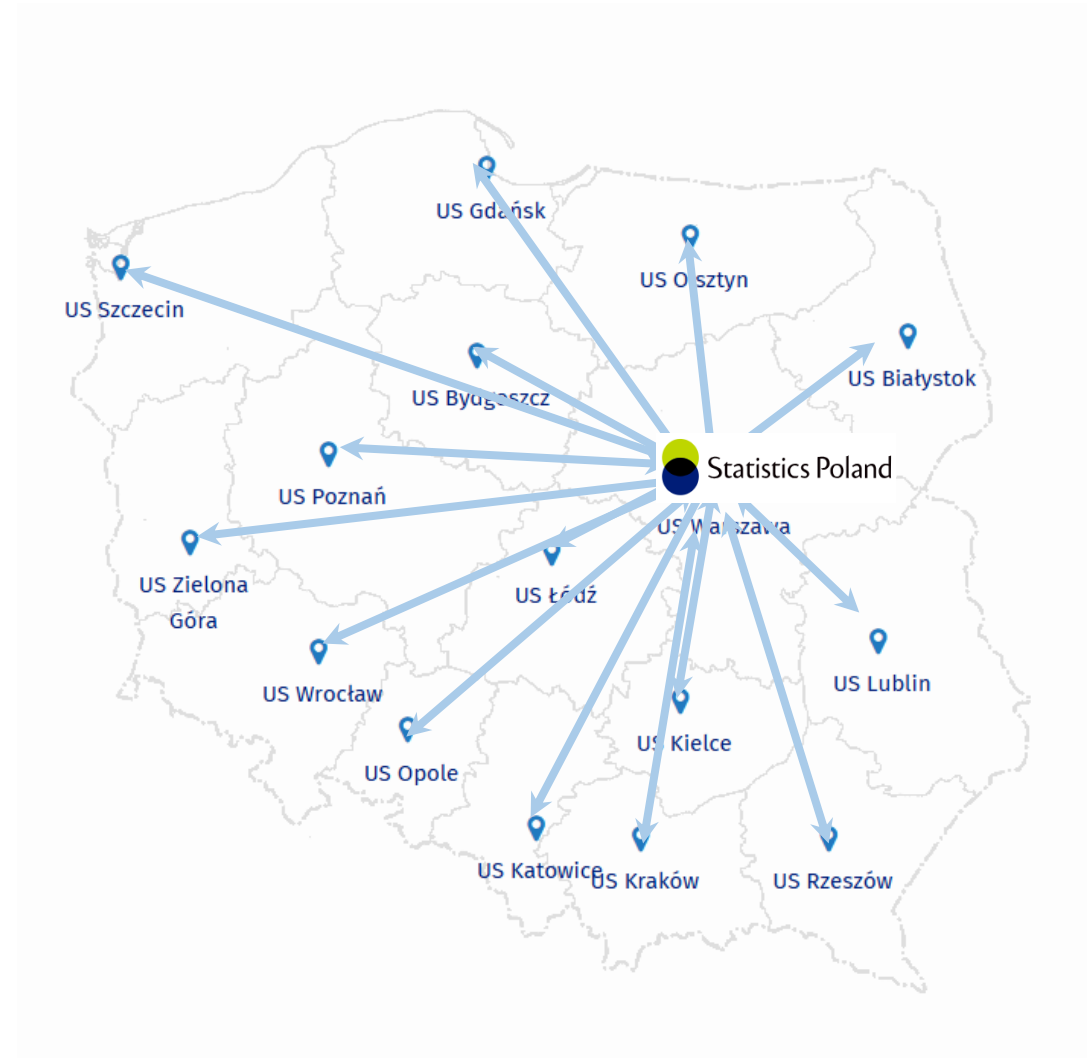


# Data Quality and Its Growing Value in the Modern World

Experience from Statistics Poland  
– Data Value Chain perspective

Dominika Rogalińska  
Vice-President

# Organizational Structure at Statistics Poland



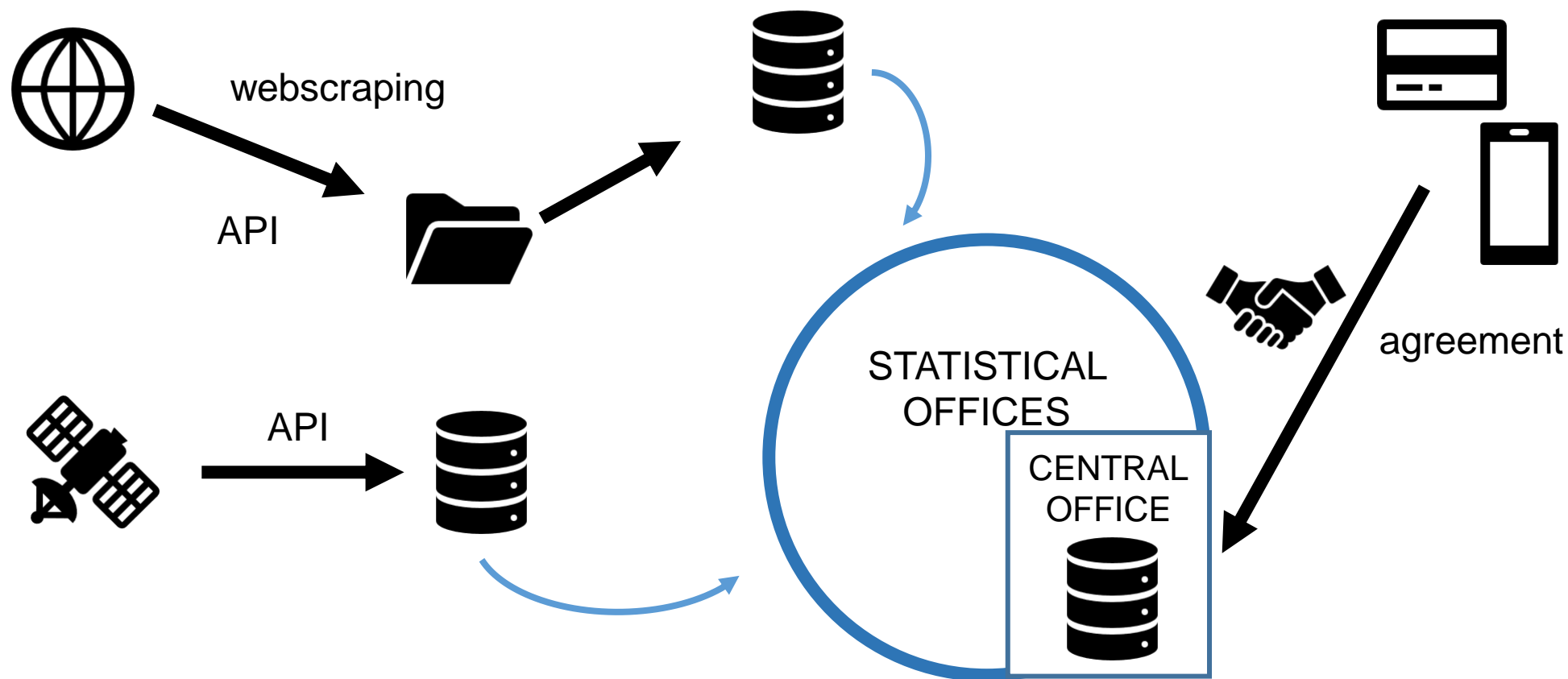
**16 regional** statistical offices

and

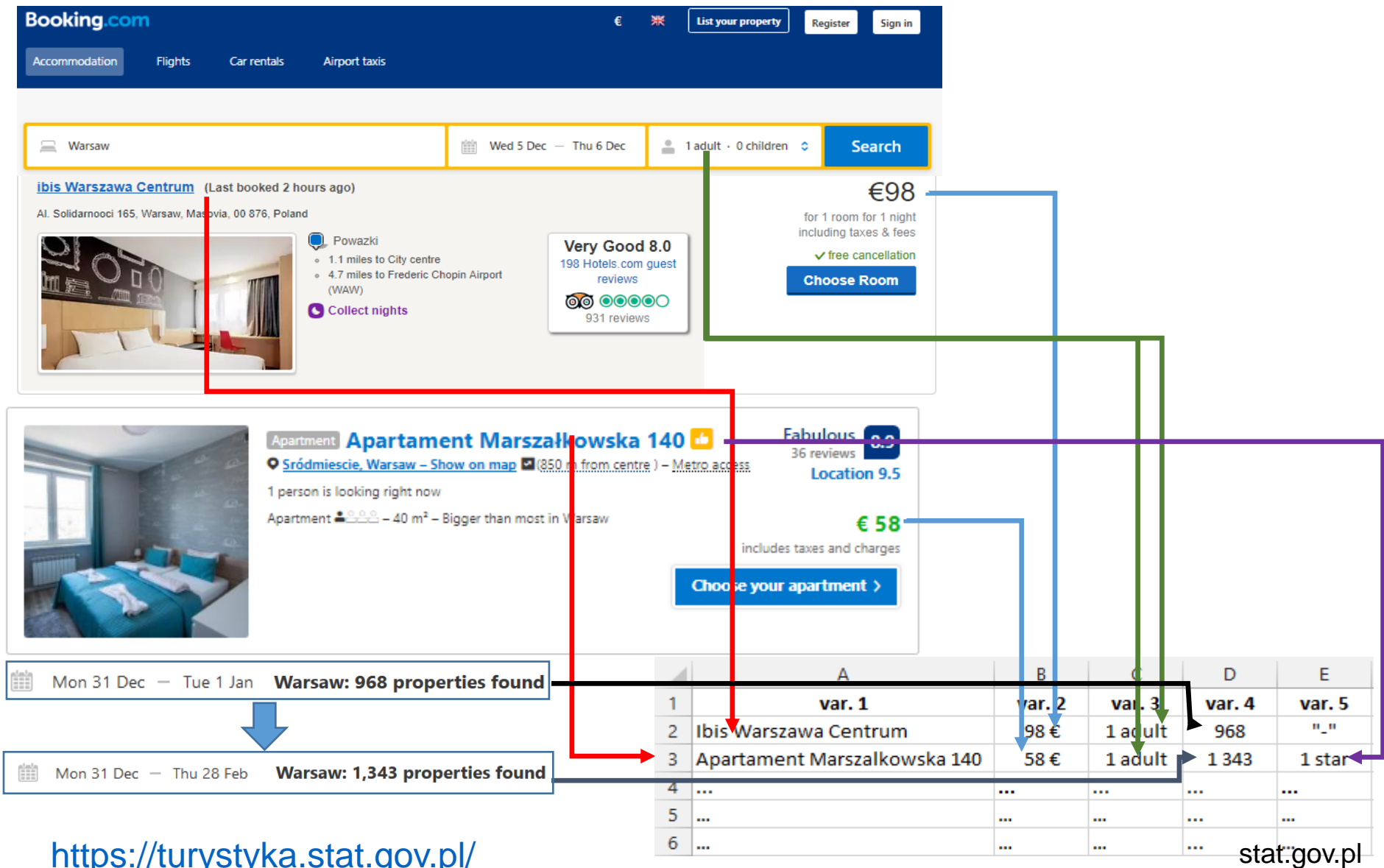
**1 central** statistical office in Warsaw

# Ensuring Data Quality at Each Step of the Data Value Chain

From webscraping and APIs to private data partnerships – each step includes validation, metadata control and quality monitoring

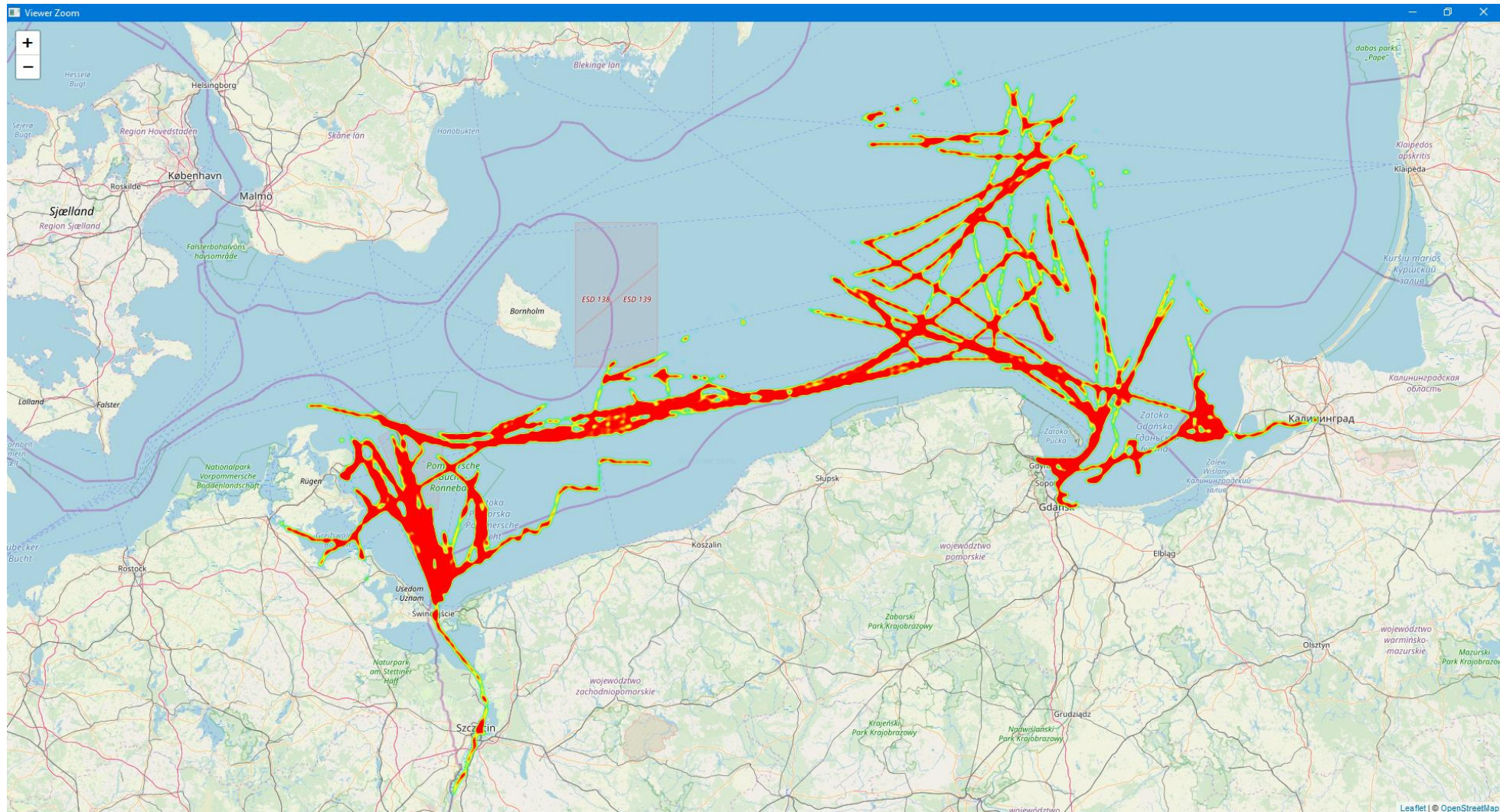


# Diverse Data Sources – Diverse Quality Challenges (Tourism statistics)



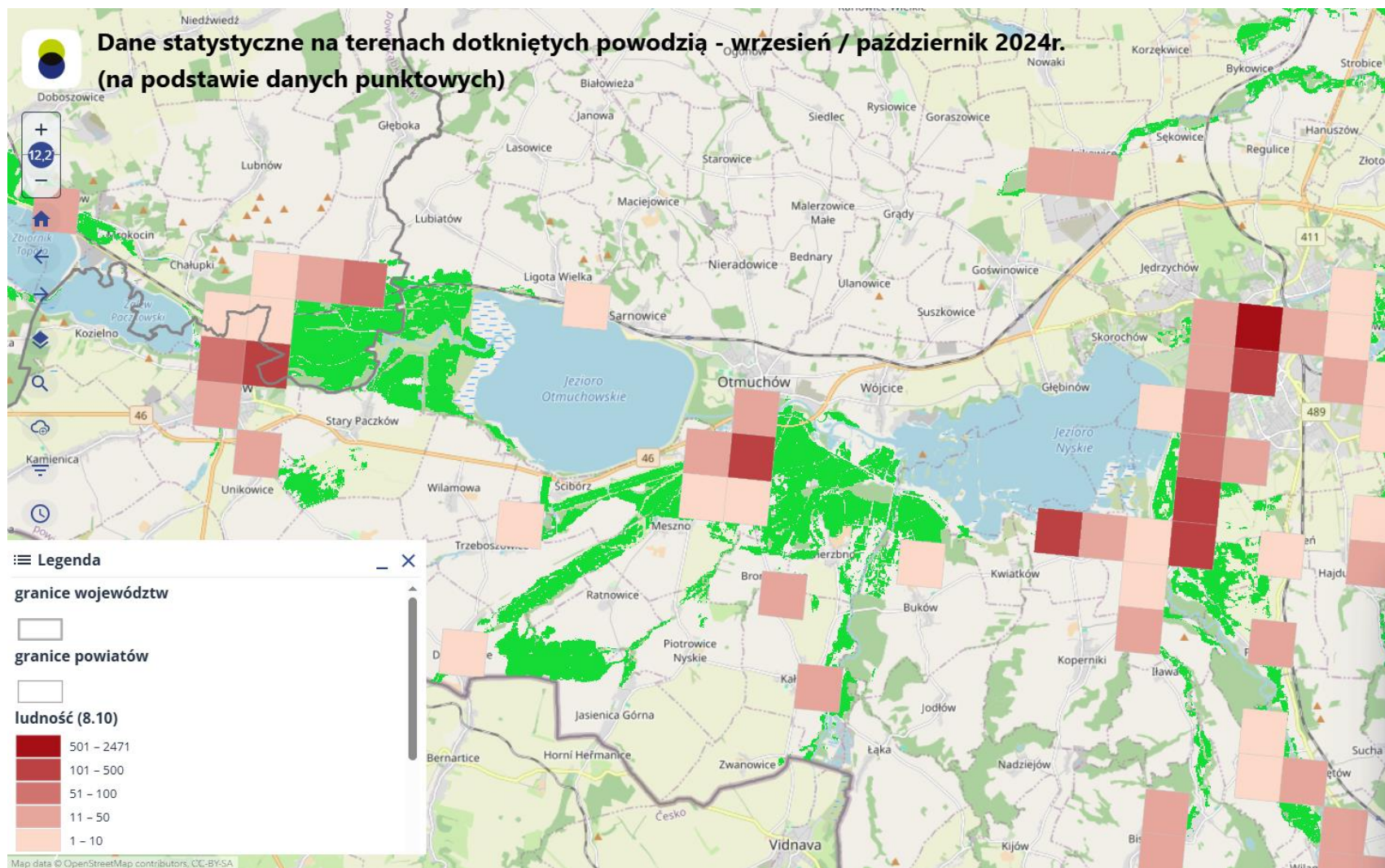


# Diverse Data Sources – Diverse Quality Challenges (Maritime Statistics)

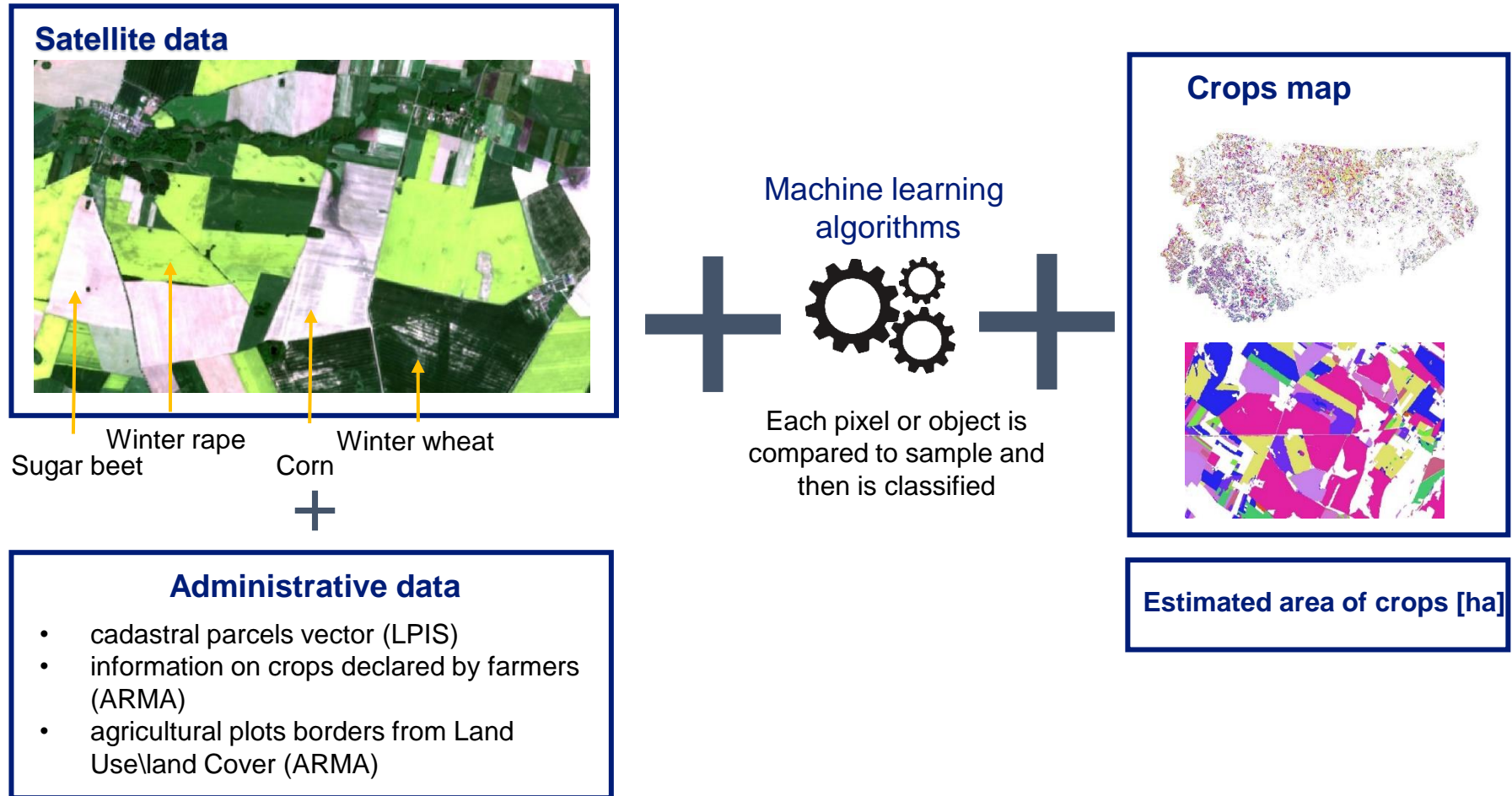




# Diverse Data Sources – Diverse Quality Challenges (Flood areas)



# Satellite imagery data – crop recognition



# Mobile network operators data



## Refugee Health in Poland

### Barriers to accessing health-care

Long waiting times to receive medical services

Information barrier

Other reasons

Required prescriptions, cost of medication

Cost of services for consultations, treatment

Logistics to attend facilities

Lack of access to medical care

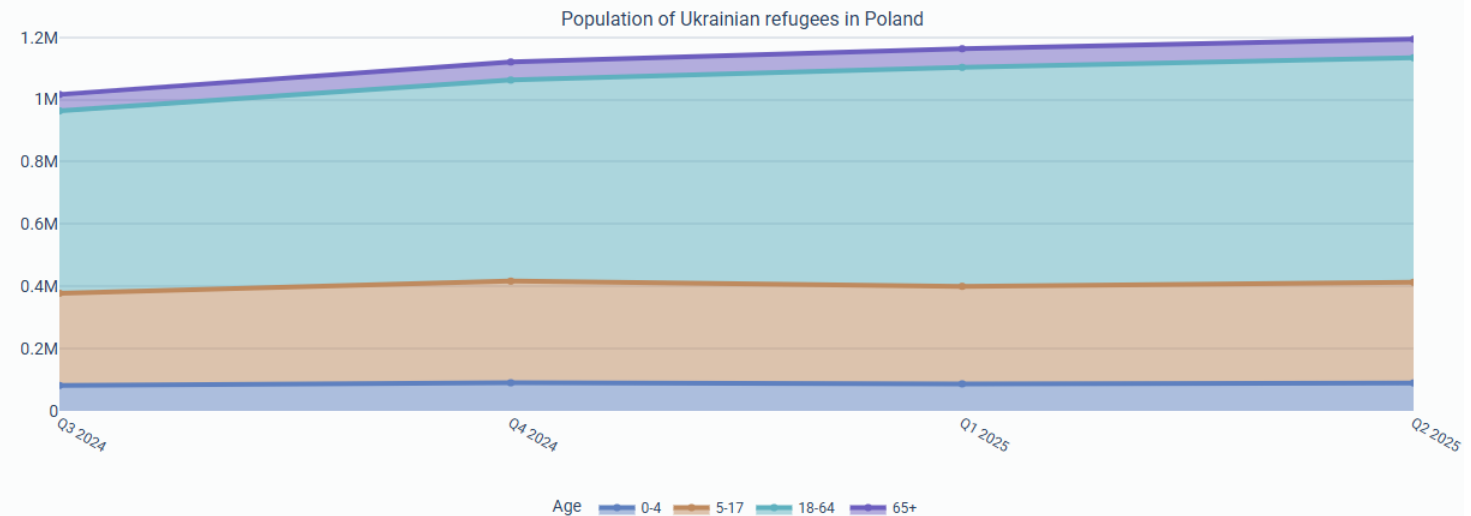
0

20

Period ● Q3 2024 ● Q4 2024 ● Q1 2025 ● Q2 2025



**Very nice!**



Statistics Poland

<https://healthofrefugees.stat.gov.pl/>  
<https://www.who.int/europe/publications/i/item/9789289062381>

stat.gov.pl



# Big Data and Data Quality in Official Statistics

From volume to value – quality as the cornerstone of trustworthy innovation

## Data sources

- web data
- satellite imagery
- IoT
- sensors

## Techniques

- webscraping
- machine learning
- autocoding and classifying
- imputation
- estimates

## Usage

- online job advertisement
- enterprise characteristics
- agriculture statistics
- urban areas
- tourism statistics
- trusted smart statistics
- data quality
- covid-19

# Big Data: Opportunities, Challenges and Risks for Data Quality

## PROFITS

- Reducing research costs and respondent burden
  - Shortening data delivery times, even in (near) real-time analyses
  - New information, new competencies in public statistics
- 

## CHALLENGES

- Collaboration with private sector data providers
  - Legal regulations lagging behind technological advances
  - Integration with existing public statistics systems, data comparability
  - Ensuring methodological soundness of new sources
  - Adapting IT infrastructure
- 

## THREATS

- Data access fees
- Statistical confidentiality, privacy protection
- Data quality and stability, insufficient information coverage
- Exodus of specialists to the private sector
- Quality uncertainty and lack of transparency of private data sources

# Big Data: Opportunities, Challenges and Risks for Data Quality

**DASHBOARDS**

## PROFITS

- Reducing research costs and respondent burden
- Shortening data delivery times, even in (near) real time
- New information, new competencies in public statistics

## CHALLENGES

- Collaboration with private sector data providers
- Legal regulations lagging behind technological advances
- Integration with existing public statistics systems, data comparability
- Ensuring methodological soundness of new sources
- Adapting IT infrastructure

## THREATS

- Data access fees
- Statistical confidentiality, privacy protection
- Data quality and stability, insufficient information coverage
- Exodus of specialists to the private sector
- Quality uncertainty and lack of transparency of private data sources

# Big Data: Opportunities, Challenges and Risks for Data Quality

## PROFITS

- Reducing respondent burden
- Shortening data delivery (near) real time
- New information, new content, new data sources

**DATA SCIENCE  
ACADEMY**

**DASHBOARDS**

## CHALLENGES

- Collaboration with private sector data providers
- Legal regulations lagging behind technological advances
- Integration with existing public statistics systems, data comparability
- Ensuring methodological soundness of new sources
- Adapting IT infrastructure

## THREATS

- Data access fees
- Statistical confidentiality, privacy protection
- Data quality and stability, insufficient information coverage
- Exodus of specialists to the private sector
- Quality uncertainty and lack of transparency of private data sources





# Big Data: Opportunities, Challenges and Risks for Data Quality

## PROFITS

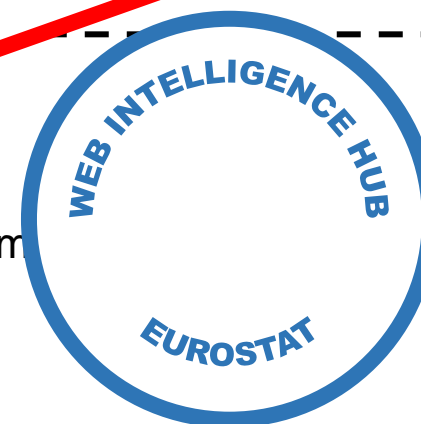
- Reducing respondent burden
- Shortening data delivery time (near) real time
- New information, new content, new data sources

**DATA SCIENCE  
ACADEMY**

**DASHBOARDS**

## CHALLENGES

- Collaboration with private sector data providers
- Legal regulations lagging behind technological advances
- Integration with existing public statistics systems, data compatibility
- Ensuring methodological soundness of new sources
- Adapting IT infrastructure



## THREATS

- Data access fees
- Statistical confidentiality, privacy protection
- Data quality and stability, insufficient information coverage
- Exodus of specialists to the private sector
- Quality uncertainty and lack of transparency of private data sources



# Big Data: Opportunities, Challenges and Risks for Data Quality

## PROFITS

- Reducing respondent burden
- Shortening data delivery time (near) real time
- New information, new content, new statistics

**DATA SCIENCE  
ACADEMY**

**DASHBOARDS**

## CHALLENGES

- Collaboration with private sector data providers
- Legal regulations lagging behind technological advances
- Integration with existing public statistics systems
- Ensuring methodological soundness of new data sources
- Adapting IT infrastructure

**DATA  
SHARING  
FORUM**

WEB INTELLIGENCE HUB

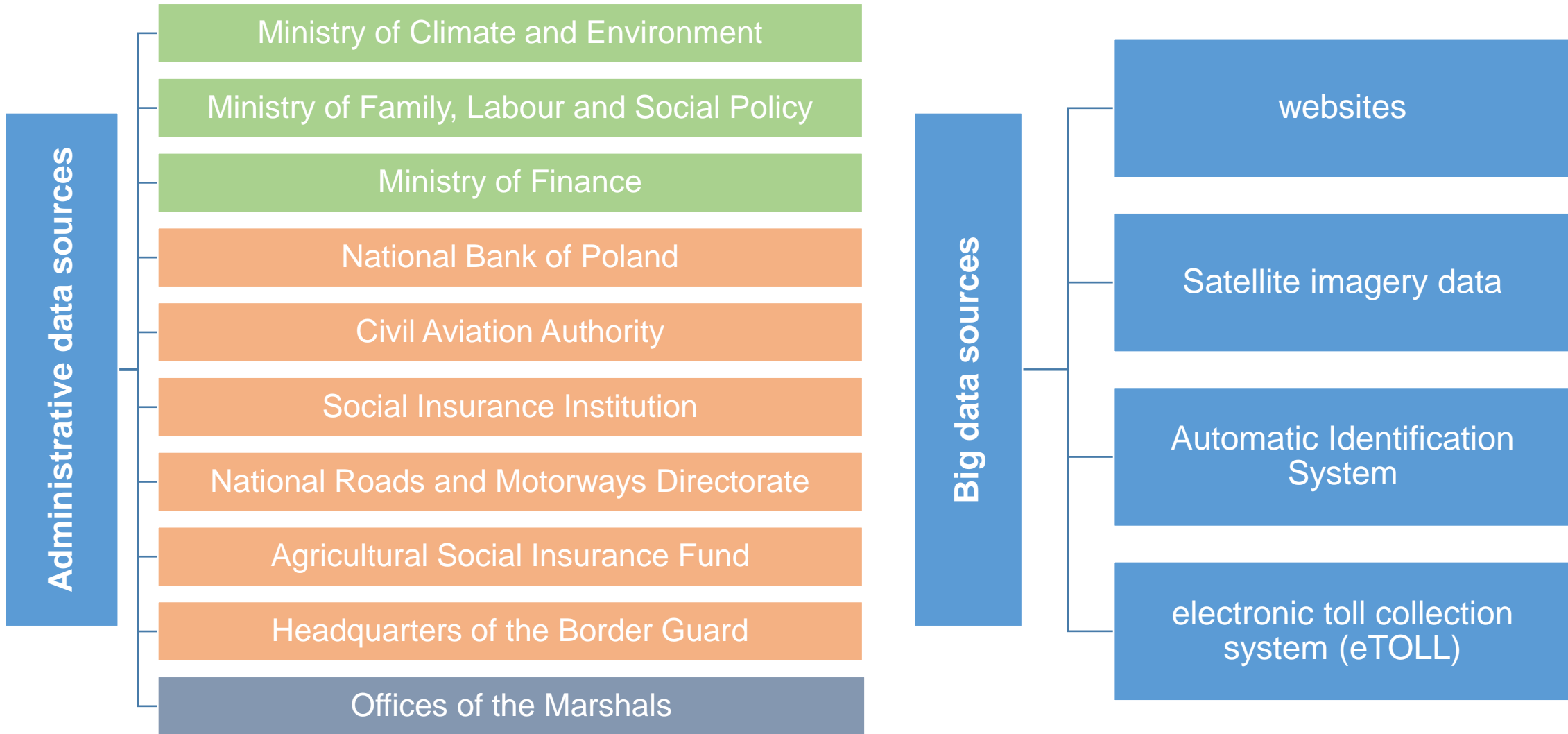
EUROSTAT

## THREATS

- Data access fees
- Statistical confidentiality, privacy protection
- Data quality and stability, insufficient information coverage
- Exodus of specialists to the private sector
- Quality uncertainty and lack of transparency of private data sources



# Data sources



# Key dimensions of data quality in the data value chain

Accuracy and  
reliability

Timeliness  
and  
coherence

Transparency  
and metadata  
standards

Ethical use  
and privacy  
assurance

Reusability  
and long-  
term value



# Looking Ahead – the future of data quality in the modern world

- AI ready data
- Methodology for quality of bigdata and ML models.
- Continuous development of the IT system while carrying out ongoing tasks
- From collection to communication – quality is embedded at every step of the data value chain.

Thank you  
Сіздерге рақмет

Dominika Rogalińska  
[D.Rogalinska@stat.gov.pl](mailto:D.Rogalinska@stat.gov.pl)

**High-quality data  
builds trust.  
Trust builds value.**