



## Integrated Data Processing System (EAR)

Session number: 8

Date: June 1<sup>th</sup> 2016

Hajnalka Debreceni  
Hungarian Central Statistical Office  
[Hajnalka.Debreceni@ksh.hu](mailto:Hajnalka.Debreceni@ksh.hu)

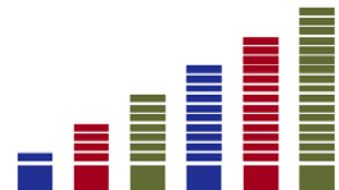


Madrid, May 31 - June 3

# Outline

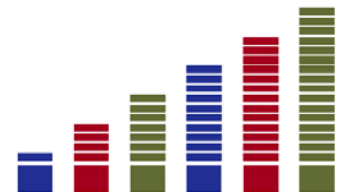


<p><b>Methodological aspects</b></p> <ul style="list-style-type: none"> <li>▪ Hungarian GSBPM (ESTFM)</li> <li>▪ Place of this system in data-production process</li> <li>▪ Quality impacts of EAR</li> </ul>	<p><b>Functionality</b></p> <ul style="list-style-type: none"> <li>▪ Framework</li> <li>▪ Workflow diagram</li> <li>▪ Set of elements</li> <li>▪ EAR user interface</li> </ul>
<p><b>Technical aspects</b></p> <ul style="list-style-type: none"> <li>▪ The milestones of system development</li> <li>▪ Security management</li> <li>▪ Operational environments</li> </ul>	<p><b>Users and the staff</b></p> <ul style="list-style-type: none"> <li>▪ Statisticians</li> <li>▪ Methodological experts</li> <li>▪ System administrators</li> <li>▪ Developers</li> </ul>

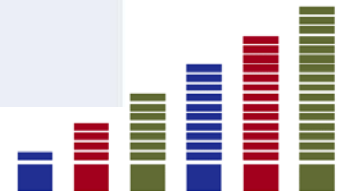
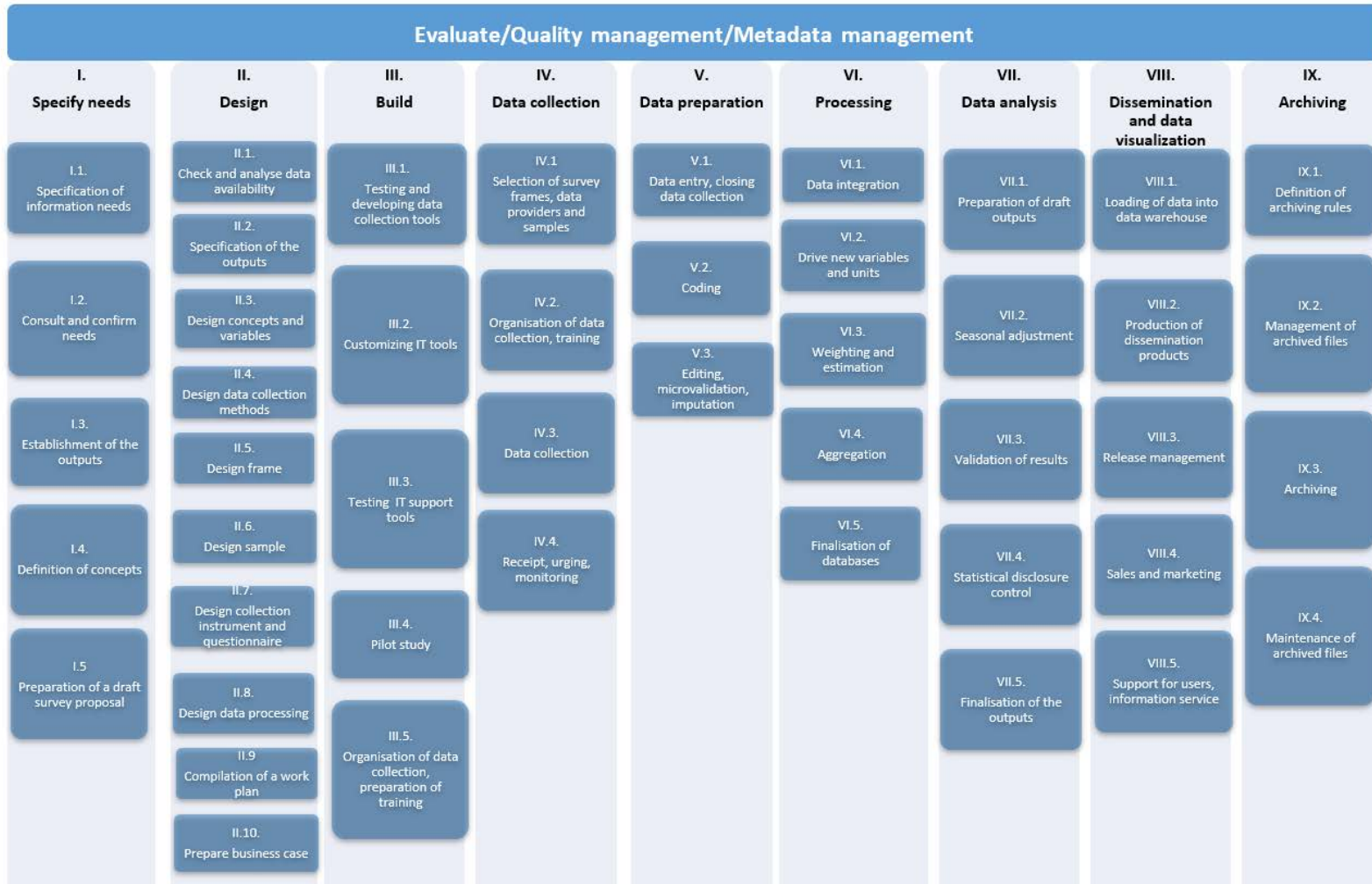


# Methodological aspects

- The main purpose of the EAR system is to increase the quality and efficiency of data processing and to support these activities in general, using standard protocols for any statistical domain.
- Processing is a crucial phase of the Hungarian Generic Statistical Business Process Model ([ESTFM](#)). The ESTFM is the Hungarian adaptation of GSBPM, consists of several phases and sub-processes. The Processing phase covers 5 sub-processes.
- Related sub-processes, besides the Processing phase:
  - II.8. Design data processing
  - V.3. Review and validate
  - VII.2. Seasonal adjustment
  - VII.3. Validate outputs
  - VII.4. Apply disclosure control
  - IX.3. Archiving
- There are overarching processes characterizing the whole of statistical data production processes (Metadata management, Product quality measurement).
- Since the EAR is an IT tool, the Build (III.) process phase is relevant, not to the functionality but the deployment.



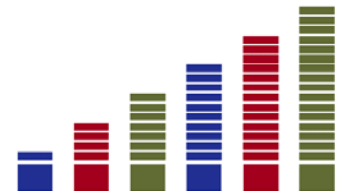
# Hungarian Generic Statistical Business Process Model (ESTFM)



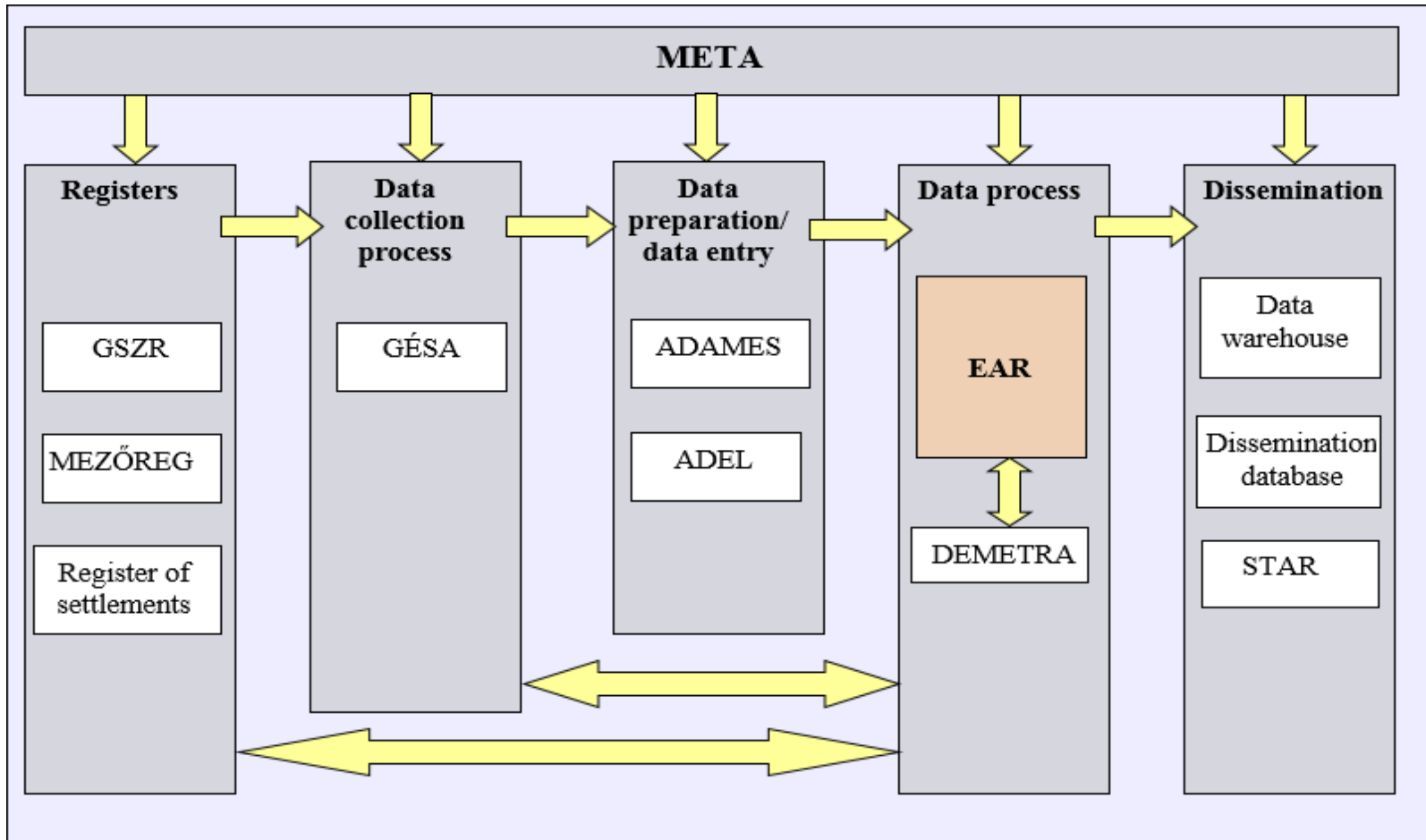
# The advantages of EAR

Why  
useful?

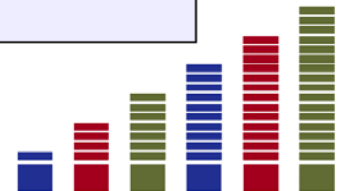
- Metadata-driven application
- Transparent workflow
- Reproducible data processing phase
- Repeatable and modifiable process plans
- Documented processes
- Support the data processing activities in general using standard protocols for any statistical domain
- Support the production process through more data production periods for a number of data collections
- IT system controlled by statisticians, they design, redesign and manage their data processing activities
- Cooperation with other integrated IT systems



# Place of the EAR in data-production process



Connection to other integrated IT systems



# From dream to reality

## *Technical aspects*

### Requirements

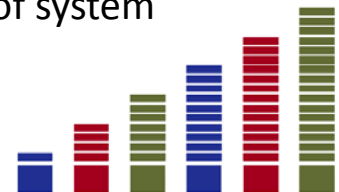
- 2008 – Specification of requirements (extensive on-demand survey within the HCSO)
- 2009 – Requirements analysis and system engineering with contracted partner

### Development

- 2010 – Development and testing (external development, a 3-tier application, Java based, underlying Oracle database)
- Two environments: production and test
- The security management (authentication and authorization)

### Implementation

- End of 2010 – Start of implementation (Agriculture and annual calculation of GDP)
- Maintenance of system



# Functionality of EAR system

## Framework

- It supports the design and the management of data processing activities, using standard methods and data process definition structure.
- The display objects - like menu, icons, tags with different functionality and colours - help the users in planning of a processing cycle.

## Set of elements

- The standalone standards, defined in a common way to support the definition of data processing activities.
- The hierarchical structure of these elements:
  - Process stages
  - Process methods
  - Procedures and functions which execute the above methods





# List of standard process stages

---

ST01 – Loading data into processing environment

---

ST02 – Outlier treatment

---

ST03 – Imputation

---

ST04 – Deriving new variables, making aggregates

---

ST05 – Calculating weights, estimation of population parameters,  
standard errors and variances

---

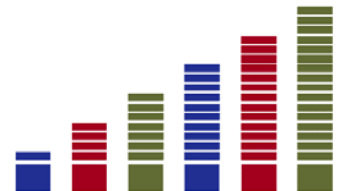
ST06 – Seasonal adjustment

---

ST07 – Statistical disclosure control

---

ST08 – Finalizing output



# Snapshot of the EAR user interface

Egységes Adatfeldolgozó Rendszer - Debreceni Hajnalka - KSH éles környezet

Tervezés Feldolgozás Adminisztráció Eszközök Nézet Súlyó X Kilépés

Üzenetek x Táblázatszerkezetek x Feldolgozás irányítás x Feldolgozás tervezés x Adatnyilvántartás x

Gyorskeresés:  Ok

Frissítés Útvonal: /ALT/Mintafolyamatok/ST04/ST04.3/ST04.3.1/ST04.3.1\_AGGR/ST04.3.1\_AGGR\_00/FOLY

ALT - Általános mappák

- Mintafolyamatok - Mintafolyamatok mappa
  - ST01 - Kiinduló adatok beolvasása EAR-ba
  - ST02 - Outlier szűrés
  - ST03 - Pótlás
  - ST04 - Mutatóképzés, aggregálás
    - ST04.1 - Indexek és indexesorok képzése
    - ST04.2 - Egyedi mutatószámok képzése
    - ST04.3 - Adatok aggregálása
      - ST04.3.1 - Szerkezeti aggregálás
        - ST04.3.1\_AGGR; Aggregálás; Támcsu Izabella Ildiko
        - ST04.3.1\_AGGR\_00; TERVEZES\_ALATT; Állatállomány aggregálás
          - Paraméterek
            - VEV - Vonatkozási év
            - VHO - Vonatkozási hó
            - ALLAT\_ELEMI - Elemi állatállomány adatok
            - ALLAT\_FAJTA - Állatállomány adatok aggregátumai
            - FOLYAMAT - Folyamatdefinió
              - start
              - 1. szakasz - ST04 - Adatok aggregálása
                - start
                - 1.1 lépés - ST04.3.1 - Adatok aggregálása állatfajtákra országos szinten
                  - Állatfajtánkénti aggregátum előállítás
                  - stop
                - stop

- ST04.3.2 - Időbeli aggregálás
- ST04.3.3 - Területi aggregálás
- ST04\_MKOD KEPZES - Mkod képzés
- ST04\_ERTEKOSSZEG - Lefelé kumulált értékösszeg számítása
- ST04\_NEP\_KAT - Népességnagyság kategória képzés
- ST04\_TERULETI\_JELLEMZOK - Megye (MT03) és régió (MT50) beállítása Településszámból
- ST04\_ADATTARHAZ - Adatelőállítás és adatátadás Adattárháznak
- ST04\_TARS - TARS-ba töltés előkészítése
- ST05 - Súlyképzés, becslés, hibaszámítás
- ST06 - Szezonális kiigazítás
- ST07 - Felfedés elleni védelem
- ST08 - Eredmény adatok átadása EAR-ból
- Mintafolyamatok\_TANF - A tanfolyami mintafolyamatok mappája
- Számjelek
- Interfészek
- Adminisztratív\_adatok - Adminisztratív forrásból kapott adatok szerkezete
- O - Ágazati gazdaságstatisztika
- Q - Általános gazdaságstatisztika
- T - Területi statisztika

Művelet: Állatfajtánkénti aggregátum előállítás x Művelet: AGGREGALAS\_SZERKEZET x

Művelet: AGGREGALAS\_SZERKEZET ?

Paraméterek

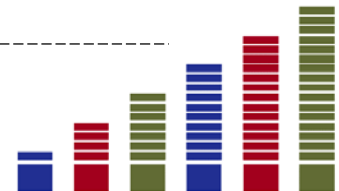
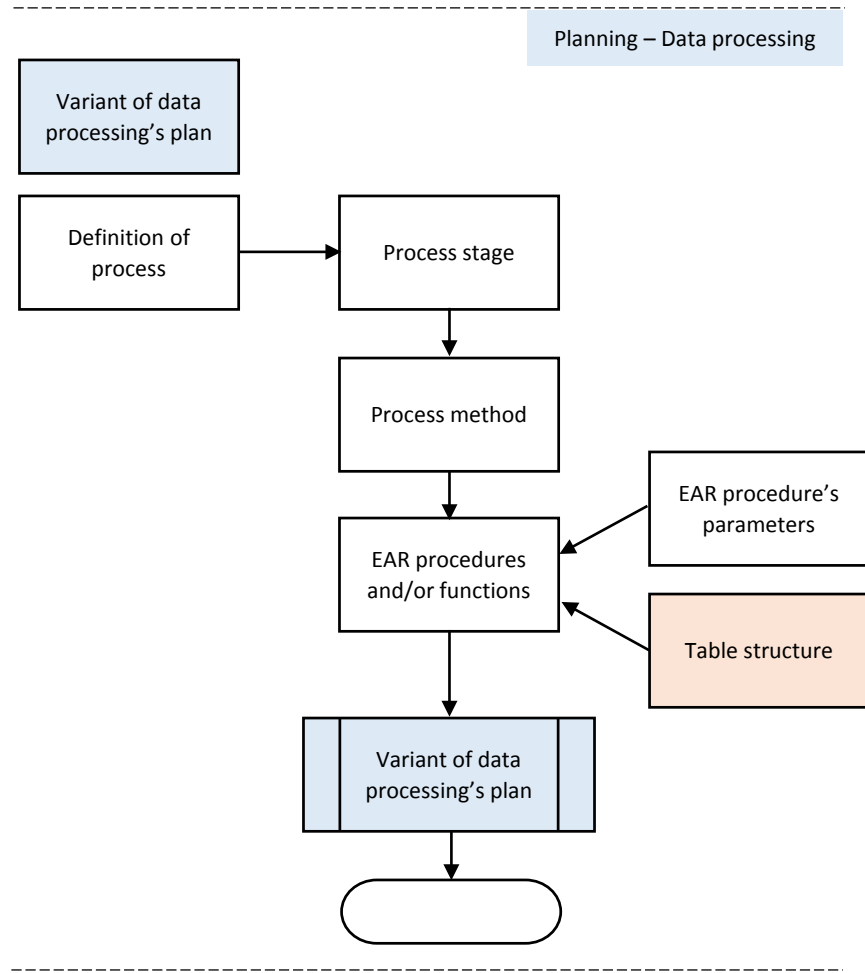
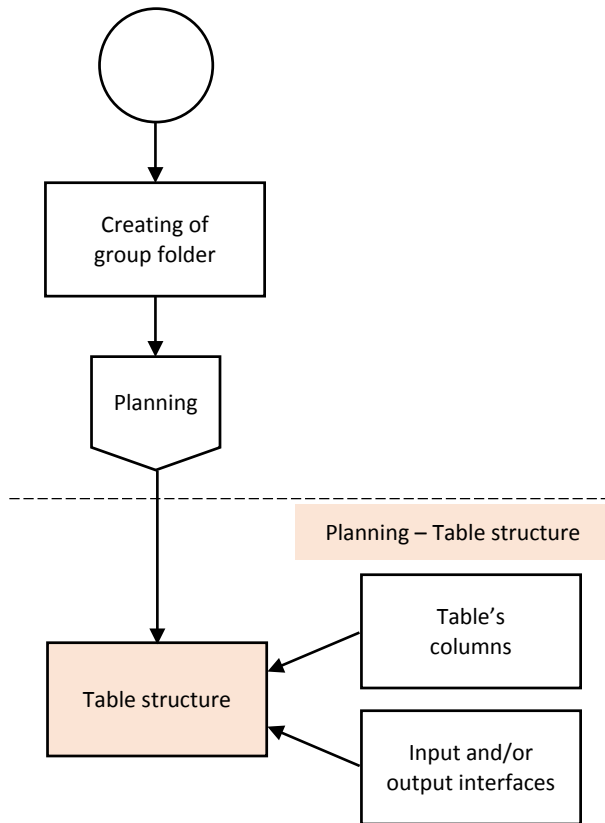
| Paraméter név                       | Paraméter érték  |
|-------------------------------------|--|
| Céltáblázat CN *                    | ALLAT_FAJTA  |
| Forrás táblázat FN *                | ALLAT_ELEMI  |
| Csoportképző ismérvek listája C...  | Választható érték<br>ALLAT_ELEMI.TEV<br>ALLAT_ELEMI.MHO<br>ALLAT_ELEMI.MM562   |
| Oszloplista OL (opcionális)         | Választható érték<br>MAAA004 := OSSZESEN( ALLAT_ELEMI.MAAA004 ,MIND)<br>MAAA005 := DARAB( ALLAT_ELEMI.MAAA004 ,ELTERO) |
| Szűrőfeltétel a forrásra SZ1 (op... | Érték  |

Módosítások visszairása a műveletbe Bezárásk mentés nélkül

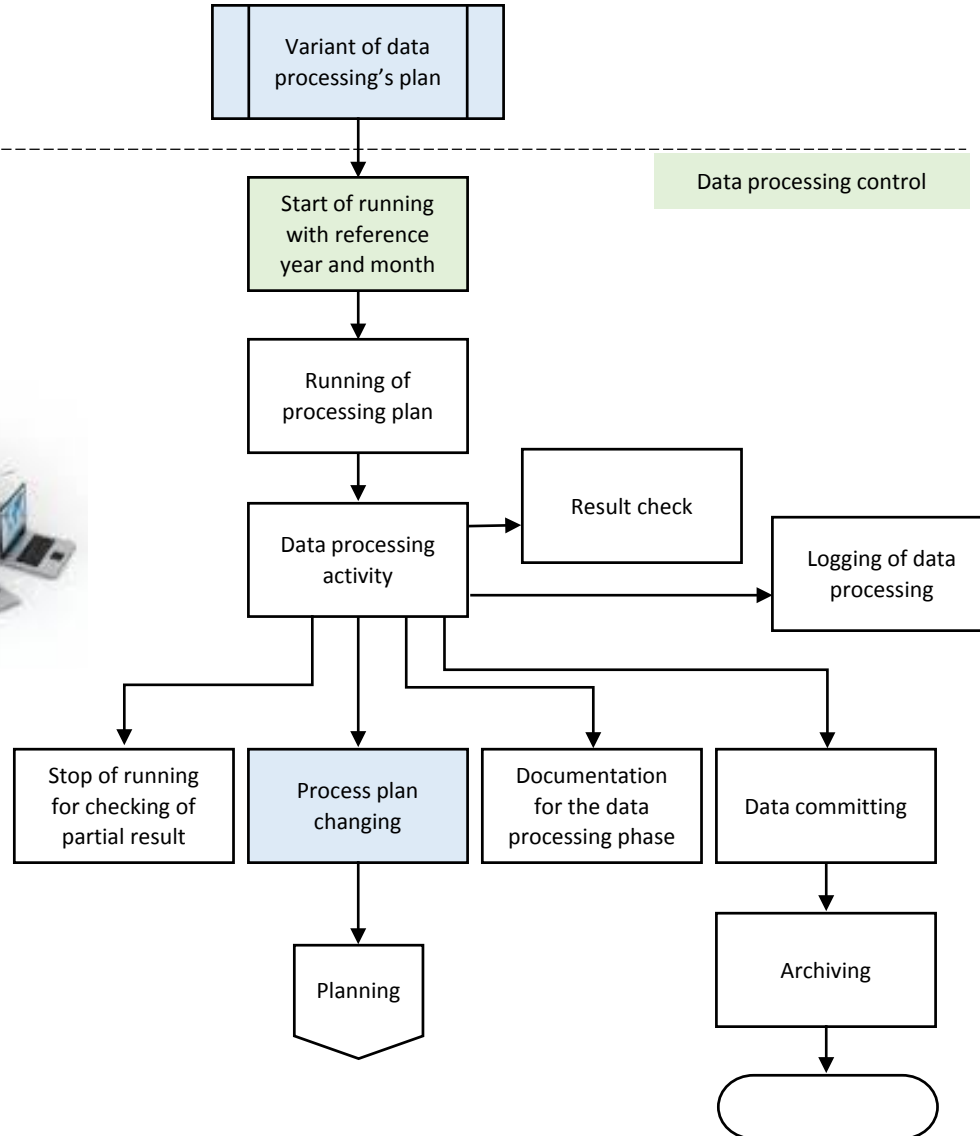
Parameterised  
aggregate  
procedure



# EAR workflow - Planning



# EAR workflow - Processing



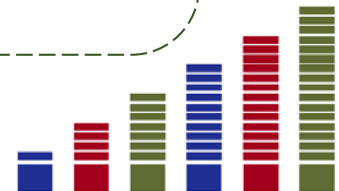
# Conclusions and future plans

## Conclusion

- EAR is one of the key integrated IT systems of the HCSO for its business processes with a lot of functionality to support statisticians in the design and management of the data processing tasks.
- The main outcomes of the EAR is to increase the quality of data processing tasks due to standardization, transparency, full integration with the HCSO metainformation system and the automatically generated documentation of data processing activities.

## Next steps

- Fine-tuning of EAR system according to user needs following methodological and technical changes (e.g. applying product quality indicators in the processing)
- The EAR is in use at several subject-matter domains – like environment, tourism, transport, social, health care, culture and internal trade – for the present. EAR is currently being extended to all production processes at all subject-matter domains, to fully extend the EAR system for all data processing activities.



# Thank you for your attention

**Hajnalka Debreceni**

Head of Section

Hungarian Central Statistical Office

IT Department

Data Processing Development Section



✉ [Hajnalka.Debreceni@ksh.hu](mailto:Hajnalka.Debreceni@ksh.hu)

☎ (+36 -1) 345-1173

🌐 [www.ksh.hu](http://www.ksh.hu)

