

3. The HMIS Approach

In response to Congress' directive, HUD elected not to use the traditional point-in-time approach, but opted instead to develop and introduce national data and technical standards for locally situated computer systems that collect, process and share details of each client's utilization of service related to homelessness. These are termed Homeless Management Information Systems ("HMIS"), which are described in terms of the parties to and from which data flows and the data elements that constitute information flow. At this writing, the initial data elements had already been altered to better protect the privacy of domestic violence shelter clients from intimate abusers, but other privacy concerns remain which are addressed herein.

3.1 Data flow in HMIS

Using a HMIS, information does not flow directly from a homeless service provider to HUD. Instead, a HMIS introduces an intermediary (termed a "planning office" in this writing and referred to as a "continuum of care" or "CoC" in HUD documents)³ that is local to a group of homeless service providers (e.g., shelters). The purpose of the planning office is to establish an HMIS for a group of service providers. Information flows from clients to service providers, who in turn, provide visit information to their local planning office. Because clients are expected to consume services from multiple providers, the planning office associates visits across providers over time to provide an unduplicated accounting to HUD for the services delivered in their geographical region.

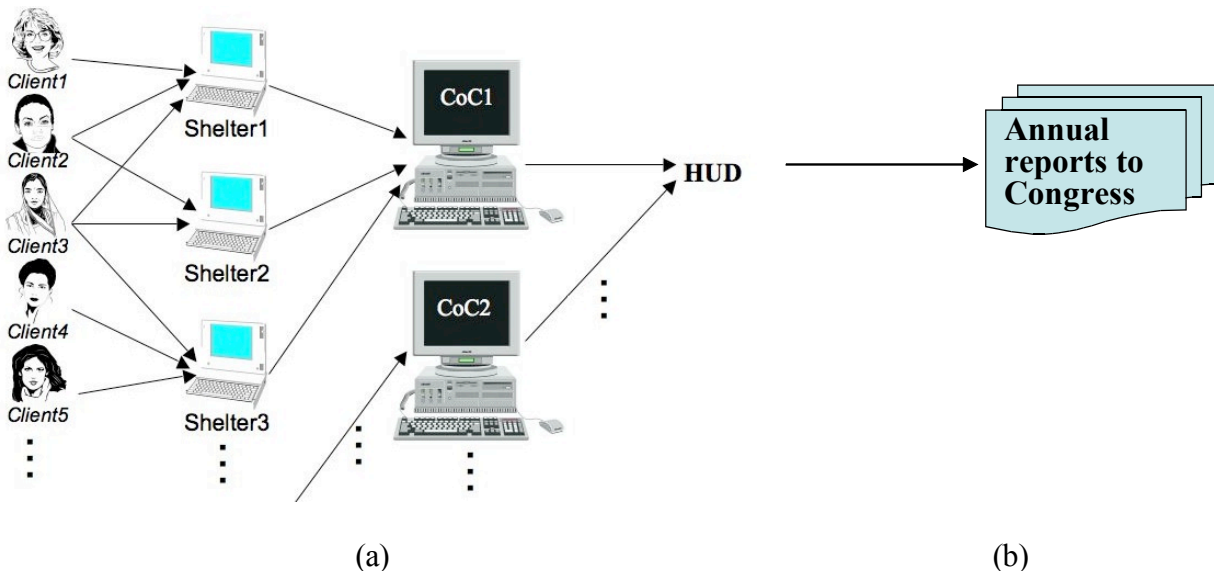


Figure 3. Flow of information from Clients to HUD: (a) Clients give information to Shelters, which report information to Planning Offices (CoCs), which in turn provide non-identifiable unduplicated count information to HUD, (b) which becomes the source data for annual homeless service utilization reports to Congress.

3 The purpose of a regional planning office predates and is broader than HMIS, but for the purposes of this writing, planning offices are examined narrowly in their HMIS context.

While a HMIS includes lots of different services for many types of homeless clients, the work reported herein is specifically focused on clients who visit domestic violence shelters. Hereinafter, unless otherwise noted, references to “Shelters” are exclusively domestic violence shelters and may generally apply to a suite of homeless service providers. Similarly, references to “Clients” are homeless persons serviced by Shelters and to “Planning Offices” are the CoCs servicing Shelters.

Figure 3(a) depicts the flow of information from Clients to Shelters through Planning Offices to HUD. A Client visits one or more Shelters. Each Shelter provides information to one Planning Office. HUD uses non-identifiable information from Planning Offices to provide annual reports on the utilization patterns of homeless people to Congress; see Figure 3(b).

3.2 Comparing HMIS to point-in-time approaches

Because Client demographics and specific visit data are captured on each visit, many of the shortcomings found with point-in-time studies may potentially be resolved by the HMIS approach.⁴

For example, a HMIS seeks to record sufficient information to allow the same Client to be identified on subsequent visits to the same or other Shelters, thereby thwarting the potential for double counting. Associated date and length of stay information may be recorded to identify seasonal, climate and temporal visit patterns. Recording the reason given for each visit may help identify utilization characteristics related to different kinds of homelessness, and tracking Clients across the same and different shelters can provide recurrence and duration rates.

3.3 Concern about selecting planning offices

It is understood that a Client may visit one or more Shelters, which is why de-duplication across Shelters is necessary, but if the same Client visits Shelters reporting to different Planning Offices, then the de-duplication effort can be thwarted.

For example, consider Shelters servicing Boston and Cambridge Massachusetts. These are two cities between which people regularly walk and ride multiple times a day. If each of these cities has their own Planning Office, then a single Client being serviced by a Shelter in Cambridge and by another Shelter in Boston, would be counted twice—once by the Planning Office for Cambridge and again by the Planning Office for Boston. Similar situations can exist with Planning Offices located in close proximity to one another irregardless of city, county, or state boundaries. To combat this problem, the following recommendation is made.

4 One shortcoming of both the survey used by the Urban Institute and the HMIS approach is the sole reliance on service providers. Homeless people who are not using shelters or covered services are not captured. These include homeless people who may live in automobiles, make-shift housing (such as cardboard boxes or tents), or doubled-up situations.

Recommendation #1: Coordination of privacy protection schemes is necessary across Planning Offices that service a geographical region in which Shelters within the region report to different Planning Offices but service some of the same Clients. Lack of coordination can distort the unduplicated accounting.

In 2006, HUD funded about 400 Planning Offices. This funding extends beyond HMIS to the coordination and funding of homeless services at the local level. A Planning Office defines its own geographical service area and competes to receive HUD funds for homeless programs. Because geographical service areas are not dictated by HUD, cooperative coordination of privacy protection schemes in overlapping areas allows a Client's utilization pattern to be determined without compromising the identity of the Client.

3.4 Removal of explicit identifiers from HMIS

Almost as soon as the first HMIS standards were announced, privacy concerns emerged over the need for protections for clients of domestic violence shelters [14][15]. Tracking victims of intimate domestic violence who seek refuge in Shelters may be necessary for HMIS accounting, but many feared HMIS data collection and sharing might become a vehicle to further endanger a victim whose information would appear in HMIS data as a result of her attempting to remove herself from a harmful situation.

A privacy protective action taken by HUD involved changing HMIS standards to allow Shelters to provide Client information without making reference to any client explicit identifiers (e.g., name and Social Security number). Instead, an approved proxy, coded, encrypted, hashed, or other alternative termed a “**unique identification number**” (or “**UID**”) is to be used by Shelters to provide client information to Planning Offices, provided each Planning Office has the ability to recognize the occurrence of the same clients in the same and different shelters (including shelters that are not domestic violence provider shelters) over time.

Section 4 examines the nature of privacy threats in detail. Section 5 provides a method for assessing technologies for creating and using UIDs. Section 6 and Section 7 report on assessments of UID technologies initially considered. The remainder of this section examines the data elements collected and shared in a HMIS.

3.5 Details of HMIS data elements

HUD requires certain data elements be sent from Shelters to Planning Offices. The data elements that HUD requires Shelters to provide to Planning Offices are termed the “Universal Data Elements,” and consists of a record for each Client's visit to a Shelter and includes the Client's UID. The original data elements were modified to use UIDs, in lieu of explicit identifiers, as shown in Figure 5. Shelters participating in a HMIS must collect the Universal Data Elements and share them with the Planning Office at least once a year in a privacy-preserving manner that includes replacing name and Social Security number with UIDs.

“Program-Specific Data Elements” are additional fields of information that Shelters may be required to provide on each Client visit. All McKinney Vento funded Shelters that are required to complete an Annual Progress Report are required to collect and share certain Program-Specific

Data Elements with the Planning Office⁵. Figure 6 lists the Program-Specific Data Elements and identifies which data elements are required for the Annual Progress Report.

HUD places no further restriction on the information collected between Clients and Shelters. Beyond the noted data elements, Shelters may elect to collect additional information for their own purposes. A Unique Person Identification Number (“PIN”) is included among the Universal Data Elements. This field allows a Shelter to store its internal reference number for a Client. However, care must be taken to share only when the PIN is sufficiently privacy-protecting, as noted in the following recommendation.

Recommendation #2: A Shelter may assign a unique person identification number (PIN) to internally identify a client, but it should not share the client's PIN externally. PINs that include the Client's name, Social Security number, or other characteristic may be used alone or in combination with other data elements to re-identify a Client. Any characteristic not allowed as a data element or a UID, should not be used as an externally shared PIN.

In summary, Figure 4 shows the flow of information from a Client through the Planning Office to HUD using the Universal and the Program-Specific Data Elements.

Hereafter, the information transmitted from a Shelter to a Planning Office is collectively termed the “Dataset” in this writing and refers to the Universal Data Elements unless otherwise stated.

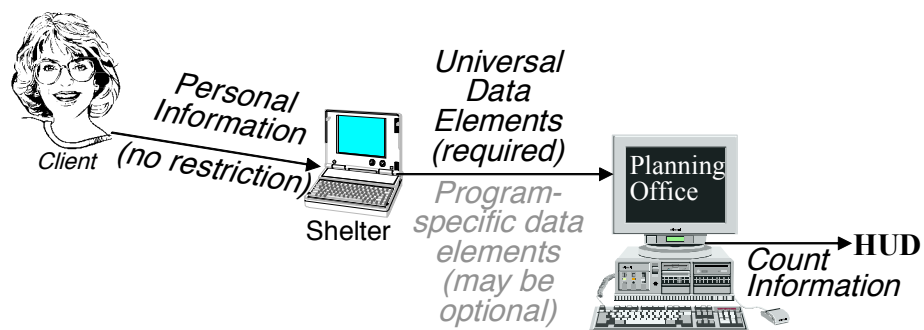


Figure 4. Flow of information: Client gives explicit personally identifying information to the Shelter, which provides the Universal Data Elements and Program-Specific Data Elements to the Planning Office, which in turn provides to HUD, non-identifiable, unduplicated count information of Client visits across all Shelters in the Planning Office's region.

⁵ See <http://www.hud.gov/offices/cpd/homeless/apr.doc>.

UNIVERSAL DATA ELEMENTS		
#	Description	Comments and Possible Values
1	Name	DV shelters collect but not share; use UID instead
2	Social Security Number	Domestic violence (DV) shelters collect but not share.
3	Date of Birth	Month, day and year of birth
4	Ethnicity and Race	Hispanic/Latino or not; American Indian, Asian, Black, Pacific Islander, White
5	Gender	Male or female
6	Veteran Status	Yes, no, don't know, refused
7	Disabling Condition	Yes, no, don't know, refused
8	Residence Prior to Program Entry	<p>Part I: Type of Residence Emergency shelter, transitional house for homeless, permanent housing for former homeless, psychiatric facility, substance abuse treatment facility, hospital (non-psychiatric), legal incarceration, rental unit, home ownership, family member's home, friend's home, emergency shelter voucher at hotel, foster care home, place not intended for habitation, other, don't know, refused</p> <p>Part II: Length of Stay in Previous Place Emergency shelter, transitional house for homeless, permanent housing for former homeless, psychiatric facility, substance abuse treatment facility, hospital (non-psychiatric), legal incarceration, rental unit, home ownership, family member's home, friend's home, emergency shelter voucher at hotel, foster care home, place not intended for habitation, other, don't know, refused</p>
9	ZIP Code of Last Permanent Address	5-digit code, don't know, refused
10	Program Entry Date	Month, day, year
11	Program Exit Date	Month, day, year
12	Unique Person Identification Number	"PIN" Shelter's internal reference number for Client.
13	Program Identification Number ("Shelter ID")	<p>Part I: FIPS code identifying geographic location of shelter</p> <p>Part II: Identification code for shelter, including HUD assignment</p> <p>Part III: Program Type Code: Emergency shelter, transitional housing, permanent supportive housing, street outreach, homeless prevention service, other service</p>
14	Household Identification Number	Constructed number to identify clients receiving services as a household

Figure 5. HMIS Universal Data Elements includes the generated unique identification number (UID).

PROGRAM-SPECIFIC DATA ELEMENTS			
#	Description	Need for Annual Progress Report	Comments and Possible Values
1	Income and Sources	Yes	Part I: Source of Income Earned income, unemployment insurance, supplemental security income (SSI), Social Security disability (SSDI), veteran's disability, private disability insurance, worker's compensation, temporary assistance for needy families (TANF), general assistance program (GA), Social Security retirement income, veteran's pension, former job pension, child support, alimony, other source, no financial resources. Part II: Total monthly income in dollars
2	Non-cash benefits	Yes	Food stamps, MEDICAID health insurance, MEDICARE health insurance, state children's health insurance, women-infants-children program (WIC), veteran's medical services (VA), TANF child care, TANF transportation services, other TANF services, public housing, other source.
3	Physical Disability	Yes	No, yes
4	Developmental Disability	Yes	No, yes
5	HIV/AIDS	Yes	No, yes
6	Mental Health	Yes	Part I: Mental health problem – no, yes Part II: Expected indefinite duration – no, yes
7	Substance Abuse	Yes	Part I: Problem: none, alcohol, drug, dully diagnosed Part II: Expected indefinite duration – no, yes
8	Domestic Violence	Yes	Part I: Experience –no, yes Part II: Time of experience past 3 months, 3-6 months ago, 6 to 12 months ago, more than a year ago, don't know, refused.
9	Services Received	Yes	Part I: Date of service – month, day, year Part II: Type of Service Food, housing, material goods, financial aid, transportation, consumer assistance, legal services, education, health care, HIV/AIDS services, mental health care, substance abuse services, employment, case management, day care, personal enrichment, outreach, other.
10	Destination	Yes	Part I: Destination Emergency shelter, transitional housing, permanent housing for formerly homeless, psychiatric facility, substance abuse treatment center, hospital (non-psychiatric), legal incarceration, rental unit, home own, family home, friend's home, hotel paid by shelter voucher, foster care, place not meant for habitation, other, don't know. Part II: Tenure Refused, permanent, transitional, don't know, refused Part III: Subsidy Type None, public housing, Section 8, S+C, HOME program, HOPWA program, other housing subsidy, don't know, refused.
11	Reasons for Leaving	Yes	Housing opportunity, completed program, non-payment of rent, non-compliance with project, criminal activity, reached maximum

PROGRAM-SPECIFIC DATA ELEMENTS			
#	Description	Need for Annual Progress Report	Comments and Possible Values
			time allowed, needs could not be met, disagreement with rules or people, death, disappeared, other
12	Employment	No	Part I: Employed – no, yes Part II: If employed, number of hours worked past week Part III: If employed, tenure --permanent, temporary, seasonal Part IV: If not employed, looking for work – no, yes
13	Education	No	Part I: In school – no, yes Part II: Received vocational training – no, yes Part III: Highest Level of School Completed No schooling, nursery school to 4 th grade, 5 th or 6 th grade, 7 th or 8 th grade, 9 th grade, 10 th grade, 11 th grade, 12 th grade with no diploma, high school diploma, GED, post-secondary school. Part IV: Post-Secondary Education If high school diploma or equivalent, earned Associated Degree, Bachelor's, Masters, Doctorate, other graduate/professional degree.
14	General Health Status	No	Excellent, very good, good, fair, poor, don't know
15	Pregnancy Status	No	no, yes
16	Veterans Information	No	Part I: Military Service Era Persian Gulf, post Vietnam, Vietnam era, between Korean and Vietnam wars, Korean war, between WWII and Korean war, World War II, between WWI and WWII, World War I. Part II: Duration of active duty in months Part III: Served in a war zone – no, yes Part IV: If served in War Zone, Specify Zone Europe, North Africa, Vietnam, Laos and Cambodia, South China Sea, China-Burma-India, Korea, South Pacific, Persian Gulf, other. Part V: If served in war zone, number of months served Part VI: Received hostile or friendly fire –no, yes Part VII: Branch of the Military Army, Air Force, Navy, Marines, other. Part VIII: Discharge Status Honorable, general, medical, bad conduct, dishonorable, other.
17	Children's Education	No	Part I: Current enrollment status – no, yes Part II: Name of School (explicitly stated) Part III: Type of School – public, parochial-private Part IV: Last date of enrollment –month, day, year Part V: If not enrolled, Identify Problem Residency requirements, availability of school records, birth certificate, legal guardian requirements, transportation, lack of preschool program, immunization requirements, physical examination requirements, other.

Figure 6. Program Specific Data Elements are supplemental information that may be made available to planning offices.

3.6 The unduplicated accounting

The motivation for HMIS data collection and sharing are the annual reports HUD will provide to Congress, which will report on homeless demographics, utilization patterns, and service availability. These reports are termed the “**Annual Homeless Assessment Report**” (“**AHAR**”). To produce the AHAR, Planning Offices use HMIS data to provide aggregate count information to HUD.

HUD provided the first AHAR to Congress in 2006 using HMIS data collected in 2005. An initial draft of the data analysis for the 2006 AHAR shows how HMIS data elements contribute to the AHAR [19]. Basic questions addressed by the AHAR focus on emergency shelters and transitional housing for individuals and for households. Figure 7 has a sample of the kinds of questions answered by the AHAR using HMIS data elements. The sample questions pertain to individuals at emergency shelters, but similar questions exist for transitional housing and for households. Notice that all the data elements are used except UID and PIN (recall name and Social Security number had already been removed). A Planning Office provides HUD with answers to these questions, which are aggregated counts and not the raw data used to compute the counts.

A Planning Office can generate a “De-identified Dataset”⁶ to perform the de-duplication and compute the unduplicated count information needed for the AHAR by linking Client demographics to Shelter utilizations using Client UIDs. The resulting data, which does not itself have to further include Client UIDs and PINs, is de-identified.

The UID is used to identify data relating to the same Client. Once the visit records are grouped by Client, the UIDs are no longer needed. A sequentially assigned Client number from 1 to the total number of distinct Clients appearing in the dataset can be used to reference Clients in the De-identified Dataset.

PINS are not needed in the De-identified Dataset. If a data problem occurs, the Planning Office has the originally received data for communicating with a Shelter using the Shelter’s PIN.

Similar to UIDs, once Clients belonging to the same households are linked together, the Household Identification Number can be replaced with a sequentially assigned number from 1 to the total number of distinct households appearing in the dataset.

Figure 8 shows an example for a single Client. The Client’s utilizations relate to her demographics but not to her explicit identity. Clients belonging to the same household are linked by sharing the same Household Identification number. Figure 9 provides an example of four clients, two of which are in the same household. The de-identified data can be used to compute values necessary to forward to HUD for the AHAR. Removing PINs and replacing UIDs and Household Identification numbers adds privacy protection to the De-identified Dataset, though more privacy protections are needed, as discussed in the remainder of this writing.

6 While the De-identified Dataset is sufficient for computing the aggregate unduplicated count information that is forwarded to HUD, Planning Offices are not required to use the exact de-identified dataset described above.

Recommendation #3: If a Planning Office produces a De-identified Dataset from the HMIS data collected from Shelters, the De-identified Dataset should not include any original Personal Identification Numbers (PINs), Unique Identification numbers (UIDs), or Household Identification numbers.

Universal Data Elements	Question #
Date of Birth	3,5
Ethnicity and Race	3
Gender	3,5
Veteran Status	3
Household Identification Number	2,3
Disabling Condition	3
ZIP Code of Last Permanent Address	4
Residence Prior to Program Entry	4
Program Entry Date	1,5
Program Exit Date	1,5
Program Identification Number	1,2,3,4,5

(a)

Question #	AHAR Questions: Emergency Shelter -Individuals
1	How many people used emergency shelters at ___ time?
2	What is the distribution of family sizes using emergency shelters?
3	What are the demographics of individuals using emergency shelters?
3	distribution by gender?
3	distribution by race and ethnicity?
3	distribution by age group?
3	distribution by household size?
3	distribution by veteran status? By disabling condition?
4	What was the living arrangement the night before entering the emergency shelter?
4	within/outside geographical jurisdiction?
5	What is distribution of the number of nights in an emergency shelter?
5	distribution by gender?
5	distribution by age group?

(b)

Figure 7. Data elements from Figure 5 above (a) associated with sample questions answered by the AHAR (b). Planning Offices provide HUD with aggregated unduplicated count information as answers to the questions.

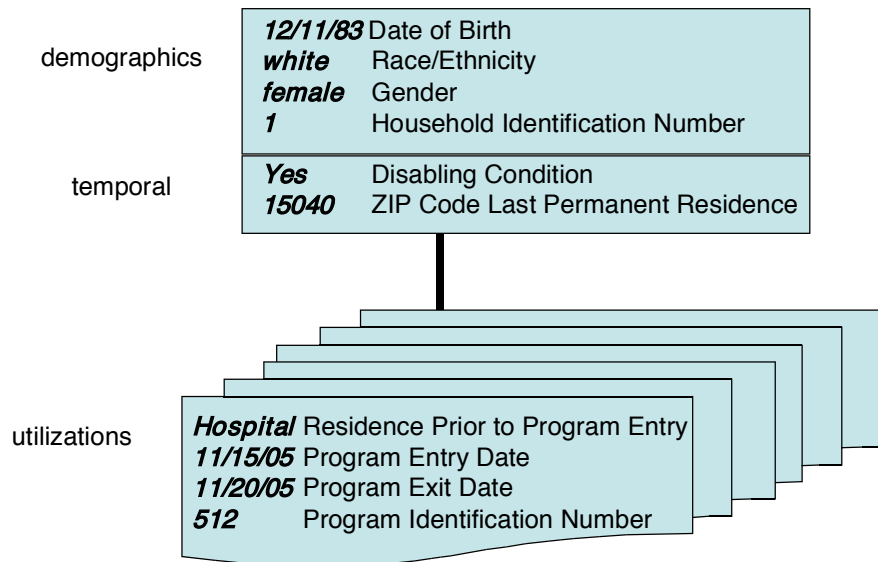


Figure 8. De-identified data for a Client includes demographics, some information that may change over time (disabling condition and ZIP of last residence), and program utilizations.

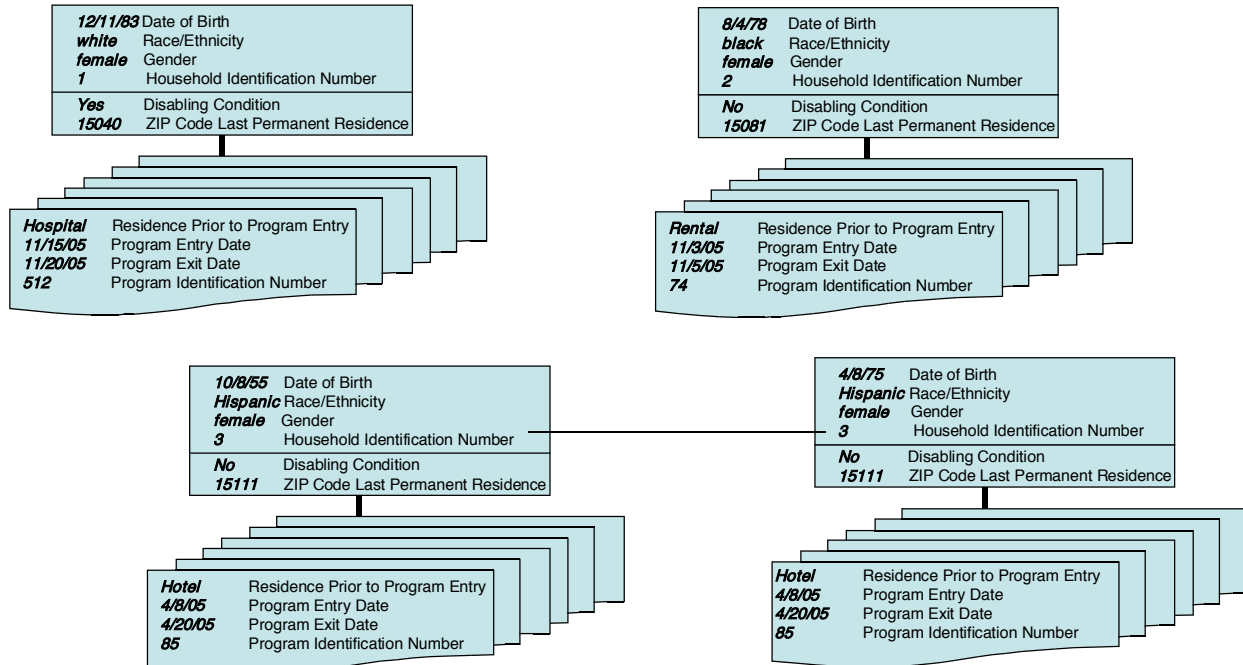


Figure 9. De-identified data for Clients includes utilization patterns. Some Clients are linked together by sharing the same Household Identification Number (depicted by the link between the bottom Clients).