



Metadata and Integration: IPUMS-International Methods and Practices

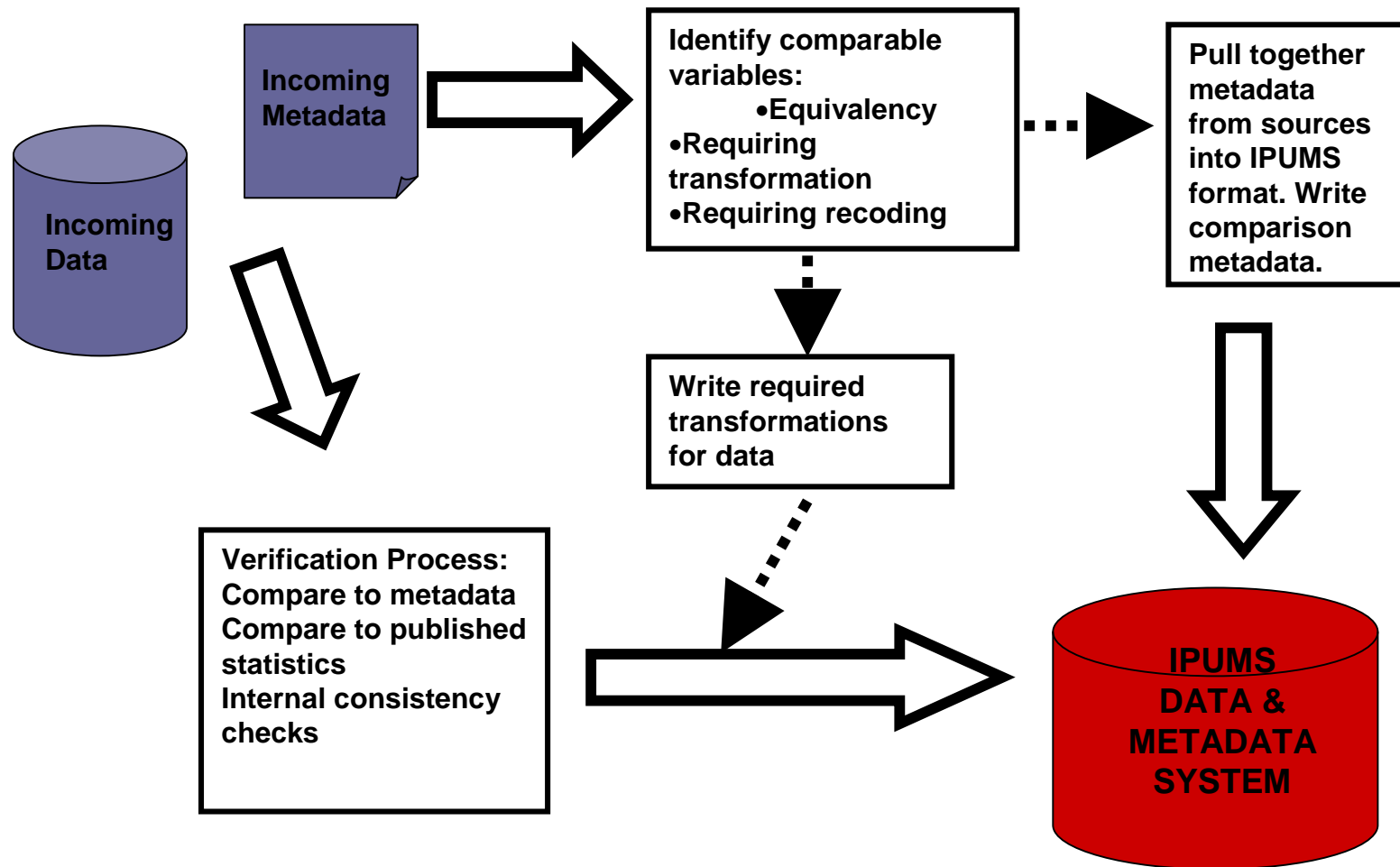
Wendy Thomas

20 June 2008

RatSWD 4. Konferenz für Sozial- und Wirtschaftsdaten

Critical Components

- Codebook for current data
 - Geographic codes
 - Occupation / Industry codes
- Enumeration forms
- Enumerator instructions
- Data processing
- Additional materials



Harmonization

- **Equivalent variable**
 - Add country specific definitions to IPUMS metadata
 - Transfer data as is
- **Equivalent categories – different codes**
 - Record code transformation in translation tables and apply when translating data
- **Different categories**
 - Map to current IPUMS categories
 - Add new categories to IPUMS as needed
 - Recode to reflect appropriate hierarchies

200	MARRIED/IN UNION		
210	Married (not specified)		
211	Civil	7=only civil marriage	12=only civil marriage
212	Religious	8=only religious marriage	13=only religious marriage
213	Civil and Religious	6=civil marriage	11=civil marriage
220	Consensual union	9=living maritally	14=consensual or other

Original Metadata

- **Cataloging**
 - creating records for over 22,000 items
- **Scanning**
 - full or first page scan
- **Preservation**
 - transfer to acid free boxes/files
- **Dissemination**
 - selected scanned materials going to U of MN Digital Conservancy
 - create on-line catalog with links to scanned items

Case Study: Germany

- Coverage
- Translation
- Support / Question response
- Related materials

BEVÖLKERUNG
UND ERWERBSTÄTIGKEIT

FACHSERIE

1

Volkszählung vom 27. Mai 1970

Heft 25

Methodische und praktische Vorbereitung sowie
Durchführung der Volkszählung 1970



HERAUSGEBER: STATISTISCHES BUNDESAMT WIESBADEN
VERLAG: W. KOHLHAMMER GMBH STUTTGART UND MAINZ

Anonymisierung der Volkszählung von 1971 der Deutschen Demokratischen Republik zur Verwendung als Public-Use-File

I. Basismaterial

Das Basismaterial der Volkszählung 1971 umfasst zwei Datenfiles: die Personendaten und die Wohnungsdaten. Die Personendaten enthalten demografische Informationen sowie Angaben zu den Quellen des Lebensunterhalts, der Bildung, der Erwerbstätigkeit und der Haushaltszusammensetzung. Die Wohnungsdaten geben Auskunft über den baulichen Zustand der Gebäude, die Wohnungsbelegung und ihre Ausstattung (Heizung, Warmwasserversorgung, sanitäre Anlagen).

Das Basismaterial umfasst 16,4 Mio. Personen, 6,2 Mio. Haushalte und 6 Mio. Wohnungen.

I. Base material

The 1971 census data are comprised of two sets of data: person data and dwelling data. The person data contain demographic information as well as information about the source of livelihood, education and training, employment and the composition of households. The dwelling data inform about the condition of buildings, occupation and facilities (heating, water supply, toilettes). The base data material comprises 16.4 million persons, 6.2 million households and 6 million dwellings.

GDR 1971 Census
 DE 2 MD/27
 Census
 Integrated description of record layout, labels and codes - Public-Use-File person, building and dwelling data 1971
 PH = Private
 Household
 GE = Institution/Collective dwelling [Gemeinschaftseinrichtung]

field identification	field start	field end	field length	variable	codes, values, comments	Recoding
vp = person data						
vp1	1	3	3	type of record	P71	
Regional- und Ordnungangaben						
vp2	4	5	2	Province [Land]	11 = Berlin 12 = Brandenburg 13 = Mecklenburg-Vorpommern 14 = Sachsen 15 = Sachsen-Anhalt 16 = Thüringen	

Support

- Metadata in multiple formats to facilitate processing within IPUMSI
- Links to online sources of metadata in other formats
- Quick response to specific questions regarding translation and/or interpretation

Vielen Dank

Wendy Thomas
wlt@pop.umn.edu