ACF Data Governance Consulting and Support Project

Formative Data Collections for Program Support

0970 – 0531

Supporting Statement

Part B

July 2021

Submitted By:
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

4th Floor, Mary E. Switzer Building
330 C Street, SW
Washington, D.C. 20201

Project Officers:
Brett Brown
Valeria Butler

Part B

B1.   Objectives

Study Objectives

The objectives of this study are 1) to examine the current state of the field of technical assistance (TA) supporting states, localities, and Tribes interested in the development and use of administrative data linking and integrated data systems (IDS) for research purposes by conducting an environmental scan, and 2) to identify areas of strength in current TA efforts for data linkage and integration and areas of opportunity and unmet need by conducting a needs assessment.

Generalizability of Results

This study is intended to present internally-valid description of available TA supports for administrative data linking and IDS for research purposes as well as areas of unmet need in chosen sites, not to promote statistical generalization to other sites or service populations.

Appropriateness of Study Design and Methods for Planned Uses

As noted in Part A2, this information collection effort is to contribute to the body of knowledge on ACF programs. It is also a way to assess the TA needs of agencies and organizations related to their data literacy and data skills. Conducting interviews with stakeholders is critical because this information is not already available in published literature or existing datasets. In addition, the research questions are exploratory and warrant the collection of qualitative data to better understand the existing context. The resulting information from this collection will allow the study team to provide options to guide the Administration for Children and Families Office of Planning, Research and Evaluation (ACF/OPRE) and the field on the needed TA to enhance and strengthen the skills of these agencies and organizations in data linkage and integration efforts. Also, refer to Part A2 for details on the study design and the process for the environmental scan and needs assessment. The table in that section presents the data collection activities, instruments, respondents, content, mode, and duration for each of the four instruments to be used.

The SRI study team will use the findings from the environmental scan and needs assessment to identify gaps between what supports are available and needed and will then produce a set of options for how the ACF/OPRE Division of Data and Improvement (DDI) can strengthen the existing TA system in ways that lead to more and better integrated data systems to support research. Assessing the findings will meet the goal of finding out what type of information and training staff need to better support their work. It will also fulfill HHS' requirement to assess the current staff data literacy and data skills as per the Federal Data Strategy 2020 Action Plan.

A limitation of the study design is that the results are not designed to be representative of or generalizable to a given population, and this key limitation will be included in written products associated with this study. As noted in Supporting Statement A, this information is not intended to be used as the principal basis for public policy decisions and is not expected to meet the threshold of influential or highly influential scientific information.

**Alternative Supporting Statement for Information Collections Designed for
Research, Public Health Surveillance, and Program Evaluation Purposes**

B2.   Methods and Design

Target Population

For the environmental scan, the study team will collect information from 12 TA center directors who are responsible for overseeing the provision of TA related to data linkage and IDS for research purposes, and a sample of 5 TA funders [e.g., Contracting Officer's Representative (COR) or Federal Program Officer (FPO) for a federally funded TA center, or a program officer for a philanthropically funded TA center]. For the needs assessment, SRI will collect information from a sample of 7 TA leads/specialists who provide direct support to states, localities and Tribes, and 8 TA recipients who use the supports and services from the TA centers identified in the environmental scan. The team will use non-probability purposive sampling to identify potential respondents who can provide information on the study's key constructs. Because participants will be purposively selected, they will not be representative of the population of TA center directors, TA leads/specialists, TA recipients, or TA funders who oversee, provide, use, or fund supports related to data linkage and IDS for research purposes, respectively.

Respondent Recruitment

The study team will interview **12 TA center directors**, that is, the person with the most responsibility for overseeing the provision of TA related to data linkage and integration.

- To select TA center directors, we will first identify appropriate TA centers for this study by conducting a web search. We are defining a "TA center" as a center, program, organization, or university partner that meets the following criteria:
    - Is federally or non-federally funded
    - Includes a focus on data linkage or data integration for research or statistical purposes
    - Provides TA to states, localities, or Tribes
        - May provide TA to various entities, a specific audience (e.g., Head Start grantees, state education agencies), or to larger communities or consortia
    - Offers TA services related or connected to data integration or data linking for research and statistical purposes
        - TA services may include individual TA, group TA (e.g., webinars, workshops, communities of practice), and/or products (e.g., toolkits, practice briefs)
- Factors that we will take into consideration when choosing a sample of TA center directors include variation in the types of TA centers they oversee by:
    - Target TA recipients
    - Sector of human services (e.g., early childhood, education, housing)
    - Funding source (e.g., federal, philanthropic)
    - Type of TA provided (e.g., intensive, targeted, universal)
    - Types of published resources (e.g., toolkits, practitioner brief)

We will select a purposive sample of **5 TA funders** that includes variability by the following factors:
- Funding source (e.g., federal, philanthropic)
- Sector of human services (e.g., early childhood, education, housing)
- Years of experience funding data linkage and integration TA

We will also identify alternates for each role recruited. For example, for a philanthropic funder, we would have a primary and a secondary philanthropic partner option to ensure we maintain variation across participants. The secondary participant will be used as a back-up in case the primary participant declines or does not respond to our invitation to participate in an interview.

We will ask TA center directors and TA funders to recommend the TA leads/specialists we should interview. Factors that we will take into consideration when selecting a purposive sample of **8 TA leads/specialists** include variation in:
- Target TA recipients
- Sector of human services (e.g., early childhood, education, health, housing)
- Funding source (e.g., federal, philanthropic)
- Type of TA provided (e.g., intensive, targeted, universal)
- Types of published resources (e.g., toolkits, practitioner brief) they support TA recipients to use

We will ask TA center directors to recommend individuals who use their TA services to participate in an interview to learn more about their needs. The study team will select a purposive sample of **8 TA recipients** that includes variability by the following factors:
- Whether the TA recipient is a representative from a state, locality, or Tribe
- The type of TA center(s) (i.e., federally funded, philanthropically funded) from which the TA recipient has received support
- Sector of human services (e.g., early childhood, education, housing, employment)
- Level of experience with linking and integrating data (e.g., novice, experienced)
- Type of TA they have received and used (e.g., universal, targeted, intensive)

TA recipients will also need to meet the following criteria:
- Individual has received TA related to data linkage or IDS for research purposes
- Individual has worked directly with TA providers
- Individual has first-hand experience using TA resources (e.g., toolkits, training modules)

B3.   Design of Data Collection Instruments

Development of Data Collection Instruments

The study team used a multistep process to develop interview protocols that maximize relevance, accuracy, and completeness of the information collected while minimizing respondent burden. The research questions provided the structure for the interview protocol, and the literature on key practices for effective TA (Trohanis TA Projects at Frank Porter Graham Child Development Institute, 2020) and frameworks for data linkage and integration (Actionable Intelligence for Social Policy, 2020; Center for the Integration of IDEA Data, 2018; Coffey et al., 2014) informed the constructs and wording of interview probes.

SRI researchers developed four separate semi-structured interview protocols for each type of respondent (i.e., TA center directors, TA funders, TA leads/specialist, and TA recipients) designed to address the research questions (see Instruments 1-4). To ensure alignment with the research questions, we developed a cross-walk of each research question and the corresponding items on the interview protocols (see Appendix A). In advance of using the interview protocol with study participants, Julie Quaid, the project's Tribal consultant, also reviewed the interview protocol to ensure it is culturally appropriate for respondents from Tribes. This review and feedback helped to improve the clarity and comprehensiveness of the interview questions.

B4.   Collection of Data and Quality Control

The study team will refine a list of study participants (i.e., TA center directors and TA funders) and will email selected TA center directors and TA funders to invite them to participate in this study (see

Appendix B). If TA center directors and funders do not respond within 1 week, a study team member will follow-up by phone and email (if they leave a voicemail). If a TA center director or funder does not respond 2 weeks after the initial invitation, an SRI researcher will contact them by phone and/or email to invite them to the study and answer any questions. If we do not receive a response, we will use our identified secondary options created for each type of respondent. The study team will ask TA center directors and TA funders to provide recommendations about TA leads/specialists and TA recipients who meet the study inclusion criteria to participate in an interview. We will review the recommendations and select a proposed sample of TA leads/specialists and TA recipients based on the factors described in section B2 and will finalize this list in collaboration with DDI CORs.

Contracted researchers from SRI will participate in an environmental scan protocol and/or needs assessment protocol training before beginning any interviews. The training will include an overview of the environmental scan or needs assessment objectives; documents to review in advance of the interview (e.g., information from the environmental scan about a specific TA center that the TA recipient has used); and data collection procedures such as obtaining consent, audio recording, submission of audio files for transcription, a careful review and discussion of the semi-structured interview protocol questions and probes; the post-interview internal debriefing process; and a brief overview of the analysis procedures. Researchers will also learn the procedures for tailoring interview questions to a specific respondent based on the information the team has collection from the environmental scan resource/document review.

Two trained data collectors will conduct the semi-structured one-on-one interviews online using the Zoom for Government platform. Each interview will be no longer than 60 minutes, and interviews will be audio recorded for the purpose of transcription with the participant's permission. A primary interviewer will lead the conversation and a secondary interviewer will provide technical support (e.g., ensuring the interview is recorded), serve as a timekeeper and active listener, and assist with follow-up questions and probes, as needed Julie Quaid, the project's Tribal consultant, will lead the interviews of any selected interviewees from Tribes to help ensure the study is culturally and linguistically responsive to the needs of the Tribes. After each interview, the primary and secondary interviewers will debrief and document emerging topic areas to inform coding scheme development.

B5.   Response Rates and Potential Nonresponse Bias

Site/Respondent selection
The (interviews/focus groups/case studies) are not designed to produce statistically generalizable findings and participation is wholly at the respondent's discretion. Response rates will not be calculated or reported.

NonResponse
As participants will not be randomly sampled and findings are not intended to be representative, non-response bias will not be calculated. We will qualitatively assess non-response by documenting the number of interview refusals.

B6.   Production of Estimates and Projections

The data will not be used to generate population estimates, either for internal use or dissemination.

B7. Data Handling and Analysis

Data Handling

All interviews will be audio recorded with participant's permission and transcribed to minimize errors with respect to documenting the information study participants provide during interviews.

Data Analysis

A team of two researchers will engage in first cycle coding to "initially summarize segments of data" and then pattern coding, as a second cycle method, for "grouping those summaries into a smaller number of categories, themes, or concepts" (Miles et al., 2020, p. 79). For first cycle coding, the team will use provisional coding, which uses a 'start' list of a priori researcher-generated codes based on what the team anticipates might appear in the data based on the research questions asked and our understanding of the TA topics and approaches that are important for supporting data linkage and integration based on the literature and the team's TA expertise (Miles et al., 2000). Provisional codes can be revised and modified; the team will delete or expand on this list of codes after reviewing all the transcripts.

A team of two researchers who participated in the interviews will read all the interview transcripts to modify the list of provisional codes based on what topics emerge from the interview data. A senior staff member who led interviews will also provide feedback on the preliminary list of codes. The team will apply the provisional codes to two randomly selected transcripts, and then will meet to discuss whether the preliminary coding scheme is sufficient or what additional codes may be needed. The team will then revise the preliminary coding scheme to apply to the remaining transcripts.

Transcripts will be analyzed using Dedoose, a cross-platform app for analyzing qualitative data and mixed methods research. We will apply O'Connor & Joffe's (2020) suggested procedure for intercoder reliability assessment, which involves first making preliminary decisions regarding the number of coders, amount of data to code, unit of coding, reliability measure, and threshold of acceptable reliability. The team will use two coders; the unit of coding will be conceptually meaningful "chunks" of text; coders will double code 25% of the transcripts; Cohen's Kappa that is calculated via Dedoose will be the reliability measure; and Cohen's Kappa of 0.80 higher will be deemed acceptable. Once the coding scheme is established, a researcher designated as Coder 1 will segment the data into data units (conceptually meaningful "chunks" of text) and will label them with the relevant codes. Using the training center test capabilities in Dedoose, a researcher designated as Coder 2 will use the coding scheme to independently code the same transcript file with Coder 1's pre-determined segmented data units. The team will then compare reliability for each code to clarify any code misinterpretations and revise the coding scheme (if needed) before formal reliability evaluation begins. The team will then double code a randomly selected subsample of 25% of the transcripts using the process described above. Once the team achieves a Cohen's Kappa of 0.80 of higher, the remaining transcripts will be coded by one researcher.

As part of the second cycle coding process, the team will review the excerpted text for each code to develop a smaller number of themes. These themes will help to condense the codes into "more meaningful and parsimonious units of analysis." (Saldaña, 2021, p. 322). The team will develop a table of themes, frequencies, and illustrative quotes to summarize the information collected.

All team members who participated in the interviews will engage in in-depth discussions about the data collected (separate discussions for the environmental scan and needs assessment data). In advance of these meetings, team members will review the qualitative data tables and identify any questions or

comments (e.g., whether certain findings resonate or seem incongruent with what they heard during interviews). The primary aim of this discussion is to serve as a validity check to ensure that the preliminary findings align with other data collectors' perceptions of key themes from the interviews.

The interview data will be used in concert with information we collect from a website and document review of publicly available information on TA center websites to understand existing TA supports related to data linkage and integration for research purposes.

Data Use

SRI will use the environmental scan and needs assessment findings to produce a set of options for how the field in general, and DDI in particular, can strengthen the existing TA system in ways that lead to more and better integrated data systems to support research. Findings will be disseminated via an oral briefing and final detailed report to DDI. Other possible dissemination activities may include a webinar to federal program office staff and TA center directors that oversee TA focused on data linkage and integration to foster coordination and collaboration, and blog posts targeted to states, localities, and Tribes interested in resources to support their use of administrative data linking and IDS for research purposes.

B8. Contact Person(s)

Missy Coffey
Missy.coffey@sri.com

Kathi Gillaspy
kgillaspy@anlar.com

Attachments

Instrument 1: TA Center Director Interview Protocol for Environmental Scan

Instrument 2: TA Funder Interview Protocol for Environmental Scan

Instrument 3: TA Lead-Specialist Interview Protocol for Needs Assessment

Instrument 4: TA Recipient Interview Protocol for Needs Assessment

Appendix A: Research question cross-walk

Appendix B: Study recruitment emails

References

Actionable Intelligence for Social Policy (AISP). (2020). *Introduction to data sharing and integration*.
    https://www.aisp.upenn.edu/wp-content/uploads/2020/06/AISP-Intro-.pdf
Actionable Intelligence for Social Policy (AISP). (n. d.). *About data sharing*.
    https://www.aisp.upenn.edu/about-data-sharing/
Center for the Integration of IDEA Data (CIID). (2018). *CIID data integration toolkit*.
    https://ciidta.grads360.org/#communities/pdc/documents/12575

Coffey, M., Chatis, C., Sellers, J., and Taylor, R. (2014). SLDS Early Childhood Integrated Data System Guide. U.S. Department of Education. Washington, DC: National Center for Education Statistics. https://slds.ed.gov/#communities/pdc/documents/8968

Groenewald, T. (2008). Memos and memoing. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 506-506). SAGE Publications, Inc., https://www.doi.org/10.4135/9781412963909.n260

Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook*. Sage.

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 1-13. DOI: 10.1177/1609406919899220

Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage.

Trohanis TA Projects. (2020). Effective technical assistance practices. https://trohanis.fpg.unc.edu/docs/trohanis-effective-ta.pdf