Survey of Connected Data Efforts

Draft – Do Not Share

April 2022

INTRODUCTION

This survey is the second data collection activity of a study commissioned by the U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation in partnership with the Children's Bureau. The goal of this survey is to learn how and to what extent child welfare agencies link, integrate, and share child maltreatment data with data from other systems. [IF Username IS FOR Puerto Rico or Washington, DC: For the purpose of this survey, when we refer to "state," we are including Puerto Rico and DC.]

In the first survey for this study, you were identified by a child welfare agency as the primary contact for [IF Username IS FROM Q19a IN THE FIRST SURVEY: non-child welfare data that is integrated into [SYSTEM]. For the purpose of this survey, we refer to integrated data as data that has been systematically incorporated into [SYSTEM], through direct entry by staff or a data exchange.] [IF Username IS FROM Q19b IN THE FIRST SURVEY: a data sharing agreement that supports a linked dataset that contains child welfare data. For the purpose of this survey, we define linked data as a set of records containing information combined from multiple data sources based on a common identifier or other matching criteria.]

In this survey, we will ask you about [IF Username IS FROM Q19a IN THE FIRST SURVEY: your integrated data] [IF Username IS FROM Q19b IN THE FIRST SURVEY: the linked datasets associated with the data sharing agreement [NAME OF DATA SHARING AGREEMENT] involving your state's child welfare agency and [AGENCIES IN DATA SHARING AGREEMENT]]. The survey will ask about the types of data that are connected, how the data are connected, how the data are used, how often the connected data are updated, who can use the connected data, and how users gain access to the data.

The survey has two sections. You may forward this to a designee(s) The survey is expected to take you and your designee(s) 35 minutes to complete. Please respond to the survey by no later than [DATE].

Participating in this survey will help us better understand states' experiences sharing and linking child welfare administrative data. Your participation is voluntary and you may choose not to answer any questions you don't want to answer. Your state may be identified in public reports of study findings based on this survey data. However, public reports will not identify you by name. The data may also be archived for restricted use by other researchers. Archived survey data may identify states but will not include the names or contact information of individuals.

If you have any questions or problems accessing the survey, please contact [NAME] at [PHONE] or [EMAIL]. We thank you in advance for your participation.

A. CONTACT INFORMATION FOR LINKED DATA FROM [DATA SHARING AGREEMENT NAME]

1. **a. Please give the name and contact information for the person who will be completing most or all of this survey. The information will allow us to follow up if we have any questions. Survey questions can be shared with additional people if you do not know the answer to a question(s).**

    Contact name: _____

    State: _____

    Agency/organization name: _____

    Position: _____

    Email: _____

    Phone number: _____

    [IF Username IS FROM Q19a IN THE FIRST SURVEY, SKIP TO Q23, QUESTION SET FOR INTEGRATED DATA]

2. **Please tell us about all linked analytic datasets associated with the [DATA SHARING AGREEMENT NAME]. A linked analytic dataset is a dataset that is used for analysis and not solely for use in providing services. For each linked analytic dataset, indicate (Y/N) whether the dataset is relevant to understanding child maltreatment incidence.**

    a. Name of linked analytic dataset: _____
    Please choose a name that can distinguish this dataset from others associated with this data sharing agreement. We will only use this name in the survey to reference this specific dataset.   This does not need to be an official name.

    b. Relevant to child maltreatment incidence?: $_1\square$ Yes $_0\square$ No
    Please use your best judgement as to whether the dataset has information about children who are (or may be) abused and neglected and/or referred to the state child welfare agency. A clear example of a time when the answer should be yes is when there are records on children in the linked dataset, including the population of children who are referred or indicated by the state child welfare agency as child maltreatment victims. Another example is when the linked dataset does not contain maltreatment information but contains hospital records with diagnosis codes indicative of child abuse.

    c. Is there another linked dataset associated with the [DATA SHARING AGREEMENT NAME]?

    $_1\square$ **Yes** $_0\square$ **No**

    [IF 2C IS YES, REPEAT 2A-2C. MOVE TO 2D WHEN 2C IS NO.]

    d. Which dataset be the focal dataset for this survey?
    Please select the dataset for which at least one of the following is true:

    -   *this is the only dataset relevant to child maltreatment incidence,*
    -   *this dataset is the one that is most utilized among those that are relevant to understanding child maltreatment incidence*
    -   *this dataset is the one that is most utilized and none of the datasets are relevant to understanding child maltreatment incidence.*

    [LIST NAMES OF DATASETS IDENTIFIED IN 2A AS RESPONSE OPTIONS; SELECT ONE ONLY]

B. DATA GOVERNANCE AND DETAILS ABOUT THE [LINKED DATASET NAME] DATASET
Data governance includes the processes and procedures guiding the collection, storage, and use of the data set. The questions in this section will give us details about the data and data governance for the [LINKED DATASET NAME] dataset.

3. **Who manages linkages for the [LINKED DATASET NAME] dataset?**

   a. Provide the name of the organization that manages the linked dataset (i.e., the organization that oversees the data linking process and administers access to the dataset):

   _____

   b. Select the type of organization:

   SELECT ONE ONLY
   1 ☐   State child welfare agency or division
   2 ☐   Another state agency (including state research organizations affiliated with a state agency)
   3 ☐   College/university
   4 ☐   Research organization (not affiliated with a state agency)
   5 ☐   Other *(specify):* _____

4. **What type of agreement governs the sharing and use of the data?**

   SELECT ALL THAT APPLY
   1 ☐   Data Use Agreement/Data Sharing Agreement/Memorandum of Understanding/Memorandum of Agreement/Letter of Agreement
   2 ☐   Business Associates Agreement
   3 ☐   Other *(specify):*_____

5. **a. Do all data sources in the [LINKED DATASET NAME] cover the same time period (i.e. the start date/year is the same for each data source?**

   1 ☐   Yes
   0 ☐   No

   b. Are all data sources updated on the same schedule, such as monthly, quarterly, or annually?

   1 ☐   Yes
   0 ☐   No

6. **What are the sources of data included in [LINKED DATASET NAME]?**
*We are interested in knowing what, if any, **source data** are part of your linked dataset. Data that a case manager may or may not collect as part of an interview protocol, but which are not the official source data, are not source data. For example, if the case management system has a field to input whether parents receive Temporary Assistance for Needy Families (TANF) services, and this is entered by the child welfare case manager who does not provide that service, then these are not source data.*

    a. **Select all that apply**
       [IF Q5.a=Yes, SKIP 6.b & 6.c]
    b. **Indicate the first year represented in the data source**
    c. **Indicate the most recent year represented in the data source**
    d. **Indicate whether there is data for all years between the first (b) and most recent (c) years**
    e. **Summarize the sample restrictions of the data. Parameters may relate to subpopulations based on age, program status, a date range, or other criteria (e.g., "All state birth records where there is a link to a child placed in foster care").**
    f. **Indicate whether data source contributes demographic information to [LINKED DATASET NAME] (e.g., race, ethnicity, gender, and age)**

	A. DATA SOURCE	B. START YEAR	C. RECURRING OR END YEAR	D. DATA COVERS RANGE (B) TO (C) OR PRESENT	E. SAMPLE RESTRICTIONS (SPECIFY)	F. PROVIDES DEMOGRAPHIC INFORMATION
1	☐ Child abuse and neglect investigations		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
2	☐ Foster care placements		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
3	☐ Child welfare case management after investigations		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
4	☐ Supplemental Nutrition Assistance Program (SNAP) records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
5	☐ Other social assistance records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
6	☐ Social Security benefits		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
7	☐ Child support systems		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
8	☐ Birth records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
9	☐ Death records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
10	☐ Medicaid eligibility records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
11	☐ Medicaid claims records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
12	☐ Intellectual or developmental disability services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
13	☐ Mental health services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
14	☐ Substance misuse or substance use services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
15	☐ Other child protective services (CPS)-contracted services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
16	☐ Education records		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
17	☐ Judicial/court services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
18	☐ Housing assistance		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
19	☐ Juvenile justice services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
20	☐ Child Care services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
21	☐ Domestic violence/Intimate partner violence services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
22	☐ Refugee assistance services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		
23	☐ Employment/labor services		☐ Recurring ☐ End year: ___	☐ yes ☐ no		

24	☐ Services defined in the state's Prevention Plan for the Family First Prevention Services Act		☐ Recurring ☐ End year: ___	☐ yes ☐ no	
25	☐ Other (specify): _____		☐ Recurring ☐ End year: ___	☐ yes ☐ no	

7. **What level are the data linked at?**

SELECT ALL THAT APPLY

1 ☐ Person/individual

2 ☐ Family

3 ☐ Household

4 ☐ Geographic location (e.g., county or region)

5 ☐ Organization

6 ☐ Other *(specify)* _____

8. **What linking method(s) are used to merge the different sources of data?**

SELECT ALL THAT APPLY

1 ☐ Deterministic or exact linking method (a common identifier)

2 ☐ Open source probabilistic linkage software (such as ChoiceMaker) *(specify)*
_____

3 ☐ Proprietary probabilistic linkage software (such as LinkageWiz) *(specify)*
_____

4 ☐ Other *(specify):* _____

[IF 8=1, ANSWER 9.A AND SKIP 9.B. ELSE SKIP 9.A.]

9. **a. What common identifiers are used to link the data?**

SELECT ALL THAT APPLY

1 ☐ Social Security number

2 ☐ State ID

3 ☐ Medicaid ID

4 ☐ Other *(specify):* _____

b. What data elements are used to link the data?

SELECT ALL THAT APPLY

1 ☐ Child/individual name

2 ☐ Name of mother

3 ☐ Name of father

4 ☐ Date of birth

5 ☐ Address

6 ☐ Sex

7 ☐ Race and/or ethnicity

10. **What population level is covered by the dataset?**

SELECT ONE ONLY

1 ☐ Statewide

2 ☐ County

3 ☐   Other sub-state level

11. **Please identify which, if any, of these population characteristics pertaining to child welfare involvement are contained in the linked analytic dataset:**

    SELECT ALL THAT APPLY

    1 ☐ Children reported for maltreatment (even if screened out)

    2 ☐ Children investigated for maltreatment

    3 ☐ Parent(s) named in child maltreatment report

    4 ☐ Siblings of children reported for maltreatment

    5 ☐ Children in foster care

    6 ☐ Children/families receiving in-home child protection case management services

    7 ☐ Children/families enrolled in child maltreatment prevention programs (other than child protection case management services)

    8 ☐ Others (specify): _____

12. **What year did [NAME OF MANAGING ORG/Q3] begin using this linked dataset?** [RANGE: 1950 – 2022]

    |__|__|__|__|

13. **How often is the information in [LINKED DATASET NAME] updated?**

    SELECT ONE ONLY

    1 ☐ Annually

    2 ☐ Semi-annually

    3 ☐ Quarterly

    4 ☐ Monthly

    5 ☐ Daily

    6 ☐ Other frequency *(specify):* _____

14. **Are there documented data governance procedures for the [LINKED DATASET NAME] dataset?**

    1 ☐ Yes

    2 ☐ No

[IF Q14=YES, ASK Q15, ELSE, SKIP TO Q19]

15. **What is the data governance structure?**

    SELECT ONE ONLY

    1 ☐ There is a joint-governance oversight team/committee/review process

    2 ☐ Data governance primarily resides with one of the agency partners

    3 ☐ Data governance primarily resides with another state organization (including research organizations affiliated with a state agency)

    4 ☐ Other *(specify):*_____

16. **Are there documented methods and procedures for linking the data?**

    1 ☐ Yes

    0 ☐ No

17.   Are data protected from unauthorized access and data corruption)?

1 ☐ Yes

0 ☐ No

18.   Are there processes for how individuals can gain authorization to access/use the data?

1 ☐ Yes

0 ☐ No

19.   What are the linked data used for?

SELECT ALL THAT APPLY

1 ☐ Practice (e.g., in casework, case management, service delivery, etc.)

2 ☐ Continuous quality improvement (CQI)

3 ☐ Program evaluation research (e.g. to build evidence)

4 ☐ State policy analysis research

5 ☐ Legislative report or mandate

6 ☐ Other research

7 ☐ Performance monitoring (government agency and/or contractor performance)

8 ☐ Predictive modeling

9 ☐ Other internal purpose(s) *(specify):* _____

10 ☐ Other external purpose(s) *(specify):* _____

20.   Who, in addition to staff from [NAME OF MANAGING ORG/Q3], actively use the linked dataset?

SELECT ALL THAT APPLY

1 ☐ (Other) State government staff

2 ☐ (Other) Researchers

3 ☐ Others *(specify):* _____

4 ☐ None (e.g., permission to use the linked data is limited to extracts from the linked dataset)

21.   Who, in addition to staff from [NAME OF MANAGING ORG/Q3], may receive permission to use the [LINKED DATASET NAME] dataset or extracts from it?

SELECT ALL THAT APPLY

1 ☐ Staff from an organization that is part of the data sharing agreement

2 ☐ State officials from state entities that are not part of the data sharing agreement

3 ☐ Researchers who are granted access to the data

4 ☐ Other *(specify):* _____

22.   Which of the following methods can be used to access the [LINKED DATASET NAME] dataset?

SELECT ALL THAT APPLY

1 ☐ Complete a paper application

2 ☐ Fill out an online (website) request

3 ☐ Direct request (email/phone) to a designated contact

4 ☐ Other *(specify):* _____

[END SURVEY – SKIP REMAINING QUESTIONS]

C. DATA GOVERNANCE AND DETAILS ABOUT DATA INTEGRATED INTO [SYSTEM]

Data governance includes the processes and procedures guiding the collection, storage, and use of the data set.

In this section, we ask questions in order to learn details about the data and data governance for your state's integrated data.

23. What type of agreement(s) governs the sharing and use of the data?

SELECT ALL THAT APPLY

1 ☐ Data Use Agreement/Data Sharing Agreement

2 ☐ Memorandum of Agreement/Memorandum of Understanding

3 ☐ Business Associates Agreement

4 ☐ Letter of Agreement

5 ☐ Other *(specify)*: _____

24. Do all data sources integrated into [SYSTEM] cover the same time period (i.e., the start date/year is the same for each data source, and all data sources are updated on the same schedule)?

1 ☐ Yes

0 ☐ No

25. What are the external sources of data that are systematically connected with [SYSTEM]?
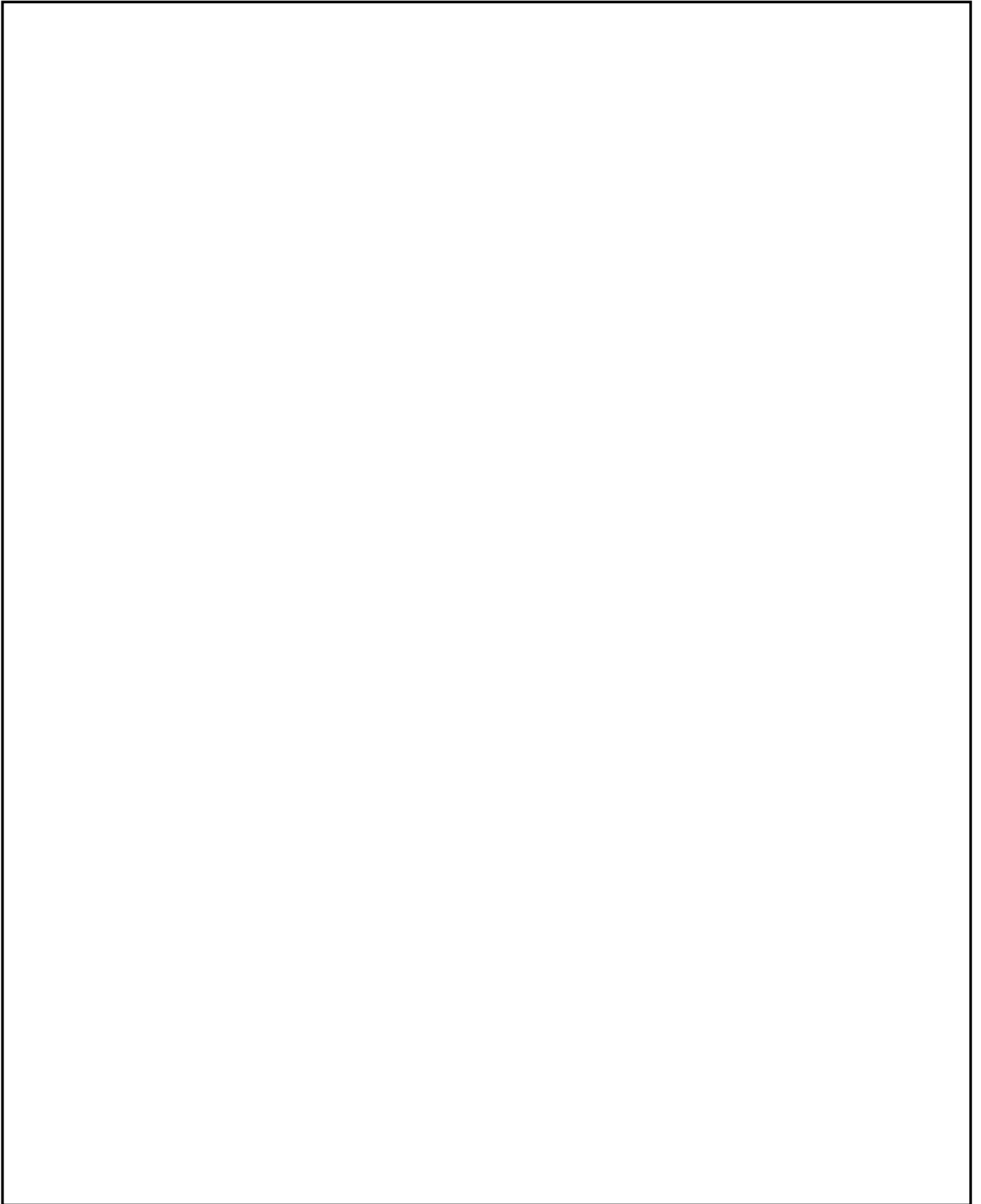
*We are interested in knowing what external **source data** are integrated into [SYSTEM], either through a data exchange, through a direct entry portal where external parties can enter data, or a data warehouse. Data that a state/county child welfare case manager may or may not collect as part of an interview protocol, but which are not the official source data, are not source data. For example, if the case management system has a field to input whether parents receive Temporary Assistance for Needy Families (TANF) services, and this is inputted by the child welfare case manager who does not provide that service, then these are not source data.*

a. **Select All That Apply**
   [IF Q25a=5 or 6, SKIP TO NEXT ROW. IF Q24=Yes, SKIP 25.b & 25.c]
b. **Indicate the first year represented in the data source**
c. **Indicate the most recent year represented in the data source**
d. **Indicate whether there are data for all years between the first (b) and most recent (c) years**
e. **Summarize briefly what are the sample restrictions of the data. Parameters may relate to subpopulations based on age, program status, a date range, or other criteria (e.g. "All state birth records where there is a link to a child placed in foster care").**
f. **Indicate whether demographic information from the data source populate records in [SYSTEM]**

	DATA SOURCE	A. ARE THE DATA FOR THIS SERVICE CONNECTED TO [SYSTEM]?	B. START YEAR	C. RECURRING OR END YEAR	D. DATA COVERS RANGE (C) TO (D) OR PRESENT	E. SAMPLE RESTRICTIONS (SPECIFY)	F. PROVIDES DEMOGRAPHIC INFORMATION
	Format/response options	1☐ Yes, [SYSTEM] is the system of record for this service 2☐ Yes, data for this service is entered directly 3☐ Yes, through a data exchange 4☐ Yes, through a data warehouse 5☐ Not now, but there is a formal plan and/or we are actively working on it 6☐ No 7☐ Other (specify): ___	*YYYY*	☐ Recurring ☐ End year:	☐ yes ☐ no		☐ yes ☐ no
1	Child abuse and neglect investigations						
2	Foster care placements						
3	Child welfare case management after investigations						
4	Supplemental Nutrition Assistance Program (SNAP) records						
5	Other social assistance records						
6	Social Security benefits						
7	Child support systems						
8	Birth records						
9	Death records						
10	Medicaid eligibility records						
11	Medicaid claims records						
12	Intellectual or developmental disability services						
13	Mental health services						
14	Substance misuse or substance use services						
15	Other Child protective services (CPS)-contracted services						
16	Education services						
17	Judicial/court services						
18	Housing assistance						
19	Juvenile justice services						
20	Child Care services						
21	Domestic violence/Intimate partner violence services						
22	Refugee assistance services						
23	Employment/labor services						
24	Services defined in the state's Prevention Plan for the Family First Prevention Services Act						
25	Other (specify): _____						

26. **What level do the integrated data connect with [SYSTEM] at to populate [SYSTEM]?**

SELECT ALL THAT APPLY

₁ ☐ Person/individual

₂ ☐ Family

₃ ☐ Household

₄ ☐ Geographic location (e.g. county or region)

₅ ☐ Organization

₆ ☐ Other *(specify):* _____

27. **What linking method(s) is used to populate [SYSTEM[with the integrated data?**

SELECT ALL THAT APPLY

₁ ☐ Deterministic or exact linking method (a common identifier)

₂ ☐ Open source probabilistic linkage software (such as ChoiceMaker) *(specify)*
_____

₃ ☐ Proprietary probabilistic linkage software (such as LinkageWiz) *(specify)*
_____

₄ ☐ Other *(specify):* _____

[IF 27=1, ANSWER 28.A AND SKIP 28.B. ELSE SKIP 28.A.]

28. **a. What common identifiers appear in the different data sources and are used to connect them to [SYSTEM]?**

SELECT ALL THAT APPLY

₁ ☐ Social Security number

₂ ☐ State ID

₃ ☐ Medicaid ID

₄ ☐ Other *(specify):* _____

b. What data elements appear in the different data sources and are used to connect them to [SYSTEM]?

SELECT ALL THAT APPLY

₁ ☐ Child/individual name

₂ ☐ Name of mother

₃ ☐ Name of father

₄ ☐ Date of birth

₅ ☐ Address

₆ ☐ Sex

₇ ☐ Race and/or ethnicity

29. **What population-level is covered by the data brought into [SYSTEM]?**

SELECT ONE ONLY

₁ ☐ Statewide

₂ ☐ County

₃ ☐ Other sub-state level

30. **How often is the integrated data integrated or pulled into [SYSTEM]?**

   SELECT ALL THAT APPLY

   1 ☐ Annually

   2 ☐ Semi-annually

   3 ☐ Quarterly

   4 ☐ Monthly

   5 ☐ Daily

   6 ☐ Real-time

   7 ☐ Other frequency *(specify)*_____

31. **Who, in addition to staff in [CW AGENCY], may receive permission to access/use the integrated data?**

   SELECT ALL THAT APPLY

   1 ☐ Staff from an organization that is part of the data sharing agreement

   2 ☐ State officials from state entities that are not part of the data sharing agreement

   3 ☐ Researchers who are not part of the data sharing agreement

   4 ☐ Others *(specify)* _____

32. **What are the integrated data used for?**

   SELECT ALL THAT APPLY

   1 ☐ Practice (e.g. in casework, case management, service delivery, etc.)

   2 ☐ Continuous quality improvement (CQI)

   3 ☐ Program evaluation research (e.g. to build evidence)

   4 ☐ State policy analysis research

   5 ☐ Legislative report or mandate

   6 ☐ Other research

   7 ☐ Performance monitoring (government agency and/or contractor performance)

   8 ☐ Predictive modeling

   9 ☐ Other internal purpose(s) *(specify):* _____

   10 ☐ Other external purpose(s) *(specify):* _____

[END SURVEY]

Thank you for taking the time to complete this survey! Your participation will help us better understand states' experiences sharing and linking child welfare administrative data. If you have any questions or problems accessing the survey, or any other concerns about the survey, please contact [NAME] at [PHONE] or [EMAIL].