

# Business microdata dissemination at Istat

Daniela Ichim  
Luisa Franconi

ichim@istat.it  
franconi@istat.it

- Released products
- Microdata dissemination
- Business microdata dissemination
- Documentation of microdata files
- Further work

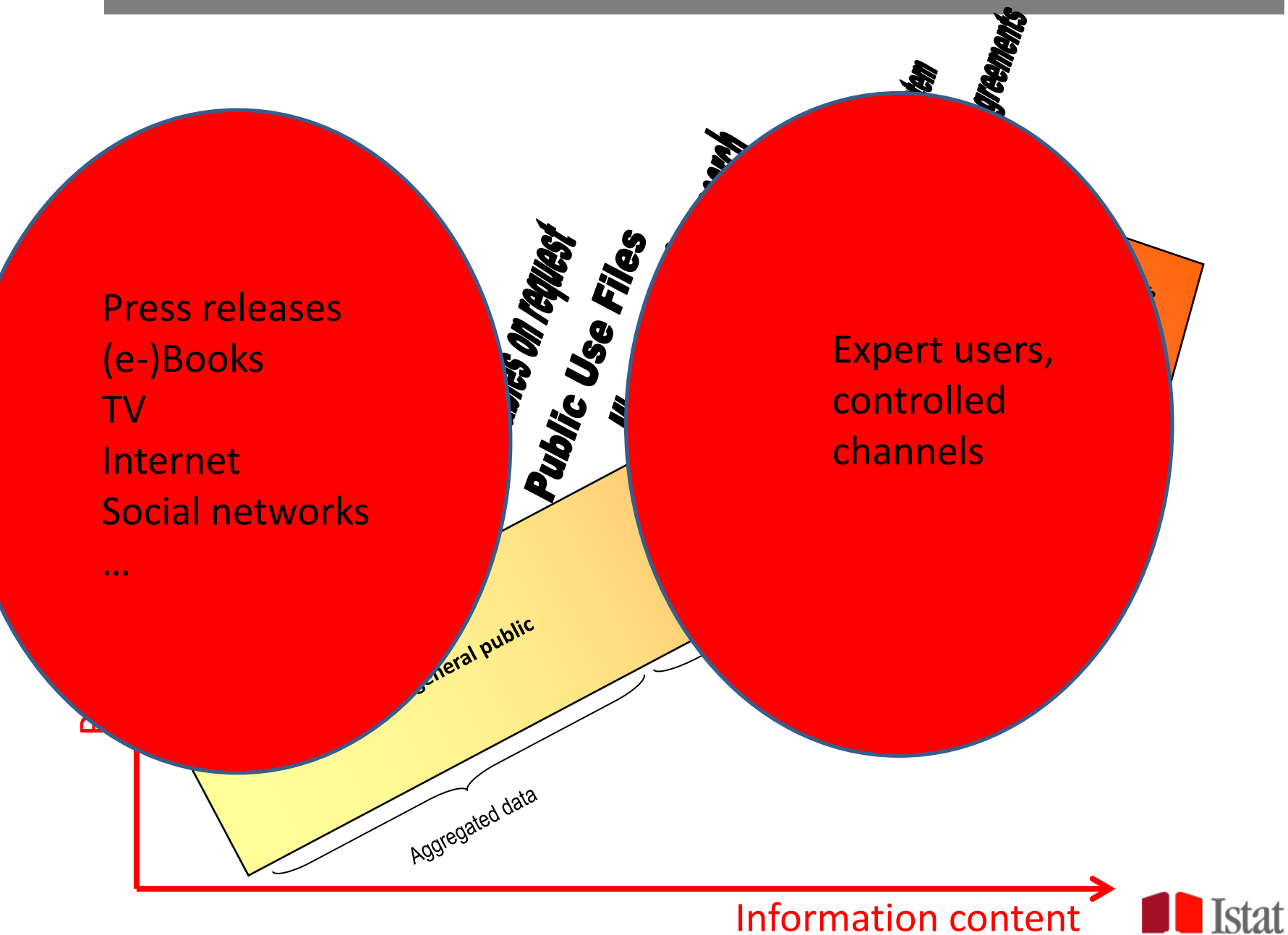
The mission of National Statistical Institutes (NSI) is to produce and disseminate:

reliable  
impartial  
transparent  
accessible  
pertinent

} information

The dissemination of this information should be performed in full compliance with the legislation pertaining to the privacy and confidentiality of respondents.


# Different users, different needs



The demand of analysis of microdata is steadily increasing:

- a) infrastructure advances (computational power, software availability)
- b) more information is available (internet)
- c) need to analyze more “localized” phenomena

## Advantages of dealing with microdata

- data processing is unlimited and unrestricted:
  - data selection
  - models and methods
  - prioritisation of variables and/or sources
- training (and experience) on ‘real’ data, complex datasets
- transparency, neutrality and impartiality
- reproducibility of research and Official Statistics  Istat

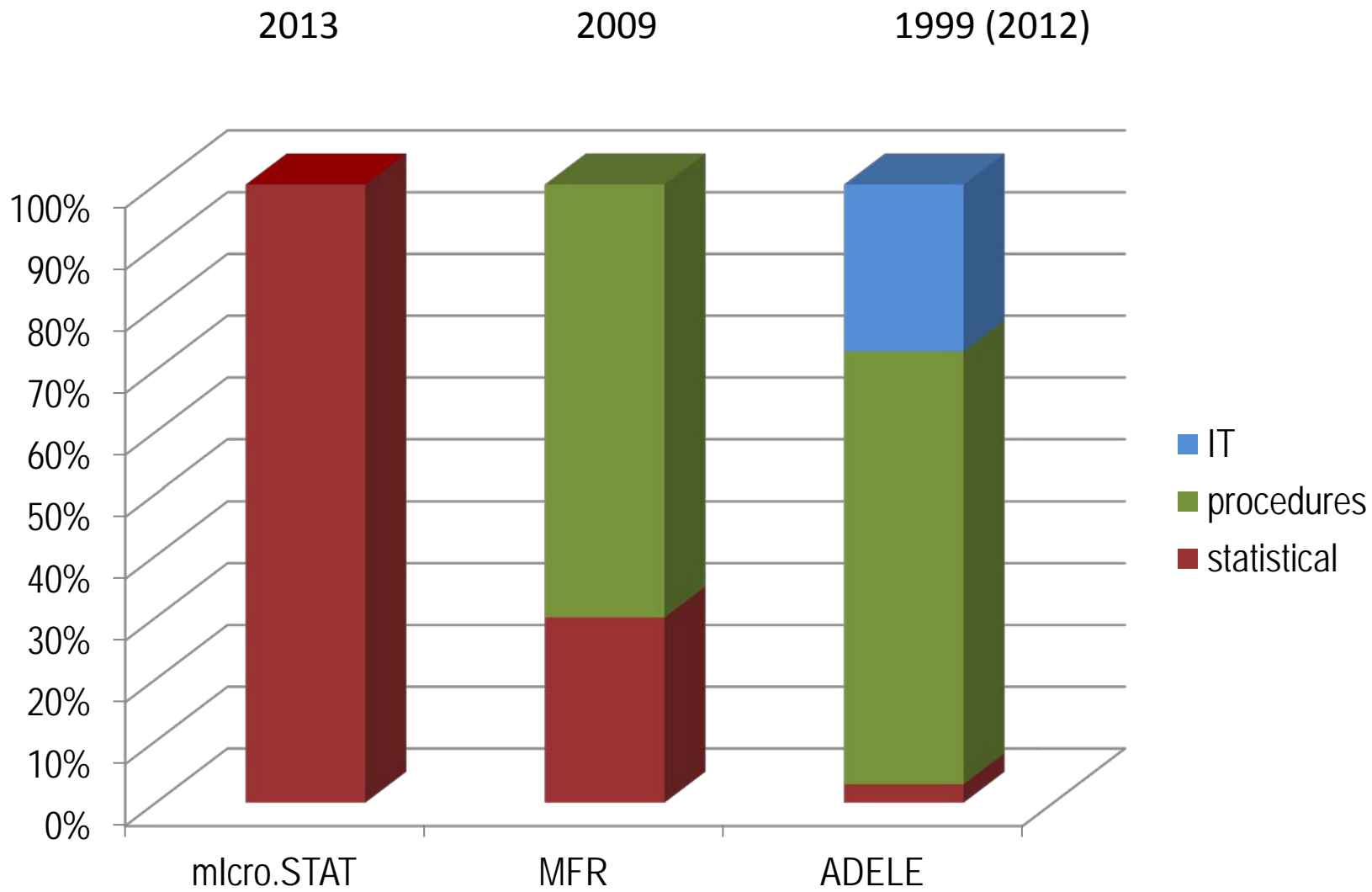
The demand of analysis of microdata is steadily increasing:

- a) infrastructure advances (computational power, software availability)
- b) more information is available (internet)
- c) need to analyze more “localized” phenomena

Disadvantages - microdata are NOT user-friendly

- software tools are required
- knowledge (IT, statistical, methodological, subject-matter) is required
- privacy and confidentiality
- controlled access and dissemination

# Microdata dissemination at Istat



## ADELE

- accredited researchers
- scientific research projects
- data analysis only in the secure rooms
- output is checked by expert staff before its transmission to the users

## MFR

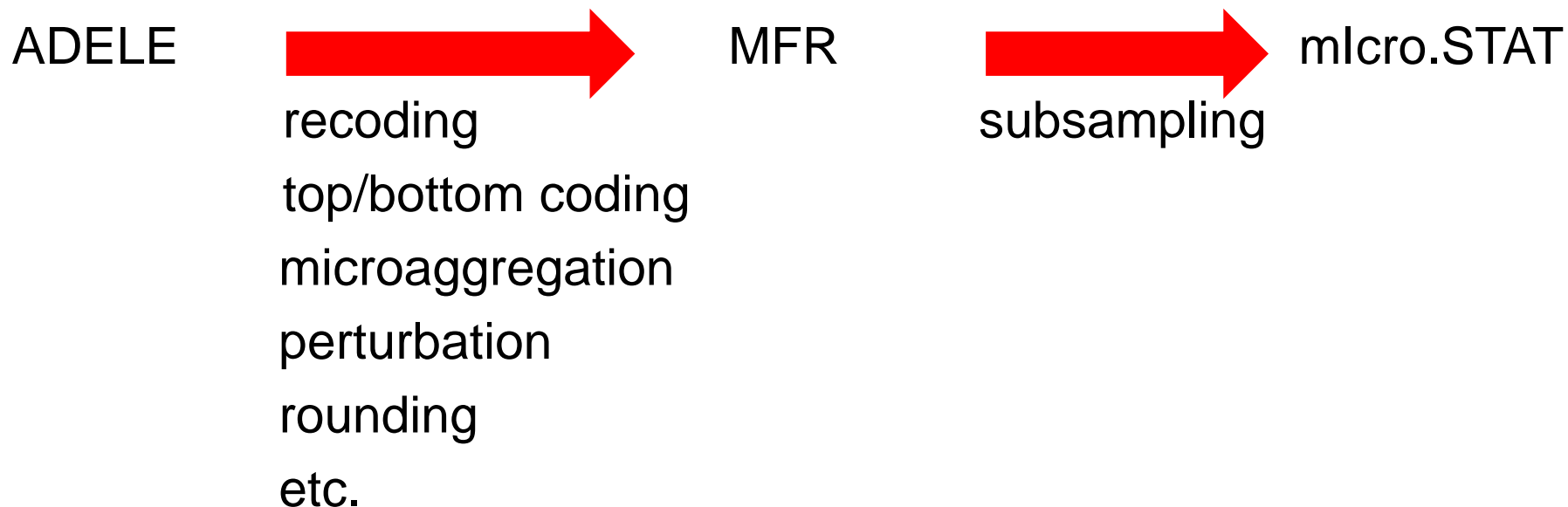
- accredited researchers
- scientific research projects
- no statistical or IT restriction on the analyses

## micro.STAT

- registered users (only a valid e-mail is necessary)
- no statistical or IT descriptions



Integrated system (microdata files share the same structure)



Multiple releases from the same survey.

## Legal aspects:

- ✓ access to social, business, registers and integrated microdata
- ✓ access is independent on nationality
- ✓ no consent is required, but we have to inform respondents

Access is free of charge.

# Microdata dissemination at Istat

ADELE: any Istat survey

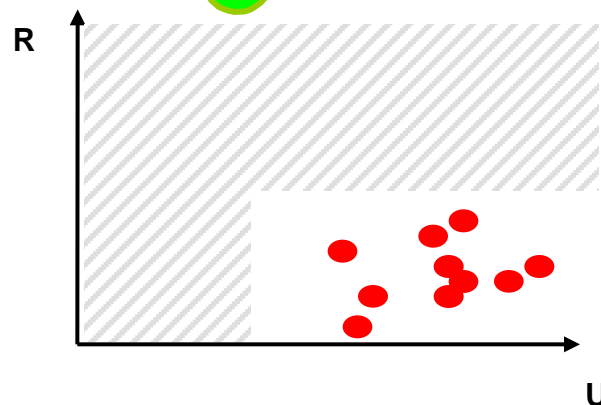
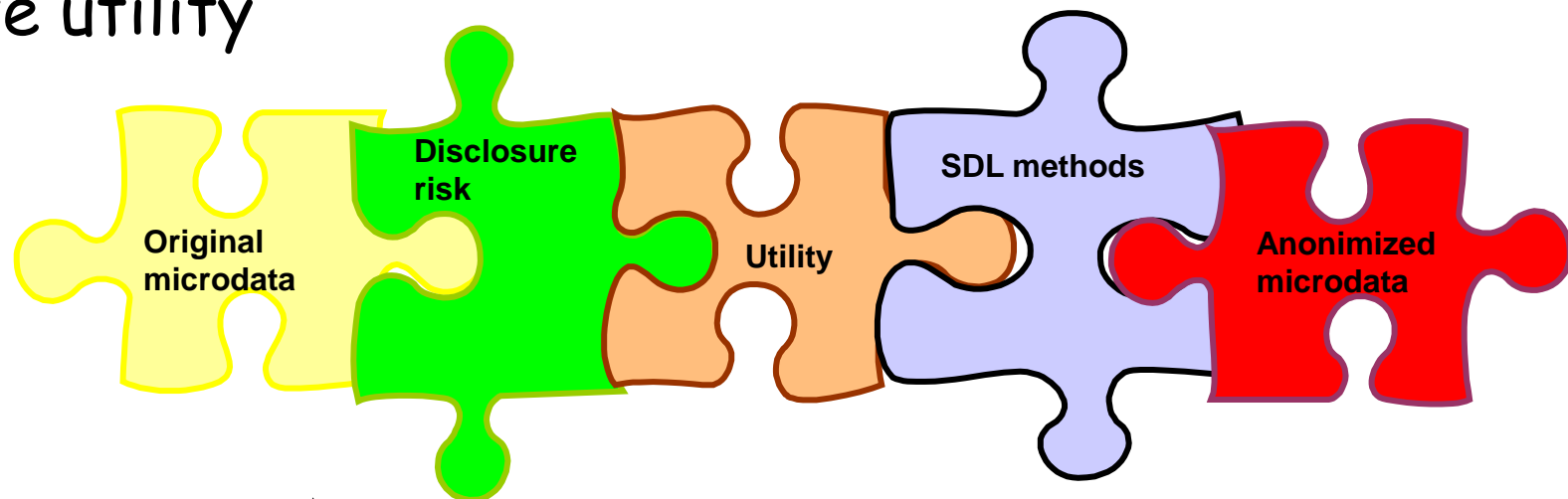
MFR:

Survey	EU	Survey	EU
Continuous Vocational Training Survey (CVTS)	YES	Road Accidents Resulting in Death or Injury	NO
Factors of Business Success (FOBS)	NO	Structure of Earnings Survey (SES)	YES
Farm Structure Survey (FSS)	NO	Survey on Doctorate Holders' Vocational Integration	NO
Graduates' Transition	NO	Italian Innovation Survey (CIS)	YES
Labour Force Survey - cross-sectional quarterly	YES	University Graduates Census	NO
Labour Force Survey-12 months longitudinal data	NO	University Graduates' Vocational Integration	NO
Population Census 2001	NO		

More information: <http://www.istat.it/it/prodotti/microdati>

Apply SDL to reduce risk maintaining some utility

Evaluate utility



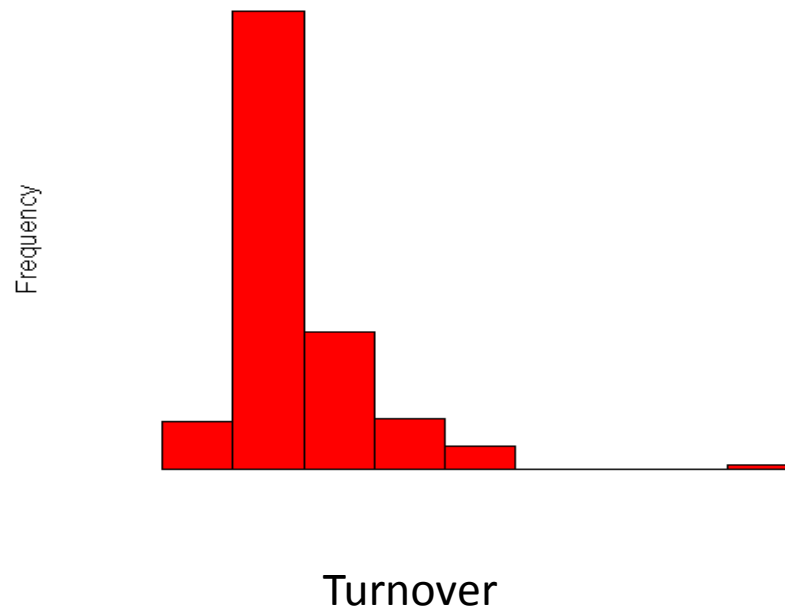
Utility: analytical validity

Particular issues from an SDC point of view:

- smaller reference population
- (known) take-all strata
- large enterprises are well-known (recognizable)
- large enterprises are dominating
- outliers, (extremely) skew distributions
- there might be some «economic» interest in identifying some businesses
- there might be some real (measurable) harm if a business is identified
- both continuous and categorical variables
- continuous variables → each record is an unique case

## Disclosure scenarios:

- categorical variables → external registers
- continuous variables → outliers (data driven approaches)



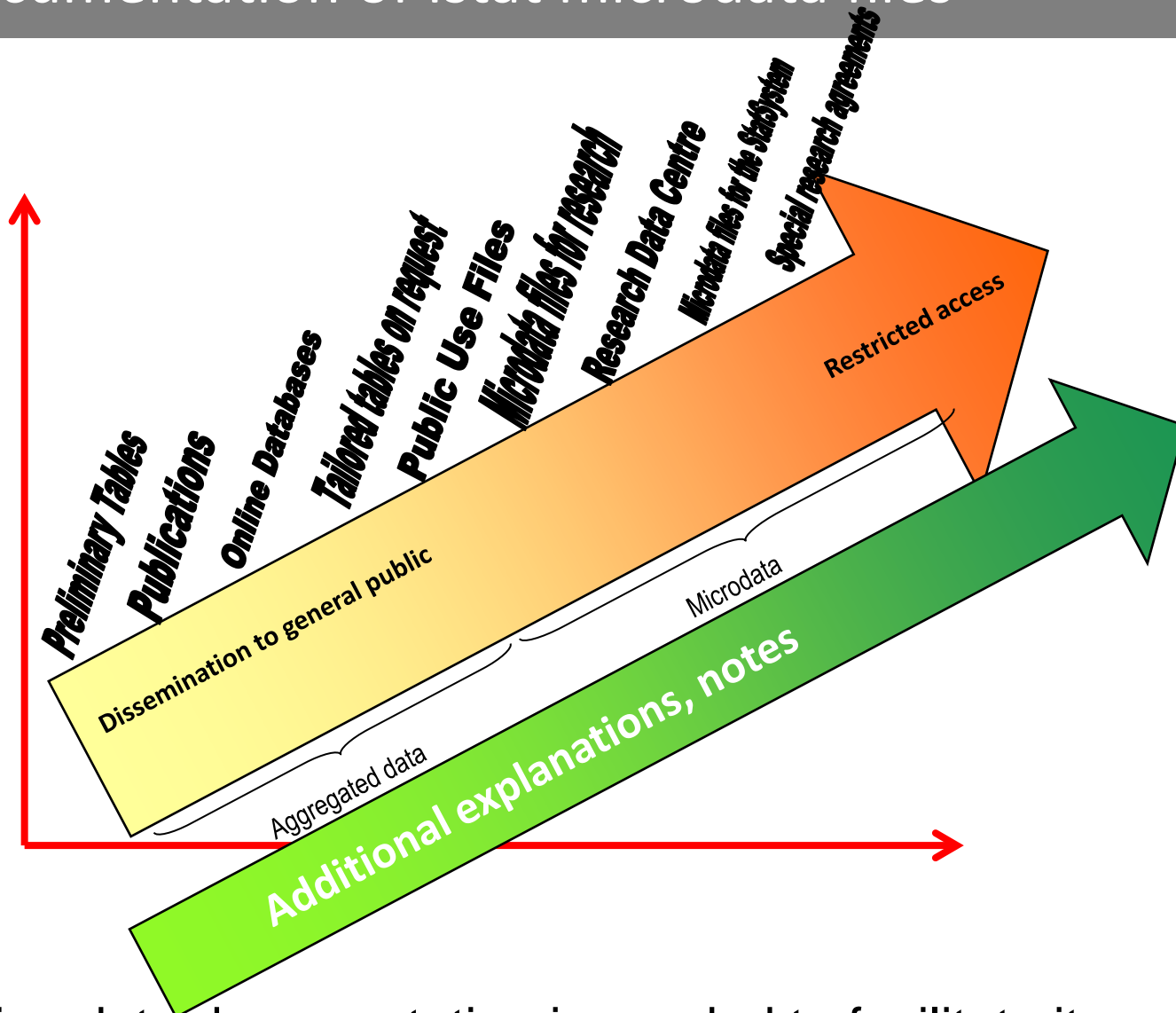
# Business microdata dissemination at Istat

Istat approach: statistical disclosure control methods

Survey\SDC	Variable suppression	Rounding	Individual ranking	Recoding	Perturbation
CIS	X	X	X	X	X
SES	X		X	X	X
CVTS	X	X		X	X
FSS	X	X	X	X	X
FOBS	X	X		X	

- suitable to the scenario
- perturb only the units at risk
- suitable to the data analysis (research potential), including comparability and harmonisation at EU level
- ensure coherence with already published information
- apply the same methodology to subsequent waves

# Documentation of Istat microdata files



Microdata documentation is needed to facilitate its use.



The microdata products share the same documentation, freely downloadable from the Istat web-site:

- a) survey methodology (sampling design, data collection, data calibration, etc)
- b) SDC methodology (disclosure scenarios, disclosure limitation methods, data utility evaluations)
- c) Survey questionnaire
- d) Layout description (list of variables and their characteristics: labels, length, type [categorical or continuous])
- e) Classifications
- f) Routines to load the data in R, STATA, SPSS and SAS
- g) A toy microdata file, an example of structure file

Istat microdata documentation is available also in English (EU).

a) Other microdata products may be developed, but it could be better to focus on «microdata are not user friendly»:

develop instruments to process microdata:

- tools:

- faster computation
- faster visualization
- faster interpretation
- faster communication

- services:

- searchability
- documentation, metadata

**standards (SDMX or DDI)**



b) Coherent multiple releases from multiple surveys (integrated statistics)

**THANK YOU FOR YOUR ATTENTION!**