

# Statistical Institutes and the need to make their data available In a safe and controlled way

**Freddy Maetens** – Programme manager VSA S-DWH – [freddy.maetens@vlaanderen.be](mailto:freddy.maetens@vlaanderen.be)

**Samuel De Klerck** – Practice Lead LACO – [samuel.deklerck@laco.be](mailto:samuel.deklerck@laco.be)

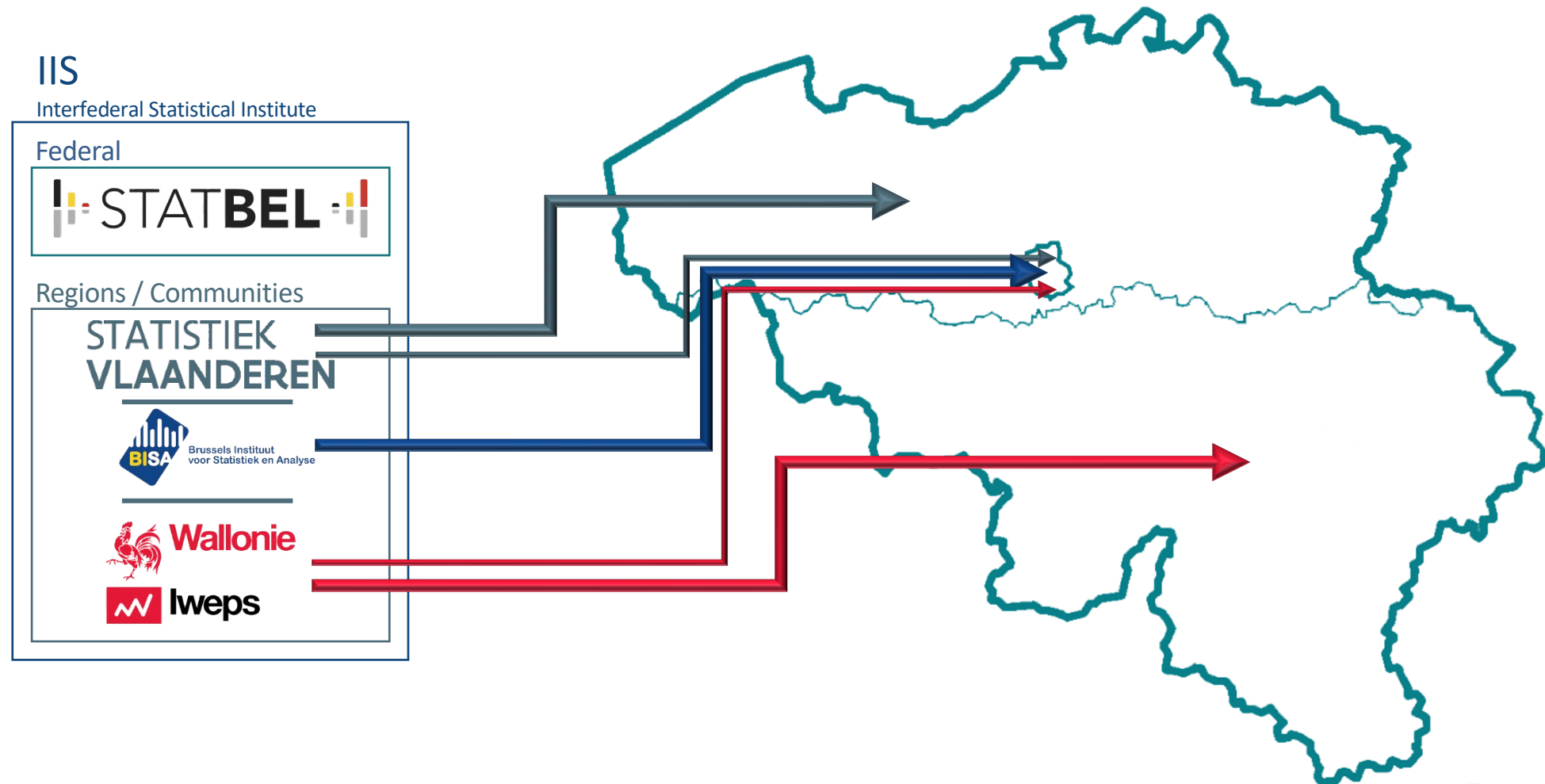


## Agenda

- ⦿ Introducing Flanders Statistics Authority (FSA)
- ⦿ The data exchange challenge
- ⦿ Comparative analysis – Approach
- ⦿ Comparative analysis – Key Outcomes



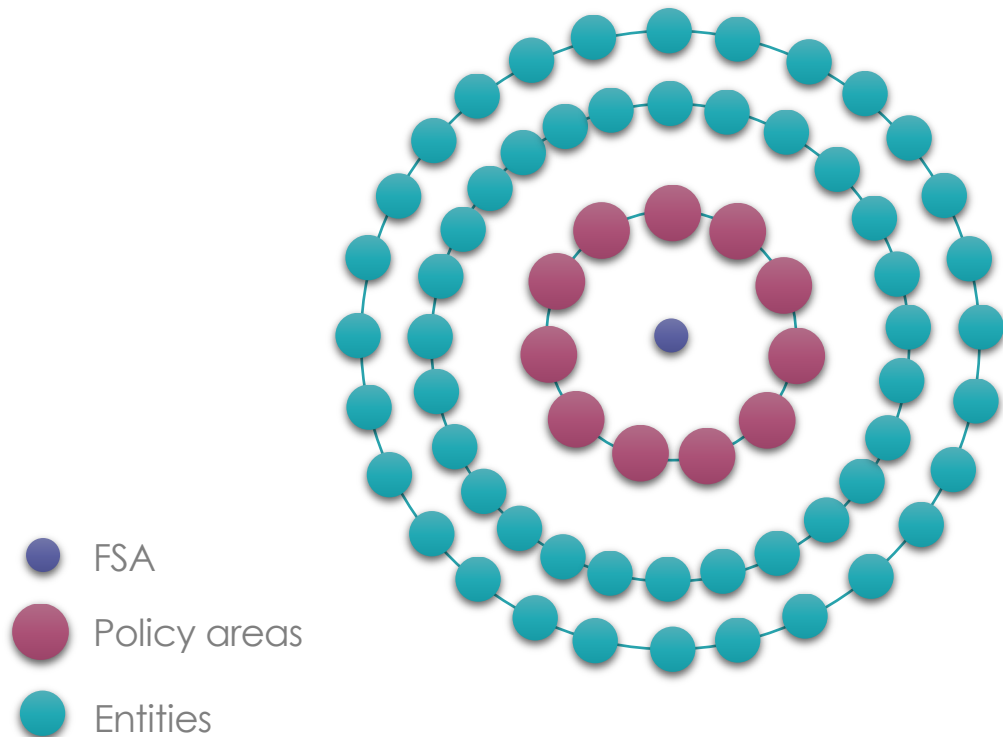
## Context FSA in Belgium



Simplified representation of the Belgian statistical landscape.

## FSA: Coordination of a decentralized network

Network Statistics Flanders



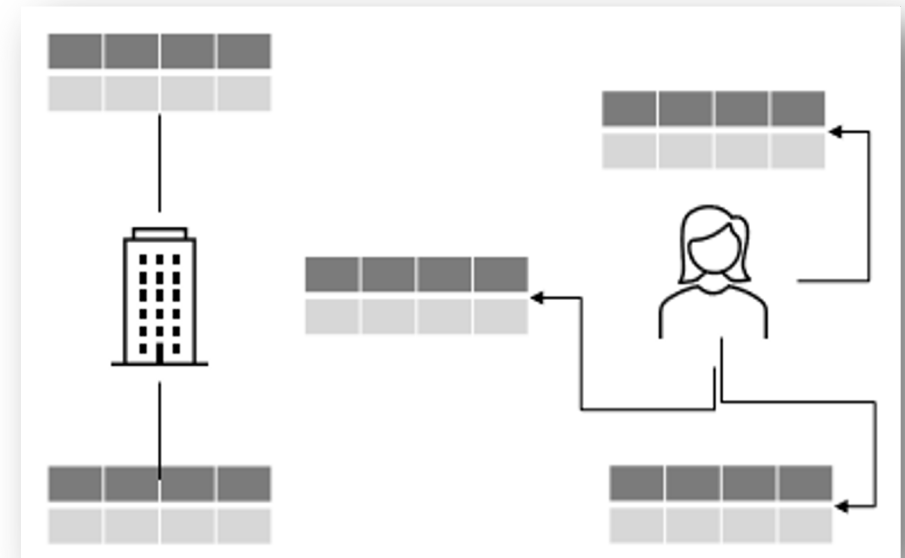
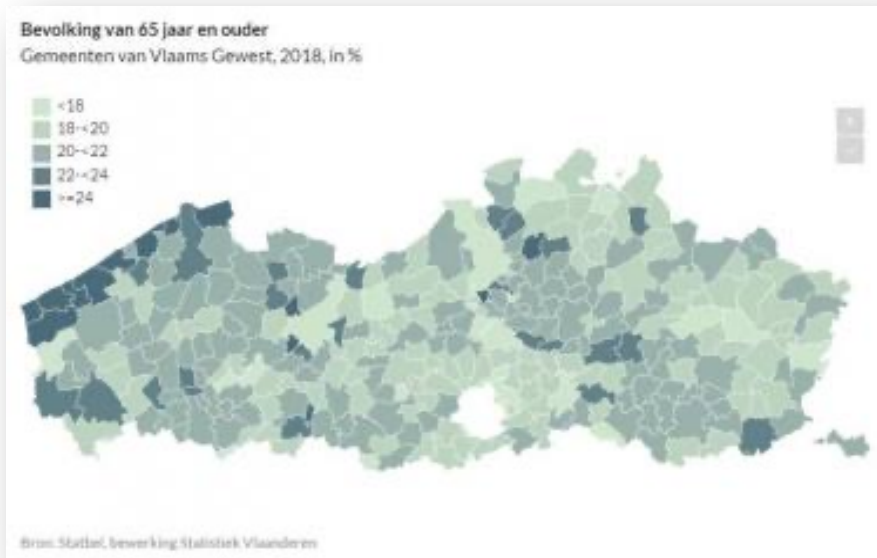
**Decentralized production** between:

- 11 policy areas
- 55 entities
- VGC: Flemish Community Commission
- Municipalities & provinces

**Sources:**

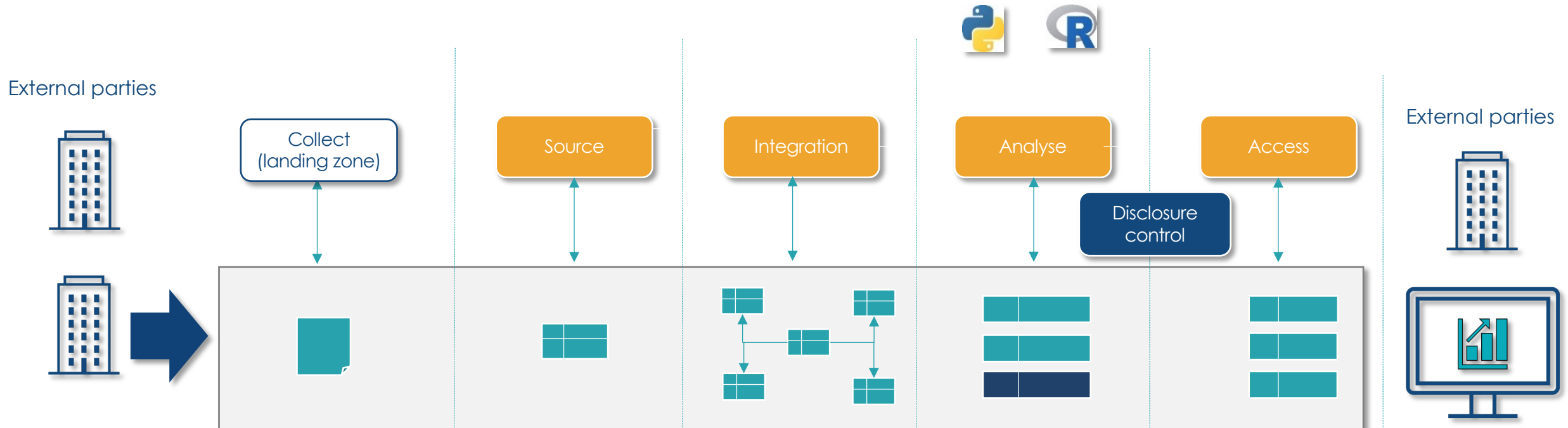
- Regional, local and federal sources and data owners
- Also sources from outside the government

## Data exchange challenge: From “statistics” to “data”

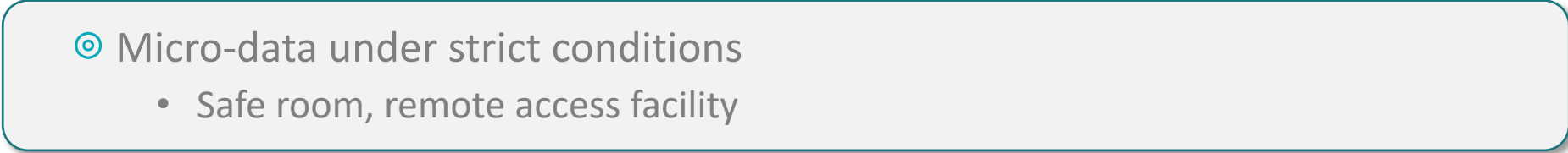


# Functional architecture of our Statistical Datawarehouse

## 4-layer model of Eurostat



## Portfolio of our offering

- ⊙ Static Aggregated 'open' data
  - Key numbers wrapped up in tables and graphics available on websites
  - Open data tables published on the Flemish open data platform
  
- ⊙ Dynamic Aggregated open data
  - Interactive (filters) data tooling available via website
  
- ➔ 
  - ⊙ Micro-data under strict conditions
    - Safe room, remote access facility
  
- ⊙ Trusted third party
  - Intermediate between governmental registers and end users
  - Adding, when necessary, additional services by linking data

## Comparative Analysis

### ⊙ Approach

- Desktop research
- Reference calls

### ⊙ Scope

- Platform
- Governance

### ⊙ Target

- 18 institutions
- Selection based on
  - Geographical spread
  - Innovative reputation
  - References

Institution
Statbel
Office of National Statistics (UK)
UK Data Service
Statistics Denmark
Statistics Estonia
Statistics Sweden
Northern Ireland Statistics and Research Agency
Research Data Centres of the Statistical Offices of the Federation and the Federal States
The National Institute of Statistics and Economic / CASD
Australian Bureau of Statistics
Eurostat
Statistics Canada
Stats New Zealand
Statistics Netherlands / CBS
National Statistics Office (Malta)
Statistical Office of the Republic of Serbia
Statistics Finland
Statistics Norway

## Outcomes – Platform

### ⊙ 4 types of platform

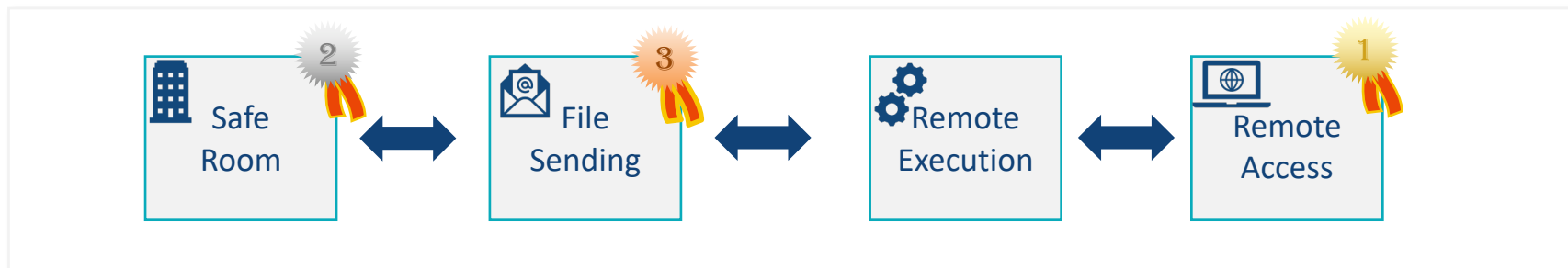
- Safe room
- Sending files
- Remote execution
- Remote access facility

### ⊙ Two popular options

- **Remote Access Facility** as new standard
- **Safe room** still widely spread

### ⊙ Two more ‘exotic’ options

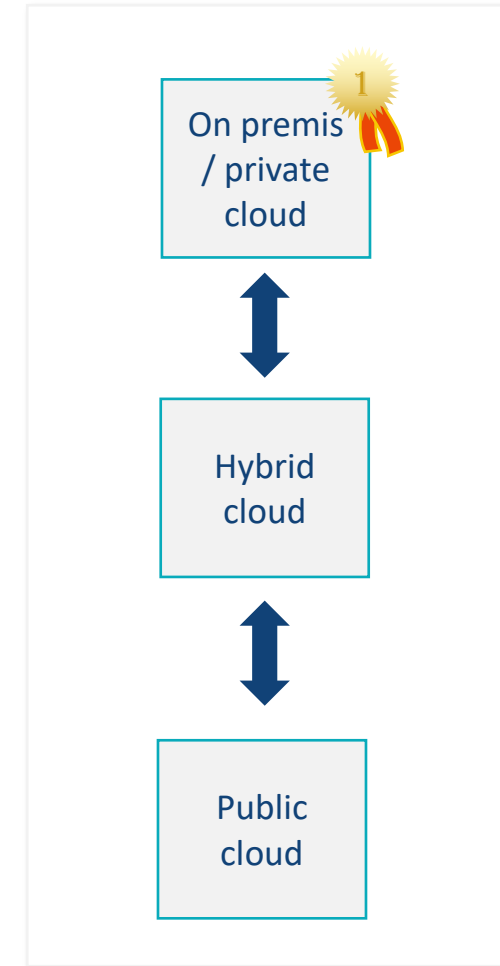
- **File sending** still used frequently
  - Non sensitive data
 or
  - Legal context of delegation
- **Remote execution** is an exception
  - Not the only option
  - For exceptional cases



## Outcomes – Platform

### Use of cloud in the Remote Access Facility context

- ⊙ All organizations chose an **on-premise** platform
  - Limited typical ‘private cloud’ functionality today
  - Due to ‘public perception’ of cloud data protection & security
  - Indications that this might evolve
- ⊙ Only **one public cloud**
  - ‘on top of’ their on-premise platform
  - In an experimental context
  - No exposure of micro data / sensitive data



## Outcomes – Platform

### Statistical Software

- ⊙ +/- De-facto standard
  - SAS / SPSS / SATA
- ⊙ Other popular options
  - R / Python
- ⊙ Limited use
  - GAUSS / MLWin

### Other Software

- ⊙ GIS software products are also often made available
  - ArcGis / gGIS
- ⊙ Other ‘supporting’ software
  - *Office software* (sometimes open source)
  - *Compression applications* – 7zip, WinZip
  - *Text Editing* - Notepad++, Latex, Ghostscript
  - *Reading PDFs* – Acrobat Reader



## Outcomes – Process – General

### ⊙ Application phase

- Make sure research request complies with the objectives and applicable laws, regulations and guidelines of the NSO

### ⊙ Preparation phase

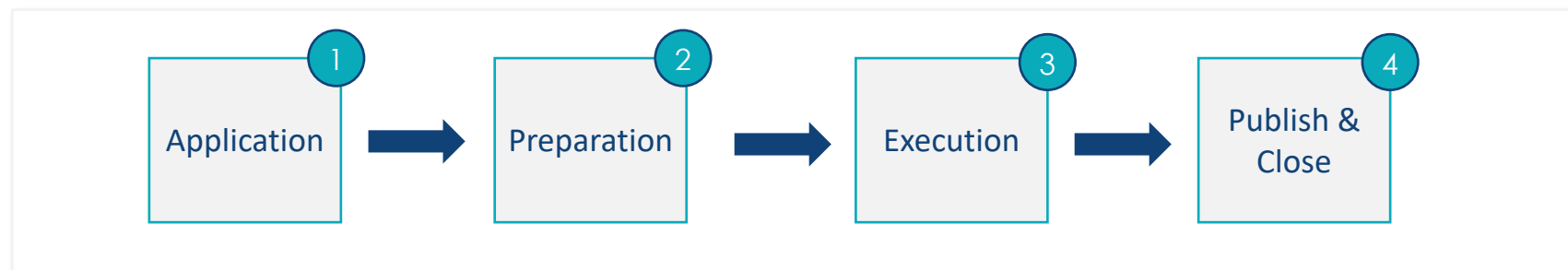
- Practical preparation

### ⊙ Execution phase

- Execution of the research

### ⊙ Publish & close phase

- Publication / export of results
- Official project closure



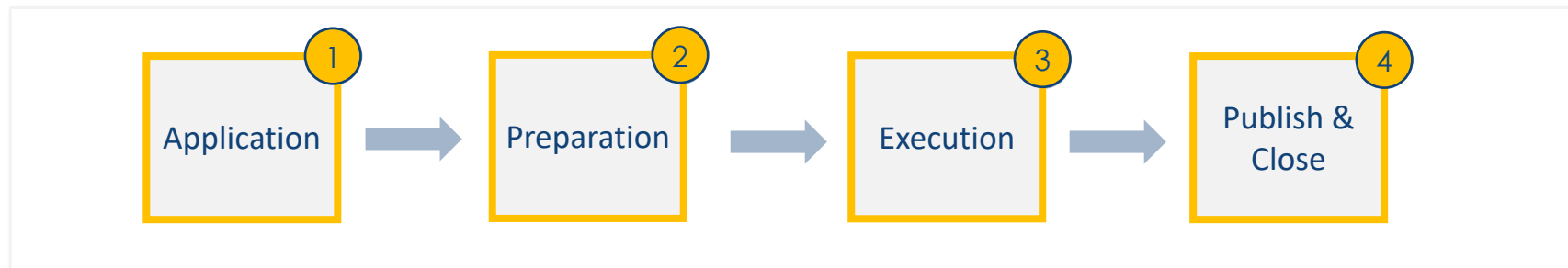
## Outcomes – Process – Remarks

- ⊙ Application supported by a **Data Catalog** (+/- half of the analysed offices)

Offers opportunities to automate hence speed-up the application & preparation phases

- ⊙ **Accreditation** process in two third of the cases also at the researcher and/or institution level
- ⊙ The role of the **DPO** in the process...

- ⊙ The importance of **training the researcher**, including a test
- ⊙ The importance of a **change request process**
- ⊙ Very useful guidelines by Eurostat for checking output data
- ⊙ To **archive** or not...



## Outcomes – Various Observations

- ⊙ Significant difference in the *'popularity' / usage*
  - Ranging from a few (4) per year to hundreds (1.300)
  - Reflected in the size of the support organization (technical & process support)
  - Not only related to the size of the territory/population
  - Some countries are better prepared from a legal point of view than other
  
- ⊙ Different *funding* approaches
  - Ranging from free usage in most cases to funding based on sizing parameters
  - No profit model – more a (partial) coverage of the cost
  
- ⊙ Significant difference in *maturity* of both platform and process
  - Initiatives started at different moments in the past
  - Different choices have been made – different 'speed' of evolution
  - Evolutive approach vs re-design approach

