

OF COMBINING ADMINISTRATIVE SOURCES AND DIRECT DATA COLLECTION

Janusz Dygaszewicz

Director

Department of Programming and Coordination of Statistical Surveys

Central Statistical Office of Poland

Member of Executive Committee of the UN-GGIM: Europe President of European Forum for Geography and Statistics EFGS

Doha, Qatar, 11 December 2017





ورشة عمل بشان تحديث الإحصاءات الرسمية في دولة قطر Workshop on Modernization of Official Statistics in Qatar

> الاثنين اا ديسمبر ۲۰۱۷ – فندق روتانا, سيتي سنتر Monday 11 December, 2017 - Rotana City Center Hotel - Doha

Long-term history (tradition) of Polish statistics

- 1789 first countrywide population census in Poland (based on registers!! ②)
 - The register of nobility and burghers
 - The parish register of births and deaths

Methodology in 2011 - Mixed Model for Population and Housing Census

Combining Census

a combination of data from administrative sources (full survey covering basic demographic variables)

with data acquired from

ad-hoc 20% sample survey.

Data collection channels in 2011 Census Round

Administrative Sources

 Including spatial data reference registers

Self-enumeration by Internet

• **CAWI (CAII)** – Computer Assisted Web (Internet) Interview

Telephone Interview

• **CATI** - Computer Assisted Telephone Interview (Call Center)

Face-to-face Interview with respondents executed by the census enumerators

 Registered on hand-held terminals with usage GPS and GIS service
 CAPI - Computer Assisted Personal Interview

CAXI





CAXI

- CAII/CAWI Computer
 Assisted Internet Interview,
- CAPI Computer Assisted Personal Interview,
- CATI Computer Assisted Telephone Interviewing.

Census organisation

The project timetable

National Population and Housing Census 2011

2007 Start of preparation

IV-V.2010 Trial census NSP 2011

IV-VI.2011 NSP 2011













IX-X.2009 Trial Census PSR 2010 **IX-X.2010** PSR 2010

2013 End of project

Agriculture Census 2010

Organization structure

Over 10 external companies optoviding on going support

16 central controllers

32 managers of the WCZS and the WCC

571 voivodship controllers

642 statistical interviewers

around 2800 enumerator leaders on gmina level

around 18000 census enumerators

ADMINISTRATIVE REGISTERS

The use of administrative sources in censuses

- The usage of administrative sources in the census:
 - direct source of research data ,
 - source of information to create a list of entities
 covered by the census frame (address-housing survey) ,
 - in addition, a source of information for :
 - imputation,
 - data estimation,
 - comparison the quality of the data.

Registers - data acquisition

Data Owners:

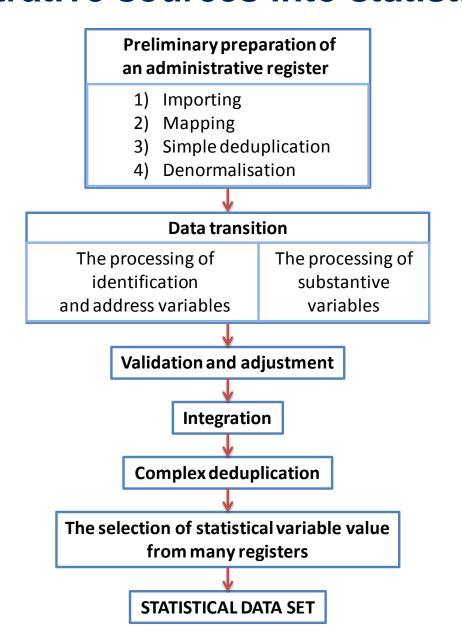
- Ministry of Finance,
- Ministry of Interior and Administration,
- Ministry of Justice,
- Agricultural Social Insurance Fund,
- National Health Fund,
- Agency for Restructuring and Modernisation of Agriculture,
- Agricultural and Food Quality Inspection,
- Agency for Geodesy and Cartography,
- State Fund for Rehabilitation of Disabled Persons,
- County Offices,
- Commune Offices,
- Regional Offices,
- Telcoms,
- Energy Suppliers,
- Office For Foreigners,
- Social Insurance Institution,
- Housing Managers,



Key administrative sources used in the Polish official statistics

- population registration system
- tax system
- economic activity information system
- farming activity information system.
- social security system
- social insurance system
- health insurance system
- real estate information system
- education information system
- vehicle and vehicle owners information system

The data transformation model from administrative sources into statistical data sets



Data quality -measures-

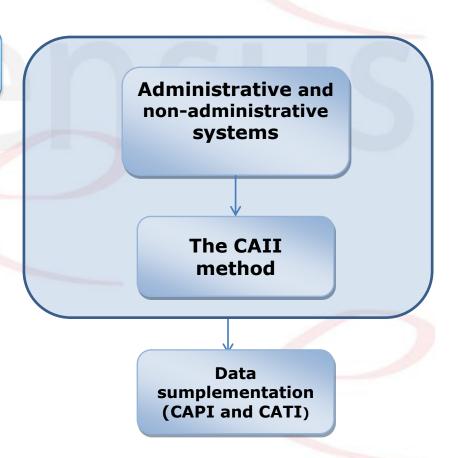
- 1. Measuring the quality of administrative registers
 - timeliness of data
 - methodological compatibility
 - completness
 - identification standards used in the registry
 - usefulness
 - compatibility of data in administrative sources to data obtained in the study/survey
- Measuring the quality in processing of data registers
 excessive coverage error rate
 incomplete coverage error rate subjective indicator of completness
 objective indicator of completness
 imputation rate
 data correction index
 integration data from various sources index

Census startegy

Full scale survey (short form - 15 questions covered by admin data)

Full-scale survey:

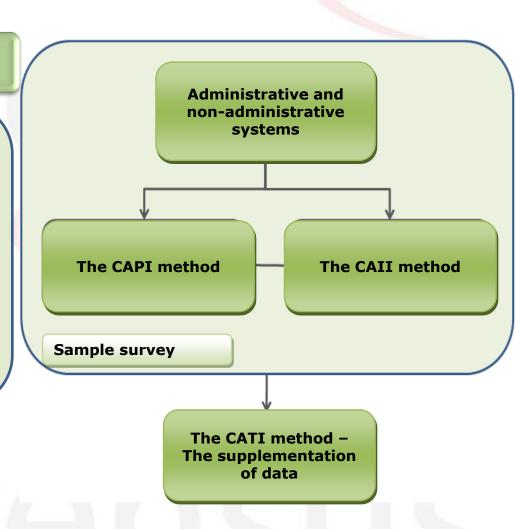
- Data from administrative register Master Record
- Data acquired using the CAII method
- Data supplemented using CATI and CAPI method



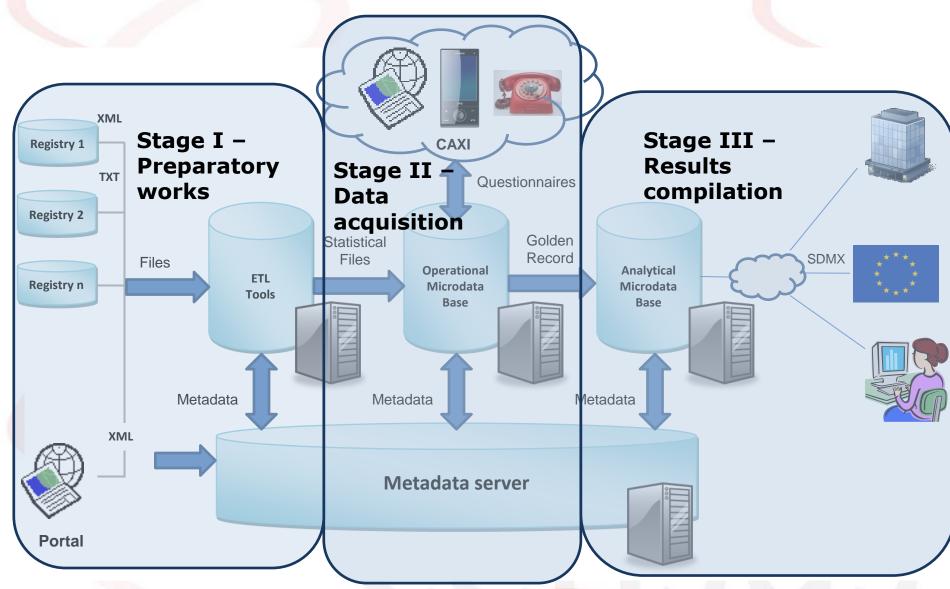
20% sample survey (long form - about 100 questions)

Sample survey:

- Data from administrative register Master Record
- 2) Data acquired using:
 - The CAII method
 - The CAPI method
- Data supplemented using CATI method



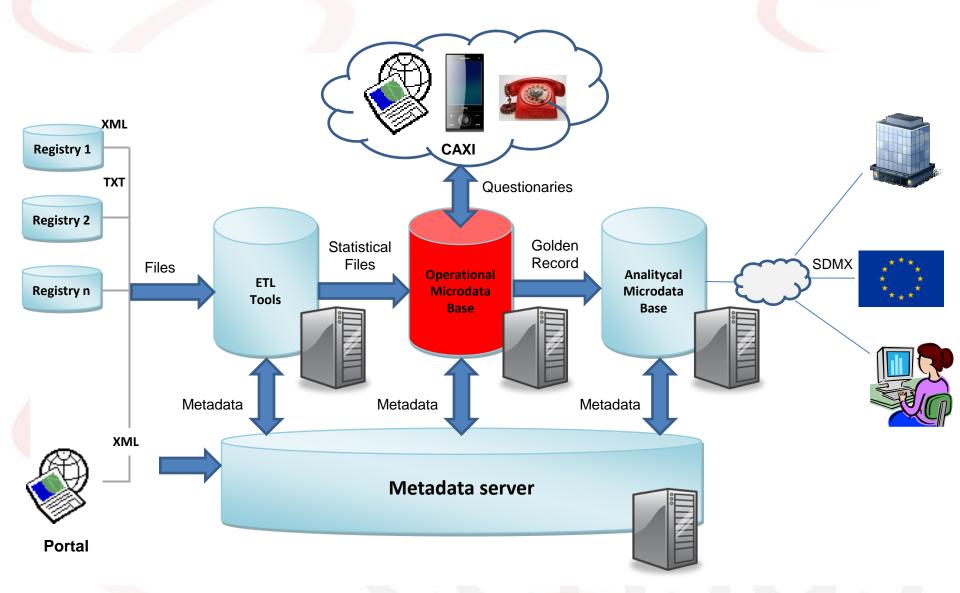
Architecture solution



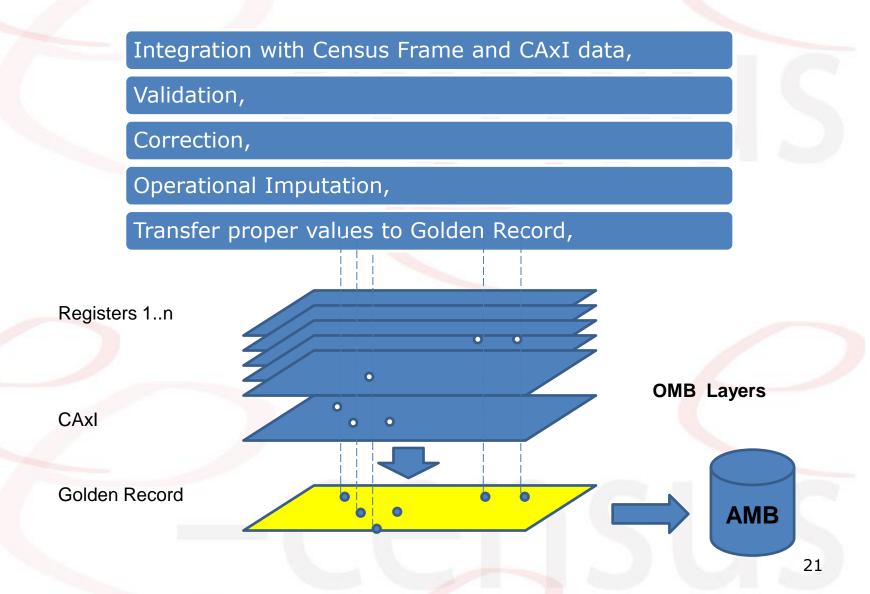
GOLDEN RECORD

-census

Golden Record generation

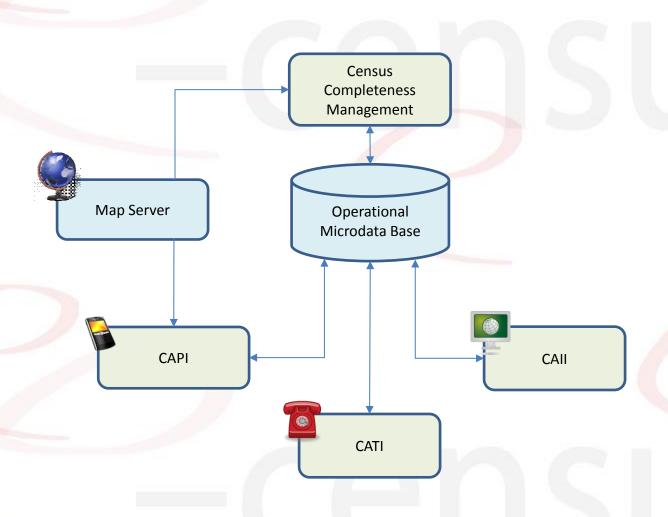


Golden Record generation

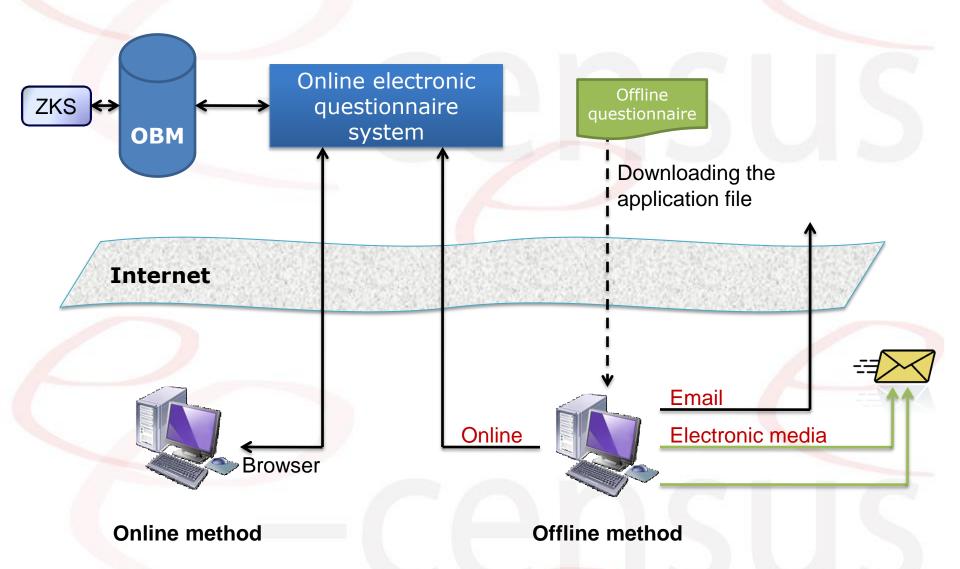


DATA ACQUISITION

On-line channels for data collection System Architecture



The architecture of a CAII/CAWI method



Self-enumeration by Internet filling the questionnaire by the respondent

Identification

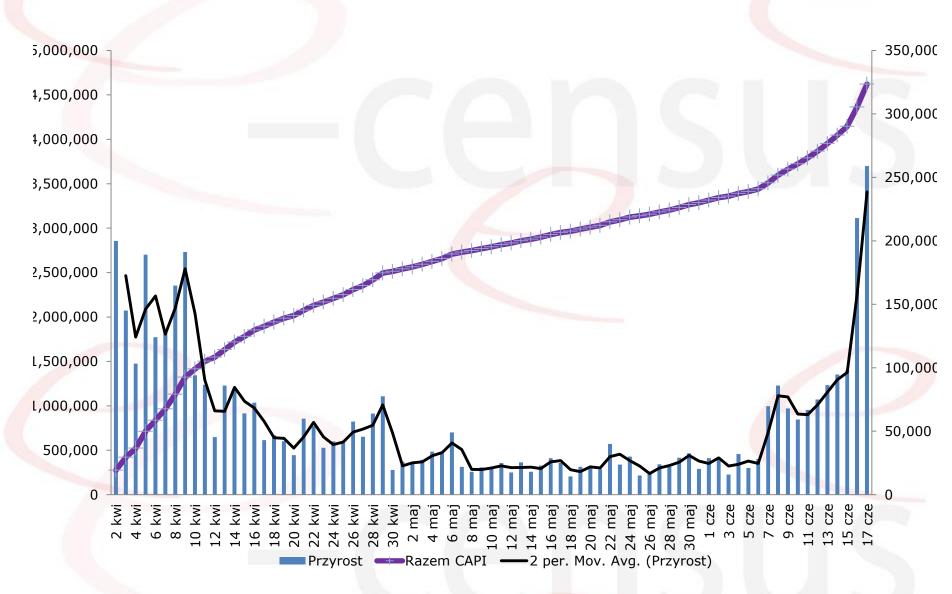
Used to confirm the identity of the respondent.

- Entering identification data in a questionnaire(f.ex.: PIN, NIP, first name, last name) or additional authentication qualities (f.ex. a place of birth, mother's maiden name)
- Establishing a password which jointly with PIN was the basis of authentication within 14 days

Typical person using selfenumeration

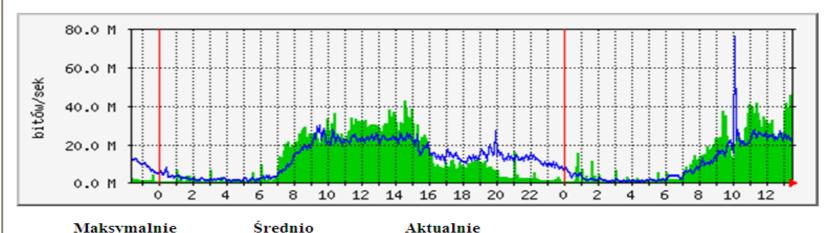
- □ A man
- □ Age: about 24 years old
- □ A city inhabitant
- □ Secondary degree

Self enumeration trend



Ostatnie uaktualnienie statystyki Środa dzień 25. Maj 2011 o godzinie 13:30, gdy 'r-ext1' był włączony przez 6:50:23.

`Dzienny' Graf w ciągu (5 Minut/y - Średnio)

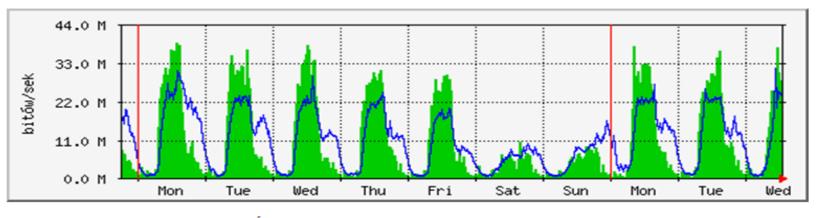


Maksymalnie Do 45.1 Mb/s (4.7%) Z 75.5 Mb/s (7.9%) **Średnio** 11.4 Mb/s (1.2%) 11.1 Mb/s (1.2%)

35.2 Mb/s (3.7%)

23.1 Mb/s (2.4%)

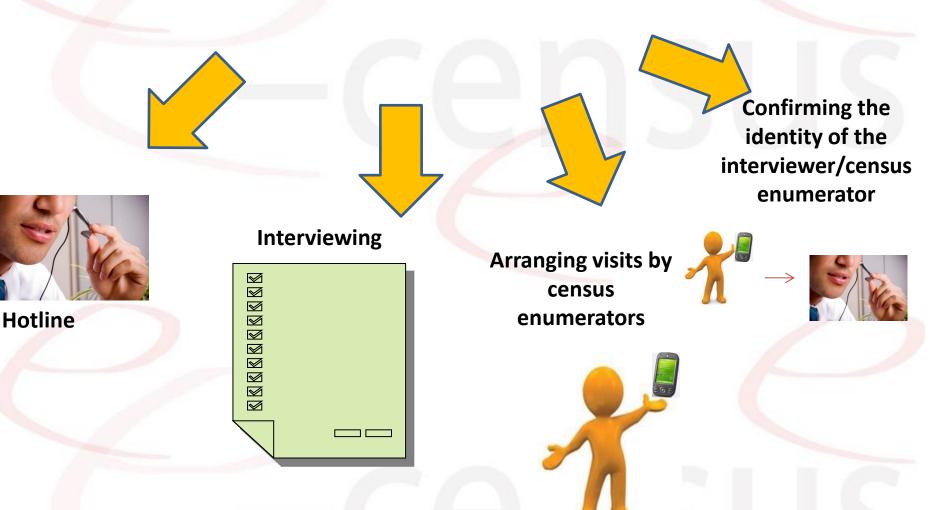
`Tygodniowy' Graf w ciągu (30 minut - Średnio)



CATI – Computer Assisted Telephone Interview

- scheduled as the second (following CAII) channel of collecting data;
- working posts of telephone interviewers located in separated Call Center studies;
- telephone interviewers provided with professional equipment.

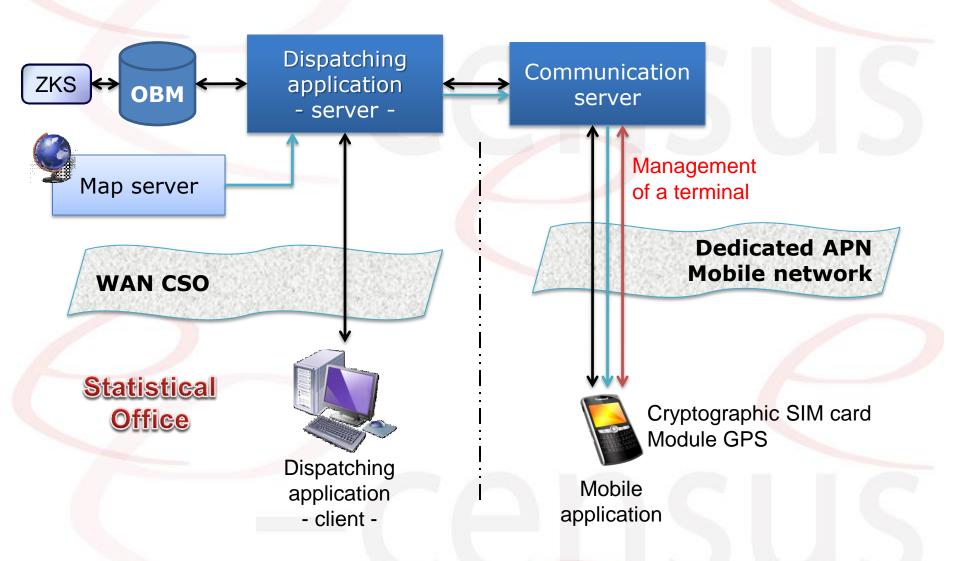
The most significant functionality of Call Center



CAPI – Computer Assisted Personal Interview

- the third channel of data collection in the case of failure to obtain a complete set of data via CAII and CATI channels
- direct interviews in households (first or second channel)
 - where such a way of proceeding results from adopted methodology or
 - whose members has not expressed consent for a telephone survey

The architecture of a CAPI method



HH - Mobile terminal with GPS

- HTC Touch Pro2
- Screen
 - touch-screen
 - size 3,6"
 - resolution 480 x 800 pixels
 - sliding, tilting convenient usage
- sliding, 5-rows QWERTY keyboard
- GSM/GPRS/EDGE/UMTS/HSPA
- GPS module
- camera 3,2 MP
- Windows Mobile® 6.5

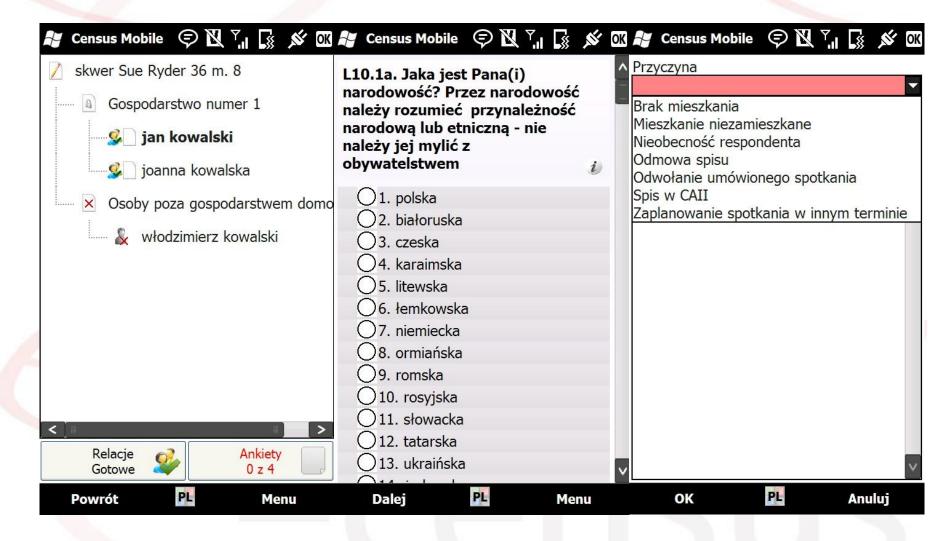




Enumerator

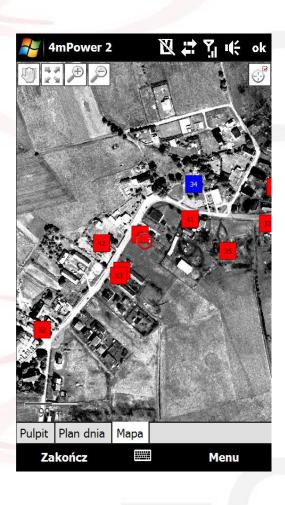


Enumerator



The GIS application for field operations - handheld devices

Enumerator – GIS technology



Map module - GIS

- Ortophotomap
- Cadastral Data
- Assigned Tasks
- Started Tasks
- Completed Tasks

Enumerator



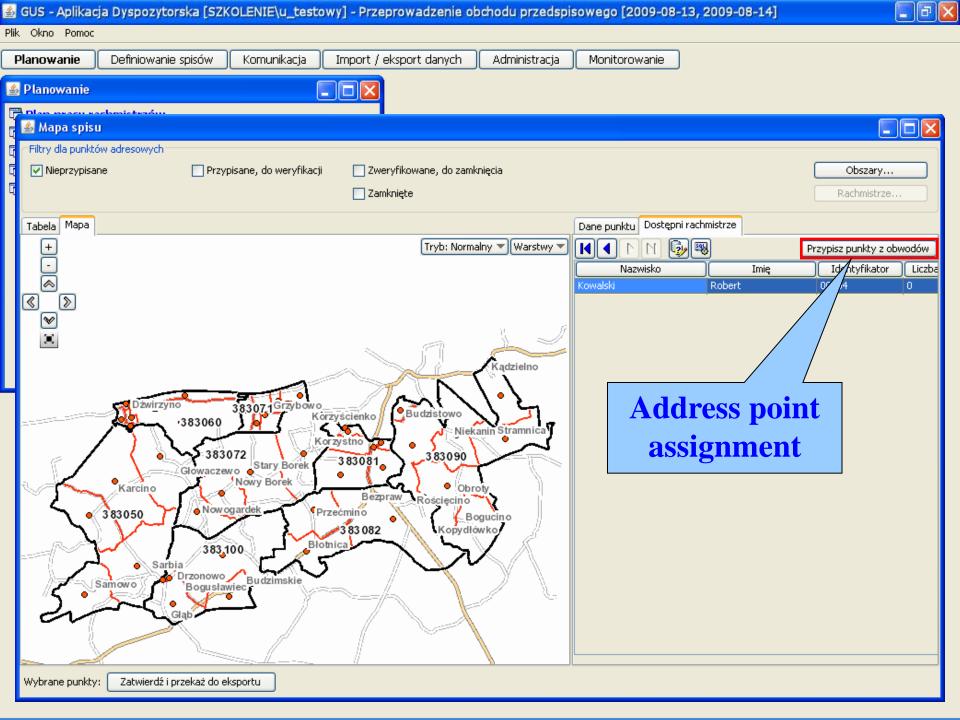
Alarm procedure

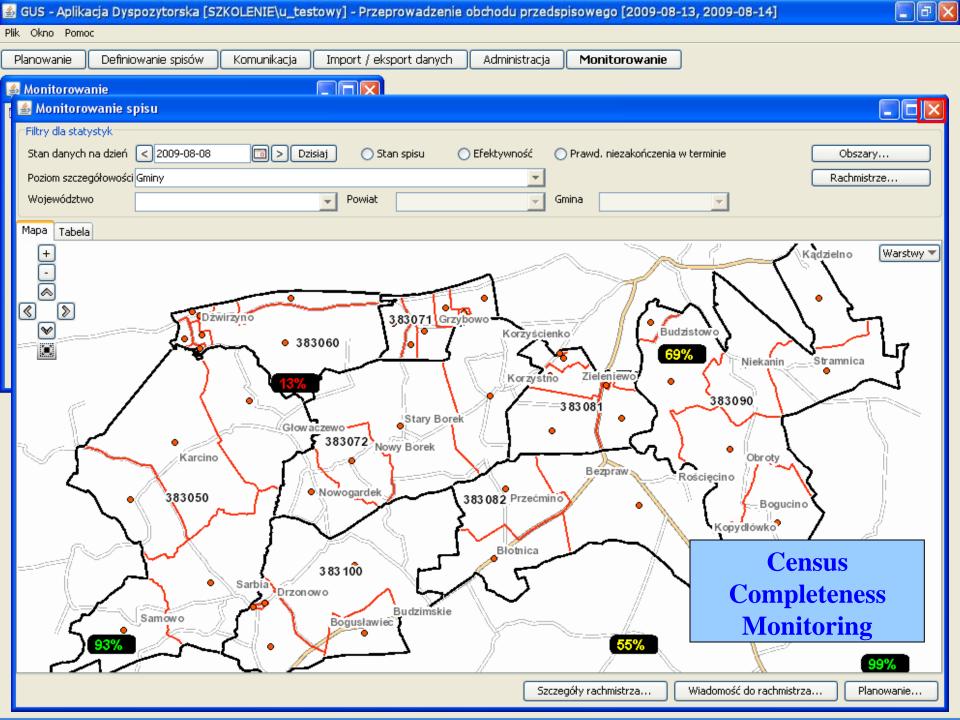
- In emergency situations,
 enumerators have a possibility
 of sending an alarm signal to
 their supervisors
- Alarm notice is sent
 to the supervisor application
 and via SMS to the supervisor

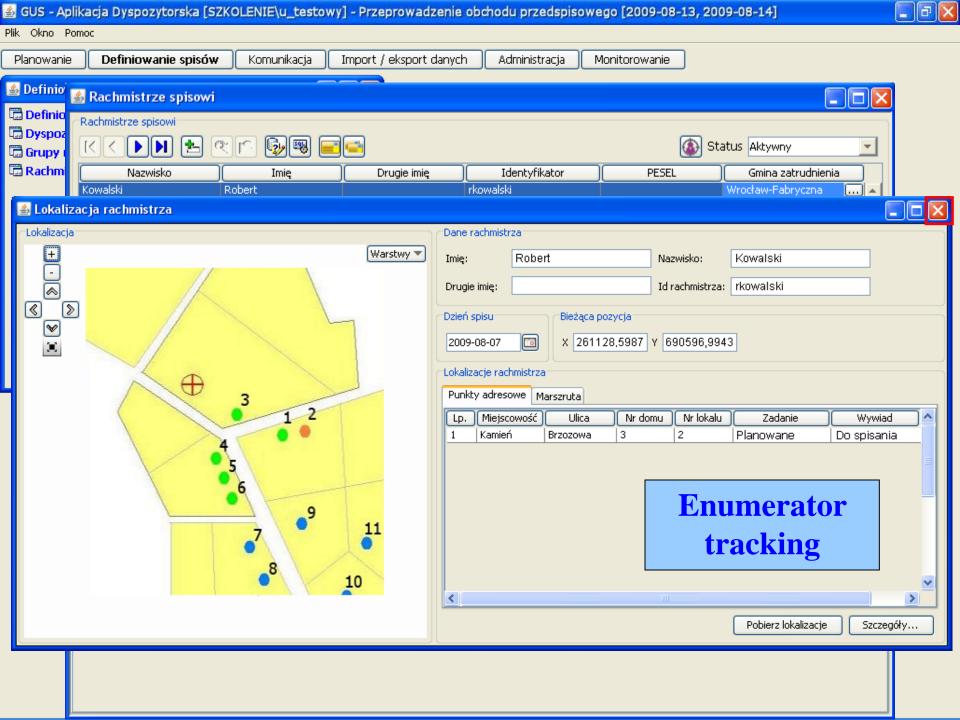
Regional Supervisor (NUTS2) level Field enumeration management

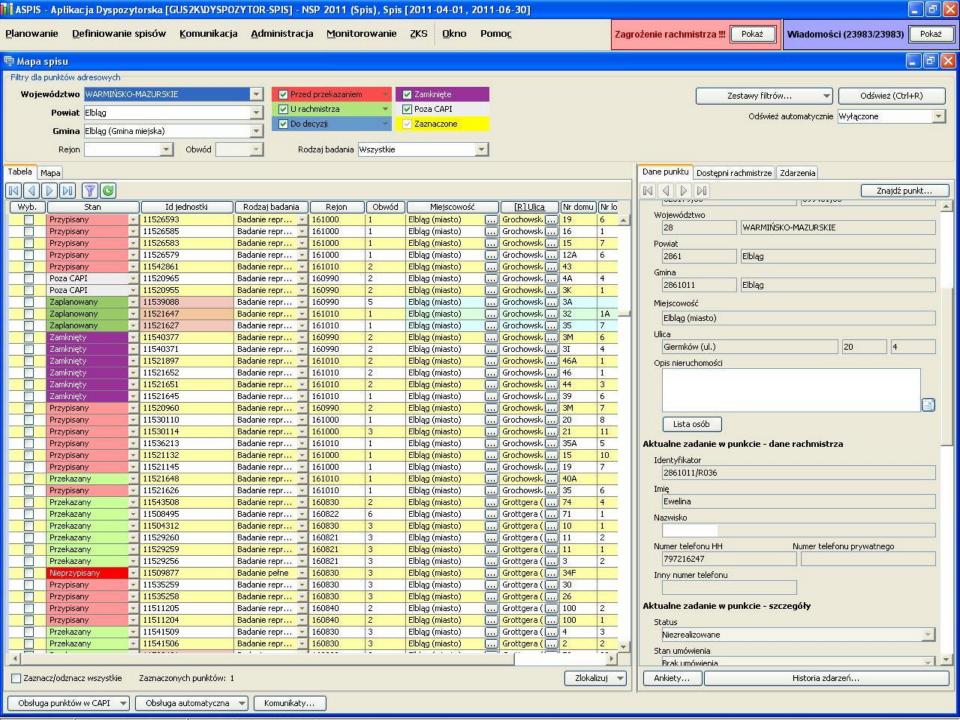
Responsibilities:

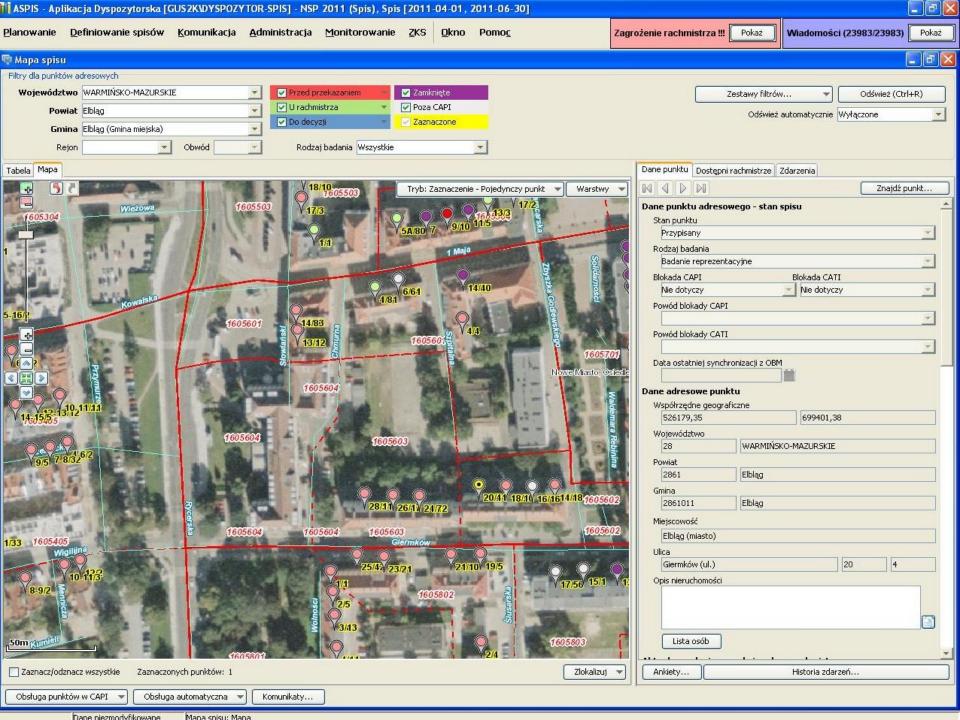
- Address Point and Census Area management
- Enumerator monitoring
 - Census Progress
 - Localization and trail
- Emergency situation management
 - Providing help for enumerators
- Providing necessary information to enumerators

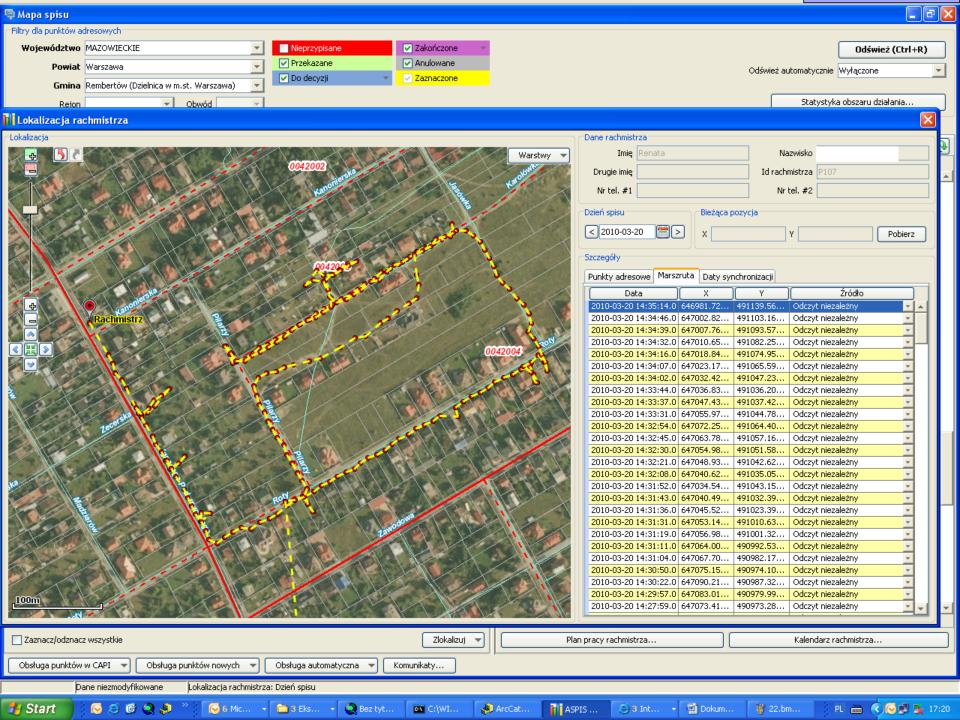


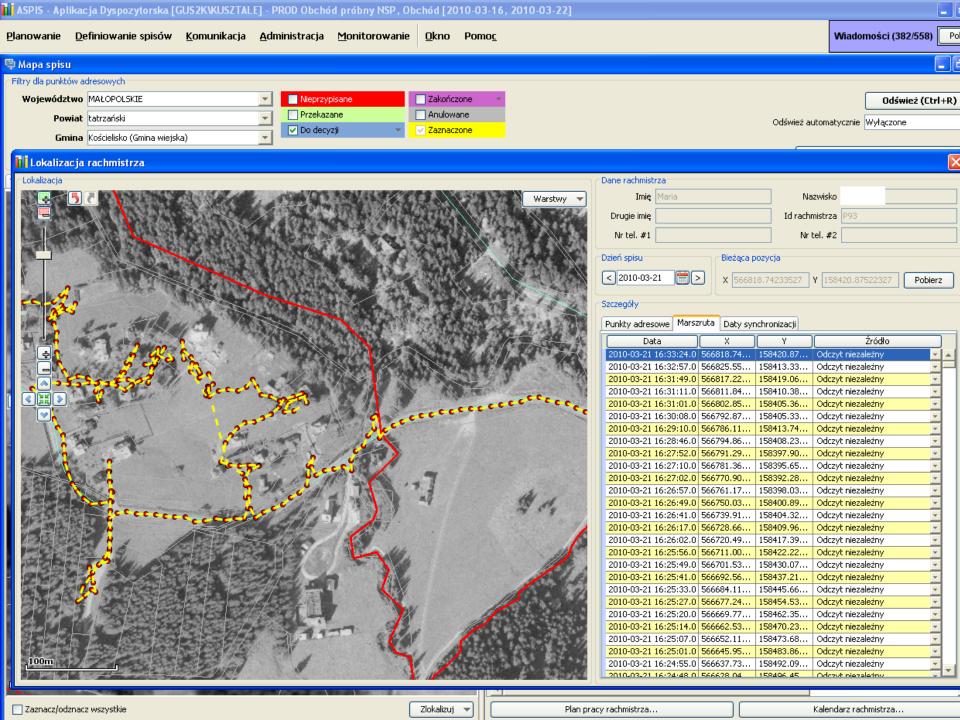


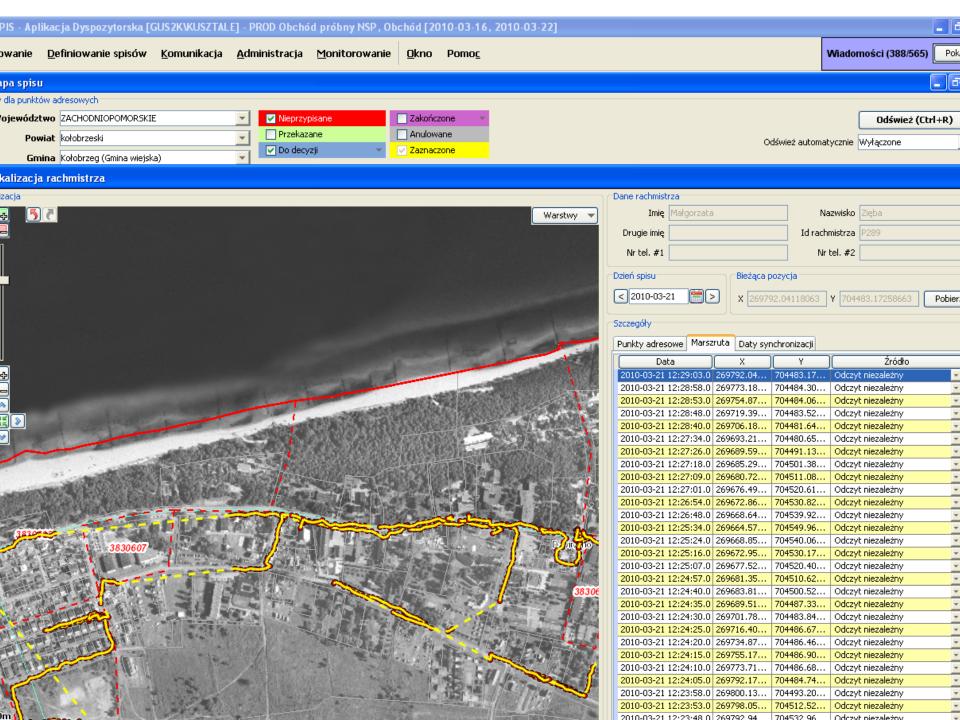


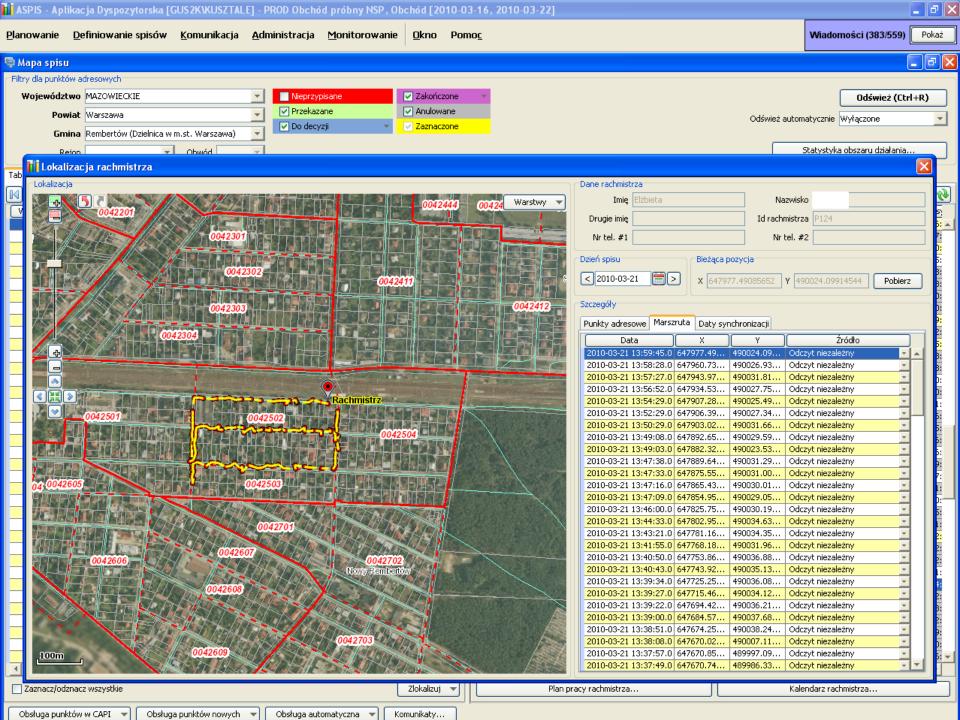












Thank you for your attention

e-mail: j.dygaszewicz@stat.gov.pl