

The IPUMS Bibliography: Harnessing Organizational Knowledge

September 2025

DOI: <https://doi.org/10.18128/IPUMS2025-03>

[†]Address correspondence to Diana Magnuson: magn0031@umn.edu.

[illegible]

The IPUMS Bibliography: harnessing organizational knowledge

Diana L. Magnuson¹

Abstract

The IPUMS Bibliography tool plays a crucial role harnessing over three decades of IPUMS organizational knowledge to support and strengthen institutional goals, demonstrating the impact of IPUMS datasets and the Institute for Social Research and Data Innovation (ISRDI, the home of IPUMS), and strengthening broader social science infrastructure. The IPUMS Bibliography captures, organizes, and makes discoverable: research analyses on tens of thousands microdata and aggregate variables over time; processes for developing and maintaining well-defined harmonized classifications; and technological innovation and advancement supporting IPUMS systems. The challenge for the IPUMS Bibliography team is threefold: refine bibliographic data capture methods; develop an efficient, scalable, and sustainable workflow for managing the expanding bibliographic database; and improve the user interface for internal and external users to discover and link to publications using IPUMS data. This article will trace the IPUMS Bibliography development from 1994 to 2025, demonstrate the benefit of harnessing organizational knowledge within an organization, and suggest the potential for leveraging organizational knowledge to support robust social science research.

Keywords

bibliography, organizational knowledge, data capture, IPUMS

Introduction

IPUMS is one of four centers at the Institute for Social Research and Data Innovation (ISRDI) at the University of Minnesota, and its datasets comprise the world's largest accessible database of census microdata, with over two billion records. The IPUMS suite contains nine harmonized census and survey microdata and aggregate geographic data collections. The signature activity of IPUMS is harmonizing variable codes and documentation to be fully consistent across data sets. Since the launch of its microdata harmonization work in 1991, IPUMS has expanded from historical U.S. census data to cover international censuses, health data, labor force data, and educational surveys. In addition, IPUMS integrates and disseminates area-level census data and electronic boundary files describing U.S. and international geography.² IPUMS data thus support research analysis on billions of microdata and aggregate cases and data points over time.

For over three decades, the IPUMS Bibliography has played a crucial role in supporting and strengthening institutional goals, demonstrating the impact of IPUMS datasets, and contributing to

broader social science research efforts. The expansion and growth of IPUMS since 1993 and the concurrent reliance by the research community on these harmonized data collections signals the significance of maintaining an accurate bibliographic record of IPUMS research analysis on billions of cases over time, on internal process for developing and maintaining well-defined harmonized variable classifications, and on technological innovation and advancement supporting IPUMS systems.

This article traces the development and user interface transformation of the IPUMS Bibliography from 1994-2025, describes the current metadata processing system that supplies the information underlying the bibliographic system, demonstrates the benefit of harnessing organizational knowledge across an organization, and suggests the potential curating IPUMS bibliographic metadata has supporting robust social science research.

IPUMS Bibliography user interface development

IPUMS Bibliography curation began in 1994, three years after the effort to convert the extant US public use microdata samples (1880, 1900, 1910, 1940, 1950, 1960, 1970, 1980 and 1990) into a single consistent data and documentation format began (Ruggles 1991). The first IPUMS harmonized U.S. microdata set was downloaded using an anonymous FTP site on November 19, 1993 (Magnuson and Ruggles 2022). Two articles citing IPUMS data were published in peer reviewed journals in 1994, and these citations were the beginning of the IPUMS Bibliography (Ruggles 1994a and Ruggles 1994b). The Bibliography began as a hand-curated list maintained by a single individual (S. Ruggles, personal communication, November 6, 2024).

Access to the Internet and use of the World Wide Web did not come into common use until the mid- to late 1990s (Kumar and Perakis 2022). The IPUMS Bibliography user interface development from 1995-2025 follows the contours of web design capacity and limitations by decade.³ The first IPUMS website appeared in March 1995, one of the first 15,000 websites created (Magnuson and Ruggles 2022). The earliest retrievable screen capture of the online IPUMS Bibliography is from 1998 (Wayback Machine 1998). The 1998 landing page is titled “IPUMS Research” and is qualified as “very

preliminary” and directs researchers using “the IPUMS” to submit their published and unpublished research citations (Figure 1). The 1998 version of the IPUMS Bibliography is a static list with limited functionality for the user. The list is organized by publication type (with a hyperlink to jump down to that type in the list), author last name, and publication date. Publication types using “the IPUMS” included: published research, completed dissertations, publications about the IPUMS, working papers, and research in progress. There is no ability to access the text of the articles from the 1998 IPUMS Research user interface.

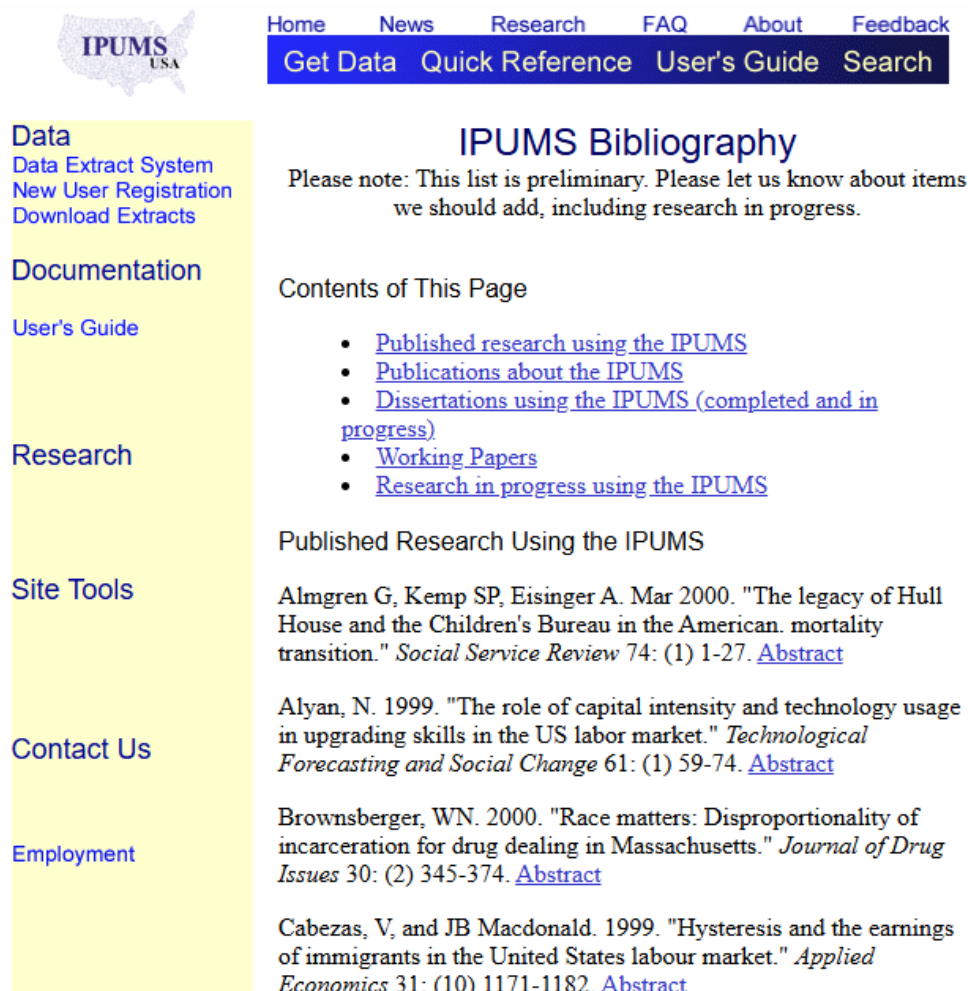
Figure 1. IPUMS Research user interface 1998 (Wayback Machine 1998)



“IPUMS Research” was renamed “IPUMS Bibliography” by 2000. The user interface layout was more colorful, including a left sidebar with hyperlinks to access other pages on the IPUMS website and, significantly, offering a hyperlink to the abstract of each identified citation in the bibliography (Figure 2). While certainly useful for users, the addition of an abstract for each citation undoubtedly increased person labor around metadata creation. Users of the 2000 version of the IPUMS Bibliography were still limited, however, to scrolling down the citation list, organized by publication type. While the inclusion of citation abstracts was an important functional upgrade, the user still had

to use an alternative mechanism to locate the text of the article. During this period of data dissemination, IPUMS was a product, “the IPUMS,” not a brand, “IPUMS.”


Figure 2. IPUMS Bibliography user interface 2000 (Wayback Machine 2000)



The 2003 IPUMS Bibliography added important functional refinements (Figure 3). First, links took users to a form for submitting a new citation for entry consideration or viewing recently submitted citations not yet integrated into the bibliography. This type of user crowdsourcing continues to be employed by the IPUMS Bibliography. A second functional refinement encouraged users to provide feedback through a link “If you would like to see added functionality to this system.” New features included a hyperlinked list with “award-winning IPUMS-related research,” a keyword search with “Match ANY” and “Match ALL” radio buttons, and the option to view by publication type and/or “topic/subject” classifications. Publication types now included: books, journal articles, edited

volume chapters, dissertations, working papers, conference papers, newspaper articles, and “other” (e.g., courses, CD-ROMS). Topic/subject classifications included key demographic research areas: migration and immigration, labor force and occupational structure, family and marriage, fertility and mortality, race and ethnicity, housing and segregation, education, poverty and welfare, aging and retirement, methodology and data collection, gender, crime and deviance, and “other” category. These external facing user interface developments increased the internal complexity of managing the bibliography. As the IPUMS Bibliography user interface evolved, so too did the bibliography workflow.

Figure 3. IPUMS Bibliography user interface 2003 (Wayback Machine 2003)



[Home](#)
[News](#)
[Research](#)
[FAQ](#)
[About](#)
[Feedback](#)

[Get Data](#)
[Quick Reference](#)
[User's Guide](#)
[Search](#)

Data

- [Data Options](#)
- [Create a New Extract](#)
- [Revise an Old Extract](#)
- [New User Registration](#)
- [Download Extracts](#)

Documentation

- [What is the IPUMS?](#)
- [User's Guide](#)
- [IPUMS Variables](#)
- [Download Documentation](#)
- [Revision History](#)

Research

- [Citation and Use](#)
- [Bibliography](#)
- [A Note to Genealogists](#)

Site Tools

- [Site Map](#)
- [Search This Site](#)
- [FAQ](#)
- [Links to Related Sites](#)

Contact Us

- [Feedback](#)
- [IPUMS Staff](#)
- [Mailing Address](#)
- [Employment](#)

IPUMS Bibliography

Welcome to the IPUMS Bibliography. If you would like to [add a new citation](#), [click here](#). If you would like to see added functionality to this system, please give us [feedback](#).

- [View all bibliography entries](#) (1773 citations total)
- [View award-winning IPUMS-related research](#)

• Search for words or phrases:
(type any combination of words or phrases; phrases must be surrounded by "quotation marks"):

Keyword(s):

☒ Match ANY of the Term(s)
 ☐ Match ALL of the Term(s)

• View by publication type:
(choose as many publication formats as you want from the list below)

- ☐ Books
- ☐ Journal Articles
- ☐ Edited Volume Chapters
- ☐ Dissertations
- ☐ Working Papers
- ☐ Conference Papers
- ☐ Newspaper Articles
- ☐ Other (Courses, CD-ROMs, etc.)

• View by topics/subject classifications:
(choose as many topic headings as you want from the list below)

- ☐ Migration and Immigration
- ☐ Labor Force and Occupational Structure
- ☐ Family and Marriage
- ☐ Fertility and Mortality
- ☐ Race and Ethnicity

More curation effort was required, which in turn necessitated expanding the training and supervision of undergraduate student workers. Developments to the IPUMS Bibliography workflow are addressed below.

Between 2005 and 2007 the IPUMS Bibliography was rebranded “The Minnesota Population Center Bibliography,” a shift reflecting the organization of the Minnesota Population Center as “a university-wide interdisciplinary cooperative for demographic research” at the University of Minnesota (Magnuson and Thomas 2023). The IPUMS USA and IPUMS International harmonized microdata products were organizationally subsumed under the Minnesota Population Center (MPC). This rebranding also marked a visual refresh for the bibliography interface and some functional simplification. Radio buttons “Match ANY” and “Match ALL” were dropped in the keyword search field and “publication year” with date range was added (Figure 4). Requests for adding citations or additional functionality to the bibliography were removed from the user interface but captured through email submissions.

Figure 4. Minnesota Population Center Bibliography user interface 2007 (Wayback Machine 2007)

Welcome to the Minnesota Population Center Bibliography!

This bibliography contains references to publications, presentations, and works in progress that use the Integrated Public Use Microdata Series projects (IPUMS) or any other data resource created at the Minnesota Population Center. We appreciate your help in keeping the bibliography accurate and up to date.

Keywords:

Title:

Author's Name:

Type: All Citation Types ▾

Publication Year: to

Show All Search

Page: 1 2 3 ... 113 > 2260 total citations in bibliography

Sort results by: Click the column heading to sort.

	Type	Pub. Year ↑
What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital Among Immigrants •Bleakley, Hoyt	journal	2008
10*(10*6) Worlds and Beyond: Efficient Representation and Processing of Incomplete Information •Antova, Lyublena •Koch, Christoph •Olteanu, Dan	conference paper	2007
493 Quantitative Historical Analysis •Shammas, Carole	misc	2007
A Century of Women in Science and Engineering •Norling-Ruggles, Abby	misc	2007
A Holistic Approach to the Geography of Social Distress: A Typology of U.S. Counties •Rogers, Richard Lee	conference paper	2007
A Measure of Segregation Based on Social Interactions •Echenique, Federico •Fryer, Roland	journal	2007
A Practical Approach to Using Multiple-Race Response Data: A Bridging Method for Public-Use Microdata •Lieber, Carolyn A. •Halpern-Manners, Andrew	conference paper	2007
An Accounting Exercise for the Shift in Life-Cycle Employment Profiles of Married Women Born Between 1940 and 1960 •Buttet, Sebastien •Schoonbroodt, Alice	working paper	2007
An Empirical Analysis of Exploitation in the Labor Market Using a Weberian Approach: Manufacturing Industries in the U.S., 1971-1996 •Sakamoto, Arthur •Kim, Changhwan	conference paper	2007
An International Perspective on Scandinavia's Historical Censuses •Thorvaldsen, Gunnar	journal	2007
Are Immigrant Youth Faring Better in U.S. Schools? •Fry, Richard	journal	2007

By 2011 users could filter their queries to search the MPC Bibliography for research using seven IPUMS harmonized data “projects” (products) (Figure 5). This reflects not only the growth of

the bibliography but the diversification of the IPUMS collections. In terms of the interface, the user is presented with a search tool with filtering options rather than a lengthy list of citations. The number of IPUMS harmonized data products in 2011 totaled seven: IPUMS-USA, IPUMS-International, National Historical Geographic Information Systems (NHGIS), Current Population Survey (CPS), American Time Use Survey (ATUS), Integrated Health Interview Series (IHIS), and North Atlantic Population Project (NAPP). The 2014 version of the bibliography interface was more colorful and the search tool more streamlined with the same functionality as the previous iteration (Figure 6). The diversification and scale of IPUMS harmonized data products was beginning to challenge the capacity of the IPUMS Bibliography Team and the infrastructure they relied on (Table 1).

Figure 5. Minnesota Population Center Bibliography user interface 2011 (Wayback Machine 2011)

MPC Bibliography

[new search](#) | [add your publication](#)

Advanced Search

Enter search terms in at least one of the fields below

Keyword:

Title:

Author:

Limit results by (optional)

Journal:

Type:

Publication Year: to

Restrict to topics:

<input type="checkbox"/> Aging and Retirement	<input type="checkbox"/> Crime and Deviance	<input type="checkbox"/> Education
<input type="checkbox"/> Family and Marriage	<input type="checkbox"/> Fertility and Mortality	<input type="checkbox"/> Gender
<input type="checkbox"/> Health	<input type="checkbox"/> Housing and Segregation	<input type="checkbox"/> Labor Force and Occupational Structure
<input type="checkbox"/> Methodology and Data Collection	<input type="checkbox"/> Migration and Immigration	<input type="checkbox"/> Poverty and Welfare
<input type="checkbox"/> Race and Ethnicity	<input type="checkbox"/> Other	

Restrict to projects:

<input type="checkbox"/> ATUS	<input type="checkbox"/> IHIS	<input type="checkbox"/> IPUMS-CPS	<input type="checkbox"/> IPUMS-International
<input type="checkbox"/> IPUMS-USA	<input type="checkbox"/> NAPP	<input type="checkbox"/> NHGIS	

Figure 6. Minnesota Population Center Bibliography user interface 2014 (Wayback Machine 2014)

The screenshot shows the Minnesota Population Center (MPC) Bibliography web interface. At the top, there is a blue header with the MPC logo and the tagline "bringing you the world's population data". To the right of the tagline are several icons representing different data sources: IPUMS, NHGIS, NAPP, IHIS, ATUS, and TerraPop. Below the header is a yellow banner with the word "Bibliography" in bold, followed by the text "Publications, working papers, and other research using data resources from the [Minnesota Population Center \(MPC\)](#)". Below the banner is a navigation bar with three links: "New Search", "Add A Publication", and "Admin". The main content area is a light gray box containing a "Search" section. This section has a light blue background and contains several search filters: "Keywords" (text input), "Title" (text input), "Author" (text input), "Journal" (text input), "Citation types" (dropdown menu with "All Citation Types" selected), "Year published" (text input), "Topic" (dropdown menu with "All Topics" selected), and "Project" (dropdown menu with "All Projects" selected). At the bottom of the search section are two buttons: "Search" and "Clear".

The user interface in 2018 returned to “IPUMS Bibliography,” updated its branding to conform to a total refresh of the IPUMS websites’ visual design, and changed the search filter “projects” to “data collections” (Figure 7). In all other respects, the functionality for the user remained the same as the 2014 version.

Figure 7. IPUMS Bibliography user interface 2018 (Wayback Machine 2018)

IPUMS

ABOUT | SUPPORT | PROJECTS | BIBLIOGRAPHY | BLOG

BIBLIOGRAPHY
Publications, working papers, and other research using data resources from IPUMS.

NEW SEARCH | ADD A PUBLICATION | ADMIN

Search

Keywords

Title

Author

Journal

Citation types

Year published

Topic

Data Collections

Search Clear

The IPUMS IT group undertook an overhaul of the IPUMS Bibliography user interface in 2022. The primary goals of the overhaul were useability from the perspective of internal and external users and long-term maintainability from the perspective of the IPUMS IT Product Team responsible for supporting the system. IPUMS Bibliography Team requests for feature updates were delayed by an IT reorganization and COVID disruptions in 2020. Updates to the codebase underpinning the IPUMS Bibliography raised the software team’s confidence that they could maintain the tool and provide future enhancements as needed. IPUMS IT also added features to enhance the user experience and provide back-end support for the IPUMS Bibliography Team during the metadata creation stage (Fabrizio 2022). The results of the user enhancements are visible in a sample search results page (Figure 8). The updated user interface displays project tags, links to abstract, full citation, and Google Scholar (for project tag-controlled vocabulary, see Table 3).

Figure 8. IPUMS Bibliography user interface results page with project tags, 2025

BIBLIOGRAPHY
Publications, working papers, and other research using data resources from IPUMS.

[NEW SEARCH](#) | [ADD A PUBLICATION](#) |

Showing Results for:
Minimum Year Published: 1994
Maximum Year Published: 2025
[Modify Search](#)

Results per page: Sort by: [Apply](#)

1 2 3 4 5 ... [Next >](#) [Last >](#)

Total Results: 29111
[Show all abstracts](#)

Pedroza, Juan Manuel; Losada, Elena
2025. *It's About Time: Parent's Direct Care for Children in Hispanic Noncitizen Households Across State Immigration Policy Contexts.*
[Abstract](#) | [Full Citation](#) | [Google](#)
ATUS

Taalbi, Josef
2025. *On the gendering of the early American electric car.*
[Abstract](#) | [Full Citation](#) | [Google](#)
NHGIS

Jones, Hannah; Hale, Danielle
2025. *Housing Supply Gap Reaches Nearly 4 Million in 2024.*
[Abstract](#) | [Full Citation](#) | [Google](#)
CPS

Luna, Isidro Manuel Victor
2025. *Migration perspectives during Trump's second term in office.*
[Abstract](#) | [Full Citation](#) | [Google](#)
CPS

Modji, Komi K.S.; McCoy, Katherine E.; Creswell, Paul D.; Tomasallo, Carrie D.; Bedno, Sheryl
2025. *Snow or ice-related injury claims in cold months during 2016-2021.*
[Abstract](#) | [Full Citation](#) | [Google](#)
USA

IPUMS Bibliography workflow

IPUMS harmonized data product releases have driven the growth of the IPUMS Bibliography (Table 1).⁴ As the suite of IPUMS products and the number of users has expanded, so too has the challenge of capturing and managing bibliographic metadata. The previous section showed the evolution of the user interface in response to changing user and organizational needs. This section

describes the current metadata processing system that supplies the information underlying the bibliographic system.

The IPUMS Bibliography team manages three ongoing challenges to its curation work: improving bibliographic data capture methods; developing efficient, scalable, and sustainable workflows for managing the expanding bibliographic database; and refining a usable user interface for internal and external users to discover and link to publications using IPUMS data. The context of these challenges is the limited resources allocated to this work, necessitating the use of student labor, which in turn required a workflow suited to that workforce.

Table 1. IPUMS data product releases and bibliography citation count

IPUMS data product	IPUMS data product released	IPUMS Bibliography citation cumulative count
USA	1993	[19]
International	2002	[1,098]
NHGIS	2004	[1,954]
CPS	2006	[3,093]
ATUS	2008	[4,511]
IHIS, NAPP	2011	[7,745]
USA Full Count*	2013	[10,512]
NHIS [formerly IHIS]	2014	[12,095]
AHTUS, HIGHER ED	2015	[13,779]
DHS, MTUS	2016	[15,548]
PMA, MEPS, NAPP	2018	[19,460]
IHGIS, MLP, USA Full Count**	2020	[23,092]
MICS, USA Full Count***	2024	[28,651]
<i>Current</i>	2025	[29,268]

*Without IPUMS codes, restricted version only. **With IPUMS codes. ***Preliminary 1950 data included.

IPUMS Bibliography data capture workflow steps are outlined in Table 2. There are three major steps to the workflow. Step one consists of organizing and allocating citations captured through Google Scholar, Google News, and Web of Science. The IPUMS Bibliography administrator receives email citation alerts from the three search engines. Citations are then copied from the email alerts and pasted into a Google document (“batch sheets”), organized by IPUMS data product and month, and then allocated to undergraduate student workers. This step is labor intensive, and it is challenging to eliminate citation duplication. The use of Google documents to organize and distribute batch sheets

to undergraduate student workers has three benefits: work is distributed and monitored in manageable chunks; the use of the comment function facilitates timely communication between the student and administrator around a troubling citation; and the batch sheets provide reference points for training and refining the bibliography workflow. Student worker turnover often occurs within academic years, so it is essential that the workflow seamlessly transfers to succeeding student workers. The IPUMS Bibliography and IT Product Teams are continually working to identify solutions to automate and increase accuracy and efficiency around identifying, organizing, and allocating captured citations.

Step two in the IPUMS Bibliography workflow consists of populating citations with metadata using the Mendeley Desktop and Mendeley Web Importer applications. Student workers verify the accuracy and completeness of the automatically ingested metadata and manually tag each citation with appropriate IPUMS Bibliography controlled vocabulary (Table 3). The development of formal IPUMS Bibliography technical and process documentation for onboarding and training has been essential and preserves institutional memory around IPUMS Bibliography process development. The third step in the IPUMS Bibliography workflow involves manual and automated review of each new citation before it is uploaded to the IPUMS Bibliography. As part of the bibliography expansion work in 2022, the IPUMS IT team created a tool for running a report to identify inaccurate Mendeley data, including invalid project, top, or country tags (Fabrizio 2022). Validated citations are moved into an “approved” folder and ingested by the IPUMS Bibliography tool.⁵

Table 2. IPUMS Bibliography workflow

Step	Source
Step One	Source One
Identify citations using IPUMS harmonized data products	Google Scholar, Google News, Web of Science [ProQuest], and user-submitted citations
Organize IPUMS citations into “batch sheets” by data product and month.	Google Scholar, Google News, Web of Science emails to IPUMS Bibliography email account, and user-submitted citations
Allocate batch sheets to undergraduate student members of IPUMS Bibliography Team	Batch sheets collated and assigned by IPUMS Bibliography project manager
Step Two	Source Two
Populate citation with automated metadata ingest and manual supplementation	Mendeley Desktop and Mendeley Web Importer apps
Tag citation with IPUMS projects (p-), topics (t-) and countries (c-)	IPUMS Bibliography controlled vocabulary
Step Three	Source Three
Validate bibliographic metadata through manual and automated review	Mendeley “unapproved,” “forthcoming,” and “approved” folders of new tagged citations; IPUMS IT generated error report
Disseminate IPUMS Bibliography	IPUMS Bibliography bibliography.ipums.org

Table 3. IPUMS Bibliography controlled vocabulary (sample)

Publication types	Mendeley Desktop reference management tool
book, chapter, journal, thesis (honor, master, doctorate), conference paper/proceeding, working paper (must be official working paper series with number), generic (e.g., report)	Mendeley Desktop app has twenty publication types of which IPUMS Bibliography uses seven in its controlled vocabulary
Project tags [p-ipums-]	IPUMS harmonized data products
atus, ahtus, cdoh, cps, dhs, higher-ed, ihgis, international, meps, mics, mtus, nhis, nhgis, pma, terra, usa, usa-full-count	See ipums.org for links to each IPUMS data product
Topic tags [t-]	Keywords
aging-and-retirement	aging, retirement, life course, intergenerational, gerontology
crime-and-deviance	trial, law, legal, crime, police
education	schooling, educational attainment, college, human capital
family-and-marriage	marriage, family, cohabitation, family structure, family formation
fertility-and-mortality	fertility, mortality, births, reproductive
gender	gender, feminization, masculine, women
health	anything health related
housing-and-segregation	home ownership, housing tenure, housing security, rent, living arrangements, redlining, affordable housing, neighborhood

labor-force-and-occupational-structure	labor force, employment status, occupation, industry, profession, labor supply, wage, earnings
land-use-and-urban-organization	land use, agriculture, urbanization, urban planning, transportation, property rights
methodology-and-data-collection	method, methodology, data collection, enumeration, questionnaire, forms, sample, modeling, data infrastructure, harmonization
migration-and-immigration	migrants, immigrants, nativity, foreign-born, cultural assimilation, cultural identity
natural-resource-management	agriculture, rainfall, water, food supply, extreme weather
population-data-science	longitudinal analysis, methodology, modeling, large-scale data, linking, harmonization
population-health-and-health-systems	public health, healthcare, health behaviors, care providers, health insurance, disability, access to care, healthcare utilization
population-mobility-and-spatial-demography	mobility, urbanization, neighborhood, local, displacement, spatial, mapping
poverty-and-welfare	poverty, minimum wage, income, social safety net, social security, inequality, material hardship, program participation
race-and-ethnicity	race, ethnicity, Hispanic, Latino/a/x, ethnic, culture, assimilation, minority, Great Migration
reproductive-and-sexual-health	reproductive, childbearing, maternal, female genital cutting, sexuality, contraceptives, family planning, pregnancy, postpartum, maternal morbidity and mortality,
work-family-and-time	parenting, motherhood, fatherhood, work-life-balance, single parthood, family structure, childcare, caregiving, paid leave, women's work
other	focus of publication not elsewhere classified (e.g., voting behavior)
Country tags [c-]	ISO Country List
Afghanistan, Albania, Algeria ... Yemen, Zambia, Zimbabwe	Updated as necessary

Conclusion

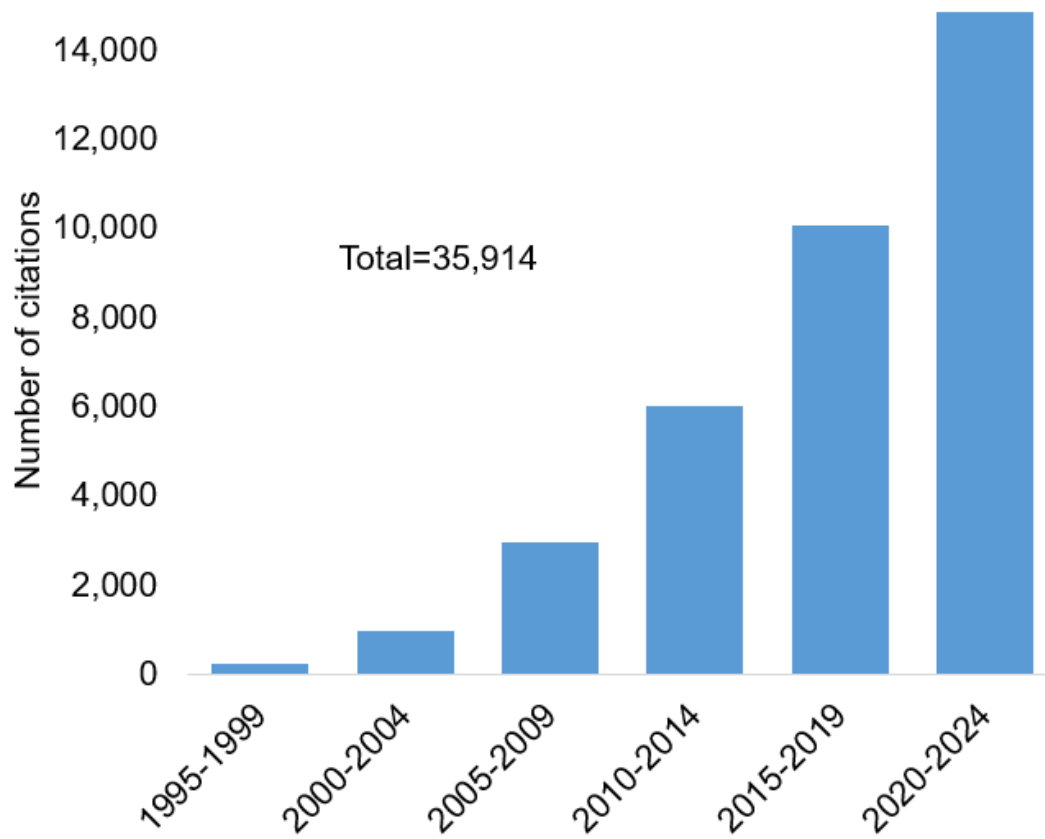
Curating, maintaining, and disseminating a bibliographic database requires significant short- and long-term investment in human, operational, and technical activities. The robust and creative use of IPUMS harmonized data products for academic research, policy, and teaching over thirty years and counting has presented an exciting challenge to curating the IPUMS Bibliography database. Internally, naming conventions around IPUMS harmonized data products were not standardized until all products were rebranded with the IPUMS prefix in 2016. Citation capture was greatly aided by

adoption of digital object identifiers (DOI) in 2016 (Magnuson and Thomas 2023). Prior to 2016, the effort to collect citations of IPUMS harmonized data products was significantly more challenging. The steady expansion and increasing diversity of IPUMS continue to present challenges of scale for the IPUMS Bibliography Team (Table 1).

From the earliest days of IPUMS, we recognized the need to compile a bibliography for both external and internal uses. Although they have evolved and we use them in our process, proprietary search engines have never proved sufficient for our purposes. For example, the use of Google Scholar can support but not replace our organization's knowledge database (Chaddaway et al. 2015). Long term curation, preservation, and access to traditional scholarly citations and citations of "gray literature" using IPUMS is the responsibility of the IPUMS Bibliography Team. Nevertheless, Google Scholar search results provide a useful tool for the IPUMS Bibliography Team to audit and measure data capture progress (Figure 9).

The IPUMS Bibliography tool is crucial for harnessing and making discoverable over three decades of IPUMS organizational knowledge. Institutional goals are supported and strengthened by the IPUMS Bibliography through capturing, organizing, and making discoverable: research analysis on tens of thousands of microdata and aggregate variables over time; processes for developing and maintaining well-defined and evolving harmonized classifications; and technological innovation and advancement supporting IPUMS systems. Intentional curation of an organizational knowledge database supports internal goals and strengthens social science infrastructure more broadly. IPUMS administration and data project managers use the IPUMS Bibliography to demonstrate the impact of IPUMS harmonized datasets to current and potential funders and to keep abreast of new directions taken by the research community using the data. IPUMS users can explore potential topics for investigation and conduct literature reviews of quantitative social science research. The IPUMS Bibliography workflow suggests a model to units situated in an organizational setting where the bibliography work is vital but secondary to the main product.

Figure 9. Google Scholar citation capture of articles using IPUMS, 1995-2024



References

- Fabrizio, F. (2022) "Bibliography Expansion PDS," General Information, ISRDI Archive.
- Chaddaway, N.R., Collins, A.M., Coughlin D, and Kirk, S. (2015) "The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching," PLoS ONE 10(9), p. e0138237. <https://doi:10.1371/journal.pone.0138237>
- Kumar, S. and Perakis, G. (2022) "A Brief History of the Internet and the World Wide Web," Management and Business Review, Vol. 2, No. 1, pp. 67-72. <https://doi.org/10.1177/2694105820220201001>
- Magnuson, D.L. and Ruggles, S. (2022) "Challenges of Large-Scale Data Processing in the 1990s: The IPUMS Experience," IEEE Annals of History and Computing, Vol. 44, Issue 4, pp. 71-83. DOI: [10.1109/MAHC.2022.3214736](https://doi.org/10.1109/MAHC.2022.3214736)
- Magnuson, D.L. and Thomas, W.L. (2023) "Expanding our perspective: Building a sustainable metadata culture," IASSIST Quarterly, 47(2). <https://doi.org/10.29173/iq1046>
- Ruggles, S. (1994a) "The Transformation of American Family Structure," The American Historical Review, Vol. 99, No. 1, pp. 103-128. <https://doi.org/10.2307/2166164>
- Ruggles, Steven. (1994b) "The Origins of Africa-American Family Structure," American Sociological Review, Vol. 59, No. 1, pp. 136-151. <https://doi.org/10.2307/2096137>
- Ruggles, Steven (1991) "The U.S. Public Use Census Microdata Files as a Source for the Study of Long-Term Social Change," IASSIST Quarterly, 15(2). <https://doi.org/10.29173/iq703>
- Wayback Machine. (2018) "IPUMS Bibliography," June 28. <https://web.archive.org/web/20180628045310/https://bibliography.ipums.org/>
- Wayback Machine. (2014) "Minnesota Population Center Bibliography," August 31. <https://web.archive.org/web/20140831071059/https://bibliography.ipums.org/>
- Wayback Machine. (2011) "Minnesota Population Center Bibliography," May 11. <https://web.archive.org/web/20110511120130/http://bibliography.ipums.org/>
- Wayback Machine. (2007) "Minnesota Population Center Bibliography," June 30. <https://web.archive.org/web/20070630231455/http://bibliography.ipums.org/>
- Wayback Machine. (2003) "IPUMS Bibliography," June 5. <https://web.archive.org/web/20030605043120/http://www.ipums.org/usa/research.php>
- Wayback Machine. (2000) "IPUMS Bibliography," August 16. <https://web.archive.org/web/20000816044420/http://www.ipums.umn.edu/~pipums/research.html>
- Wayback Machine. (1998) "IPUMS Research," December 12. <https://web.archive.org/web/19981212030023/http://www.ipums.umn.edu/>

Endnotes

- ¹ Diana L. Magnuson is curator and historian at the Institute of Social Research and Data Innovation (ISRDI) at the University of Minnesota. She can be reached by email at magn0031@umn.edu.

-
- ² "IPUMS by the Numbers," <https://www.ipums.org/by-the-numbers>
- ³ <https://www.webdesignmuseum.org/web-design-history>
- ⁴ <https://www.ipums.org/>
- ⁵ <https://bibliography.ipums.org>