5. Aggregálás, becslés, hibaszámítás	VII.5. Elemzés	
I.6. Adatfelvételi javaslattevő készítése	II.6. Adatbegyűjtést követő folyamatok tervezése	III.6. Adatelőállítási folyamat informatikai eszközeinek véglegesítése			VI.6. Feldolgozás lezárása	VII.6. Adatelőállítási folyamat outputjainak véglegesítése	

2. ábra: Egységes Statisztikai Folyamatmodell

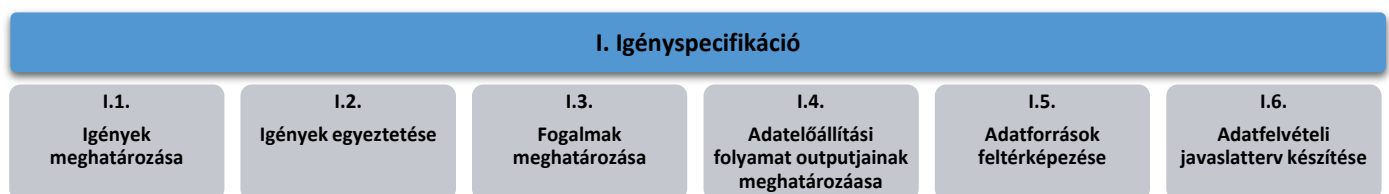
Az Egységes Statisztikai Folyamatmodell (ESTFM) tartalmi leírása

Az alábbi leírás tartalmazza az ESTFM 8 folyamatszakaszának kifejtését (megfeleltetve a KSH-ban fellelhető folyamatszakaszokat a GSBPM v5.0 egyes folyamatszakaszainak), a hazai sajátosságokhoz igazítva.

Az egyes folyamatszakaszok és az azokhoz tartozó részfolyamatok leírását az ESTFM 3 átfogó területének leírása követi: minőségmenedzsment, statisztikai adatmenedzsment, metaadatmenedzsment.

A statisztikai adatelőállítási folyamat folyamatszakaszai és részfolyamatai:

I. Igényspecifikáció



A folyamatszakasz elsődleges célja, hogy összegyűjtse és rendszerezze a statisztikai információk terén jelentkező külső (felhasználói, adatszolgáltatói) és belső (statisztikai adatelőállítási folyamat megvalósításához szükséges) igényeket, illetve döntés szülessen azok szakmai relevanciájáról, valamint az egyeztetett igényekre épülő adatelőállítási folyamat részletes tervezésének megkezdéséről. Elméleti előkészítő folyamatszakasznak tekinthető, melynek során tisztázásra kerülnek a felmerülő külső és belső igények részletes tartalmi elemei. Tartalmazza a jelenlegi outputok felülvizsgálatát, illetve szükség esetén új outputok meghatározását, hogy azok minél jobban ki tudják szolgálni az új és változó igényeket.

Igénynek tekintjük az új statisztikai információk előállítására vonatkozó elvárásokon felül a meglévő statisztikai információk módosítására, részletesebb információk előállítására vagy minőségének (pl. időszerűségének vagy az előállítandó becslések pontosságának) javítására, az információkból való tájékoztatás változására vonatkozó javaslatokat is. A statisztikai információ változatlan tartalma mellett az adatelőállítási folyamat módosítására, hatékonyságának növelésére vonatkozó igények is megjelenhetnek ebben a folyamatszakaszban. Az igényspecifikáció folyamatszakaszban már megjelenik az adatkezelések adatvédelmi vetülete, amely magába foglalja a statisztikai adatok bizalmas kezelését és a személyes adatok védelmét is.

Ebben a folyamatszakaszban a statisztikai szervezet:

- feltérképezi, összegyűjti az igényeket;
- meghatározza a felmerült igényekhez kapcsolódó sokaságot és az azt jellemző ismérveket (fogalmak, mutatók);
- megvizsgálja, hogy az igények kiszolgálhatóak-e a KSH-ban elérhető elsődleges vagy másodlagos adatforrásokból vagy új adatforrások feltárása szükséges-e;

- meghatározza, felülvizsgálja az új igények kapcsolatát más szakstatisztikákkal (pl. vonatkozási kör, gyakoriság, alkalmazott módszertan szempontjából);
- megvizsgálja, hogy az előállítandó statisztikai információk megfelelnek-e az Európai Statisztikai Rendszer által elvárt igényeknek;
- javaslatot tesz az igények kiszolgálásának módjára (adatgyűjtés, adatátvétel, adatfelvételi módszer, adatelőállítási folyamat);
- az adatelőállítási folyamat módosítási igénye esetén megvizsgálja az igény általános, más statisztikai információkra való alkalmazásának lehetőségét.

I.1. Igények meghatározása

A részfolyamat célja, hogy összegyűjtse, leírja az új vagy módosuló külső és belső igényeket (függetlenül attól, hogy az igényt az adatelőállítási folyamat felelőse az eddigi tapasztalatok alapján már ismeri vagy az igények pl. külön egyeztetés keretében merülnek fel) annak megállapításához, hogy milyen statisztikák szükségesek, illetve azok minőségével kapcsolatban milyen igények merülnek fel. Az igény felmerülhet döntéshozói szinten, pl. kormányzati és Európai Statisztikai Rendszer szintű adatigényből, alapulhat hazai vagy nemzetközi jogszabályi előírásokon, illetve felmerülhet más, szakmai szervezetek (pl. szövetségek, kamarák, kutatói közösségek), közvetlen felhasználói visszajelzések (pl. felhasználói véleménykutatás, illetve igényfelmérés) vagy a hivatalon belüli más szakstatisztikák oldaláról is. Ide tartoznak azok az igények is, amelyeket a KSH a folyamatosan értékelt minőségmenedzsment keretében az adott igények kiszolgálásához használt adatelőállítási folyamatok, illetve a hazai és nemzetközi gyakorlatok, módszertanok kapcsán tár fel. Az igény egyben a konkrét statisztikai cél meghatározását is jelenti; különös tekintettel a természetes személyek személyes adatainak kezelése esetén. A részfolyamat a felmerült igények meghatározásával zárul.

I.2. Igények egyeztetése

A részfolyamat célja az összegyűjtött igények egyeztetése az érintettekkel, illetve az igények relevanciájáról szóló szakmai döntés meghozatala. E részfolyamatban történik a felhasználói igények egyeztetése, rangsorolása, prioritási lista meghatározása, illetve a nem releváns igények kizárása. Az igények egyeztetése tartalmazza a jelen gyakorlatban nem teljesülő igények megismerését, a különböző érintettek eltérő igényeinek ütköztetését, a felesleges információk kiszűrését. Itt történik az igényelt statisztikai információk célsokaságának (vonatkozási körének), vonatkozási idejének, illetve az igényelt statisztikai adatok gyakoriságának meghatározása is. Az igények pontos ismerete segít a statisztikai adatelőállítás megfelelő módjának kiválasztásában is.

A részfolyamathoz tartozik az adatelőállítási folyamatban igényelt módosítások egyeztetése a kapcsolódó folyamatszakaszok felelőseivel, illetve a más statisztikai információk adatelőállítási folyamataiban is megvalósítható adat-független megoldási lehetőségének vizsgálata is. A részfolyamat az egyeztetett igények meghatározásával, prioritizálásával zárul.

I.3. Fogalmak meghatározása

A részfolyamat célja az I.2 részfolyamatban meghatározott felhasználói igényekben felmerült sokaságot jellemző ismérvek (fogalmak, mutatók) felhasználók szempontjából történő meghatározása. Nem cél, hogy ezen a szinten az azonosított fogalmak megegyezzenek a jelenlegi gyakor-

latban használt fogalmakkal (a felhasználói igények alapján meghatározott fogalmak és a jelenlegi gyakorlatban használt fogalmak közötti megfeleltetés, illetve a fogalmak, mutatók kiválasztása a II.2. részfolyamatban valósul meg). A részfolyamat az egyeztetett igényekre épülő fogalmak megnevezésével és azok meghatározásával zárul.

I.4. Adatelőállítási folyamat outputjainak meghatározása

A részfolyamat célja, hogy meghatározza azokat az outputokat, melyeket elő kell állítani vagy módosítani kell annak érdekében, hogy kiszolgálják az I.2 részfolyamatban meghatározott felhasználói igényeket. Az outputokra vonatkozó minőségkövetelményeket is itt kell rögzíteni tekintettel az egyes outputok kapcsán alkalmazandó felfedés elleni védelmi megoldásokra is. A részfolyamat az egyeztetett igény szerinti outputok és azok minőségi követelményeinek meghatározásával zárul.

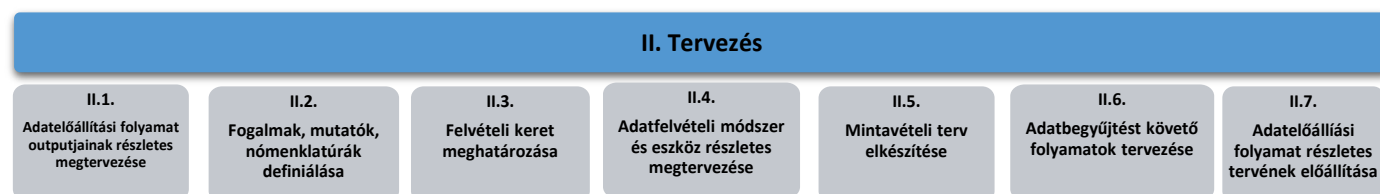
I.5. Adatforrások feltérképezése

A részfolyamat célja az adatforrások felkutatása annak céljából, hogy a megfelelő fogalmakra, sokaságra épülő információk az elvárt minőségben előálljanak. A részfolyamat során minden olyan adatforrást fel kell deríteni, amely rendelkezésre állhat a jóváhagyott igények szerinti statisztikák előállításához (pl. meglévő statisztikai információk felhasználása, statisztikai adatgyűjtésbe való bevonás, másodlagos adatforrások felhasználása). Amennyiben az elvárt információk előállítása új vagy módosuló adatfelvétellel lehetséges, akkor fel kell mérni az adatfelvétel több forrásból való kiszolgálásának lehetőségét is, az adatkörök meghatározásával és a szükséges erőforrásigény becslésével együtt. A részfolyamat keretében meg kell határozni a különböző adatforrások felhasználásának minden olyan alternatíváját, amelyekkel az elvárt igények kiszolgálása biztosítható.

I.6. Adatfelvételi javaslattevő készítése

A részfolyamat célja az információigények alapján új vagy módosuló adatfelvételek indítására vagy elérhető adatforrások másodlagos felhasználására vonatkozó javaslat összeállítása. Az adatfelvételi javaslattevő elősegíti az adatfelvétel részletes tervezésének elindítására irányuló támogató vagy elutasító döntést. Tartalmazza az igényspecifikációs részfolyamatok során felmerült, egyeztetett és jóváhagyott információigényeket, az adatfelvétel megvalósításához szükséges főbb fogalmakat, outputokat, az azokkal kapcsolatos minőségi elvárásokat, valamint a hozzáférhető adatforrásokat (adatköröket) és a felhasználásukhoz szükséges erőforrásigényt is. Az adatfelvételi javaslattevő elfogadása teszi lehetővé a részletes tervezés elindítását. A részfolyamat a statisztikai információkra vonatkozó új igény esetén az adatfelvételi javaslattevő, a statisztikai adatelőállítási folyamat módosítására vonatkozó igények esetén pedig a követelményspecifikációk elkészültével zárul.

II. Tervezés



Ebben a folyamatszakaszban kerül megtervezésre az a statisztikai adatelőállítási folyamat, amelynek felhasználói igényekre támaszkodó átfogó elméleti előkészítése az előző folyamatszakaszban (I. Igényspecifikáció) megtörtént. Új adatfelvétel esetén a teljes adatelőállítási folyamat felépítése, módosuló adatfelvétel esetén pedig jellemző módon a korábbi, már meglévő adatelőállítási folyamat tervének módosítása a feladat. Magában foglalja:

- az adatforrások részletes elemzését;
- az adatelőállítási folyamatban használtfogalmak, nómenklatúrák definiálását;
- a felvételi keretek meghatározását;
- reprezentatív statisztikai adatfelvételek esetén a mintavételi terv elkészítését;
- az adatelőkészítésben, a feldolgozásban, a tájékoztatás előkészítésében és a tájékoztatásban használandó módszertanok tervezését;
- adatfelvételi mérőeszközök tervezését;
- az adatfelvétel tervezéséhez szükséges metaadatok előállítását;
- az adatelőállítási folyamatban létrehozandó outputok definiálását;
- a folyamat munka- és költségtervezését.

A tervezési folyamatszakasz végére rendelkezésre áll a részletes adatfelvételi terv dokumentációja, mely alapján indulhat majd a megvalósítás.

II.1. Adatelőállítási folyamat outputjainak részletes megtervezése

A részfolyamat célja az I.4. részfolyamatban meghatározott outputok tartalmának részletes megtervezése (pl. statisztikai tábla, tárolandó adatok terve, mikroadat, kiadványterv, homogén adatkörök, stb.), valamint az előállításukhoz szükséges eszközök vizsgálata, beleértve a szükséges fejlesztési igények megfogalmazását is. Az output tervezése magába foglalja a személyes adatok esetében az adatvédelmi elveknek megfelelő adattárolás tervezését is. A részfolyamat a konkrét, részletes outputok (pl. logikai adatbázistervek, táblatervek, stb.) megtervezésével zárul.

II.2. Fogalmak, mutatók, nómenklatúrák definiálása

A részfolyamat célja az I.3. részfolyamatban azonosított fogalmaknak megfelelő, illetve azok alapján számítandó és publikálandó statisztikai mutatók tartalmának és előállítási módjának definiálása, valamint azoknak az alap statisztikai mutatóknak és változóknak a meghatározása, melyeket a hivatal különböző adatfelvételi eszközökkel fog begyűjteni. Ezen túlmenően meg kell határozni azokat az osztályozásokat, nómenklatúrákat, melyek az adatelőállítási folyamat során felhasználásra kerülnek. A részfolyamat keretében történik meg az igényként meghatározott fogalmak, mutatók, nómenklatúrák és az adott fogalomhoz, mutatóhoz, nómenklatúrához már korábban

alkalmazott definíciók, értékészletek egyeztetése, megfeleltetése (pl. fordítókulcsok képzése, mutatókapcsolatok definiálása) is. A részfolyamat az adatelőállítási folyamathoz szükséges fogalmak, mutatók, nómenklatúrák definiálásával zárul.

II.3. Felvételi keret meghatározása

A részfolyamat célja az I.2. részfolyamatban egyeztetettekkel összhangban az adatfelvétel célosságának (vonatkozási körének) definiálása, az azt lefedő felvételi keret (mintavételen alapuló felvétel esetén a mintavételi keret), valamint a keret vonatkozási idejének meghatározása.

A részfolyamat keretében történik meg a felvételi keretek kialakításához szükséges források meghatározása (pl. adminisztratív adatforrások, statisztikai regiszterek, illetve nyilvántartások, korábbi teljeskörű felvételek (cenzusok) és reprezentatív felvételek, térképek, stb.). A részfolyamat a felvételi keret definiálásával zárul.

II.4. Adatfelvételi módszer és eszköz részletes megtervezése

A részfolyamat célja, hogy az I.5. részfolyamat során beazonosított adatforrásokat részletesen elemezzük, és az elfogadott célkitűzéseknek és erőforrás-kereteknek megfelelően döntés szülessen az adatfelvétel módjának meghatározásáról (statisztikai adatgyűjtés, adatátvétel). A részfolyamat keretében történik meg annak ellenőrzése is, hogy az adatforrás felhasználására megfelelő jogi alap rendelkezésre áll-e. Amennyiben nem vagy nem megfelelő minőségben áll rendelkezésre jogalap, úgy a megfelelő jogi alap kialakításának kezdeményezése (pl. adminisztratív adatgazdával való együttműködési megállapodás megkötése/kiegészítése) is a részfolyamat része. Önkéntes adatgyűjtés esetén a személyes adatok kezelésénél a tájékoztatás és az érintettek hozzájárulásának tervezését is tartalmazza a részfolyamat. A részfolyamatba tartozik az adatfelvételi módok szerint a mérőeszközök megtervezése. Adatgyűjtés esetén ebben a lépésben történik a kérdőív tartalmi, vizuális megtervezése, az ezt hordozó eszközökkel (pl. laptop, elektronikus önkitöltő alkalmazás) szembeni elvárások összegyűjtése. Adatátvétel esetén pedig ebben a szakaszban történik az adatok körének, struktúrájának, formájának, gyakoriságának, beérkezési határidejének, adatszolgáltatónak/nyilvántartási egységeinek meghatározása a célhoz kötöttség és az adatvédelmi elvek figyelembe vételével. Amennyiben az adatfelvételhez segédanyagok is szükségesek (pl. kitöltési útmutatók, kódkészletek, válaszlapok, demonstrációs eszközök, oktatási segédanyagok, adatszolgáltatást ösztönző eszközök, tájékoztatók, adatvédelmi tájékoztatók, stb.), illetve az adatgyűjtés sikeréhez előzetes tájékoztatás, kommunikáció is szükséges, akkor azok megtervezése is a részfolyamathoz tartozik. A részfolyamat meghatározza az érintett adatfelvétel végleges adatforrásait, adatköröit, az adatfelvétel módját (statisztikai adatgyűjtés, adatátvétel), statisztikai adatgyűjtés esetén az adatgyűjtés módjait (pl. papíralapú kérdőív, postai úton; telefonos kikérdezés, web-alapú adatszolgáltatás; közvetlen megfigyelés, stb.), illetve a kérdőív kérdéseit, kérdésblokkjait, megtervezi az adatgyűjtési eszközt. Másodlagos adatforrás átvétele esetén egyeztet az adatgazdával az adat- és metaadatátvétel köréről, formájáról, gyakoriságáról, az átvétel módjáról (általában együttműködési megállapodásban rögzítve). A részfolyamat az adatforrás nyilvántartó lapok aktualizálásával és az adatgyűjtés-adatátvétel tervének elkészültével zárul.

II.5. Mintavételi terv elkészítése

A részfolyamat célja a II.3. részfolyamatban meghatározott felvételi kereten alapulva a mintavétel módszerének, gyakoriságának, rétegzett mintavétel esetén a rétegeképző ismérveknek, az előállítandó becslések pontossági követelményeinek, illetve a minta nagyságának meghatározása. Beletartozik a különböző statisztikai területek által használt, azonos típusú egységeket tartalmazó minták közötti konzisztencia biztosítása, valamint az átfedések elkerülése az adatszolgáltatói terhek mérséklése érdekében. A meghatározott mintát a IV.1. részfolyamatban kell kiválasztani az e részfolyamatban meghatározott módszertan alapján. A részfolyamat a mintavétel módszertanát tartalmazó mintavételi terv elkészültével zárul.

II.6. Adatbegyűjtést követő folyamatok tervezése

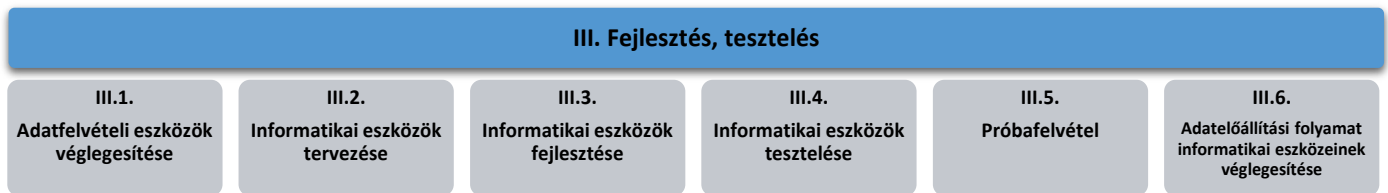
A részfolyamat célja az adatelőkészítésre, feldolgozásra, tájékoztatásra való felkészülés, illetve a tájékoztatás és az archiválás módszereinek megtervezése, melyek az V–VIII. folyamatszaka-szokban, illetve a B. átfogó területen³ kerülnek felhasználásra. Beleértjük a módszertani tervezést és a hozzá kapcsolódó informatikai megoldások tervezését, specifikálását. A részfolyamat jellemző módon magában foglalja az adatbevitel módjának meghatározását, a kódolási, validálási, editálási, pótlási, becslési, hibaszámítási, adatintegrációs, szezonális kiigazítási, felfedés elleni védelmi, személyes adatok esetében az anonimizálási, álnevesítési, illetve minden egyéb, adatállomány véglegesítési eljárás tervezését. A részfolyamat az érintett részfolyamatok specifikációinak elkészítésével zárul.

II.7. Adatelőállítási folyamat részletes tervének előállítása

A részfolyamat célja egy részletes adatfelvételi terv keretében az adat útjának leírása, a minőségbiztosítási terv elkészítése, a teljes adatelőállítási folyamat, felelősségi rend és költségek meghatározása az adatok begyűjtésétől azok archiválásáig. Az adatfelvételi folyamat összeállítása során meg kell vizsgálni az egyes részfolyamatok kapcsolódásait, ismétlődési feltételeit, döntési pontjait. A részfolyamat tartalmazza a személyek rendszerekhez való hozzáférési jogosultságának kérdését, az egyes tevékenységekhez rendelt felelősök, részfelelősök, közreműködők megnevezését, valamint ezeknek a tevékenységeknek a költség- és időigényét, a határidőket, a folyamathoz köthető kockázatok kezelését. A részfolyamat tartalmazza a minőségbiztosítási terv elkészítését, beleértve a munka-standardokat, a hivatalos kötelezettségeket és a minőség-ellenőrzési akciók rendszerét a mintavételi és nem mintavételi hibák kiküszöbölésére, monitorozására és értékelésére, valamint az adatelőállítási folyamat különböző pontjainak irányítására, továbbá jelentős mennyiségű személyes adatok nemzeti szintű kezelése vagy új technológia alkalmazása esetén az adatvédelmi hatásvizsgálatot is. A részfolyamat a részletes adatfelvételi terv összeállításával zárul, melynek részét képezik a II. folyamatszakas részfolyamataiban előállított részletes specifikációk.

³ Lásd „a statisztikai adatelőállítási folyamatot átfogó területek” résznél leírtakat.

III. Fejlesztés, tesztelés



A fejlesztés, tesztelés folyamatszakasz feladata a megtervezett adatfelvételi mérőeszközök elkészítése, az adatelőállítási folyamat háttérét biztosító technikai, informatikai eszközök és rendszerek kialakítása és éles üzem előtti próbája, egymással való kompatibilitásuk vizsgálata, valamint az éles üzemre állítása. Magában foglalja:

- az adatelőállítási folyamat metaadatainak teljes körű rögzítését;
- adatgyűjtés esetén a kérdőívek végleges tervezését, kivitelezését;
- adatfelvételi segédanyagok, tájékoztatók, oktatási anyagok elkészítését;
- az adatelőállítási folyamat informatikai alkalmazásokban való megvalósítását, tesztelését az adatgyűjtési szakasztól a tájékoztatásig;
- az adatelőállítási folyamat egészének tesztelését;
- az adatelőállítási folyamat véglegesítését.

Szintén ebben a folyamatszakaszban kerül sor az adatgyűjtés teljes folyamatára kiterjedő, a terepen zajló próbafelvételre, valamint a tesztelések és próbák tapasztalatait hasznosítva a területi végrehajtás előkészítésére. A részletes adatfelvételi terv alapján a rendszer dokumentálása és a metaadatok rögzítése ebben a folyamatszakaszban készül el.

A fejlesztés, tesztelés szakasz végén az adatfelvételi rendszer működésre kész.

III.1. Adatfelvételi eszközök véglegesítése

A részfolyamat célja a II.4. részfolyamatban rögzített követelményeknek megfelelő eszközök kiválasztása, beszerzése, fejlesztése, tesztelése, illetve az adatfelvételi mérőeszközök és segédanyagok tartalmi tesztelése, véglegesítése. Az adatszolgáltatói teher mérséklése, a begyűjtendő adatok megbízhatóságának és minőségének növelése céljából a tervezett adatfelvételi módok használhatóságának tesztelése és korrekciója is itt történik. A bizalmas adatok kezeléséhez, idértve a személyes adatok kezeléséhez szükséges biztonsági garanciák kiépítése ugyancsak a részfolyamathoz tartozik. A részfolyamat az adatfelvételi mérőeszközök és segédanyagok tartalmi véglegesítésével zárul.

III.2. Informatikai eszközök tervezése

A részfolyamat célja a tervezés során előállított specifikációk alapján az informatikai alkalmazások fizikai rendszertervének elkészítése, a kapcsolódó metaadatok véglegesítése, rögzítése. A részfolyamat a fejlesztéshez szükséges rendszerdokumentáció elkészültével, metaadatok rögzítésével, a szükséges informatikai beszerzések indításával zárul.

III.3. Informatikai eszközök fejlesztése

A részfolyamat célja a II. folyamatszakaszban meghatározott, illetve a III.1. és III.2. részfolyamatokban pontosított feladatok adatelőállítási folyamatot támogató informatikai eszközökben történő megvalósítása. Itt történik az adatbázis objektumok létrehozása, a részfolyamatok feladatainak általános informatikai alkalmazásokban való implementálása vagy új eszközök fejlesztése. Adatgyűjtés esetén jellemzően a kérdőívek tényleges megszerkesztése, illetve az e-kérdőívek informatikai eszközök segítségével történő tényleges megvalósítása a részfolyamat feladata. Az adatelőállítási folyamat működése céljából a különböző eszközök közötti interfész-kapcsolatok működését is biztosítani kell. A részfolyamathoz tartozik a tájékoztatási termékek előállítási eszközeinek megtervezése is a II.1. részfolyamatban meghatározott digitális, illetve papír formátumú outputok minden megjelenési formájáig, beleértve a mikroadatokhoz való hozzáférést is. A részfolyamat végén az adatelőállítási folyamathoz felhasználandó informatikai eszközök tesztelésre készek.

III.4. Informatikai eszközök tesztelése

A részfolyamat célja az adatelőállítási folyamat során alkalmazandó ITeszközök üzembe állítása előtt a feladat jellege szerint szükséges tesztek elvégzése, mint a felhasználói funkcionális teszt (a szakmai specifikációnak megfelelés), IT biztonsági teszt, terheléses teszt, integrációs teszt. Ez a részfolyamat logikailag a III.2. és III.3. részfolyamatokkal áll összefüggésben. A részfolyamat a tesztelési jegyzőkönyvek elkészítésével zárul (a jegyzőkönyveknek tartalmazniuk kell a biztonsági kockázatok felmérését is, sérülékenység vizsgálatot, stb.).

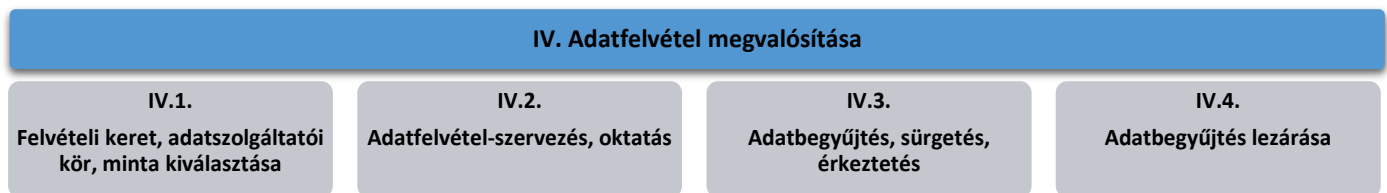
III.5. Próba felvétel

A részfolyamat célja a teljes adatelőállítási folyamat tesztelésének lebonyolítása. Általában a célsokaságnál, illetve a mintánál szűkebb körben végrehajtott adatfelvételt jelent, melynek keretein belül tesztelni lehet az adatfelvétel módjait, technikai támogató eszközeit, az adatok feldolgozását, elemzését, stb. is. A részfolyamat a próba felvétel eredményeinek elemzésével, kiértékelésével, a következtetések levonásával, dokumentálásával zárul.

III.6. Adatelőállítási folyamat informatikai eszközeinek véglegesítése

A részfolyamat célja az informatikai eszközök véglegesítése, a működtetéshez szükséges dokumentációk elkészítése. A részfolyamat keretében kerülnek elkészítésre és kiosztásra az informatikai alkalmazások futtatási dokumentációi, illetve az adatbegyűjtési, adatelőkészítési részfolyamat végrehajtási utasításai. Az informatikai eszközök tesztkörnyezetből éles környezetbe való áthelyezése is itt történik. A részfolyamat végén az adatelőállítási folyamat felhasználásra kész.

IV. Adatfelvétel megvalósítása



Ez a folyamatszakasz az elkészített terveknek megfelelően vonatkozási időszakonként az adatok begyűjtését foglalja magában, mely a statisztikai adatok közvetlen begyűjtésén kívül történhet kombinált módszerek alkalmazásával is (pl. adatleválogatás másodlagos nyilvántartásokból, illetve statisztikai regiszterekből), magában foglalva az adatok elektronikusan elérhetővé tételét, és megfelelő környezetbe (pl. adatbázisba) történő betöltését is. A folyamatszakasz keretében a begyűjtött adatok csak olyan módon kerülhetnek módosításra, amit az adatbenyújtó végez. A begyűjtött adatok bármilyen egyéb javítása, módosítása az Adatelőkészítés folyamatszakaszban történik. Magában foglalja:

- a felvételi keret létrehozását;
- reprezentatív adatfelvételek esetén a mintavételi terv alapján a minta tényleges kiválasztását, az adatszolgáltatói kör meghatározását;
- az adatfelvételek szervezését;
- az adatfelvétel-specifikus szakmai ismeretek oktatását;
- az adatok tényleges begyűjtését;
- az adatbenyújtók sürgetését;
- a begyűjtött adatok érkeztetését;
- az adatszolgáltató által közölt adatok javítás nélküli rögzítését;
- a begyűjtött adatok adatbázisba töltését.

IV.1. Felvételi keret, adatszolgáltatói kör, minta kiválasztása

A részfolyamat célja adatgyűjtés esetén a statisztikai regiszterek adott időpont szerinti („befagyasztott”) állapotára, illetve egyéb forrásokra (pl. adminisztratív adatforrások, nyilvántartások, korábbi felvételek, térképek, stb.) épülve a felvételi keret (vonatkozási kör) létrehozása, elemzése, és az adatszolgáltatói kör, reprezentatív felvétel esetén a minta meghatározása. A minta kiválasztása a felvételi keret elemzését követően a II.3. és a II.5. részfolyamatokban meghatározott tervek szerint történik. Adatátvétel esetén adatgazdánként az adott időszakra szóló adatátvételi feladatok meghatározása történik. A regiszter befagyasztott állapotának, a teljes számjelnak az elkészítése nem a regiszterre épülő adatgyűjtés, hanem az adott regiszter, illetve számjel adatelőállítási folyamatának a része. A részfolyamat statisztikai adatgyűjtés esetén a felvételi keret kiválasztásával, reprezentatív felvétel esetén a minta tényleges kiválasztásával, adatátvétel esetén pedig az adatgazdánkénti adatátvételi feladatok meghatározásával zárul.

IV.2. Adatfelvétel-szervezés, oktatás

A részfolyamat célja az adatfelvételi tervekben rögzített személyi és technikai erőforrások rendelkezésre állásának biztosítása, illetve az adatszolgáltatók, adatgazdák statisztikai adatszolgáltatási kötelezettségükről való értesítése, illetve adatszolgáltatásra való felkérése. Ennek megfelelően megtörténik a III.2. részfolyamatban véglegesített kérdőív és segédanyagainak előállítás, címjegyzékek előállítása, expedálás, szervezők, összeírók adatfelvételhez kapcsolódó szakmai képzése. Összeírás típusú, elsősorban lakossági adatgyűjtésnél ide tartozik a felvétel előzetes kommunikációja. A részfolyamat az adatszolgáltatók tájékoztatásával, illetve összeírás típusú adatgyűjtéseknél az adatfelvételre kész összeírók rendelkezésre állásával zárul.

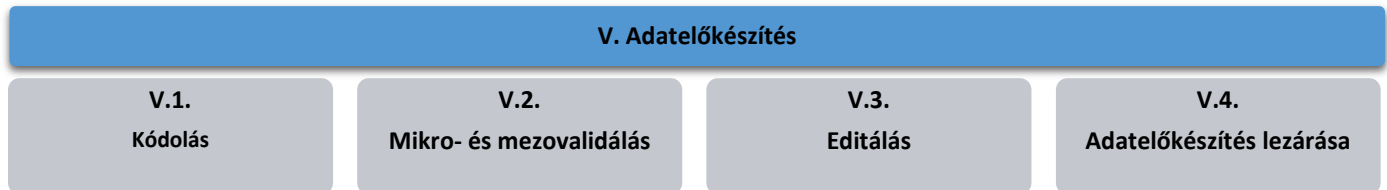
IV.3. Adatbegyűjtés, sürgetés, érkeztetés

A részfolyamat célja a statisztikai adatok és a hozzájuk tartozó metaadatok és paraadatok begyűjtése, illetve az adatszolgáltatók, adatgazdák sürgetése, a begyűjtött adatok érkeztetése. Az adatokat természetes személyek vagy gazdasági szervezetek szolgáltatják, gyűjtésük a tervezés során meghatározott módokon és eszközzel történik, mint pl. az összeírás (interjú, kikérdezés), adatszolgáltatás (papír- vagy web-alapú bevallás), közvetlen megfigyelés vagy adatátvétel. Összeírás típusú adatfelvételek esetén a folyamat a terepen történő összeírást, valamint az érkeztetést foglalja magában. Ennek megfelelően adatgyűjtés során az összeíró által végrehajtott kódolási tevékenység is a részfolyamat része. Önkitöltős adatfelvételek esetén az adatszolgáltatókkal való kapcsolattartást, az érkeztetést és sürgetést a részfolyamat részének tekintjük; az adatrögzítés az elektronikus adatgyűjtések során az adatgyűjtéssel egy időben történik. Elektronikus adatgyűjtés esetén a kérdőíven belüli alapvető összefüggések beépítésre kerülnek az adatgyűjtés eszközébe, ezért az adatszolgáltatónak a hibákat már a kérdőív beküldése előtt javítania kell. Hibás kérdőívek, adatállományok újraküldése (az adatok adatszolgáltató általi javítása) a részfolyamat része. A részfolyamat keretében történik meg a hiányzó kérdőívek vagy adatátvételek okainak kódolása és rögzítése, a hiányzó kérdőívek pótlásra jelölése, valamint az adatszolgáltatást megtagadók szankcionálása is. A részfolyamat az érkeztetés lezárásával végződik.

IV.4. Adatbegyűjtés lezárása

A részfolyamat célja az adatfelvétel során begyűjtött adatok begyűjtéskori állapotának elektronikus formában való eltárolása. A nem elektronikus formában begyűjtött adatokat begyűjtött állapotukban (javítás nélkül) rögzítjük és az adatfelvétel többi begyűjtött adatával együtt, jellemzően adatbázisban tároljuk. Ezzel történik meg az adatbegyűjtés lezárása. A részfolyamat az adatelőkészítésre kész adatállomány rendelkezésre állásával zárul.

V. Adatelőkészítés



Ez a folyamatszakasz az adatelőkészítésre kész mikroadat-állományban lévő szöveges mezők kódolását, a mikroadatok elemi adat szintű ellenőrzését, javítását tartalmazza a mikroadatok feldolgozásra kész állapotának előállítására céljából. A nem elektronikus formában begyűjtött, adatbegyűjtési folyamatszakaszban rögzített adatok ellenőrzése is ebben a folyamatszakaszban történik. A folyamatszakasz eredményeként előáll a feldolgozásra alkalmas adatállomány.

A folyamatszakasz magában foglalja:

- a szöveges mezők kódolását;
- az elemi adatok azon célú ellenőrzését, hogy azok a megtervezett validálási szabályoknak megfelelnek-e;
- a validálási szabályoknak nem megfelelő elemi adatok (pl. outlier-ek) minősítését;
- a hibás adatok tényleges javítását és a javított adatok megjelölését;
- az előkészített, feldolgozásra kész adatok adatbázis-tervezési ajánlásnak (szabványnak) megfelelő szerkezetű adatbázisba töltését.

V.1. Kódolás

A részfolyamat célja a II.6. részfolyamatban megtervezetteknek megfelelően az érintett, begyűjtött adatok különböző nomenklatúrák, osztályozások szerinti kódokkal való ellátása.

Ez megtörténhet automatikusan, vagy végezhetik szakértők is manuálisan, IT eszköz támogatással. Ennek keretein belül, jellemzően a kérdőíveken vagy átvett adatokban a nyitott kérdésekre adott szöveges válaszokat kell megfeleltetni egy adott nomenklatúra elemeinek, vagy az alkalmazott kódkészletet kell a KSH által alkalmazandó kódkészletre átkódolni. Papír alapú adatfelvételnél az összeíró által az adatgyűjtés folyamata során végzett kódolási tevékenység a IV.3. részfolyamat részét képezi. A részfolyamat a kódolt mikroadat-állomány rendelkezésre állásával zárul.

V.2. Mikro- és mezovalidálás

A részfolyamat célja annak vizsgálata, hogy a rendelkezésre álló elemi adatok megfelelnek-e a kitűzött validálási szempontoknak. Mikrovalidálás során több adatforrást is felhasználhatunk az elemi adatok érvényességének vizsgálatára. A vizsgálat mindig az elemi adatok szintjén zajlik. Ennek során történik a kódolás ellenőrzése, az elemi adatok adatforráson belüli és adatforrások közötti összefüggésének vizsgálata. Az adatok összehasonlítása, ellenőrzése azonos, illetve más (pl. statisztikai, adminisztratív) adatforrásokból, regiszterekből származó adatokkal vagy előző időszaki adatokkal. Ebben a részfolyamatban történik meg többek között az outlier-ek azonosítása, valamint a hibás kódolás kezelése is. Ide sorolható a hivatalban alkalmazott mezovalidálási eljárás is, amely az elemi szintű adatok mikrovalidálásnál átfogóbb, több adatfelvétel alapján végzett összefüggés-ellenőrzését jelenti. A részfolyamat a feltárt hibák összegyűjtésével és azok súlyosság szerinti minősítésével és a validált mikroadat-állomány rendelkezésre állásával zárul.

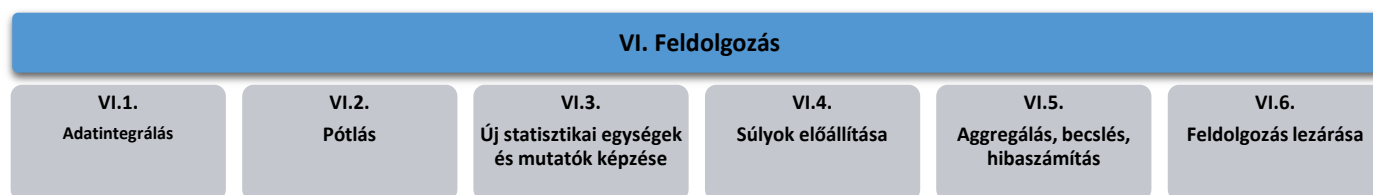
V.3. Editálás

A részfolyamat célja a validálás során feltárt hibák tényleges kezelése, javítása meghatározott szabályok szerint, automatikus vagy manuális eljárások segítségével. A hibák javítása történhet adatszolgáltatótól, összeírótól érkezett javított kérdőív, szóbeli egyeztetés, illetve ennek hiányában szakmai szempontok alapján is. A részfolyamat a hibás értékek javításával, illetve a javított adatok megjelölésével és az editált mikrodát-állomány rendelkezésre állásával zárul.

V.4. Adatelőkészítés lezárása

A részfolyamat célja az ellenőrzött, javított, adatok eltárolása a minősítésükkel együtt. A részfolyamat keretében kell gondoskodni az állomány metaadatbázisban leírt adatainak, adatbázis-tervezési szabályoknak megfelelő szerkezetben való tárolásáról is. Ezzel történik meg az adatelőkészítés lezárása. A részfolyamat az ellenőrzött, elfogadott adatokat tartalmazó, feldolgozásra kész (kódolt, mikrovalidált, editált mikrodát) állomány rendelkezésre állásával zárul.

VI. Feldolgozás



A feldolgozási folyamatszakasz célja, hogy előálljon minden olyan adatbázisba szervezett, feldolgozott adat, amely a tájékoztatási termékek előkészítéséhez szükséges, illetve amely további, ún. másodlagos feldolgozások inputjává szolgál. A folyamatszakasz inputját az adatelőkészítésen átesett, illetve feldolgozásra kész adatállományok, illetve más feldolgozási folyamatok eredményei alkotják. A folyamatszakasz magában foglalja:

- különböző adatállományok összekapcsolását;
- üres cellaadatok egység-, illetve tételszintű pótlását;
- rendelkezésre álló információk alapján új mutatók, indexek, új statisztikai egységek létrehozását;
- becsléshez szükséges súlyok előállítását;
- becslések és hozzájuk tartozó hibák kiszámítását;
- aggregátumok képzését;
- az adatok minősítését (javított, pótoltt, stb.).

A folyamat a feldolgozott adatokat tartalmazó másodlagos feldolgozásra és tájékoztatás előkészítésére kész adatállományok előállításával zárul.

VI.1. Adatintegrálás

A részfolyamat célja az egy vagy több forrásból származó adatállományok összekapcsolása. A részfolyamat inputja lehet szervezeten belüli, illetve azon kívüli adatállomány, vagy ezek együttese, kombinációja. Amennyiben az integrálni kívánt adatforrások egységei részben átfedik egymást és rendelkeznek egyedi, korrekt azonosítókkal vagy kulcsváltozók alkalmas kombinációjával,

akkor ezeken keresztül történik az adatintegráció. Ha nincs ismeretünk arról, hogy van-e átfedés az adatforrások egységei között, továbbá vannak olyan változók, mutatók, amelyek nem lettek megfigyelve mindkét adatforrásban, akkor közös változók felhasználásával kapcsoljuk össze a leginkább hasonló egységekre vonatkozó rekordokat. Az integrálás az adatvédelmi elvek figyelembe vételével történik. Személyes adatok esetén ide tartozik az adatok anonimizálása is. A részfolyamat az érintett állományok összekapcsolásával zárul.

VI.2. Pótlás

A részfolyamat célja a feldolgozást megelőző folyamatszakaszokban pótlásra jelöltek, illetve a II.6. részfolyamat során rögzítetteknek megfelelően az üres cellák adatokkal való feltöltése. Ide tartozik az egység szintű és a tétel szintű hiányzó értékek pótlása. A részfolyamat a pótlandó értékek tényleges pótlásával, illetve a pótolta adatok megjelölésével zárul.

VI.3. Új statisztikai egységek és mutatók képzése

A részfolyamat célja azon változók és statisztikai egységek előállítása, melyek az adatok begyűjtésével közvetlenül még nem állnak rendelkezésre (pl. fajlagos mutatók, képzett (számított) mutatók, feldolgozási eredményként előállítandó indexek, indexsorok), de az előre meghatározott output létrehozásához elengedhetetlenül szükségesek. A tevékenység elvégzése során az adattállományban lévő változókból kerülnek előállításra a szükséges mutatók egy előre meghatározott algoritmus, eljárás, képzési szabály szerint, a II.6. részfolyamatban rögzítetteknek megfelelően. Az új statisztikai egységek létrehozása történhet a beszámolási egységek aggregálásával vagy becslési eljárással. Példaként szolgál erre a háztartási szintű adatok létrehozása azokban az esetekben, amikor az adatgyűjtés megfigyelési egysége a természetes személy; vagy a megfigyelési egység a jogi személy, de a tájékoztatási egység a vállalatcsoport. A részfolyamat az új statisztikai egységek és mutatók előállításával zárul.

VI.4. Súlyok előállítása

A részfolyamat célja a statisztikai egységekhez tartozó súlyok előállítása. A végső súlyok kialakítása magában foglalja a nemválaszolás figyelembe vételét, a kiugró értékek kezelését, az ismert sokasági értékekhez való igazítást (kalibrálást), valamint a súlyok évek közti változásának kezelését, konverzióját (homogenizálását) is. A súlyokat a teljes sokaságra vonatkozó becslés előállításához használják fel. A részfolyamat a végleges súlyok rendelkezésre állásával zárul.

VI.5. Aggregálás, becslés, hibaszámítás

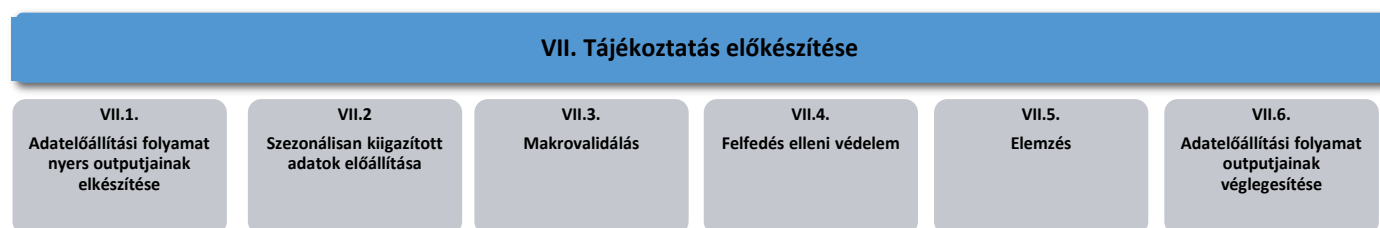
A részfolyamat célja az elemi (mikro) adatok alapján előállítandó aggregált adatok, illetve „mindösszesen”-ek képzése, a becslések előállítása, valamint azok hibájának kiszámítása a II.6. részfolyamatban meghatározott módszerek szerint. Az aggregálás történhet valamely csoportosító ismérv alapján (pl. területi dimenzió mentén), illetve időben (pl. havi adatok összesítése negyedévre, évre). A becsléseket a végleges súlyok segítségével állítjuk elő.

Itt számítjuk a becslések hibáját is. A részfolyamat az aggregált értékek, indexek, becslések és azok hibájának rendelkezésre állásával zárul.

VI.6. Feldolgozás lezárása

A részfolyamat célja a feldolgozott adatok eltárolása (beleértve az előállított súlyokat, hibákat és minősítéseket is). Ezzel történik meg a feldolgozás lezárása. A részfolyamat a feldolgozott adatokat tartalmazó másodlagos feldolgozásra és tájékoztatás előkészítésére kész, adatvédelmi szabályoknak megfelelő állományok rendelkezésre állásával zárul.

VII. Tájékoztatás előkészítése



A folyamatszakasz célja, hogy a feldolgozott adatok alapján előkészítse a tájékoztatásra szánt termékeket a II.1. pontban specifikált outputok figyelembe vételével. Ebben a folyamatszakaszban a tájékoztatásra szánt adatok állnak elő, melyek teljes mélységükben ellenőrzésre kerülnek, így közölhetővé válnak. A folyamatszakasz magában foglal minden olyan tevékenységet, mely a külső és belső felhasználók számára lehetővé teszi az előállított adatok megértését.

Ennek érdekében a folyamatszakasz keretében történik:

- az adatelőállítási folyamat első, nyers outputjainak előállítása;
- szezonális kiigazítás alkalmazása esetén a kiigazított adatok előállítása;
- az előállított információk makro szintű validálása;
- a végleges tájékoztatási termékek előállításához szükséges felfedés elleni védelmi módszerek alkalmazása;
- a végleges tájékoztatási termékek előállításához szükséges eredmények értelmezése, szöveges elemzése és kiértékelése;
- a tájékoztatási célú leíró metaadatok előállítása.

A folyamatszakasz végén állnak elő a végleges tájékoztatási tartalmak.

VII.1. Adatelőállítási folyamat nyers outputjainak elkészítése

A részfolyamat célja, hogy a rendelkezésre álló adatokból létrejöjjön az adatelőállítási folyamat outputjainak első nyers változata a II.1. részfolyamatban megtervezett formában. Emellett magában foglalja a nyers outputok kiértékeléséhez szükséges minden segédinformáció (pl. indexek, trendek, minőségre vonatkozó jellemzők) előállítását is. A részfolyamatban történik meg az adatok adattárházba töltése, valamint a tájékoztatásra szolgáló adatbázisokba való töltés előkészítése. A részfolyamat a nyers outputok és azok kiértékeléséhez szükséges segédinformációk előállításával zárul.

VII.2. Szezonálisan kiigazított adatok előállítása

A részfolyamat célja a rögzített paramétereknek megfelelő, szezonálisan kiigazított idősorok előállítása a II.6. részfolyamatban meghatározott módszerek szerint. Ide tartozónak tekintjük a szezonálisan kiigazított idősorok módszertani ellenőrzését, illetve szükség esetén az idősorokon lévő modellek javítását. A részfolyamat az elfogadott, szezonálisan kiigazított adatok előállításával zárul.

VII.3. Makrovalidálás

A részfolyamat célja annak vizsgálata, hogy a rendelkezésre álló aggregált adatok megfelelnek-e a II.6. részfolyamatban rögzített makroszintű validálási szempontoknak. A részfolyamat keretében azt vizsgáljuk, hogy az adott szakstatisztikával kapcsolatos információk (pl. tükörstatisztika, más szakstatisztikák releváns információi) alapján az outputok megfelelnek-e a validálási szempontoknak. A makrovalidálási tevékenység tipikusan az alábbiakat tartalmazza:

- a lefedettség arány és a válaszadási hányad vizsgálata;
- a statisztikai adatok összevetése egy korábbi időszak adataival, idősor-elemzés;
- a statisztikai adatok összevetése más adatforrásokból származó adatokkal;
- inkonzisztenciák feltárása más statisztikákkal;
- a statisztikai adatok összevetése a várható értékekkel.

A részfolyamat a feltárt hibák összegyűjtésével és azok javítási módjáról, korábbi folyamat-szakaszra (adatelőkészítés, feldolgozás) való visszatérésre vonatkozó döntéssel zárul.

VII.4. Felfedés elleni védelem

A részfolyamat célja az előállított statisztikai outputok felfedés elleni védelmi szempontú ellenőrzése és védelme a II.6. részfolyamatban meghatározott módszerek szerint (pl. táblázatos adatok védelme, anonimizált mikroadatok előállítása, az adatállományok biztonságos környezetbeli használatra történő előkészítése, nyilvános mikroadat-állományok felfedés elleni védelme), valamint a biztonságos környezetben előállított kutatási eredmények utólagos adatvédelmi szempontú ellenőrzési feladatainak (output checking) ellátása. A részfolyamat a felfedés elleni védelemmel ellátott outputok előállításával, illetve kutatási eredmények ellenőrzésének befejezésével zárul.

VII.5. Elemzés

A részfolyamat célja az adatelőállítási folyamat outputjainak értelmezése, szöveges elemzése, grafikus ábrázolása, megjelentetésre előkészítése a tájékoztatási szempontok és a különböző tájékoztatási csatornák eltérő igénye szerint. A részfolyamat feladata az adatelőállítási folyamat információinak felhasználásával annak kiértékelése (szöveges leírása) is, hogy az előállított outputok mennyire felelnek meg az induló elvárásoknak. A részfolyamat része az adatelőállítási folyamat és az előállított termékek minőségértékelése is. A részfolyamat az adatelőállítási folyamat outputjaihoz tartozó elemzések, szakstatisztikai információk, leíró metaadatok és minőségjelentés elkészítésével zárul.

VII.6. Adatelőállítási folyamat outputjainak véglegesítése

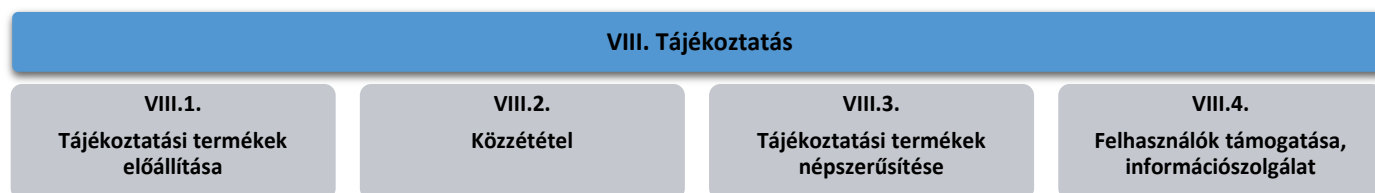
A részfolyamat célja annak megállapítása, hogy a statisztikai adatok és a hozzájuk kapcsolódó metaadatok a kitűzött céloknak, terveknek megfelelnek-e, teljesítik-e a minőségi követelményeket, tehát felhasználásra alkalmasak-e.

Ez magában foglalja:

- az előállított végleges outputok konzisztencia-vizsgálatát;
- a közölhetőség szintjének meghatározását;
- a belső használatra szánt dokumentációk véglegesítését;
- a folyamat- és termékminőség jelentések elfogadását;
- a tájékoztatásra szánt termék végső tartalmi egyeztetését a megfelelő szakstatisztikai területekkel;
- a tájékoztatási termék tartalmának jóváhagyását;
- a jóváhagyott adatok adattárházban való véglegesítését.

A részfolyamat a tájékoztatásra szánt termék tartalmának, valamint a kapcsolódó elemzéseknek és leíró információknak (statisztikai adatok és metaadatok) a véglegesítésével zárul.

VIII. Tájékoztatás



Ebben a folyamatszakaszban a megfelelő tájékoztatási csatornák útján és közlési formák alkalmazásával eljuttatjuk a statisztikai eredményeket (adatokat) a felhasználókhoz a különböző tájékoztatási csatornáknak megfelelő adatvédelmi szabályok figyelembe vételével.

Jellemző tevékenységek:

- tájékoztatási termékek (elemzések, adatösszeállítások, grafikonok) szerkesztése;
- a közlésre szánt tájékoztatási adatok kész tábláinak honlapra, webes megjelenítő alkalmazásba, Tájékoztatási adatbázisba töltése;
- adatátadások teljesítése, azaz a tájékoztatási termékek átadása külső és belső partnereknek;
- az előállt tájékoztatási termékek népszerűsítésével kapcsolatos feladatok;
- felhasználók számára nyújtott információszolgálati tevékenységek ellátása.

VIII.1. Tájékoztatási termékek előállítása

A részfolyamat célja a II.1. részfolyamatban rögzített és a VII.6. részfolyamatban véglegesített tényleges tájékoztatási termékek közzétételének előkészítése, szerkesztése a tájékoztatásra kész adatok alapján. A tájékoztatási termék típusától függően ide tartozik a közzétételre szánt adatok, metaadatok megjelenítő felületre, webes alkalmazásba, Tájékoztatási adatbázisba töltése, az online és a nyomtatott termékek (pl. gyorstájékoztatók, elemzések, évkönyvek, médiaanyagok, közlemények, térképek, adatvizualizációk) szerkesztése, megjelenítése, illetve az adatkérési igényeknek adott adathozzáférési csatornának megfelelő kiszolgálásához szükséges adatállományok előállítása. A részfolyamat a tájékoztatási termékek tényleges előállításával, hozzáférhetővé tételével zárul (hozzáférhetővé tehető adatok).

VIII.2. Közzététel

A részfolyamat célja a tájékoztatási termékek rendelkezésre állásának biztosítása, tényleges közzététele; beleértve a publikálás időpontját is. Magában foglalja a publikálás előtt álló embargós kiadványok előzetes hozzáférhetővé tételét (sajtószoba, illetve előzetes kiküldés a hatályos belső szabályok szerint). Ennek keretében szükséges biztosítani a tájékoztatási termékek partnereknek, adatigénylőknek történő átadását, az adathozzáférési csatornákon történő tényleges adathozzáféréseket, illetve a tájékoztatási csatornák fenntartásával (tartalomszolgáltatással) összefüggő feladatokat is. A részfolyamat a tájékoztatási termékek felhasználók számára való elérhetőségének biztosításával zárul (közzétett adatok).

VIII.3. Tájékoztatási termékek népszerűsítése

A részfolyamat célja, hogy az adatelőállítási folyamat során előállított tájékoztatási termékek és azok hivatali szolgáltatásokon keresztüli elérhetősége a felhasználók lehető legszélesebb körében váljanak ismertté. Magában foglalja az egyes felhasználói csoportok oktatását (pl.: felhasználói fórumok keretein belül), rövid tájékoztatását. A tájékoztatási termékek népszerűsítését a média és marketing széles és hatékony eszközrendszerére támaszkodva kell megvalósítani a hivatal honlapjára, mint a legfőbb tájékoztatási és kommunikációs eszközre építve. A részfolyamat a tájékoztatási termékek népszerűsítési akcióinak végrehajtásával zárul, melynek a sajtóval való kommunikáció mellett fontos részét képezi a közösségi médiában való megjelenés.

VIII.4. Felhasználók támogatása, információszolgálat

A részfolyamat célja a felhasználói adat- és információigények rögzítése és határidőre, elvárt minőségben történő teljesítése. Az információszolgálatot ellátó szervezeti egységek tevékenysége, illetve a szakértők által nyújtott szakértői támogatás hivatott biztosítani ezt a célt. A részfolyamat keretében kell döntést hozni az egyes felhasználói igények kiszolgálásáról (pl. rendelkezésre álló adatok alapján teljesíthető, az adatok alapján új termék létrehozása szükséges vagy az igény nem kiszolgálható). A részfolyamat az felhasználói igények kiszolgálásával, a felhasználói elégedettségméréssel, illetve az adatelőállítási folyamat feldolgozási vagy tájékoztatás-előkészítési részfolyamataiba, valamint következő időszak tervezési folyamataiba való visszacsatolással zárul.

A statisztikai adatelőállítási folyamatot átfogó területek

Minőségmenedzsment (A.)

A. Minőségmenedzsment

A minőségmenedzsment célja, hogy kereteket alakítson ki és működtessen a statisztikai termékek, statisztikai adatelőállítási folyamatok kiváló minősége, valamint az ezeket támogató szervezeti keretek biztosítása érdekében. Mindennek végső célja a felhasználói igények minél teljesebb körű kiszolgálása, hiszen a statisztikai termékek minőségét alapvetően a felhasználói igényekből kiindulva definiáljuk: „A minőség az egység (termék vagy szolgáltatás) azon jellemzőinek összessége, amelyek befolyásolják azt a képességét, hogy meghatározott és elvárt igényeket kielégítsen.” (MSZ EN ISO 8402 – 1996). A termékek a statisztikai adatelőállítási folyamatának eredményeként állnak elő, a termékminőségre tehát jelentős befolyással bír a statisztikai adatelőállítási folyamat, illetve a folyamatminőség. Az egyes folyamatszakaszok minőségéhez kapcsolódó elvárásainkat a minőségi irányelvek írják le.

A minőség vonatkozhat a szervezetre, a folyamatokra és a termékekre. A minőségmenedzsment keretrendszere a folyamatos fejlesztés egyik eszközének, az ún. PDCA⁴-ciklusnak a logikájára épül. A PDCA-ciklus értelmében felállítunk egy, a minőségre vonatkozó tervet, illetve az elvárásokat alapul véve célokat tűzünk ki annak érdekében, hogy minél kiválóbb termék- és/vagy folyamatminőséget érjünk el (pl. jogszabályok, standardok, minimumkritériumok a termékminőség-összetevőkre, felhasználó-specifikus elvárások, minőségi irányelvek). Ezután végrehajtjuk a statisztikai adatelőállítási folyamatot, közben dokumentálunk és mérünk (pl. minőségjelentések, minőségindikátorok, felhasználói elégedettségmérés). Ezután összevetjük a terveket a dokumentációval és értékeljük a terveknek való megfelelést (pl. önértékelés, belső-külső audit, felhasználó-szakértő fórum). Ha az értékelés alapján szükséges, fejlesztési intézkedéseket határozunk meg, melyek beépülnek a következő ciklus elvárásaiba.

Statisztikai adatmenedzsment (B.)

B. Statisztikai adatmenedzsment

A statisztikai adatok kezelése a statisztikai adatelőállítási folyamat valamennyi folyamatszakaszát átfogó szabályozott területe. A statisztikai adatvagyon áttekinthetősége érdekében az adatbázisba töltésre, névkonvencióra, az adatbázistáblák (állományok) szerkezetére, a karbantartás, hozzáférés jogosultságára vonatkozóan egységes elvek alkalmazása szükséges. Az adatkezelésnek kell biztosítani a hozzáférés hatékonyságát, és az adatok konzisztenciájának fenntartását is. A statisztikai adatvagyon részben a metaadatbázisból, részben az alkalmazott szoftvereszközök katalógusaiból (adatszótáraiból) ismerhető meg és követhető nyomon. Az adott adatelőállítási folyamat

⁴ PDCA – plan – tervezés, do – cselekvés, check – ellenőrzés, act – beavatkozás.

adatkezelésének előkészítése a III.1-es részfolyamatban az adatelőállítási folyamat fizikai tervének elkészítésével és az adatbázistáblák létrehozásával történik. Az adatok kezelése, elérése pedig a IV–VIII. folyamatszakaszok átfogó, valamennyi lépést érintő területe.

Az adatok tárolásának, kezelésének, hozzáféréseinek tervezése, működtetése során fontos szempont az adatvédelmi elvek érvényesítése. Ez különösen fontos a személyes adatok kezelése során, ahol az azonosítók kezelésére, felfedés elleni védelmére, az azonosítókat jellemző ismérvek azonosítókkal való kapcsolatának biztosítására eltérő szabályok vonatkoznak az egyéb mikroadatokat és aggregált adatok kezeléséhez képest (pld. természetes és mesterséges azonosítók külön tárolása, adatintegritás szabályai, stb.).

Az egységes statisztikai adatelőállítási folyamat része az archiválás. Az archiválás rendjét archiválási politika és az arra épülő egységes archiválási szabályzat határozza meg, amely javaslatot tesz a különböző típusú eredménytermékek (gyakoriságtól, feldolgozottsági foktól függő) online elérési időszakot követő archiválására. Az eredménytermékek közül az adatelőállítási folyamat inputját, az adatelőkészítés eredményét (azaz a feldolgozási szakasz inputját) és a feldolgozási folyamat végén előálló kész adatokat (elemi és aggregált adatokat), esetleg a tájékoztatásra kész adatokat kell archiválni. Az archiválás magában foglalja mind az adatok, mind pedig az értelmezésükhöz szükséges egyéb információk, metaadatok, dokumentációk, programok tárolását. Az adott folyamathoz kapcsolódó archiválási tervet a tervezés során a II.6. részfolyamatban kell megalkotni, figyelembe véve többek között az archiválásra szánt adatok és metaadatok formátumát, rendelkezésre állásának idejét, az archívumba töltés menetét, az archivált adatok katalogizálását, az adatok archívumból való visszanyerésének módját. Maga az archiválás így több, a tervezés során meghatározott folyamatszakaszhoz is kapcsolódik. A folyamat lezárásakor vagy bizonyos, előre meghatározott idő múltán az archiválás megtörténtének ellenőrzését is végre kell hajtani.

Metaadat-menedzsment (C.)

C. Metaadat-menedzsment

A metaadat-menedzsment mint átfogó terület célja a megfelelő metaadatok biztosítása az adatelőállítási folyamat tervezése és végrehajtása során. A metaadat-menedzsment a metaadatok mindhárom típusának (leíró, strukturális, vezérlő) kezelését magában foglalja. A metaadatok minden egyes folyamatszakaszban megjelennek függetlenül attól, hogy az adott folyamatszakaszban álltak-e elő, vagy az előző folyamatszakaszból kerülnek felhasználásra.

A metaadatok menedzsmentje a leíró, strukturális és vezérlő metaadatok megtervezését, elkészítését és karbantartását jelenti, amelyek a folyamatszakaszok, részfolyamatok leírásához és vezérléséhez adnak paramétereket.

A folyamatba beleértendő a metaadatok kezelésének egységes szabályozása is. A szükséges metaadatok az érintett részfolyamattal egyidőben, ahhoz kapcsolódóan állnak elő és a további szakaszokban ezek kerülnek felhasználásra. A folyamat az adatelőállítási folyamathoz szükséges megfelelő minőségű metaadat, valamint a metaadatok menedzsmentjére vonatkozó szabályozók biztosításával tekinthető teljesnek.

Melléklet – Fogalmak

A melléklet az ESTFM-ben használt kifejezések fogalmi meghatározásait tartalmazza az ESTFM közös nyelvezetének megteremtése, használata céljából. A fogalmak mögött a metaadatbázisban már szereplő fogalmak azonosítószáma szerepel.

- **Adatelőkészítés folyamatszakasz:** Az Egységes Statisztikai Folyamatmodell (ESTFM) ötödik folyamatszakasza, amely során az adatgyűjtés vagy adatátvétel segítségével begyűjtött adatok kódolása, ellenőrzése és javítása után rendelkezésre áll az elfogadott adatokat tartalmazó, feldolgozásra kész adat.
- **Adatelőkészítés:** A statisztikai célra begyűjtött, illetve a más intézménytől átvett adatok rögzítése, ellenőrzése, javítása.
- **Adatfelvétel eszköze:** Összefoglalja az adatfelvétel végrehajtásához szükséges technikai eszközöket, beleértve a szerkesztett kérdőívet (e-kérdőív vagy papír) és a kikérdezés eszközt is (pl. laptop, PDA, stb.).
- **Adatfelvétel megvalósítása folyamatszakasz:** Az Egységes Statisztikai Folyamatmodell (ESTFM) negyedik folyamatszakasza, amely során az adatfelvételek vonatkozási időszakaikra az adatfelvételi módszerek megfelelően összegyűjtik a felvételi sokaságot jellemző adatokat.
- **Adatfelvételi javaslat:** Az adatfelvétel tervezés kiinduló dokumentuma, döntéselőkészítő anyag egy adatfelvétel indításhoz, módosításához. Javaslatot tesz az új vagy módosuló felvétel céljára, tartalmára, a felvétel végrehajtásának módjára, outputjára, költségére.
- **Adatfelvételi terv:** Az adatfelvétel-tervezés eredménye, amely meghatározza az adatfelvétel tartalmát, módját, eszközét, specifikálja a begyűjtött adatok ellenőrzésének, feldolgozásának, tájékoztatásának feladatát.
- **Adatforrás:** A statisztikai adatfelvétel megvalósításának, a megfigyelendő adatok begyűjtésének forrása. A statisztikai megfigyeléshez szükséges adatok összegyűjtése több adatforrásból történhet.
- **Adatgyűjtés-szervezés:** A statisztikai adatelőállítási folyamatnak az adatfelvétel tervezését követő, az adatok begyűjtésével foglalkozó végrehajtási szakasza. Az adatgyűjtés-szervezés feladata a felvételi keret (vonatkozási kör) meghatározása, mintavétel, az adatszolgáltatói kör kijelölése, a nyomtatványok előállítása, adatszolgáltatóhoz/összeíróhoz eljuttatása (expediálás, pótexpediálás), címbejárás, a kérdőívek beérkezésének regisztrálása, a hiányzó kérdőívek sürgetése, a hiány okának regisztrálása, az adatgyűjtésben részt vevők munkájának szervezése, az adatgyűjtés-szervezés folyamatának monitorozása. A begyűjtött adatokat az adatgyűjtés-szervezés az adat-előállítási folyamat következő szakaszának, az adatelőkészítésnek, a hiányinformációkat az adatfeldolgozásnak adja tovább.
- **Adatintegrálás:** Két vagy több forrásból származó adatok összekapcsolásának eljárása statisztikai output előállítása céljából.
- **Adatkör:** A statisztikai adatfelvétel által megfigyelt ismérvek egy adott témára és adott megfigyelési szintre vonatkozó, azonos logikai egységbe tartozó csoportja.
- **Adatminősítés:** Az adatok előkészítése és feldolgozása során az adatokat minősítésekkel láthatjuk el, melyek az adatok további feldolgozása, minőségének értékelése során fontos információt jelentenek. Ilyen például, hogy egy adat javított, pótol, outlier, védendő.

Minősítéssel el lehet látni az adattábla egy-egy sorát (egység szintű minősítés), illetve egy-egy celláját (tétel szintű minősítés) is.

- Adatszolgáltatói kör: A statisztikai adatfelvételi rendszerben a vonatkozási kör adatszolgáltatásra kijelölt azon egységeinek összessége, amelyekről a vizsgált statisztikai, beszámolási és megfigyelési egységekre vonatkozó, a célsokaságot jellemző adat nyerhető.
- Adminisztratív adatforrás: Olyan másodlagos adatforrás, amelyben tárolt adatok gyűjtését vagy azokról nyilvántartás vezetését jogszabály írja elő az adminisztratív adatforrás kezelője számára.
- Anonimizálás: Olyan eljárás, amellyel az érintett statisztikai egység (jellemzően természetes személy vagy gazdálkodó szervezet) nem vagy többé nem azonosítható.
- Archiválás: Tervszerűen megvalósuló, szabályozott tevékenység és folyamat, mely során a statisztikai adatelőállítási folyamatok során előálló végleges outputok tárolása előre meghatározott, jellemzően hosszú ideig visszakereshető és visszaállítható módon, biztonságosan, az eredménytermékek elsődleges keletkezési helyétől fizikailag elkülönített helyen történik.
- Álnevesítés: A személyes adatok olyan módon történő kezelése, amelynek következtében további információk felhasználása nélkül többé már nem állapítható meg, hogy a személyes adat mely konkrét természetes személyre vonatkozik, feltéve hogy az ilyen további információt külön tárolják, és technikai és szervezési intézkedések megtételével biztosított, hogy azonosított vagy azonosítható természetes személyekhez ezt a személyes adatot nem lehet kapcsolni.
- Becslés: Egy ismeretlen sokasági paraméter értékének közelítése a teljes sokaság egy mintájára vonatkozó információk alapján.
- Célsokaság: Azon egységek összessége, melyek jellemzőiről információt kell nyújtani, becslést kell készíteni.
- Editálás: Statisztikai tevékenység az adatok, és az adatok összefüggése hibájának felderítésére, hihetőségének vizsgálatára és a hibák javítására.
- Elsődleges adatforrás: Olyan adatforrás, amelyből a statisztikai adatfelvétel által leírt sokaság adott időszakra vagy időpontra vonatkozó ismérvei teljes mértékben vagy részben statisztikai adatgyűjtéssel meghatározhatók.
- ESTFM átfogó terület: A folyamatszakaszok mindegyikére kiterjedő általános tevékenység, melynek részei: minőségmenedzsment, statisztikai adatmenedzsment, metaadat-menedzsment.
- ESTFM folyamatszakasz: A statisztikai adatelőállítási folyamat részét alkotó egyik fő lépés, tevékenység. A folyamatszakaszok az alábbiak: igényspecifikáció; tervezés; fejlesztés, tesztelés; adatfelvétel megvalósítása; adatelőkészítés; feldolgozás; tájékoztatás előkészítése, tájékoztatás.
- ESTFM részfolyamat: A folyamatszakasz azon egysége, melynek keretében a többi részfolyamattól különböző, meghatározott outputot áll elő.
- Egységes Statisztikai Folyamatmodell (ESTFM): Az ENSZ statisztikai egysége (UN Economic Commission for Europe: UNECE) által létrehozott GSBPM (Generic Statistical Business Process Model – Általános Statisztikai Folyamatmodell) magyar adaptációja, a teljes statisztikai adatelőállítási folyamatra vonatkozó folyamatmátrix.

- **Érkeztetés:** Az adatszolgáltatói kör és a ténylegesen beérkező kérdőívek összevetése. Szűkebb értelemben a beérkező kérdőíveket az érkeztetés regisztrálja. Tágabb értelemben az érkeztetés magában foglalja a hiányzó kérdőívek adatszolgáltatóinak megkeresését, sürgetését, a hiány okának felderítését és regisztrálását is.
- **Fejlesztés, tesztelés folyamatszakasz:** Az Egységes Statisztikai Folyamatmodell (ESTFM) harmadik folyamatszakasza, amely során az adatfelvételhez, a statisztikai adatok előállításához szükséges eszközök (kérdőívtől, illetve adatátvételi tervtől az alkalmazási rendszereig) elkészülnek, tesztelésük, dokumentálásuk után működőképes az adatfelvételi folyamat.
- **Feldolgozás folyamatszakasz:** Az Egységes Statisztikai Folyamatmodell (ESTFM) hatodik folyamatszakasza, amely során az feldolgozásra kész adatokból tájékoztatásra kész adatokat állítunk elő, az adatok összekapcsolása, pótlása, becslése, aggregálása, új mutatók, új statisztikai egységek révén.
- **Felvételi keret:** A felvételi sokaság egységeit, azok jellemzőit leíró nyilvántartás adott időpontra vonatkozó állapota.
- **Felvételi sokaság:** A célsokaság megfigyelésére alkalmas felvételi egységek összessége.
- **Fordítókulcs:** Két nomenklatúra elemei közötti összefüggés, megfeleltetés.
- **Hibaszámitás:** A mintavételi és a nem mintavételi hiba számítása.
- **Igényspecifikáció folyamatszakasz:** Az ESTFM első folyamatszakasza, amely során a külső és belső igények elemzésére épülve döntés-előkészítő adatfelvételi javaslat készül.
- **Imputálás:** A hiányzó vagy nem használható adathelyekhez értéket rendelő eljárás.
- **Kódolás:** Technikai eljárás, melynek révén verbális információk kóddá konvertálhatók.
- **Makroadat:** Statisztikai feldolgozási folyamat eredményeként elemi (egyedi) adatokból létrejött aggregált adat, vagy az eleminél magasabb szinten gyűjtött adat.
- **Makrovalidálás:** Statisztikai adatok hibájának felderítése aggregált adatok ellenőrzésével, illetve a teljes adatállomány vagy egy adott részhalmoz statisztikai módszerekre épülő ellenőrzésével.
- **Másodlagos adatforrás:** Olyan adatforrás, amelyből a statisztikai adatfelvétel által leírt sokaság adott időszakra vagy időpontra vonatkozó ismérvei teljes mértékben vagy részben statisztikai adatátvétellel meghatározhatók. A másodlagos adatforrások begyűjtésének két típusát különböztetjük meg: az adminisztratív forrásból történő adatátvételeket (adminisztratív adatátvétel) és az egyéb másodlagos forrásból történő adatátvételeket.
- **Metaadat:** Olyan adat, amely meghatároz, leír más adatokat.
- **Mezovalidálás:** Olyan mikrovalidálás, ahol a statisztikai adatok más statisztikai témákhoz tartozó adatokkal való összefüggésének ellenőrzése, hibájának felderítése történik.
- **Mikroadat:** Olyan, rekordok sorozatából álló állomány, amely megfigyelési egységek adatait tartalmazza. Az aggregált adatok előállításának elsődleges forrása.
- **Mikrovalidálás:** Elemi statisztikai adatok vagy azok összefüggésének vizsgálata az adat hibájának kiderítése érdekében.
- **Minőségbiztosítási terv:** Az adatfelvételi terv részeként létrejött terv, amely tartalmazza a munka-standardokat, a hivatalos kötelezettségeket (mint a törvények és a belső szabályok) és a minőség-ellenőrzési akciók rendszerét a mintavételi és nem mintavételi hibák kiküszöbölésére, monitorozására és értékelésére, valamint a statisztikai adatelőállítási folyamat részfolyamatainak irányítására.

- Minőségi követelmény: A statisztikai termékre, statisztikai adatelőállítási folyamatra vagy a statisztikai szervezet egészére vonatkozó olyan konkrét, mérhető követelmények, elvárások, melyek alapján a statisztikai termék, a statisztikai adatelőállítási folyamat vagy a statisztikai szervezet minősége megítélhető.
- Minta: A mintavételi keret egyedei közül kiválasztott elemek halmaza.
- Mintavételi keret: A felvételi sokaság (vonatkozási kör) elemeit, vagy azok rétegeképző ismérveit leíró nyilvántartás, jegyzék adott időpontra állapota, amely alapján a reprezentatív statisztikai adatgyűjtés mintája kiválasztásra kerül, illetve amely alapján a sokaságra vonatkozó becslés készül.
- Mintavételi terv: A mintakiválasztás folyamatára vonatkozó szabályok, specifikációk (minta-méret, rétegek, kiválasztási szempontok) összessége egy adott mintavételi keret vonatkozásában.
- Mutató (mutatóismérv): Valamely rendszeresen ismétlődő társadalmi, gazdasági, környezeti jelenség statisztikai ábrázolására szolgáló ismérv.
- Mutatócsoport: Valamely rendszeresen ismétlődő társadalmi, gazdasági, környezeti jelenség statisztikai ábrázolására szolgáló ismérvek adott témához tartozó, adott szinten és gyakorisággal megfigyelt csoportja.
- Mutatóérték: Egy statisztikai mutató adott statisztikai egységhez (a statisztikai egységek egy csoportjához) tartozó adott időszakra vonatkozó értéke.
- Mutatóváltozat: A mutató adott sokaságra jellemző megvalósulása, figyelembe véve a mutató vonatkozási körét, statisztikai egységét (dimenzióit, aggregáltsági szintjét).
- Nómenklatúra: Valamely ismérv által felvehető értékek azonosított halmaza, rendszerezett jegyzéke.
- Osztályozás: A sokaság elemeinek olyan meghatározott szempontok szerinti csoportosítása, mely teljes körűen lefedi a sokaságot, melyben a csoportok átfedésmentesek, a sokaság minden egyes egysége egy és csakis egy kategóriába sorolható, továbbá egy kategóriába hasonló egyedek tartoznak.
- Output: Mindazon termékek, amelyek az adatfelvétel során létrejönnek. Ide értjük az adatfelvételtől származó statisztikai adatokat, a statisztikai táblákat, mikroadatokat, elemzéseket, az adatokból betöltött adatbázistáblákat, kiadványokat, más tájékoztatási termékeket.
- Paraadat: A statisztikai felvétel folyamatát, körülményét, eredményét leíró, jellemző adat.
- Próba felvétel: Az adatfelvétel tervezése során megtervezett lépések és eszközök alkalmasságának próbája, mely információval szolgál többek között arról, hogy egyértelműek-e a kérdések és a kitöltési utasítás előírásai; betartható-e a tervezett beküldési határidő, illetve mennyi időt vesz igénybe az összeíró által végzett kikérdezés; rendelkeznek-e az adatszolgáltatók a kérdések megválaszolásához szükséges információval. Mintavételen alapuló adatfelvétel esetén célszerű előre megvizsgálni a valószínűségi feltételek várható teljesülését is.
- Regiszter befagyasztott állapota: A regiszter befagyasztott állapota a regiszter egységeit és az azokat jellemző ismérvértékeket rögzíti egy adott időpontra vonatkozóan.
- Reprezentatív statisztikai felvétel: A megfigyelni kívánt objektumok vagy azok adott osztályai mint teljes sokaság (célsokaság) egésze helyett egy abból megfelelő módon kiválasztott részsokaság, minta képezi a megfigyelés tárgyát és az eredményeket a teljes célsokaságra vonatkoztatják.

A kiválasztási, mintavételi eljárások szerinti reprezentatív felmérési módszerek fajtái:

- véletlen mintavétel (egyszerű véletlen, egylépcsős, többlépcsős, rétegezett);
- nem véletlen mintavétel;
- szisztematikus mintavétel.

- Statisztikai adat: A valós világ egyedeinek tulajdonságaira vonatkozó statisztikai megfigyelések, illetve további statisztikai műveletek eredménye.
- Statisztikai adatfelvétel: Egy adott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek gyűjtése egy társadalmi, gazdasági vagy környezeti jelenség statisztikai megfigyelése céljából, különböző adatforrások felhasználásával.
- Statisztikai adatgyűjtés: A statisztikai adatfelvétel által meghatározott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek összegyűjtése a kiválasztott adatszolgáltatók megkérdezésével vagy közvetlen megfigyeléssel.
- Statisztikai felfedés elleni védelem: A statisztikai célra összeállított adatállományok módosítását célzó olyan módszerek összessége, melyek a lehető legnagyobb mértékben csökkentik a felfedési kockázatot. Az adatállományok statisztikai felfedés elleni védelmének legfontosabb célja annak megakadályozása, hogy egyedi adatok jelenjenek meg bennük.
- Statisztikai fogalom: A statisztika tárgyához, előállítási, működtetési rendszeréhez kapcsolódó önálló jelentéssel bíró gondolati egység.
- Statisztikai mérőszám: Egy adott csoportba tartozó statisztikai egységeket jellemző egyedi numerikus változóértékek összesítése (pl. átlag, módusz, összesen, index, stb.).
- Statisztikai nómenklatúra: Valamely statisztikai sokaság elemeiből egy vagy több ismérv szerint képzett csoportok felsorolása. Nevek, megnevezések rendszerezett jegyzéke.
- Statisztikai regiszter: A statisztikai adatelőállítási folyamat támogatása és egységesítése céljából a Hivatalos Statisztikai Szolgálat tagjai által vezetett olyan statisztikai nyilvántartás, amelyből elérhetők a különböző adatfelvételek sokaságainak aktuális és történeti állapotai, a változások oka, hatálya, forrása, és amely a sokaság egyedeire (egységeire) vonatkozó adatokat strukturáltan, adatbázisban tárolja.
- Statisztikai változó: A mutatóváltozat egy adatállományban való konkrét megvalósulása.
- Sürgetés: Adatgyűjtés-szervezés adatbenyújtási határidőhöz kötött lépése az adatbenyújtók (adatszolgáltatók) késedelemről való értesítésével. Ide értjük a határidő előtti proaktív emlékeztetést is.
- Szakstatisztika: A valóság (gazdaság, társadalom, környezet) egy logikailag összetartozó részterületének megismerése érdekében végzett statisztikai tevékenység és annak eredménye, beleértve az adatforrásokat, az adatbegyűjtés, adatelőkészítés és feldolgozás módszereit, az eredménytermékeket, azok metaadatait, valamint ezek tartalmára, illetve minőségére vonatkozó dokumentációit.
- Szakstatisztikai fogalom: Az egyes szakstatisztikák által megfigyelt társadalmi, gazdasági, környezeti fogalmak statisztikai megjelenítése, értelmezése.
- Személyes adat: Azonosított vagy azonosítható természetes személyre (érintett) vonatkozó bármilyen információ.
- Szezonális kiigazítás: Azon eljárás, amelynek során az eredeti adatsorból kiszűrjük a szezonális hatást és a naptárhatást. A szezonális kiigazítási eljárásba beleértjük továbbá az outlierok kezelését, és a véletlen tényezőtől mentes trend előállítását is.

- Tájékoztatás előkészítése folyamatszakasz: Az Egységes Statisztikai Folyamatmodell (ESTFM) hetedik folyamatszakasza, amely a feldolgozott adatokból előállítja, ellenőrzi, felfedés ellen védi, kiértékeli, dokumentációval látja el a tájékoztatáshoz szükséges adatokat.
- Tájékoztatás folyamatszakasz: Az Egységes Statisztikai Folyamatmodell (ESTFM) nyolcadik folyamatszakasza, a statisztikai adatelőállítási folyamat során létrehozott információk közzétételének komplex rendszere, illetve a felhasználók igényeinek, elégedettségének mérése és visszacsatolása.
- Teljes számjel: A felvételi keret törzs a felvételi keretek meghatározásának közös alapjául szolgáló nyilvántartás adott időpontra vonatkozó állapota.
- Teljeskörű felvétel (cenzus): Olyan statisztikai adatfelvétel, amely a megfigyelés tárgyai, vagy azok adott osztályai valamennyi elemére kiterjed, mindarra az objektumra, amely a célsokaságot/teljes sokaságot képezi, de gyakorlatilag a felvétel a keretsokaságon (keretpopuláción) értelmezett felvételi kereten hajtható végre. A megfigyelni kívánt sokaság teljes körére irányuló adatfelvétel.
- Tervezés folyamatszakasz: Az Egységes Statisztikai Folyamatmodell (ESTFM) második folyamatszakasza, amely során az adatfelvételi terv részletes specifikációja elkészül az adatfelvétel megvalósításától a tájékoztatásig.
- Vonatkozási idő: Az az időintervallum vagy időpont, amire a megfigyelt változó, mutató vonatkozik.