2016 AMERICAN INDIAN AND ALASKA NATIVE DATA WORKSHOP SUMMARY REPORT

Office of Policy Analysis
U.S. Department of the Interior
November 2016

Contents

I. INTRODUCTION

The Office of Policy Analysis of the Department of the Interior held a workshop on American Indian and Alaska Native (AI/AN) data on September 13, 2016. The purpose of the workshop was to bring together producers and users of AI/AN data to share how data is developed and used, limitations of the data, and data gaps in analytic applications. The workshop provided a forum for users of AI/AN data to share information, discuss options for addressing AI/AN data issues, and identify opportunities to collaborate on AI/AN data collection to leverage Federal and other resources. This report summarizes and synthesizes the presentations and discussions that took place at the workshop, including descriptions of each session, a discussion of the common themes that arose during those sessions, and a list of critical data gaps identified. It also includes a list of possible next steps for addressing those gaps.

II. AI/AN DATA GAPS AND MOU

Federal and tribal decision makers require a wide variety of data to inform their decisions. However, the quality, quantity and accessibility of data that pertain to AI/AN people and institutions are often insufficient to meet needs. Researchers have documented AI/AN under-representation, misidentification, or omission in many data collection efforts, including the U.S. Census Bureau's American Community Survey (ACS), [1, 2] official health records and death statistics, [3, 4] and data on crime and juvenile justice in Indian Country. Such data issues make it difficult to identify or quantify problems impacting AI/AN individuals and communities, which impedes policy and decision making for both tribes and federal agencies[5, 6] and can lead to detrimental effects on AI/AN citizens and communities. For example, less than half of the measures in the National Healthcare Disparities report can be applied to AIANs, most often because data on AIANs were not collected or there were too few samples for accurate

estimates and comparisons.[7] This gap makes it difficult to document disparities in health care between AIANs and other groups.

Data deficiencies can also impede federal agencies' effective management of programs that benefit tribal nations and individuals, which in Fiscal Year 2016 received over $19 billion.[8] For example, the under-representation of AI/AN people in federal government surveys may cause them to miss out on the federal funding to which they would otherwise be entitled. In FY 2009, the ten largest federal assistance programs all relied in part on census-related data, including the ACS. Those programs accounted for $478.3 billion in federal spending (including $122.7 billion from the American Recovery and Reinvestment Act).[9]

On January 13, 2016, the U.S. Census Bureau (Census) and the Bureau of Indian Affairs (BIA) within the U.S. Department of the Interior (DOI) signed a Memorandum of Understanding (MOU). Recognizing the unique legal relationship between the U.S. government and tribes, including the federal trust obligations, the parties agreed "to promote communication and collaboration between the two agencies and improve the dissemination of accurate data". In particular, they pledged to establish a workgroup to share, discuss, and resolve issues concerning AI/AN data and make efforts to share geographic information system (GIS) shapefiles or geodatabases. The MOU also recognizes a role for other federal agencies who collect and/or use AI/AN social and economic data. The Office of Policy Analysis (PPA) has taken an active role in the implementation of the MOU and made improving AI/AN data a priority.

III.   WORKSHOP

The American Indian and Alaska Native Data Workshop was held on September 13, 2016 in the North Penthouse at the Main Interior Building (Udall Building) in Washington, D.C. A total of 53 participants attended. Participants included Federal representatives from the Departments of the Interior, Commerce, Agriculture, Education, Health and Human Services,

Housing and Urban Development, Justice, Labor, Transportation and Veterans Affairs, Environmental Protection Agency, and Executive Office of the President as well as non-Federal experts.

The workshop included a mix of presentations and discussion. During the morning session, there were two presentations on AI/AN policy, a facilitated panel discussion with officials from five federal agencies, and a question-and-answer session. During the afternoon session, there were three expert presentations, one question-and-answer session and an open facilitated discussion among attendees. The complete agenda is attached to this report as Appendix 1.

In advance of the workshop, DOI compiled an inventory of federal AI/AN datasets. The inventory identified datasets managed by federal agencies that contain data describing AI/AN people or tribal lands. The purpose of the inventory was to establish a baseline level of knowledge about the scope and variety of federal AI/AN data, allowing DOI and others to identify data gaps and identify opportunities for cooperation between agencies. A draft report describing the methodology used to create the inventory and analyzing its contents was distributed at this workshop and is available separately from PPA.

MORNING SESSIONS

Presentation by Malia Villegas, Director, National Congress of American Indians (NCAI) Policy Center

Dr. Villegas discussed a range of topics related to AI/AN data, emphasizing the importance of "tribal-level" data collection, and of considering the purpose of each data collection and how the data will actually be used. While noting NCAI's work on improving the accuracy of Census data and encouraging native people to participate, she opined that decennial census enumerations and ACS samples do not provide sufficient data at the level of the individual tribe, resulting in key omissions.

Dr. Villegas stated that data could be made more useful by building tribal data collection capacity, selecting appropriate measures, and considering how different definitions of "Indian" affect decisions and resource allocation.

Dr. Villegas advocated for the use of "native-to-native comparisons," or comparing different tribes and native groups to each other, arguing that the approach could help identify successful approaches among some native groups that could be helpful to others. She noted, however, that the approach would require the ability to disaggregate data and obtain data below the national level. She also supported further work on data linkage, and on secondary data (i.e. data collected from other available sources rather than by the primary researcher). Dr. Villegas closed by calling for existing data to be better catalogued and more accessible, and for the impacts of data policies to be transparently monitored.

Presentation by Karen Diver, Special Assistant to the President for Native Affairs, former Chairwoman of the Fond du Lac Band of Lake Superior Chippewa

Ms. Diver drew on her experience as a tribal chairwoman and administrator to discuss the importance of making data collections culturally relevant and ensuring that policies actually help. As an example, she discussed the State of Minnesota's policy of distributing state housing aid based on homeless surveys, which had not been conducted in reservation communities and failed to capture the doubling- and tripling-up, or "couch surfing," that was common among homeless American Indians. She also discussed the distrust between tribal leaders and federal officials, noting that data have often been used against tribes in the past. She recognized the tension between "one size fits all" federal policy and tribal needs, noting that her own tribe was ineligible for a needed USDA grant because its per capita income was $14 too high. She pointed out that half of tribal citizens live in urban areas rather than reservations but still need appropriate services, and she argued that tribes are underfunded because those citizens are not counted. She closed by calling for participants to consider Indian needs and stay in touch with the White House Council on Native American Affairs.

Facilitated Panel Discussion

The panel was introduced by Benjamin Simon, Interior's Chief Economist, who briefly discussed the inventory of federal AI/AN data conducted by DOI. Four panelists participated, each giving a short presentation describing the data possessed and needed by their agencies.

- **Ann Marie Bledsoe Downes, Deputy Assistant Secretary for Indian Affairs at DOI and Acting Director, Bureau of Indian Education.** Ms. Bledsoe Downes outlined the data managed by DOI's Indian Affairs offices, including data on trust resources, human resources, justice services, student performance, demographic, loan, energy and mineral data. She also outlined their needs: socio-economic and business data, larger and more frequent sampling of native populations, tribal and individual business data, educational outcome and performance data, and real-time school operating cost data.

- **Victoria Velkoff, Chief, American Community Survey Office at the U.S. Census Bureau.** Dr. Velkoff provided background information on how Census collects data on AI/AN populations, primarily by tabulating data for those who self-identify as AI/AN on the decennial census, ACS and Survey of Business Owners. She discussed the data products available from Census, including population and business data. She identified several needs, including more reliable data at smaller levels of geography, ways to address the challenges of sampling in Alaska, higher response rates in AI/AN areas, and an updated code list for recognized tribes.

- **Monica Hill, Program Analyst, White House Initiative on American Indian and Alaska Native Education, Department of Education.** Ms. Hill discussed the data collected by the Department of Education, including the National Indian Education Study, as well as student performance, enrollment, school climate, educational attainment, staffing and civil rights data. All data are self-reported except for Title VII formula grant data, which uses a verified reporting form. She also discussed education data needs, including longitudinal data for matriculation, retention and graduation, and a change in

the reporting policy for ethnicity, which currently does not count students as AI/AN if they are also Hispanic or racially mixed.

- **Heidi Frechette, Deputy Assistant Secretary for Indian Affairs, Department of Housing and Urban Development.** Ms. Frechette raised the importance of data in supporting HUD programs and demonstrating their effectiveness. She explained the ongoing Native American housing needs assessment, which will include urban Indian populations, barriers to mortgage lending and native Hawaiian populations. She also raised a concern about block grants, noting that they provide more flexibility but less data about individual families and program administration. She argued that a central repository of administrative data would be beneficial.

- **Greg Bischak, Financial Strategies and Research Manager, Community Development Financial Institutions Fund, Department of the Treasury.** Dr. Bischak explained the data his office collects and publishes on native community development financial institutions (CDFIs), including both institution-level and transaction-level data and a variety of research reports. He identified a need for data from tribal governments and enterprises, especially regarding their capital credit needs, as well as information about the economic base of AI/AN regions and their institutions. He also supported more work using longitudinal data on businesses.

Discussion. Following the individual presentations, the floor was opened for 30 minutes of discussion. One notable area of focus was the issue of cross-agency collaboration and coordination, including data sharing. Dr. Villegas noted that AI/AN people interact with many different programs and systems -- such as healthcare, criminal justice and education -- all at once, not in isolation; she suggested cross-agency federal standards for AI/AN data collection. Dr. Velkoff noted that Census was collaborating with the National Center for Health Statistics to

improve death records reporting, and sharing administrative records with several other agencies.

Another much-discussed topic was AI/AN population statistics. One participant noted that population estimates for reservations or tribal areas were difficult to calculate accurately because the ACS uses county-level sampling frames. A participant from the Department of Justice noted that crime and arrest rates require a population-based denominator, and asked which measures were best for population estimates. Dr. Velkoff called ACS data the most accurate, but encouraged the use of tribe-specific data, where available. Another participant noted that many administrators focus on "service area" or "user" population, which are different from area resident population estimates.

Participants touched on the issue of business and financial data as well. One asked whether data existed to measure the impacts and outcomes of native CDFIs; Dr. Bischak responded that there were fewer transaction data, but that the existing data suggest native CDFIs were doing pretty well and performance data are gradually improving.

There were several questions related to geospatial data. One participant asked whether Survey of Business Owners data were attributable only by the owner's race or also by location on tribal lands. Census staff responded that it was primarily by race, but Dr. Richard Todd of the Federal Reserve Bank of Minneapolis added that he and Dr. Randall Akee had a plan to allow linkage to reservations. Dr. Todd also noted that it was extremely difficult to match reservation boundaries with the names of recognized tribes because the former are provided by Census while the latter are managed separately by BIA. A BIA staffer discussed BIA's ongoing work with Census to improve geospatial data, expressing hope that it will produce interactive maps that link reservations to tribes, but explaining that the records were complex.

The issue of accountability and performance was briefly raised; one participant noted that while there is demand for performance outcomes on federal programs, the structure of the Self-Governance program means that there are no data on spending for each individual

program. Dr. Villegas asked for more information from Ms. Frechette about family-level data; Ms. Frechette explained that voucher-based programs at HUD know how many families receive those vouchers and can show the impact of money spent (or not spent) on the program. She opined that similar accountability could help other programs.

AFTERNOON SESSIONS

Presentation by Richard Todd, Vice President and Advisor to the Center for Indian Country Development, Federal Reserve Bank of Minneapolis

Dr. Todd began by discussing a paper he wrote in 2012, in which he noted that Indian country research relies heavily on ACS and Census data, but major data gaps exist for business, land ownership (including trust land), government operations, and other areas. He emphasized that some AI/AN data are based on location in Indian country (i.e. "place") while others are based on the AI/AN identification of an individual (i.e. "race"). His research focuses on the former.

Dr. Todd described his ongoing research, working with Dr. Akee as well as Elton Mykerezi and Orges Ormanidhi at the University of Minnesota. Their research focuses on geocoding business microdata from Census, allowing businesses on reservations to be identified and their data to be analyzed. Dr. Todd presented initial results on business density, showing that there are significantly fewer business establishments on reservations than in nearby off-reservation county areas, especially for reservations with populations below 15,000. The average deficit is about 40%, and the difference is fairly uniform across various business sectors. Results also show a similar deficit for employment, although not as big. However, Dr. Todd noted that there is a great deal of uncertainty around the results for large reservations.

Dr. Todd also noted that other factors correlate with these patterns and may partially explain them, including population density, ruralness indicators, per capita income, poverty, and

educational attainment. However, there is not yet a clear model of causality. Future work will focus on the birth, death and growth of firms, and will incorporate additional data sources.

Dr. Todd closed with several overarching points. He emphasized the importance of leadership to encourage collaboration, which made his work (and the current workshop) possible. He noted the importance of establishing clear data ownership, so researchers know whom to contact for data. He suggested more transparency on federal funding for tribes, while acknowledging there would be tribal concerns to address. Last, he noted that the overlay of reservation and other geographic boundaries could create privacy issues in some cases.

Presentation by Randall Akee, Assistant Professor, University of California Los Angeles

Dr. Akee discussed the obstacles faced by researchers attempting to collect data on AI/AN and other indigenous people. He noted that indigenous people often make a up a very small percent of survey observations, making it hard to measure their well-being. While the annual ACS is an improvement over the decennial census long form, only the five-year ACS dataset has a large enough sample to capture small populations, so there is still no annual data. He also argued that Census or survey questions are not always culturally appropriate; for example, measures of educational attainment often ignore traditional learning practices, and income could sometimes be better measured holistically rather than solely in dollars.

Dr. Akee presented several strategies for overcoming those obstacles. Using data on bank locations and credit scores as examples, he demonstrated that data from public sources can be combined with data that is purchased from private firms. He provided examples of oversampling AI/AN populations, including data on education and asset distribution. He discussed using data linkage to disaggregate AI/AN data, demonstrating this approach with a study of race and income mobility using linked IRS and U.S. Census data. Dr. Akee concluded by encouraging participants to stay in contact with the Native Nations Institute at the University of Arizona, with which he is involved.

Presentation by Brooks Depro, Program Manager and Senior Economist, RTI

International

Dr. Depro discussed a variety of ways to analyze and interpret five-year ACS datasets. He noted that ACS data can be used for geospatial analysis as long as the relevant spatial boundaries roughly correspond to census block boundaries, and that this could be done relatively cost-effectively. He echoed Dr. Todd's call for reservation boundary data that is clearly tied to tribes. He also explained that boundary analysis could be combined with GIS or other empirical measures such as driving time, and demonstrated the use of driving time to expand analysis beyond reservation boundaries.

In addition to geospatial analysis, Dr. Depro discussed the analysis of ACS data on employment and labor force. He noted that the standard definition of unemployment may not be sufficient or understandable in every case; RTI has expanded upon it by using Census data to calculate crosstabs on unemployed youth and "disconnected youth". These measures could then be analyzed for geospatial or statistical correlation with various factors. (This study did not focus on AI/AN youth specifically, but the approach may be useful for future studies of AI/AN areas and populations.)

Dr. Depro identified three possible approaches for improving AI/AN labor force analysis. The first was to add more ACS variables, however, he cautioned that more data could make the results harder to communicate. The second was a "transfer approach," applying external unemployment metrics from another source, such as the "U-5" measure from the Bureau of Labor Statistics. The last approach would be to conduct new sample surveys on a "pilot" or "case study" basis from one or two willing tribes, then evaluate the value of additional survey work.

Facilitated Open Discussion

The final session of the workshop was a structured open discussion, led by a facilitator, intended to summarize and synthesize the points made throughout the day.

The facilitator opened by asking participants to identify common themes they heard over the course of the day. Participants noted recurring mentions of data infrastructure and knowledge transfer, particularly by tribes, while noting that tribal resources to manage data could be limited. They also raised the issue of "siloing" or isolating data among separate agencies, and several possibilities for addressing that issue: data linkage to disaggregate and analyze disparate datasets, and a catalog or map to identify available data held by a range of disparate stakeholders. Participants mentioned the issues of data ownership, trust and confidentiality, cultural competency and cultural relevance. The issue of communication was raised in several contexts, including the importance of setting clear expectations for the purpose of each data collection. The importance of geodata and defining "service area" for various purposes was raised as well, along with the need for accurate population data, representation of small subgroups, and data at the level of the individual or the family. Several participants discussed the tension between the goals of decentralized tribal autonomy and self-governance and the need for standardized, easily accessible data.

Next, the facilitator moved on to a discussion of data needs, data gaps and priorities for improvement. Several participants called for expanding or enhancing the inventory of federal AI/AN recently completed by DOI, suggesting the addition of non-federal data and indicators of data quality. There was an active discussion about potentially categorizing data according to whether it was needed for public decisionmaking or private activity, the latter including land and realty data. One participant cited a need for more regional research and data capacity, as well as more clarity and transparency in the definition of "Indian" used in various datasets. Several said that better geodata was necessary, including consistent reservation boundaries and clearer designations for reservations, and one other mentioned a need for better software.

Data Ownership and Responsibility. Speakers discussed the importance of clarifying who "owns" data about AI/AN people and institutions. For example, while a federal agency might collect data about AI/AN people for the purposes of administering a program, many other parties could argue that they have a stake in how that data is used, including the people being subject to the collection, the tribes they belong to, and the general public. It is not always clear who decides how data may be used, and who is responsible for spending time and money to maintain it -- and various stakeholders may disagree. Establishing clear lines of responsibility could help mitigate these conflicts. Dr. Todd noted that establishing clear ownership of data is also important to researchers, as it provides them with a clear point of contact in case they seek to use it for analysis.

Trust and Expectations. Tribes are sometimes hesitant to share their internal data or to allow new data to be collected, and not without reason: Ms. Diver noted that statistics collected in Indian country have frequently been used to hurt tribes and tribal people; tribal leaders wonder why certain data are needed and whether those data may come back later to hurt them. Working with tribes on data issues will require building relationships and trust over time.

Several participants also discussed the related issue of *expectations* for data use, noting that data are sometimes collected for a particular purpose and later needed or used for different purposes. Ms. Bledsoe Downes noted that data are usually collected because there is a specific statutory purpose, but federal programs are constantly seeking data for different purposes. Dr. Villegas raised the importance of talking about data use as well as accuracy, and suggested that it would be important to have discussions with tribal leaders, planners and policymakers about how data could and should be used.

Cultural Competency and Appropriateness. Several speakers discussed the importance of ensuring that data collections respect the experiences and cultural norms of AI/AN people, particularly when selecting measures and interpreting their meaning. For

example, Karen Diver recounted an anecdote from her time as tribal chairwoman in which a state-sponsored data tool intended to measure homelessness was structured in a way that presumed homeless people would be sleeping outdoors or in shelters. She noted that, among Native Americans, people without their own homes often stay with family members, sometimes rotating from relative to relative, but these arrangements that would not count as "homelessness" under the data tool. Randall Akee also noted that some native peoples define educational status differently than society at large, as apprenticeships and other traditional teaching methods do not produce diplomas even though they can involve years of intense training. The importance of culturally appropriate measures was explicitly identified as a common theme of the day's presentations during the facilitated open discussion.

  Building Tribal Capacity and Tribal-Level Work. Many participants noted the difficulty of obtaining data about tribes and other AI/AN subpopulations, particularly because of their small numbers and insufficient sampling. Dr. Villegas repeatedly stressed the value of tribal-level enumeration -- working closely with tribes (either through cooperation or delegation) to obtain data about them and their members. Ms. Bledsoe Downes also suggested increasing the tribes' role in data collection.

  While no participants opposed the idea of working more closely with tribes and increasing their role, many noted the potential challenges involved. As most tribal governments are relatively small, some participants suggested that they might need help building the institutional capacity to collect and manage their own data. Funding support, technology transfer and regional partnerships were all discussed as ways to help. Participants also noted potential challenges concerning accountability and standardization. For example, data collected by different tribes using different methods might be hard to aggregate, and while self-governance and block-grant funding mechanisms provide tribes with flexibility and autonomy, they can also increase the difficulty of measuring the impacts of those programs.

Data Access and Siloing. A frequent concern, especially among participants from federal agencies, was the question of how widely accessible existing data is and should be. While researchers, federal agencies, and tribes alike need data, individuals may have reasonable concerns about their privacy, and tribes often express concerns that if they share internal data, it may be used against them later.

Related to this concern is the issue of "siloing," or separation between data sources from different government agencies. Federal agencies, participants noted, tend to manage separate databases to fulfill specific purposes or meet the needs of individual programs, but the utility of these data can be limited by their isolation. For example, Richard Todd noted that it was extremely difficult to perform the relatively simple task of matching specific reservations to specific tribes, as BIA primarily handles tribal recognition and relations while Census provides the GIS shapefiles of reservation boundaries and other AI/AN areas.

Participants recognized the potential benefits of improving access to data. Several federal participants expressed interest in sharing data with others. Dr. Akee, in particular, highlighted the benefits of data linkage, a technique that combines records from disparate sources to fill in gaps and improve overall quality. Dr. Akee argued that broader use of data linkage could save money and improve administration, and possibly provide feedback that would prevent future problems like those that led to the *Cobell* lawsuit.

Privacy. Given that data collected from individuals and businesses can be very sensitive or personal, participants addressed the issue of privacy on several occasions. For example, Dr. Todd noted that working with small sample populations and small geographic areas increased the risk that records could be linked to specific individuals, even if the names and other sensitive information are stripped off. Sharing data between agencies could be made more difficult by the need to protect the privacy of the people who are the subjects of those data. This is particularly true for program data, which can include case files and sensitive personal information.

V.   *CRITICAL DATA GAPS*

Over the course of the day's presentations and discussions, participants highlighted many gaps in the quality and quantity of available data, relevant to a wide range of areas in the lives of AI/AN people. This section highlights several important gaps that were raised repeatedly during the workshop.

Business and economic data. Several presenters discussed AI/AN business data at length, and discussed a variety of ways that the AI/AN business sector (especially on reservations) is less well measured than the business sector in general. Dr. Todd cited a recent report by the Native Nations Institute[10] saying that it remains difficult to obtain aggregate data on the native business sector as a whole, and Dr. Bischak suggested more analysis of the economic base in Indian Country. Participants also said more data is needed on finance, consumer credit, business credit and lending, tribal bonds, and real estate finance. Also, while government data can identify businesses by both their location and the owner's race, private-sector data, by law, does not identify race -- a key limitation that affects off-reservation businesses as well as those on-reservation. Finally, several participants called for more longitudinal data and more data on the "life cycle" of businesses, to better understand the economic landscape in Indian country.

Indian land ownership and realty data. Presenters noted that more and better data were needed about land and land ownership, particularly for Indian trust lands. In most of the U.S., county title offices provide relatively reliable access to land title and ownership information, but presenters noted that this is not always the case within Indian Country. Land title and ownership records on trust lands are maintained by the BIA through the Trust Asset and Accounting Management System, or TAAMS. Ms. Bledsoe Downes suggested that TAAMS was in need of modernization, and that better survey data, ownership data and legal descriptions were needed to improve land title records. Dr. Todd argued that the bureaucratic process required to conduct a land transaction is currently burdensome, negatively affecting economic

development. Others noted that land records were one of many kinds of public data needed to support private activity, such as economic development.

Tribal government data. Several participants noted that tribal government data can be difficult to obtain. For example, Dr. Bischak noted that there was relatively little information available about the capital credit needs of tribal governments, particularly for infrastructure such as roads, schools, and energy. Dr. Todd said his research would be helped by better data sets about the administration of tribal government, including tribal laws and constitutions, regulations, courts, and budgets. There was active discussion about the willingness of tribal governments to share their financial data and other internal documents; many are uncomfortable sharing it due to concerns over how it may be used. Notably, tribal governments are the only type of government not covered by the Census of Governments conducted by the U.S. Census Bureau, further reducing the availability of administrative data.

Population and subpopulation data. While the decennial census and ACS measure the population of self-identified AI/AN people in the United States, several participants mentioned a need for more and better population data, especially on smaller AI/AN groups and geographic areas, such as individual tribes. For example, the ACS uses county-level sampling frames, complicating efforts to estimate the populations of small reservations. Dr. Velkoff noted that the Census Bureau is working to provide more reliable data at smaller levels of geography, and conducting canvasses and oversampling in certain AI/AN areas, including Alaska Native villages. She also acknowledged that many other federal and state surveys do not have large enough sample sizes to disaggregate results by tribe or subpopulation. Ms. Bledsoe Downes also called for larger and more frequent sampling of native populations, and Ms. Hill noted that the utility of the National Indian Education Study is limited by small sample sizes as well.

Education data. Several participants noted data gaps in the area of education, including longitudinal data on educational outcomes and operating cost data for schools. Also mentioned was the Department of Education reporting policy for student ethnicity, which reports Hispanic

and mixed-race students *solely* as Hispanic or mixed-race, respectively, regardless of their AI/AN or other ancestry.

VI. NEXT STEPS

By raising key issues and identifying critical data gaps, this workshop demonstrated that there is still a great deal of work to be done to improve the quality, quantity and availability of AI/AN data. The following options are presented as possible short-term "next steps" for addressing the issues raised during the workshop.

Wider access to AI/AN data inventory. DOI's recent inventory of federal AI/AN data collections is a first step towards identifying and assessing the range of AI/AN-related data across the federal government, encouraging collaboration and allowing gaps to be more easily identified and addressed. This inventory could be made more widely accessible to other agencies, non-federal stakeholders and the general public. DOI could post the inventory on a public website and notify potentially interested parties through direct outreach. This would increase awareness of available datasets and allow outside parties to play a larger role in addressing AI/AN data gaps.

Review of opportunities for linkage. Data linkage (or record linkage) is a process in which records from one source are cross-matched with records from another source, with the goal of filling in gaps, fixing errors, or disaggregating it. Several participants discussed the possibility of using linkage for AI/AN records. There are examples of this approach being used successfully to improve AI/AN health and mortality data.[11, 12] Stakeholders could begin working to identify AI/AN datasets that might be enhanced by data linkage, relying in part on the recently completed inventory of federal AI/AN datasets. These efforts could assess the potential benefits of each possible linkage as well as potential challenges, including legal obstacles, privacy issues, and concerns from tribes or the general public.

Standards or best practices for AI/AN data. Several participants, particularly Ms. Diver, noted the importance of cultural competency and sensitivity to ensure that data collections are accurate, respectful, meet the needs of AI/AN people, and do not breach their trust. In addition, remarks by Dr. Velkoff and others touched on the methodological challenges of counting small populations, such as the smaller tribes. Other participants noted that tribes can be concerned about how their own data, or data collected about their members, will be used, and express concerns that it may hurt them. To address all these concerns, stakeholders could initiate discussions on developing a set of standards or best practices for AI/AN data collection and management. This would be compatible with Dr. Villegas's suggestion to create cross-agency federal standards for data collection, but depending on the needs identified and the resources available, the document could be designed to guide non-federal parties as well.

Deeper research on identified data gaps. The workshop identified several important AI/AN data gaps, but additional discussion is needed before more detailed plans can be developed to address them. The "Critical Data Gaps" identified in the previous section could be the subject of more detailed background research and analysis, including a review of available literature on their impacts, causes, and proposals to address them. This research could be used to develop papers and other materials for discussion at future data workshops (discussed below).

Additional, in-depth workshops on specific data topics. While this initial workshop addressed a broad range of topics, many of the issues raised suggest the need for additional evaluation and discussion. Additional workshops focused on the various topics and issues could include:

- **Business and economic development.** The business sector in Indian country is generally not well understood and is not measured well. A focused discussion on economic issues could help identify specific data needs and propose tangible,

achievable ways to meet them. Better and more complete information could facilitate economic development.

- **Trust lands and realty.** Tribal and individual trust lands provide an important basis for the sovereignty and economic development of AI/AN people, but ownership and similar information on these lands can be difficult to verify or obtain, especially relative to non-trust lands. Readily accessible, reliable, and consistent land ownership is an underlying prerequisite for facilitating economic development. Improvements to TAAMS and better data on trust lands could be key foundations for economic activities, such as leasing. A workshop on trust land data could focus on identifying specific strategies to improve its quality and accessibility.

- **Collaboration and data sharing.** Several participants discussed the problem of "siloing," in which individual agencies maintain separate records for their separate purposes. Participants also encouraged federal agencies to do more work "at the tribal level," meaning in close collaboration with individual tribes, while still others noted the difficulty of accessing tribal government data. A workshop focused on collaboration and data sharing among federal agencies, among tribes, and between agencies and tribes could address this issue. The workshop would focus on identifying needs that might be addressed by collaborating and sharing data, then discussing the opportunities and obstacles to doing so.

- **Improving Population Estimates.** Participants extensively discussed the need for accurate population counts and estimates, particularly of individual tribes and other small subpopulations. Such statistics are used in many federal formula grants and have a wide variety of other policy implications. A workshop on

population data could address this issue in greater detail, allowing Census, DOI and other stakeholders to begin identifying specific solutions.

- **Service Area Boundaries.** Tribal and federal programs serve many AI/AN people who do not live within the boundaries of a reservation, and the geographic area in which services are provided is generally called the "service area". Several participants referred to the fact that service areas can be hard to define and vary by program and tribe; this ambiguity causes difficulty for federal and tribal agencies in estimating the populations they must serve. A workshop on service area boundaries could help federal agencies understand the scope and impact of this issue, and begin addressing it through coordination and dialogue.

VII.   ENDNOTES

[1] Jordan, J. and Beaghen, M. 2013. Analysis of the American Community Survey and Puerto Rico Community Survey Coverage. p. 34 et seq. U.S. Census Bureau. Published January 16, 2013; revised April 22, 2014.

[2] Deweaver, N. 2013. "American Community Survey Data On the American Indian/Alaska Native Population: A Look Behind the Numbers". URL: http://www.ncai.org/policy-research-center/initiatives/ACS_data_on_the_AIAN_Population_paper_by_Norm_DeWeaver.pdf.

[3] Stroupe, K T; Tarlov, E; Zhang, Q; Haywood, T; Owens, A; Hynes, D M. 2010. "Use of Medicare and DOD data for improving VA race data quality." Journal of Rehabilitation Research and Development, vol. 47, no. 8, pp. 781 et seq.

[4] Rosenberg, H M; Maurer, J D; Sorlie, P D; Johnson, N J; MacDorman, M F; Hoyert, D L; Spitler, J F; Scott, C. 1999. "Quality of Death Rates by Race and Hispanic Origin: A Summary of Current Research, 1999." Vital and Health Statistics, series 2, no. 128, pp. 1-13.

[5] National Congress of American Indians (2). Undated. "Federal Data Collection in American Indian/Alaska Native Communities." URL: https://www.fws.gov/nativeamerican/pdf/tek-federal-data.pdf.

[6] Administrative Office of the Courts, Judicial Council of California. 2010. "Native American Communities Justice Project — Beginning the Dialogue: Domestic Violence, Sexual Assault, Stalking, & Teen-Dating Violence." Published May 2010. Retrieved from http://www.courts.ca.gov/8117.htm.

[7] Moy, E and Johansson, P. 2006. "Gaps in Data for American Indians and Alaska Natives in the National Healthcare Disparities Report." Am Indian Alsk Native Ment Health Res. vol. 13, no. 1, pp. 52-69.

[8] U.S. Department of the Interior. 2016. "FY 2017 Federal Funding for Programs Serving Tribes and Native American Communities". URL: https://www.doi.gov/sites/doi.gov/files/uploads/FY2017NativeAmericanCrosscut.pdf.

[9] Government Accountability Office. 2009. "Formula Grants: Funding for the Largest Federal Assistance Programs Is Based on Census-Related Data and Other Factors". Report no. GAO-10-263, Published December 15, 2009. URL: http://gao.gov/products/GAO-10-263.

[10] Native Nations Institute. 2016. "Access to Capital and Credit in Native Communities." URL: http://nni.arizona.edu/publications-resources/publications/papers/2016/access-capital-and-credit-native-communities.

[11] Johnson, J C; Soliman, A S; Tadgerson, D; Copeland, G E; Seefield, D A; Pingatore, N L; Haverkate, R; Banerjee, M; Roubidoux, M A. 2009. "Tribal Linkage and Race Data Quality for American Indians in a State Cancer Registry". Am J Prev Med. June 2009, vol. 36 no. 6, pp. 549–554.

[12] Stehr-Green, P; Bettles, J; Robertson, L D. 2002. "Effect of Racial/Ethnic Misclassification of American Indians and Alaskan Natives on Washington State Death Certificates, 1989–1997." American Journal of Public Health. March 2002, vol. 92, no. 3, 443-444.