Statistics Canada's Household Survey Frames (HSF) Programme

Strategic Research Enabling a shift to increased use of Admin Data as Input to the Social statistics program

Tim Werschler - Manager

Outline

- Origin and Evolution
- HSF today
- Current paradigm
- Future of the HSF

Origin and Evolution of the HSF

- 1981 - Introduced as a post-listing coverage check in major urban areas for census.
- 2004 - Labour Force Survey started to use the Address Register for direct sampling in some areas.
- 2010 - 2013: "Common Frames" project consolidates production of a single dwelling frame for hhld surveys combining AR with telephone files and socio-economic info.
- By 2016 – National coverage. Quarterly production. Mostly maintained via admin sources, with Targeted Listing on 29%. Census Mailout for 82% of dwellings.

HSF Product Suite Today

- Dwelling Universe File (DUF)
  - List of Canadian residential addresses.
  - More than 16M records (15M) dwellings.

    83% with telephone number.
    91% with "mailable" address.

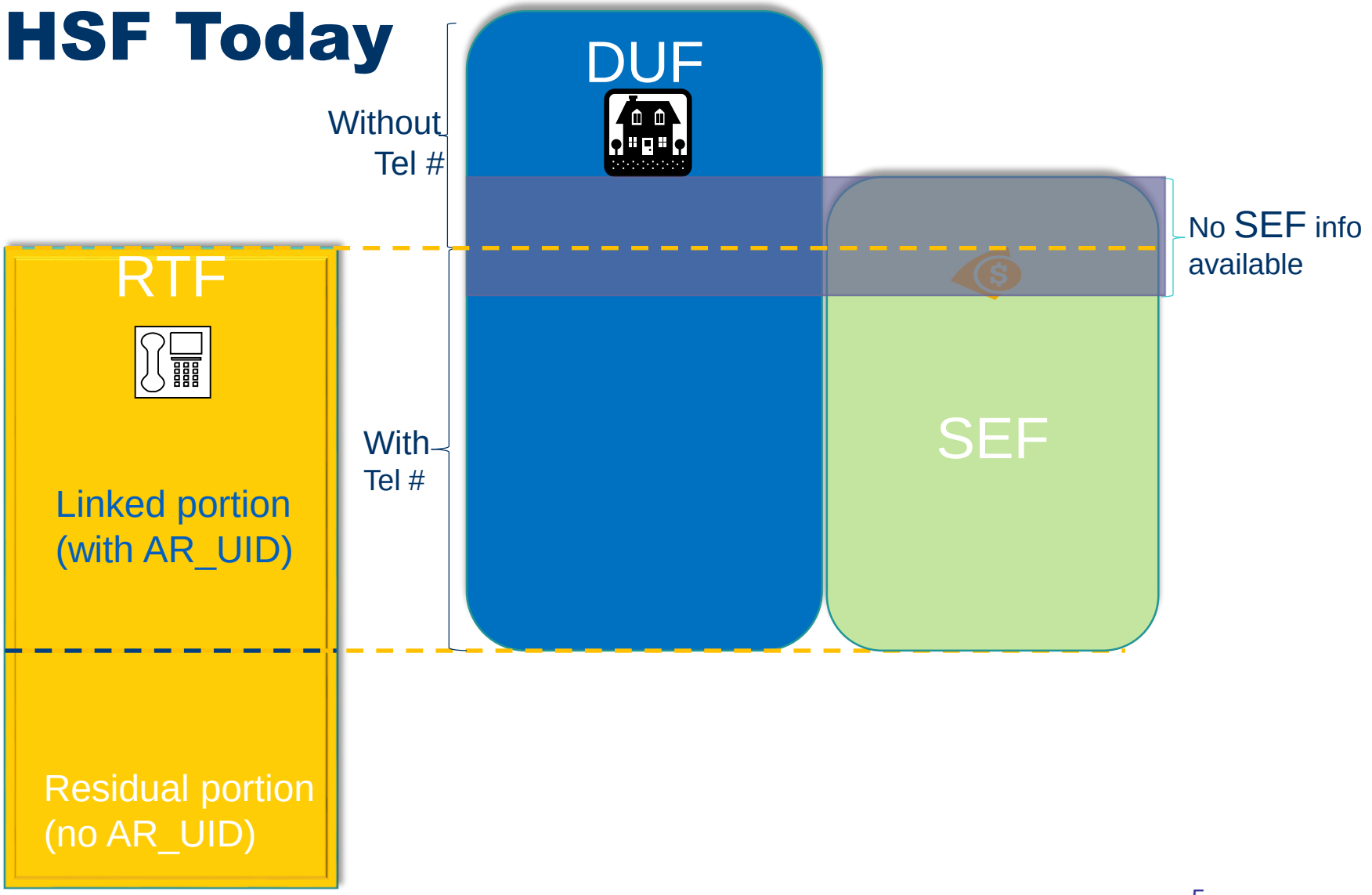
- Residential Telephone File (RTF)
  - List of residential telephone numbers valid in Canada (landline + cell).
  - As a source of contact information or as a frame.
  - More than 30M telephone numbers.

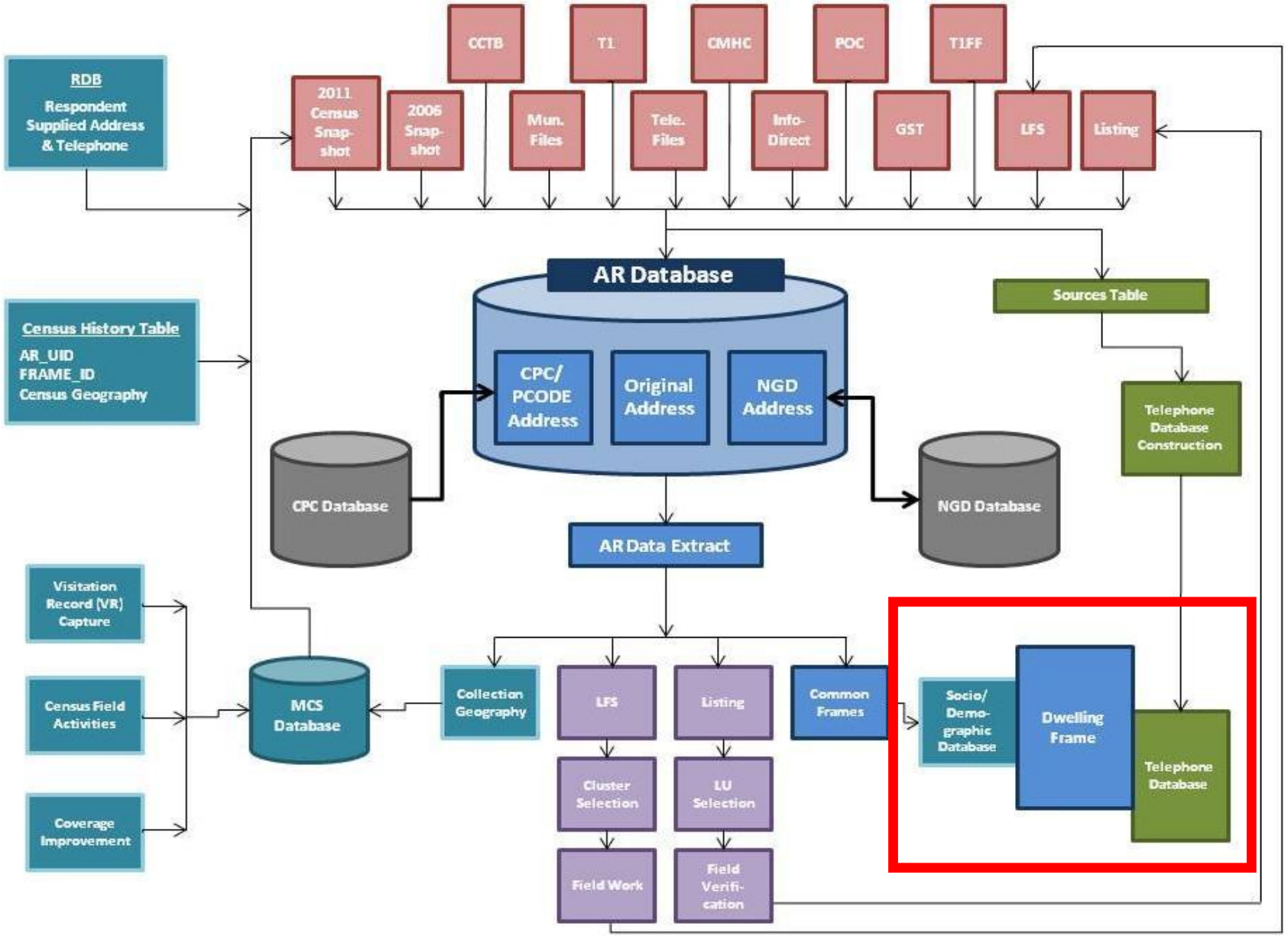
    86% are linked to a DUF address.
    55% are cell phones.

- Socioeconomic indicators File (SEF)
  - Household composition and income information associated to DUF addresses.

    88% of DUF addresses have SEF information (84% with income).

HSF Today


DUF
Without Tel #
No SEF info available
RTF
With Tel #
SEF
Linked portion (with AR_UID)
Residual portion (no AR_UID)


RDB
Respondent Supplied Address & Telephone
2011 Census Snap-shot
2006 Snap-shot
CCTB
Mun. Files
T1
Tele. Files
CMHC
Info-Direct
POC
GST
T1FF
LFS
Listing
AR Database
CPC/ PCODE Address
Original Address
NGD Address
Sources Table
Census History Table
AR_UID
FRAME_ID
Census Geography
CPC Database
NGD Database
Telephone Database Construction
AR Data Extract
Visitation Record (VR) Capture
Census Field Activities
Coverage Improvement
MCS Database
Collection Geography
LFS
Listing
Common Frames
Cluster Selection
LU Selection
Field Work
Field Verifi-cation
Socio/ Demo-graphic Database
Dwelling Frame
Telephone Database

One example of HSF use

- Canadian Community Health Survey (CCHS) – Nutrition
  - Collection mode : CAPI.
  - DUF: contact information, sampling, stratification and clustering.
    - Custom clusters built from Dissemination Areas (DA) (76% are exact DAs).
  - RTF: telephone numbers for follow-ups.
  - SEF: identification of households with members of targeted age groups:
    - 1-3, 4-8, 9-13, 14-18, 19-50, 51-70 and 71+.

Current Paradigm – Dwelling Register with contact info.

- AR evolved to serve field-based surveys (CAPI).
- Contact by mailout has been used since 2006 in Census. Most surveys use Telephone first contact, because CAPI is expensive.
- 2016 Census will provide information on mailing address (where different from civic address), email address, Canada Post epost users, landline / cell phone numbers.

Limitations Starting to Show

- 82% of dwellings can be contacted in multiple ways. Hard part is contacting the 18% of the pop scattered across 80% of the landmass.
- In rural areas, collection costs are enormous. Key factors:
  - mailing address ≠ dwelling descriptor. This precludes linkage of other useful admin data like telephone and tax-files.
  - Reticence to use telephone # at last census (even though 75% good).
  - Variation in how to use the registers in the field, and procedures.

- Collection efforts are increasing. Few surveys use contact by mail. Telephone contact getting more difficult.
- Quality of the admin data sources increases.
- These factors are pushing statistical agencies toward making more use of admin data.

Mid-Term – options for Rural-Remote Areas

- Primary objectives are to reduce collection costs.
  - Increase the range of contact info.
  - Make field-work easier when necessary.
- Extend capacity to mailout to at least 90% by 2021 by linking civic and mailing addresses in areas using PO Box, RR.
  - This will also enable linking phone numbers, tax-filer info.
- Integrate Area and List frame techniques.
  - Better control of map products, dwelling descriptions, and rules for up-dating the register.
- Remote Areas –extend dwelling descriptors where warranted with elements such as geographic coordinate, digital photo.

Mid-Term – Ideas on how to start the transition to Use of Admin Data

- Collect multi-mode contact info during collection.
- Streamline Non-response follow-up using admin input to identify unoccupied / invalid dwellings.
- Use admin data profiles to help with imputation, weighting (calibration).

What about Collective Dwellings?

- 50K dwellings of this type. 1.5M people. Cost and conceptual issues.
- Most collectives function as businesses. Business owners maintain and know their information.
- Tighter integration and coordination will enable improvements in quality, coverage, and cost.

Longer-Term (By 2026)

- Conceptual shift to registers as integrators in a Data-Centric approach. Field Collection Where Necessary.

- Geographic and dwelling registers still central, but contact info now takes on a more pronounced role in data linkage.
  - Canada Post data to help avoid allocation to businesses.
  - Hydro billing files may find sporadic dwelling growth in rural areas, and provide GPS points and mailing addresses.
  - Quality indicators for register units as input to design

- Statistical programs could shift to the use of admin data inputs for basic demographics, supplemented with survey sampling to obtain essential data not available by other means.

- Challenge is acquiring data describing the unit of observation. Then linking to registers. Then to meaningful geographic context.

Longer Term (>= 2026) Constellation of linked Registers ?

- Wallgren & Wallgren (2014) hard to argue against…
- Ability to link persons to home, work, income, government services, spending habits, etc. will go a long way to addressing small-area data needs.
- Gaps can potentially be off-set by continued calibration via field survey and other data sources.
- Important part is infrastructure planning…the sequence and timing of projects to move in this direction.
- Statistics Canada has set itself to this task. A framework for statistics, & coordinated use of admin data across gov. will not be quick, or easy, but will be beneficial.

- For more information, please contact:

- Pour plus d'information, veuillez contacter :

tim.werschler@canada.ca - Manager
edward.chen@canada.ca - Methodology
kim.charland@canada.ca – Dwellings / Register Operations
crystal.sewards@canada.ca – Contact info / Admin data / Postal Geog.
melanie.labreche@canada.ca – IT