Best Practices for Community-Based Organizations in Collecting and Storing Data

Community-based organizations (CBOs) often use multiple information systems to document the clients they serve and the services that they provide. While the primary purpose of these systems is to support service delivery, data from these systems are increasingly useful for many other purposes, such as reporting impact of services. This means that the data in an information system needs to be collected and stored in a way that allows it to be linked to other data sources, as appropriate, to inform decision making and enable more impactful reporting.

For organizations collecting person-level information electronically, this document provides recommendations for data collection and storage, as well as best practices for staff who are manually entering data into information systems. When best practices are not followed, the ability for data to be linked to other data sources, and subsequently the ability of the CBO to gain deeper understanding of the benefit of their services, will be limited. The following best practices are foundational and are not intended to be comprehensive. If a CBO has not implemented one of the best practices below, a plan to address the gap should be created.

1. **Collect Individual Identifiers:** Identifiers are necessary to link an individual's data across multiple sources; data without identifiers cannot be linked. At a minimum, organizations should collect first and last name, date of birth, phone number, and address. Collecting additional identifiers, such as middle name, secondary phone number, and email, will further improve linkage. If linking data between partner organizations, the organizations should communicate matching requirements and protocols.
2. **Train Staff**: Staff who collect or enter identifiers into an information system may perform these tasks differently, leading to variation in resulting data. For example, some staff may record date of birth as 'Feb 2' which is less complete than 02/02/1969. Training staff who collect and enter data according to a standard protocol can improve data quality.
3. **Store Identifiers in Distinct and Structured Fields**: When configuring an information system, the fields that hold identifiers should be structured, unique fields that use appropriate formats. For example, use separate fields for first name and last name rather than one field for name and collect date of birth in a date format rather than in characters.
4. **Use Predefined Values**: For fields like race, ethnicity, or sex, allow staff to select from predefined values (e.g. M and F for sex) rather than letting staff type in an unstructured free text field. Using predefined fields ensures that those values will be entered consistently and avoid multiple values that could have the same meaning such as Male, M, Boy, Man.
5. **Minimize Missing Identifiers:** Missing data is common and missing identifiers in particular can make linkage challenging or of lower quality. One way to minimize missing identifiers is to make this information required so that the information system will not save the record without data in those fields. The minimal set of identifiers required to link data should be required, with additional identifiers as optional; this approach balances data quality with data entry workflows.
6. **Maintain a Single Master Record and Identifier:** For organizations using multiple systems, establishing one master record that maintains the most complete and current identifiers for that person means that identifiers will not vary or conflict between systems. Most information systems assign a unique identifier (ID) to each person, and whenever possible, that ID should be reused across systems. Deconflicting IDs for the same person between systems can be laborious, complicated, and make determining the unique number of people served difficult.

This fact sheet was produced for the U.S. Government under Contract Number 75FCMC18D0047. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, ACL/HHS, or the U.S. Government.