



**United Nations Regional Workshop on the 2020 World
Programme on Population and Housing Censuses:
International Standards and Contemporary Technologies
Colombo, Sri Lanka
8-11 May 2018**

Session 11

Data collection with Internet

Srdjan Mrkić

United Nations Statistics Division



Data collection using the Internet

- ❑ Using Internet self-response as a collection mode
 - Advantages
 - Challenges
 - Requirements
- ❑ Implementation process
- ❑ Contact and communication strategy
- ❑ Identification and authentication of respondents
- ❑ Design of questionnaire and application features
- ❑ Support for respondents
- ❑ Managing and monitoring response
- ❑ Security



Using Internet self-response as a collection mode

□ Advantages

- Improved coverage and data quality
- Potential cost saving
- Greater convenience and lower response burden
- Improved timeliness
- Protecting privacy
- Other benefits (environmentally friendly, generating a large amount of metadata)



Using Internet self-response as a collection mode

❑ Challenges

- Providing reliable linkage between household and location
- Coordination of multi-mode collection
- Developing systems infrastructure
- Protecting data security –portal may be crashed/respondent data may be corrupted or lost
- Extended collection period (hard to capture picture as at census day)
- Mode effect and non-response bias
- High initial cost



Using Internet self-response as a collection mode

□ Requirements

- High literacy rate
- High level of access to computers and the Internet, and computer literacy
- Making reasonable assumptions about take-up rate
- Availability of address/building/dwelling list
- Legal authority
- Questionnaire design for mobile devices
- Public trust and acceptance of Internet for official business



Implementation process

□ **A first step towards the use of Internet for future census**
Feasibility study to assess the success of internet collection to obtain a measure of:

- Public acceptance
- Take-up rate
- Impact on overall response
- Distribution of returns over the collection period
- Data quality
- Impact on field management
- IT infrastructure requirements
- Impact on response burden
- Impact on under-enumeration or over count



Implementation process

□ Planning phase should include:

- Assessment of local circumstances
- Review of international experience
- Establishment of viable options
- Preferred strategy, on technology and the balance between in-house work versus outsourced contracts
- Risk analysis
- Implementation plan
- Short-term and long-term vision
- Evaluation plan
- Setting achievable targets



Contact and communication strategy

- Multi-phase contact approach
- Offering a paper questionnaire
 - at initial contact or;
 - during the reminder phase or;
 - on request
- Choosing the best approach
- Communication strategy (invitation letter/reminder letters, etc.)



Multi-phase contact strategies

JAPAN (2015)

- **PHASE 1:** Enumerators distributed internet IDs to all households
- **PHASE 2:** Enumerators distributed questionnaires to households which have not responded through the Internet. Questionnaires were submitted by handing over to enumerators or by postal mail
- **PHASE 3:** Non-response follow-up by field enumerators

KOREA (2011)

- **PHASE 1:** Internet option available for 10 days just before Census Day. Incentives included handing out gifts by drawings and giving students two-hours credits for volunteer work.
- **PHASE 2:** Field interviews during 15 days just after Census Day.



Identification and authentication of respondents

□ Benefits

- Better linkage of households to address of dwellings
- Ability to deliver different forms to different households
- Reduced risk of impersonation
- Reduced risk of duplicate responses
- Better security both actual and perceived

□ Challenges

- Confidential delivery of identifiers
- Proper linkage of identifiers to households
- Increased risk of non-response (if identifiers are misplaced)



Development of data collection application and portal

❑ Questionnaire design and application features

- Log-in screen (user experience starts here)
- Screen visibility
- Intuitive and easy to complete
- Format: matrix or sequential
- Automated skip patterns
- Response options and menus
- Validation messages
- Progress through/save/submit questionnaire
- Mobile friendly, responsive design (multiple browsers)
- Embedded support/help for respondents



Support for respondents

- Online help
- Frequently asked questions
- Embedded help text in the online questionnaire
- Questionnaire guide
- Census helpline (Call Centre)



Management and monitoring of Internet response

□ Metadata -- *During enumeration*

- Daily returns
- Cumulative returns
- Comparison with predicted returns
- Concurrent users
- Saved, abandoned or incomplete returns
- Reports on IT infrastructure stability and capacity



Management and monitoring of Internet response

□ Metadata -- *For later analysis*

- Number of edits
- Use of help pages
- Item non-response rate
- Average completion time
- Use of foreign languages
- Comparison of socio-demographic variables for households using different devices



Management and monitoring of Internet response

- ❑ **Coordination with non-response follow-up**
 - Critical in multi-mode collection
 - Central and integrated data collection operation control system,
 - Non-response follow-up procedures need to have flexibility



Security

- ❑ To maintain security of personal information:
 - Secure log-in
 - Internet application should ensure zero footprint on respondent computer
 - Timing-out after period of inactivity
 - Encryption
 - Powerful firewalls, intrusion detection
 - Strong access control procedures
 - Contingency plans for temporary service interruptions
 - Communication strategies to assure respondents



Testing

- Questionnaire -- cognitive/qualitative testing
- Planning -- experimental testing to estimate take-up rate, various metrics for planning
- IT systems -- infrastructure and system testing



User-friendly Forms - Left hand panel displays

- Allowed respondents to navigate to respective screens easily
- Allowed respondents to monitor their progress

CENSUS of population 2010

Singapore Government
Integrity • Service • Excellence
[Terms and Conditions](#) [FAQ](#) [Feedback](#) [Contact Us](#)

Statistics Act Chapter 317 Logout

open all | close all

- DN2P120
 - House Contact
 - Whereabouts
 - Household Grouping
 - HOUSEHOLD 01
 - Relationship To Head
 - Parents Of Member
 - Spouse Of Member
 - Tenancy
 - ASHLEY
 - Basic Particulars
 - Language Spoken
 - Literacy
 - Economic Status
 - Qualification Attained
 - Degree/Diploma

HOUSE DN2P120, HOUSEHOLD 01, GERALD,

Country Currently In

Where is this person currently residing in?

Australia United States of America
 China Malaysia
 United Kingdom Indonesia
 India Canada
 Others

Reason for living abroad:

Studying
 Working
 Others (e.g. staying with related persons overseas)

Regi

ards



Internet participation rate, South Korea

	2005 Census	2010 Census	2015 Census	
Year	Data collection	Key promotion	Internet Participation Rate (%)	Stage
2005	Short & Long form	none	0.9	Initial
2010	Short & Long form	Volunteer time, Gift voucher	47.9	Development
2015	Long form	Gift voucher	48.6	Mature

 The picture
be displaye



Item Non-Response by Mode

Evaluation & Challenges

