

**UCLA**

**American Indian Culture and Research Journal**

**Title**

American Indian and Alaska Native Data in Federal Data Collections

**Permalink**

<https://escholarship.org/uc/item/01v5m0gp>

**Journal**

American Indian Culture and Research Journal , 42(1)

**ISSN**

0161-6463

**Authors**

Unger, Carlann  
Simon, Benjamin  
Pattison, Malka

**Publication Date**

2018

**DOI**

10.17953/aicrj.42.1.unger

**Copyright Information**

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed



# American Indian and Alaska Native Data in Federal Data Collections

*Carlann Unger, Benjamin Simon, and Malka Pattison*



*Editor's note: the supplement to this article, "Inventory of Federal AI/AN Data Collections," is available as a free resource for researchers at [http://www.books.aisc.ucla.edu/resources/AISC\\_Unger\\_supplement\\_online\\_only.pdf](http://www.books.aisc.ucla.edu/resources/AISC_Unger_supplement_online_only.pdf).*

American Indian and Alaska Native (AI/AN) tribes need access to quality data and information to make informed decisions concerning their communities, economic development, land and resource management, and other sovereign governance decisions. The collection, dissemination, and analysis of data by the federal government is also critical to ensuring that federal agencies and programs are delivering effective services to AI/AN tribes to meet tribal needs and deliver on federal responsibilities. However, various statistical and collection issues often negatively affect the quality and availability of federally collected AI/AN data, leading to missing data points or to estimates with large margins of error relative to comparable data on non-AI/AN populations and communities.<sup>1</sup>

As a key first to identify data gaps and improve the quality and accuracy of AI/AN data in federal datasets, this article provides an inventory and preliminary analysis of current AI/AN data in federal datasets.

## TRIBAL NATIONS, THE UNITED STATES, AND AI/AN DATA INVENTORIES

The United States has a unique legal and political relationship with tribal nations established through the United States Constitution, as well as through treaties, statutes, executive orders, and judicial decisions.<sup>2</sup> Tribal nations are sovereign governments

---

CARLANN UNGER is an economist and MALKA PATTISON is a policy analyst in the US Department of the Interior Office of Policy Analysis. BENJAMIN SIMON is chief Department of the Interior economist. The views in this paper reflect the view of the authors and not the Department of the Interior.

and engage in nation-to-nation relationships with the United States. As part of this nation-to-nation relationship, federal agencies must engage in meaningful collaboration with tribes on all aspects of data governance, including the collection, use, application, and dissemination of federal data on a tribe's citizens, lands, businesses, governments, etc. Tribes in the United States are increasingly exercising tribal data sovereignty,<sup>3</sup> which is the right to control the collection, application, and use of their own data.<sup>4</sup>

The need for an inventory of AI/AN data is immediate and significant because a comprehensive inventory has never been done. The literature on AI/AN data availability is limited, revealing three inventories of AI/AN data: an inventory of data to support economic development in Indian country, an inventory of data on health and well-being of AI/ANs and Native Hawaiians and Pacific Islanders (NH/PIs), and an inventory of federal labor market data sources that include information on AI/AN populations and tribal communities. Findings from these three inventories are included below.

Richard M. Todd summarized AI/AN demographic, business, financial, land, and government data and data gaps by reviewing approximately twenty studies that compared economic outcomes across reservations.<sup>5</sup> The resulting inventory includes 274 variables from approximately seventy-four distinct data sources, both federal and nonfederal. Main findings from this inventory include:

- ✦ Three of the biggest data gaps are in the areas of tribal governments, reservation business activity, and longitudinal data on reservation residents and households.
- ✦ Data sources on reservation financial sectors are limited, with only two sources identified: the Home Mortgage Disclosure Act and dated BIA estimates of reservation credit.
- ✦ Sources are available for Indian lands data, including the US Census Bureau (Census), the BIA, and the United States Department of Agriculture (USDA), but more data on land ownership is needed.
- ✦ Data sources available on tribal governments are the most limited, with no available sources identified, and none that are published by a federal agency. One exception is data on tribal casinos that are available in Inspector General Reports.

A report prepared by the research group Westat for the Department of Health and Human Services (HHS) inventoried data sources that include information on the health and well-being of AI/AN/NH/PIs in 2006,<sup>6</sup> and a corresponding report summarizing gaps and strategies for improvement in 2007.<sup>7</sup> The resulting inventory identifies sixty-eight datasets that are organized into eleven "policy areas."<sup>8</sup> The report also examined data availability by AI/ANs combined, AI/ANs separately, NH/PI (Native Hawaiian/Pacific Islanders) combined, and NH/PI separately. The main findings from the final report include:

- ✦ For the AI/AN population (combined), there are more than two datasets for ten of the eleven policy areas (the exception is military/veterans' issues).
- ✦ For the NH/PI populations (separately), there are limited data (less than two datasets) for education, elder well-being, justice system issues, and transportation issues. No datasets on military/veterans' issues are available.

- ✦ For the AI/AN populations (separately), there are no data gaps in child well-being, elder well-being, justice system issues, and military/veterans' issues. Limited data is available for the other policy areas.
- ✦ For the NH group, there are more than two datasets available to examine health issues. Limited data are available on economic well-being, education, family well-being, housing, justice system issues, and transportation issues. No datasets were identified that allow for examining child well-being, elder well-being, or military/veterans' issues.
- ✦ For the PI group, there are data gaps for child well-being, elder well-being, justice system issues, and military/veterans' issues.

Norm DeWeaver identified four sources of federal labor market data on AI/AN populations and reservations: the Bureau of Indian Affairs (BIA) Labor Force Report, the Census's American Community Survey, the Bureau of Labor Statistics Current Population Survey, and the Bureau of Labor Statistics Local Area Unemployment Statistics.<sup>9</sup> The findings from this report on labor market information include:

- ✦ The Census's American Community Survey and the BIA Labor Force Report are the only sources that provide employment data at the reservation level.
- ✦ The BIA Labor Force Report includes estimates on the AI/AN population and employment by state, BIA region, and tribe. The last report was published in 2013, after an eight-year reporting gap. The reports released prior to the 2013 report included estimates based on unknown methodologies, prompting quality concerns.
- ✦ The American Community Survey and the most recent BIA Labor Force Report do not capture "discouraged" workers, those who have given up looking for work and are not included in the Bureau of Labor Statistics or the Census Bureau's definitions of unemployed, which apply to many living on or near reservations, and can impede an accurate assessment of the labor market conditions of a reservation.

## OBJECTIVE, SCOPE, AND AI/AN DATASET IDENTIFICATION METHODS

The primary objective of this paper is to inventory and characterize current federal AI/AN datasets. While state and tribal governments and nongovernmental organizations are important sources of AI/AN data, their inclusion was beyond the scope of this report. A secondary objective is to complete a preliminary analysis of data gaps and to promote interagency discussions on improving federal AI/AN data.

The federal AI/AN data identified in the inventory are defined as federal datasets describing AI/AN populations or tribal lands, including:

- ✦ Collections that are not primarily AI/AN, but include relevant AI/AN information that can be extracted as a subset;
- ✦ Data collected primarily, as part of an official survey or census, or secondarily as part of program administration or research; and
- ✦ Publicly and non-publicly available data.

This inventory includes data that were available in 2016 and as such is a snapshot of the data available at that point in time. Both publicly available and non-publicly available datasets are included. Publicly available and non-public administrative data

are essential for effective programs, services, and economic development, and both suffer from quality issues and data gaps.<sup>10</sup> The identification of datasets helps to establish a baseline for federal agencies to begin addressing AI/AN data gaps and quality issues. It is possible that the same microdata underlies both public and nonpublic datasets. In these instances, the datasets are counted as two distinct datasets. Datasets that may include AI/AN or tribal land data, but which cannot be identified or extracted from the larger data collection, are not included.

It is possible that some datasets include a geographic feature such as an address or zip code that would allow the user to identify data attributed to a reservation or tribe. These data collections were not included in the above counts unless they explicitly stated that they include data on tribal lands. It should be noted that this paper does not advocate for federal agencies to make public all federally collected data. Many federal datasets include personally identifiable information that federal agencies are legally required to keep private for the safety and privacy of citizens who respond to federal surveys and programs. As required by law, federal agencies should work cooperatively and in consultation with all tribes on data collections and dissemination that include data on their citizens, lands, businesses, governments, etc.

Three approaches were used to identify federal AI/AN datasets:

- Identification of datasets on Data.gov that include AI/AN data;
- Identification of databases listed on the federal IT Dashboard<sup>11</sup> that include AI/AN data; and
- A government-wide data call of federal agencies.

### *Data.gov*

Data.gov is an open data website maintained by the General Services Administration (GSA) that launched in June of 2009 to promote an open and transparent government.<sup>12</sup> The website catalogs metadata on federal datasets with the intent of allowing the public to easily find and access federal data. Since its launch in 2009, the site has grown from forty-seven datasets to 184,557, of which 146,796 are maintained by federal agencies.<sup>13</sup>

To identify AI/AN datasets, the Data.gov catalog was searched using the following search terms: (1) American Indian; (2) Alaska Native; (3) Native American; (4) tribal; and (5) tribe. Relevant information extracted for each dataset included the agency, name of data, a description of the data, and the data format.

The results from each search term were reviewed to eliminate duplicates and condense multiple entries for the same dataset. Duplicate entries occurred when search terms yielded the same dataset or multiple entries for different years; these duplicates were eliminated. For example, the ACS has separate entries in the Data.gov catalog for its release in 2014 and each release in the years before. All multiple-year entries were considered duplicates and eliminated to leave one entry. Some datasets also had multiple entries for different geographic areas or types of information. This was especially prevalent with boundary and geospatial datasets that

have multiple entries for different shapefiles or geographic layers. For example, the Environmental Protection Agency has multiple entries for the different types of facilities available in their Facility Register Service. These entries were condensed and counted as one.

Although Data.gov identified many federal AI/AN datasets, it does not include administrative data, and is limited to datasets federal agencies identify for the catalog.

### *The Federal IT Dashboard*

To address the administrative data gap in Data.gov, databases were identified through the IT Dashboard. Databases that include AI/AN data were identified using the “IT Portfolio” file. The IT Portfolio contains information on major IT investments including the agency, bureau, investment title, investment description, primary and secondary service areas, and total IT spending on the investment for the past three years. Many government agencies require major IT Investments to meet their programmatic needs.

Four methods were used to identify databases that include data on AI/AN populations and tribal lands:

- ♦ Filtering by primary or secondary service area code of 361, “American Indian and Alaska Native Relations.”
- ♦ A text search in the investment title or investment description for the same search terms used for Data.gov: “American Indian,” “Alaska Native,” “Native American,” “tribe,” or “tribal.”
- ♦ Any IT investment for the BIA, Bureau of Indian Education (BIE), and the Indian Health Service (IHS).
- ♦ A manual review of investment descriptions to identify databases potentially containing information on AI/AN individuals or lands as part of a larger, non-primarily AI/AN dataset.

### *AI/AN Dataset Information Gathered Directly from Federal Agencies*

Data.gov and the IT Dashboard identified many AI/AN datasets, but agencies may not put all datasets on Data.gov, and many data are not stored in databases that would require a major IT investment listed on the IT Dashboard. Also, as indicated above, one method of identification included a review of IT investment descriptions that incurred risk of including databases that do not contain AI/AN data. In order to identify datasets that may not be included in Data.gov or the IT Dashboard, the authors requested that federal agencies identify datasets that were not included in these sources. Data requests were sent to nineteen federal agencies. Responses were received from seventeen. Agency responses were used to review and edit information from Data.gov and the IT Dashboard to prepare this draft inventory.<sup>14</sup> The full dataset of information collected is available online.<sup>15</sup>

## RESULTS

### *Datasets by Agency and Accessibility*

Table 1 summarizes the data collected by federal agency. The table shows the total number of AI/AN datasets by agency, ranked highest to lowest by number of datasets. It also includes the number of publicly available and limited public access datasets. A dataset with “Limited Non-Federal Access” includes those that allow a limited number of individuals to view data, including datasets where a participating individual or tribal government can view their own data, or instances where high level aggregates are available in report form, but the more granular data, such as breakouts by tribe, are not available to the public. A \* symbol next to the agency name in table 1 indicates the agency did not respond to the data call, meaning all identifications were made through Data.gov and the IT Dashboard, and have not been verified by the agency. A \*\* symbol indicates the agency partial responded to the data call, and a \*\*\* indicates the agency did not receive a data call form.

The data call identified 448 AI/AN datasets from twenty agencies. The information provided varies by dataset. All datasets identify the agency, name of the data or program, and a description of the data. The information on counts is a useful starting point, but it is recognized that this type of information does not provide information about how the data are used, the frequency of data use, and data quality.

Figure 1 shows that of the 448 datasets identified, 43.3 percent (194) are publicly available. However, 23.0 percent (103) of the identified collections do not have a determined accessibility, meaning the agency did not provide that information and accessibility could not be determined with certainty.

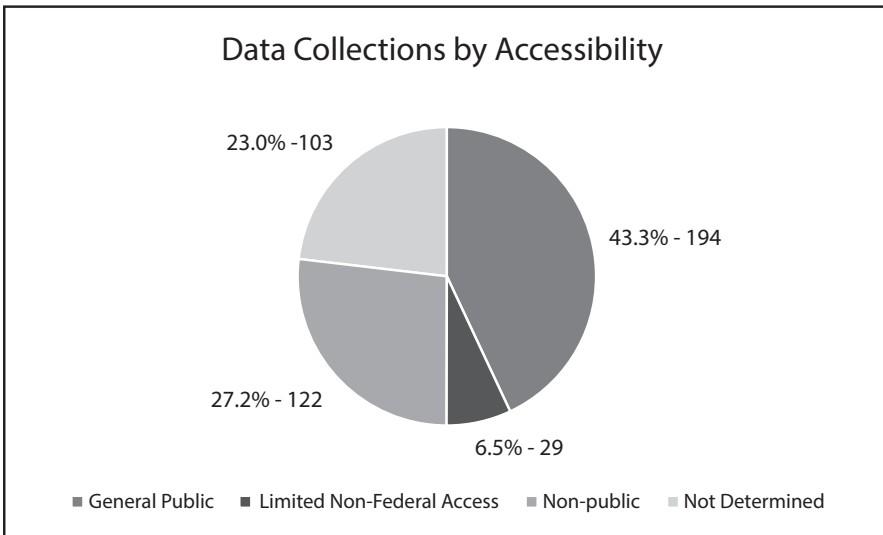


FIGURE 1. AI/AN Data Collections by Accessibility.

TABLE 1  
NUMBER OF AI/AN DATA COLLECTIONS BY AGENCY

Agency	Number of AIAN Data Collections	Number of Data Collections that are Publicly Available	Number of Data Collections with Limited Nonfederal Access
Department of the Interior	149	39	14
Department of Agriculture**	54	8	1
Environmental Protection Agency**	48	35	1
Department of Health and Human Services	40	23	2
Department of Commerce	24	21	0
Department of Justice	23	9	3
Department of Energy	20	19	1
Department of Transportation	16	11	0
Department of Housing and Urban Development*	14	3	0
Department of Homeland Security	14	2	2
Department of the Treasury	11	3	1
Department of Education	9	8	1
Department of Labor	8	7	0
Department of Defense	9	1	1
General Services Administration	3	2	1
Small Business Administration	2	0	0
Federal Financial Institutions Examination Council***	1	1	0
Nuclear Regulatory Commission	1	1	0
Social Security Administration	1	0	1
Consumer Financial Protection Bureau***	1	1	0
Department of Veterans Affairs	0	0	0
<b>Total</b>	<b>448</b>	<b>194</b>	<b>29</b>

\*Agency did not respond to data call.

\*\*Partial response to data call.

\*\*\*Agency did not receive a data call form.

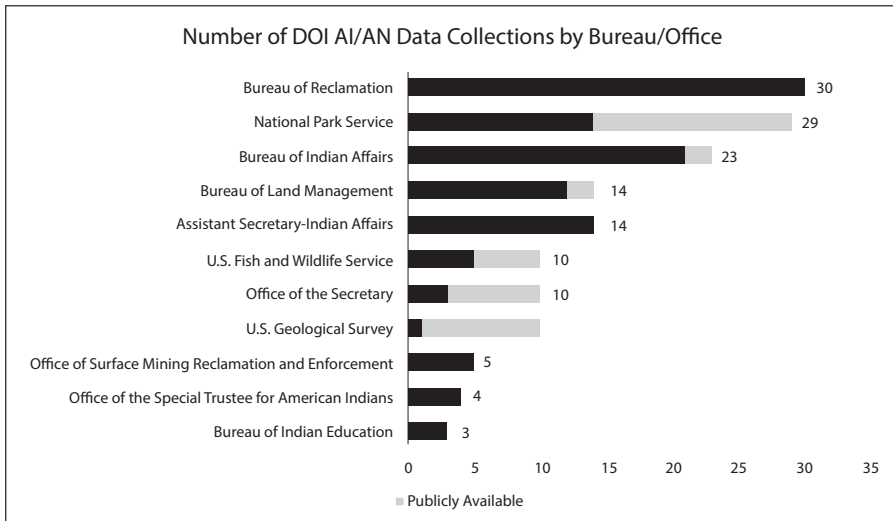


Figures 2–6 present information on the data collected by the Department of the Interior, the Department of Agriculture, the Environmental Protection Agency, the Department of Commerce, and Department of Health and Human Services. The datasets of these five federal agencies account for 70.3 percent of the total identified AI/AN datasets.

**DEPARTMENT OF THE INTERIOR**

The Department of the Interior (DOI) has 149 datasets, the highest number of AI/AN datasets identified, with more than double the number of datasets compared to the next agency. It is to be expected that DOI has a large number of AI/AN datasets, as the agency’s mission is to “protect and manage the Nation’s natural resources and cultural heritage; provide scientific and other information about those resources; and honor its trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities.”<sup>16</sup> To effectively carry out this mission, many of the DOI’s bureaus and offices provide services to, and work with, tribes.

Figure 2 shows the number of AI/AN datasets by DOI bureau and office. The two bureaus with the highest number of identified AI/AN datasets are the Bureau of Reclamation (Reclamation) and the National Park Service (NPS), followed by the BIA.



**FIGURE 2.** Number of Data Collections by DOI Bureau/Office.

The AI/AN data collected by Reclamation is for water-related projects and studies on tribal lands, including instrumentation readings at BIA dams, water delivery support, and reports prepared for needs assessments, environmental impact studies, cost estimates, etc. Reclamation also maintains data on AI/AN artifacts and archaeological sites. None of Reclamation’s datasets were confirmed as publicly available, indicating the data are primarily for program administration purposes. The

NPS also maintains data on archaeological sites, as well as historic places, geospatial data on NPS boundaries, Native American Graves Protection and Repatriation Act (NAGPRA) notices data, and others. With nine datasets identified as accessible to the public, the NPS is the DOI bureau with the highest number of publicly available datasets, including the administrative boundaries of the NPS units, the National Register of Historic Places, the National Register and Landmarks Application, the NAGPRA summaries database, and the Culturally Unidentifiable and Culturally Affiliated Native American inventories database.

While Reclamation and NPS have the highest number of datasets identified as standalone Bureaus, the DOI's AI/AN-focused bureaus and offices have the highest combined number of AI/AN datasets at forty-seven, including the BIA, BIE, the Assistant Secretary-Indian Affairs, and the Office of the Special Trustee for American Indians. These data cover the numerous programs administered by these bureaus and offices, and because many contain personally identifiable information, are not public. Of the forty-seven, two are confirmed as publicly available: the BIA Indian Lands Dataset, which depicts feature location, selected demographics, and other associated data for the 567 federally recognized tribes in the contiguous United States and Alaska; and regional budget formulation input results, which is information collected by tribes and tribal programs used to identify BIA regional priorities for the formulation of the BIA's annual budget. The DOI has forty non-publicly available datasets, indicating the majority of data are collected for program and project administration and cannot be made public due to personally identifiable information.

#### **ENVIRONMENTAL PROTECTION AGENCY**

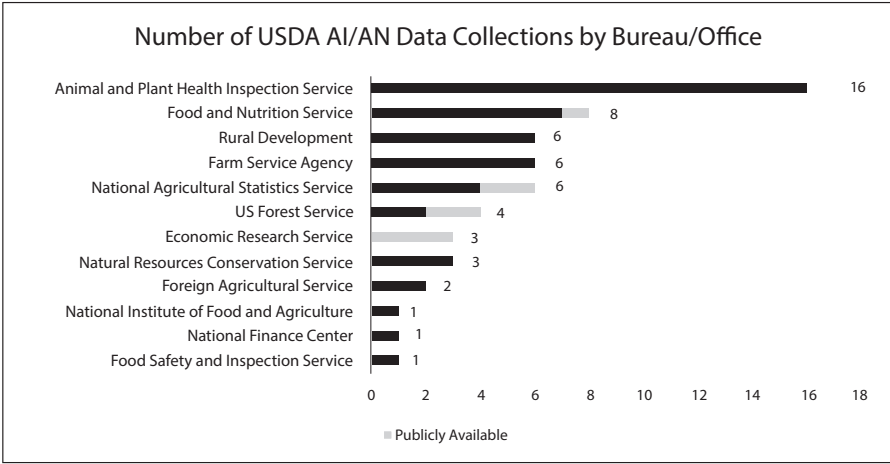
A breakout by EPA office is not currently available. However, of the forty-eight datasets identified, thirty-five are publicly available. The EPA provides data on air and water quality, EPA facilities, and EPA grants. Many of these data are included in EPA's Envirofacts data product that includes topic searches on air, waste, facility, land, compliance, water, and radiation, available at the tribal level. These data can be viewed with a geographic interface using the EnviroMapper data tool.

#### **DEPARTMENT OF AGRICULTURE**

The USDA identified fifty-four datasets as containing AI/AN data. No additional information was available on the geographic scope and accessibility of each dataset.

Figure 3 shows that the USDA bureaus/offices with the highest number of AI/AN datasets are the Animal and Plant Health Inspection Service and the Food and Nutrition Service, followed by Rural Development, the Farm Service Agency, the National Agricultural Statistics Service, and the Economic Research Service. The majority of datasets identified contain administrative data on the numerous inspection, farm, food and nutrition, and rural development programs.

The inventory identifies eight publicly available datasets. The Food and Nutrition Service provides Women, Infants, and Children participation and cost data at the state and select tribal organization level. The Economic Research Service has created two mapping tools: one that can be used to estimate the percent of non-Hispanic Native Americans in the United States, mapped by county, and a spatial overview of



**FIGURE 3.** Number of AI/AN Data Collections by USDA Bureau/Office.

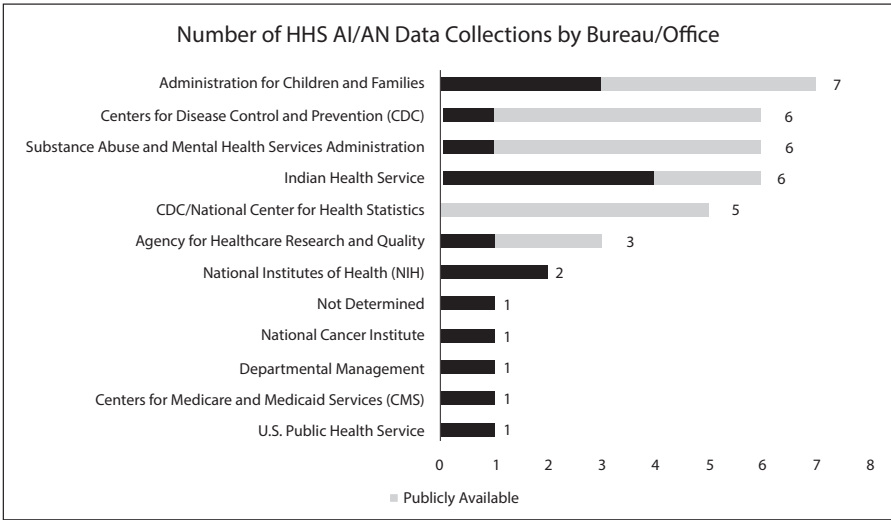
food-access indicators for low-income and other census tracts using different measures of supermarket accessibility. The USDA's National Agricultural Statistics Service conducts the Census of Agriculture every five years and is the only source of uniform, comprehensive agricultural data for every state and county or county equivalent, and for select reservations. The US Forest Service also provides publicly available geospatial data on tribal lands ceded to the United States.

**DEPARTMENT OF HEALTH AND HUMAN SERVICES**

The HHS identified forty datasets across ten different bureaus/offices. Figure 4 shows the distribution of these datasets across HHS Bureaus and offices. Of the forty identified, 57.5 percent (23) are publicly available.

The office with the highest number of datasets within HHS is the Administration for Children and Families (ACF), which collects data on the administration of two important government programs, the Head Start program and Temporary Assistance for Needy Families (TANF) programs. The TANF reporting system allows the public to query and download TANF data at the state and tribal agency level, including estimates such as number of applicants and number of approved applicants. The ACF provides data on the Head Start program at the grantee/delegate level, including demographics on children and family, information on the Head Start workforce, and program characteristics in their annual Head Start fact sheets. ACF is also currently conducting a large descriptive study of AI/AN children in Head Start. The data collection is still ongoing, but a large dataset that researchers can access will be available.

The Centers for Disease Control and Prevention, together with the National Center for Health Statistics, collects and disseminates data on various health topics, including the National Immunization Survey, National Health Interview Survey, and the National Vital Statistics System, which provides public health data including birth, mortality, fetal death, and marriage and divorce data. Data available to the public can



**FIGURE 4.** Number of AI/AN Data Collections by HHS Bureau/Office.

be accessed using the Center for Disease Control and Prevention’s data tool, CDC Wonder, that allows users to query data for reports at the national, state, and regional level, as well as by race, including AI/AN.

The IHS is the federal agency responsible for providing federal health services to American Indians and Alaska Natives. The inventory identifies six IHS datasets. The two that are publicly available are facility-mapping tools.

#### THE DEPARTMENT OF COMMERCE

While the Department of Commerce (DOC) ranks fifth in the number of AI/AN datasets identified, most likely it would rank first if AI/AN topics were identified by agency. Figure 5 shows Census has the highest number of identified datasets within the DOC. Census data is an important and widely used source of publicly available AI/AN data.

Census provides over one hundred demographic, housing, economic, and social variables at the national, state, county, metropolitan, and tribal level as part of the ACS. The inventory identifies ten distinct ACS datasets, including Estimates, Estimate Summary Files, and Public Use Microdata Sample files for 1-year, 3-year, and 5-year estimates. Higher level aggregates, such as the AI/AN population at the national and state level, are available in the 1-year estimates, while the majority of tribe-specific data are included in the 5-year estimates, as smaller areas require sixty months of aggregate data to produce reliable estimates. Estimate summary files contain estimates, margins of error, and geography files for each release. Public Use Microdata Sample files are a set of untabulated records about individual people or housing units that allow data users to create custom tables that are not available through pretabulated (or summary) ACS data products.

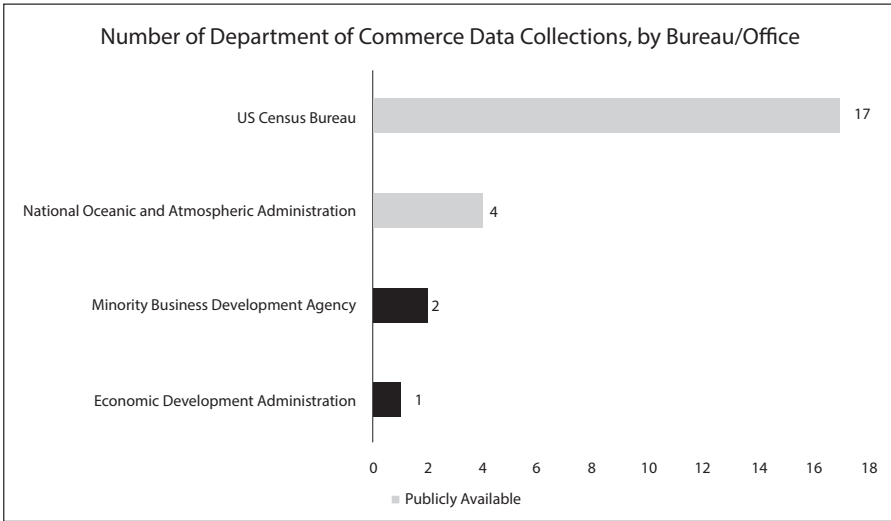


FIGURE 5. Number of AI/AN Data Collections by USDA Bureau/Office.

Census also provides population estimates at the national, state, county, and tribal level as part of the decennial census, and AI/AN-owned business data estimates through the Survey of Business Owners, including the number of firms, total sales, and annual payroll, at the national and state level.

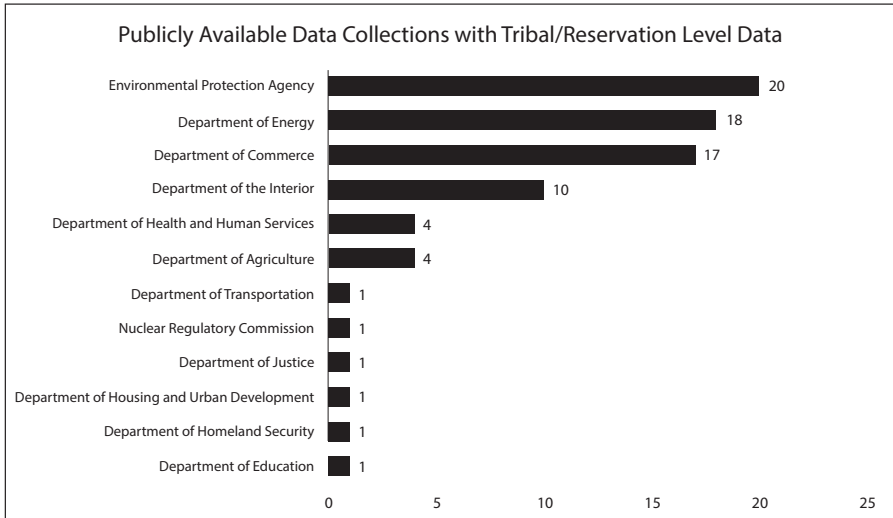
The inventory identifies four publicly available datasets from the National Oceanic and Atmospheric Administration (NOAA). However, it should be noted that they were not reviewed by the NOAA. NOAA’s datasets include Coastal Tribal Lands; Ocean Uses: Hawaii; Production Data – North Puget Sound Chinook salmon captive propagation; and Subsistence Registration Permit Program.

**PUBLICLY AVAILABLE TRIBAL- AND RESERVATION-LEVEL DATASETS**

The identified data collections include AI/AN data at the national, state, county, zip code, and tribal geographic levels. While all are important, data at the tribal and reservation levels provide an opportunity for more robust program management, as well as providing a better picture of socioeconomic conditions of tribal members and communities. It is important to make the distinction between *tribal-* and *reservation-* level data. Reservation-level data is tied to a particular geographic area, which may include multiple tribes and non-AI/AN populations. Tribal-level data may include tribal members who do not live on a reservation and may not be tied to a specific geographic area. Federal agencies often use different criteria for eligibility in AI/AN programs, and to determine tribal populations and tribal service areas.<sup>17</sup> Therefore, tribal populations may not be comparable across agencies.

Seventy-nine (40.5%) of the 195 publicly available datasets identified include tribal- and/or reservation-level data. These datasets are identified with a \* symbol next

to “General Public” in the Accessibility column of the inventory. Figure 6 below shows the breakout of these datasets by agency.



**FIGURE 6.** *Publicly Available Data Collections with Tribal/Reservation Level Data.*

The agencies collecting the data provide some insight into the type of data available at the tribal and/or reservation level. The EPA and the Department of Energy, for example, provide half of the identified datasets. The Department of Energy’s tribal-level datasets include information on renewable energy sources in Indian country, including wind speeds, solar energy, biomass, and hydrothermal energy sources in their interactive mapping tools. The data also include locations of bioenergy, natural gas, petroleum, hydrothermal, and power plants. The EPA includes a tribal land geography layer in its Enviromapper tool that allows the user to view information on many of EPA’s programs, including air, water, Superfund sites, and hazardous waste. The EPA also includes many different geographic layers depicting tribal areas by region.

Seventeen datasets include reservation-level data for the DOC, which represents 70.8 percent of the DOC’s datasets, and 100 percent of their publicly available datasets. The majority of these data come from the Census and are included in the 5-year, 3-year, and 1-year ACS estimates. However, the 1-year estimates are limited to populations over 60,000, which applies to very few reservation areas.

The Department of Justice provides one dataset, the Census of Tribal Justice Agencies, that includes information on the range of justice agencies operating in tribal jurisdictions, the services those agencies provide, and the types of information systems maintained at the tribal level. However, it has not been updated since 2002. Other data available at the tribal level include information on cultural items (DOI), tribal TANF programs (HHS), program office location and contact information, and tribes physically located within fifty miles of a nuclear power plant (National Regulatory Commission).

## CONCLUSION

This inventory of federal AI/AN data characterizes existing datasets by agency, accessibility, and tribal/reservation level availability. It is the first inventory that attempts to catalog all federal datasets that include AI/AN populations or tribal lands. The inventory identifies 448 unique datasets from twenty-one federal agencies. One hundred and ninety-four (43.3%) of these datasets are publicly available, and seventy-nine of the publicly available datasets include data at the tribal or reservation level.

A review of publicly available datasets at the tribal/reservation level showed data gaps in the areas of reservation businesses, reservation financial sectors, tribal governments, labor markets, and education. While there are publicly available data on AI/AN-owned businesses at the national, state, and county levels through the Survey of Business Owners, there are no business data available by reservation. The GSA's SmartPay captures information on tribally owned businesses, including location information that would allow a user to flag businesses on reservations. However, these data are limited to tribally owned federal agency vendors. With regard to reservation financial sectors, the Consumer Financial Protection Bureau's Home Mortgage Disclosure Tool provides mortgage lending information at the census tract and county level, including demographic information on race, sex, and ethnicity, and financial information such as approval/denial, the denial reason, and loan type, but does not provide information at the reservation level. A review of the inventory showed no data collections that include information on tribal governments. The Department of Education's 2008 report, "Status and Trends in the Education of American Indians and Alaska Natives," provides tribal level demographic information, but only at the national and state levels. These findings are similar to the findings of Todd's review of data for economic development,<sup>18</sup> the HHS review of data on the health and well-being of AI/AN/NH/PIs,<sup>19</sup> and DeWeaver's review of data on reservation level labor markets.<sup>20</sup>

This inventory may be useful for tribes, researchers, and federal agencies to locate data, identify possible data linkages, identify other data gaps, and identify possible tribal and agency collaborations. While the inventory offers a view of what data are being collected and by what agency, it does not provide information on the quality and coverage, which are important factors to consider for a complete characterization of federal AI/AN data. In the long term, this inventory could be enhanced to include this information. Importantly—given that the inventory represents a snapshot taken at a given point in time—it would be helpful if the inventory could be periodically updated, as federal agencies are likely to update existing datasets and collect new data. A web-based, more searchable tool would also be beneficial to potential users of this inventory. Most importantly, federal agencies should continue to collaborate with tribes to enhance and improve AI/AN data.

## NOTES

1. National Congress of American Indians Policy Research Center, "Data Disaggregation: The Asterisk Nation," <http://www.ncai.org/policy-research-center/research-data/data>.
2. Executive Order 13175—Consultation and Coordination with Indian Tribal Governments (Nov. 6, 2000) charges executive departments and agencies with engaging in regular and meaningful consultation and collaboration with tribal officials in the development of federal policies that might affect tribes, and strengthening the nation-to-nation relationship between the United States and AI/AN tribes.
3. Stephanie Carroll Rainie, Jennifer Lee Shultz, Eileen Briggs, Patricia Riggs, and Nancy Lynn Palmanteer-Holder, "Data as a Strategic Resource: Self-determination, Governance, and the Data Challenge for Indigenous Nations in the United States," *International Indigenous Policy Journal* 8, no. 2 (2017), <https://doi.org/10.18584/iipj.2017.8.2.1>.
4. Desi Rodriguez-Lonebear and Stephanie Carroll Rainie, "The United States Indigenous Data Sovereignty Network—About Us," Native Nations Institute, University of Arizona, <http://usindigenousdata.arizona.edu/about-us-0>.
5. Richard M. Todd, "Indian Country Economic Development: Data and Data Gaps," (August 2012), [https://www.minneapolisfed.org/community/indian-country/events/~//media/Files/community/indiancountry/Todd\\_Data\\_and\\_Data\\_Gaps\\_Paper.pdf](https://www.minneapolisfed.org/community/indian-country/events/~//media/Files/community/indiancountry/Todd_Data_and_Data_Gaps_Paper.pdf).
6. Westat, "Data on Health and Well-being of American Indians, Alaska Natives, and Other Native Americans," data catalog prepared for the Department of Health and Human Services, (December 2006), <https://aspe.hhs.gov/system/files/pdf/74856/report.pdf>.
7. The eleven categories are Child Well-being (4 datasets), Demographic and Economic Indicators (3 datasets), Economic Well-being (6 datasets), Education (7 datasets), Elder Well-being (2 datasets), Family Well-being (4 datasets), Health Policy Issues (30 datasets), Housing Issues (3 datasets), Justice System Issues (7 datasets), Military Service/Veterans' Issues (1 dataset), and Transportation (1 dataset).
8. Westat, "Gaps and Strategies for Improving AI/AN/NA Data," final report prepared for the Department of Health and Human Services (January 2007), <https://aspe.hhs.gov/system/files/pdf/74926/report.pdf>.
9. Norm DeWeaver, "Indian Workers and the Reservation Labor Market: Reality, Research, and a Way Forward" (August 2014), <https://doe.state.wy.us/lmi/LAUS/LM-dynamics-in-reservation-areas-9-1-14.pdf>.
10. Administrative data are data that are collected for program administration purposes. Examples include transaction records, application forms, and record-keeping.
11. The IT Dashboard is a federally managed website that tracks federal information technology investments by agency and is updated annually to provide information and track progress on federal information technology investments over time. Office of Management and Budget, "Frequently Asked Questions," <https://www.itdashboard.gov/drupal/frequently-asked-questions>.
12. John P. Holdren, Peter Orszag, and Paul F. Prouty, "Memorandum for Heads of Departments and Agencies re President's Memorandum on Transparency and Open Government—Interagency Collaboration" (February 24, 2009), <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2009/m09-12.pdf>.
13. Data collections fluctuate as the catalog is updated on an ongoing basis as new data collections are added and issues with duplicates are resolved.
14. For the data call, the information found through Data.gov and the IT Dashboard was used to pre-populate a collection form that requested federal agencies to: (1) review and validate the information found through Data.gov and the IT Dashboard; (2) identify AI/AN data collections their



agency collects and maintains that were not already listed; and (3) include more information on data collections that could not be found through online sources.

15. Office of Policy Analysis, Report prepared for the Department of the Interior, "Federal American Indian and Alaska Native Data Collections" (November, 2016), <https://www.doi.gov/ppa/reports-and-statistics>.

16. Department of the Interior, Mission Statement, <https://www.doi.gov/whoweare/Mission-Statement>.

17. Desi Rodriguez-Lonebear, "Building a Data Revolution in Indian Country," in T. Kukutai and J. Taylor, eds., *Indigenous Data Sovereignty: Toward an Agenda* (Canberra: Australian National University Press, 2016), 259.

18. Todd, "Indian Country Economic Development: Data and Data Gaps."

19. Westat, "Data on Health and Well-being of American Indians, Alaska Natives, and Other Native Americans."

20. DeWeaver, "Indian Workers and the Reservation Labor Market: Reality, Research, and a Way Forward."