Ensuring data confidentiality of All-Russian Population Census 2020

Oxana Dmitrieva,
Anna Troitskaya
Federal State Statistics Service,
Rosstat
What we are protecting from?

- Third-party access
- Information disclosure
Legal basis

1. Federal law “About all-Russian population Census” from 25.01.2002 № 8-FZ
2. Federal Law “On Information, Information Technologies and Information Protection” from 27.07.2006 № 149-FZ
4. “Requirements for the protection of personal data during its processing in personal data information systems” (approved by the Decree of the Government of the Russian Federation from 1.01.2012 № 1119)
Data collection during All-Russian Population Census 2020

Self-completion of electronic questionnaires in the Internet by respondents

The completion of electronic questionnaires on tablet computers with specialized software installed

The traditional completion of paper questionnaires (two-sided machine-readable Census forms)
Stages to automate

Preparation for the Census
Running
Automated processing of the materials
Results
Publication of the results
Post-Census use of the materials
Results storage
Data confidentiality in All-Russian Population Census 2020

- Organizational methods
  - contracting with interviewers on data confidentiality (paper forms and tablet computers)

- During data processing
  - technical: passwords, filters, approved access to PC and servers
  - In paper forms name doesn’t go to processing, cut off personal information
  - Data encryption

- In publication
  - methodological data protection by confidentiality algorithms (during customized requests to data bases) when the person can be identified (3+ person = disclosure) → noise method, perturbation
Data confidentiality in All-Russian Population Census 2020

Most of Census data will be collected through Internet and on tablet computers. Data management platform ensure data protection and management.

If the tablet is lost, then the information on the tablet is destroyed (but saved in data management platform).
Internet and confidentiality in All-Russian Population Census 2020

• Self-completion of electronic questionnaires in the Internet by respondents → entrance to the Portal → log in/passwords to personal account → identification of the person and household members

• Data transfers from "Unified Portal of State and Municipal Services (Functions)" (EPGU) to Rosstat IT System through SMEV channels. SMEV channels are governmental channels that guarantee data confidentiality (not only Census, but also financial services – tax payments, penalties, household running costs).
Open data confidentiality during All-Russian population Census 2020

Secure data transfer methods and protocols using SSL-certificates

PKCS № 12 form – data storage format

Personal users IDs

Passwords

Trusted OS loading

Trusted application’s loading
General solution of micro-data architecture (open data)
Data confidentiality structure in Population Census databases
Challenges of All-Russian population Census 2020

• Data consolidation from different sources: Internet Portal + tablet computers + Census office + paper Census forms (impersonal data needed to be compared by characteristics)
• Several completed questionnaires for one household
  Impersonal data challenge to consolidate the household and avoid data duplication
• Complicated methodology and technology
• Census become more expensive
Digital maps of All-Russian population Census 2020

- Results of All-Russian Population Census will be presented in digital maps from open data base
- Characteristics of each person (nationalities, socio-economic characteristics etc.) will be linked to geo coordinates.

Confidentiality issues:

When mapping a smaller piece of the territory the system should protect personal data of a person (up to 3 person), so its impossible to identify him by his social-economical characteristics
Conclusions

• New technologies aim to speed up data collection and receiving Census results
• New technologies aim to reduce the cost of Census
• New technologies make data processing easier

BUT

To ensure data confidentiality using combination of enumeration methods, Population Census faces with necessity of implementation of complex methodology and costly applications (IT systems)
Thank you!