CAPI Overview

Maria Aysa-Lastra International Programs Center Population Division

September 2022





A short contemporary history of interview modes

- PAPI: Paper-Assisted Personal Interviewing
- PMASI: Postal Mail-Assisted Self Interviewing
- CATI: Computer-Assisted Telephone Interviewing
- EMASI: Electronic Mail-Assisted Self Interviewing
- CAPI: Computer-Assisted Personal Interviewing
- CAWI: Computer-Assisted Web Interface

Coming soon:

• ARBPIR: Administrative-Records-Based Personal Information Retrieval





How did CAPI develop in the US?

- The story of CAPI is about making the process of data capture, aggregation, and tabulation more reliable and efficient.
- For the first 8 10 Censuses in the US, the process involved three things:
 - Enumerators visiting every community in the country
 - Local officials tabulating these enumerators
 - National officials tabulating the local reports





How did CAPI develop in the US?

- Between 1890 and 1970, machines and computers began to facilitate the work of aggregation and tabulation:
 - The Seaton Device and The Hollerith Machine
 - The UNIVAC I and FOSDIC
- Since 1970, computers have moved into the realm of enumeration culminating in handheld devices that allow data capture, i.e. CAPI.





Early Censuses

The first U.S. Census in 1790 was considerably different from the Census of 1890, just 100 years later.

1790 Census

- Conducted by the marshals of the judicial districts of the U.S. and any assistants they wanted to employ.
- Asked about only 6 items.
- Completed in approximately 1.5 years.

1890 Census

- Conducted by 46,804 enumerators.
- Captured four general schedules of questions (population, agriculture, manufacturing, and mortality) and eight supplemental schedules.
- Official count completed in approximately 6 months. Final bulletin report in 1895. Final published volume in 1897.





Early Censuses

1790 Census

- The marshals aggregated their districts and sent the results to the President.
- No standard sheets for enumeration or consistent methods of aggregation.
- The final report comprised just 56 pages reproduced directly from the marshals' reports with a summary at the beginning.

1890 Census

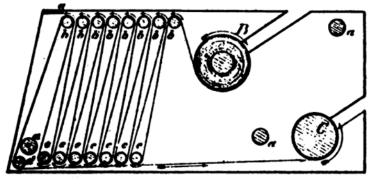
- Used the Hollerith machine to tabulate.
- The final results comprise 21,410 pages with a 3-part compendium, an abstract, and a statistical atlas.
- Illustrated with maps and diagrams.
- The Hollerith machine allowed a complete tabulation 3+ years faster than after the 1880 Census.





Tabulating Machines Make the Case for Automation

- The first tabulating machine used in the US Census, the Seaton Device, was a machine in the sense that it involved machine-like components.
 - The Seaton Device assisted hand-tallying and probably was not much more efficient than handtallying.
- The Hollerith Machine was much more like what we think of as automation. It used an electric circuit to move dials which kept track of the count.



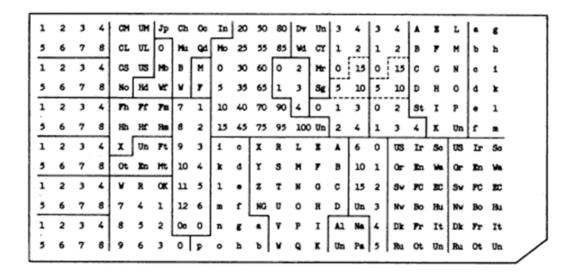
Hollerith Machine





Tabulating Machines Make the Case for Automation

• An example of a punch card coding system used with the Hollerith Machine. This would represent a single individual.







Computers Arrive in a Big Way

- The UNIVAC and FOSDIC systems started the transition to computer data capture.
- UNIVAC I (console at right) continued use of punch cards to tabulate census results. It was used to help tabulate the 1950 and 1960 censuses and to calculate the weights used in several surveys.
- The UNIVAC demonstrated that punch card creation was now the biggest limiting factor in the census process.







Computers Arrive in a Big Way

- The FOSDIC was the first step toward automating data capture. In 1970, the forms used by the census could be scanned directly by the FOSDIC.
 - The FOSDIC version of the 1960 Population and Housing Census form is at the right.

60 PH-3 (5) U.S. DEPARTMENT OF C			us 1960		ULATION AND HOUSING		article in the same is an a set of the set of the same is a set of the s				
		disambian of a Assessmented	hay .	IH. Access HS. Elders H or coulding	Condition H7. Occupancy	HR. Rouns	10. Water	H10. Tuilet	HI1. Buth	H12. Owned	H13. Visioney status
11114 NO 0557 NO 5557 NO 5550 NO 55000		and a previous along, mark below and conit all "A" little on this share factoration C	aroo ar 2 o		nend. O Becglind 0 Handle Romani. Fingl. O Second C High Burthey also O Handle C Second C	1000000 100000 100000 10000	Remaining Walter in abushung Bar wall cold. Cold uniy Rem. webe satulat Remaining weber			HI	Ter set for set of a set only for accelent on for accelent on the mass
FOR ALL UNITE AS	ND 09's (Omit in	ton Z ED'to	FOR OCCUPIED	UNITS	POR OCCUPII	ED UNITS	J/ "OWNED OR FOR SALE ONL	BEING BOL	CHT" or	VACANT-	If transcribing
HI7. Is this based the unit of the second		It stand must be A Manlage b Casking? A Manlage Min stall? Carl or state		HOS. Is flowr a telephane an which people in this pair? Mercer C		N29. (H14) Description I will, with balance.			Prom PH-2, rugo from siema H14, H15, H16 into H28, H40, H43,		
(And Hitte)	Less pain. 1998, did salas of coup.		Initial, test, or UP per 0 0 0 Bedikity 0 0 0 0 Net sil, kensen, sk 0 0 0 0 Other sil, kensen, sk 0 0 0 0 Net sil, kensen, sk 0 0 0 0 Net sil, kensen, sk 0 0 0 0		HOSe. If "Yes" in HOS C		Hell, (1475) 27 "7 unit, no husiners" About how much dryne filiak fils syngenty weekly and face an kaday's mobiler (27 eccent): What is the price solution for fils property?) [13,004-13,005] [14,004-13,005] [15,004-13			15,000-17,000 17,330-11,950 19,960-24,950 25,960-24,950 21,960 or see	
FOR ALL UNITS, OCCUPIED AND VACANT is all following questions, use "bouse", "spartoset", "fat", or "rooms", as appropriate, instead of "unit"			1933 bi Rama schaftes te Bis sull? Redec drys: sanitation: Bis sull? Redec drys: sanitation: Bis sull? Redec drys: sanitation: Bis subles: Bis subles: B		HDA. How many advanables, on avoid or regularly used by people who first in the sets O		MAN SOCOS 100				: :::::
1. How many balances In factors 0 1. Min unit 0 0 2			HD6, is many a clastice dryse in this unit? Budik		and 1 1 0 Include company core 1 0 Rept at home 1 ar man. 0		Bein C C C C 2 C 2 C 2 C 2 C 2 C 2 C 2 C 2				
120. About when wen dis lance ariginally built?	f ar suar Albert alber was file as algebrilly built? 1955 as		H25. In there intertained in this until?	1 m	h the turks and the			-	Whe	0000	() and if arrayint-
154) in 153) in 1625 au		1949 0 1949 0 1929 0	HD6. Are there any radius in this antif	1 w		TUNTS		0		0000	
01. New is this cuif bandgeff	Farm air furness.	k solls	HD7. Is there any of conditioning in this until?	kaan antis 1	No off See sugget	1 0 1 0 1 0	NHS. In addition to doors made pay for OIL and,			0000	et and if scrapiod-
1 34 A	Other means-	a ha 0	H28. Is there a hanta food fine separate from the self-garater in this unit?	···· ··· · ··· · ·			wood ar	0		0000	00000





Computers Bring the Census to You

- As computers have become better, direct entry into the computer has become feasible.
- CAPI is the culmination of a long history of automation in censuses and data collection. Along with other computer-assisted methods, it has been used by the U.S. Census Bureau in its surveys since the 1990's, particularly for economic and business surveys.
- Other computer-assisted methods include diskettes, e-mail, and web forms.





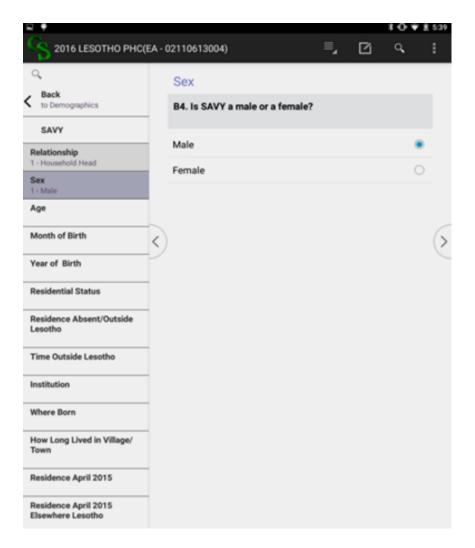
Computers Bring the Census to You

 As smartphones have matured, programs like CSPro have expanded where CAPI can be used. CSPro (screenshot at right) allows enumerators to easily capture data in the field.









Additional CAPI Characteristics

Mobile data capture allows users to take advantage of additional features that can be programmed on mobile devices. These include:

- Integrated maps and Global Positioning System (GPS).
- Computerized case management.
- Automated skip patterns.
- Concurrent processing for data collection, data capture, editing, and consistency and non-response checks.
- Automated coding.
- Ability to preload questionnaire responses and customize questions.





Advantages of using mobile data capture

- Improves data quality.
- Improves field management and real-time monitoring of enumeration activities.
- You can reduce time and costs by eliminating a separate process for data capture, automated coding, and reducing the volume of post-enumeration data verification.





Advantages of using mobile data capture

- Decreases the time between data collection and publication of results.
- It can reduce the costs of data capture, questionnaire printing, storage, and transportation. However, the increase in equipment costs can outweigh the savings.
- You can use GPS features and digital mapping capabilities.





Disadvantages of using mobile data capture

- High equipment costs with limited long-term use.
- More time is needed during the preparation stage.
- You need more qualified programmers.
- It requires enumerators with technological capabilities, more training and field support.
- The effect of questionnaire segmentation can counteract improvements in data quality.
- Technological issues could interfere with the enumeration.
- There are more infrastructure limitations to consider.





Respondent Reaction

- Varies by country/context
- If it is the first time a NSO intends to implement CAPI, care should be taken to anticipate and mitigate potential opposition
- Can profoundly impact the potential success of Census or survey





Respondent Reaction

If CAPI use is new to the context:

- Develop an introduction strategy for:
 - Local collaborators
 - Stakeholders
 - Census/survey respondents



• Involve local collaborators in census/survey design, if possible, to facilitate its adoption





Respondent Reaction

CAPI – Opportunity and Risk

- Technology, computerization can stimulate census/survey respondent interest and add legitimacy to the interviewers. Interviewers might also be more motivated to use technology in such setting.
- At the same time and in certain cultures, the use of technology can raise suspicion among respondents.



