

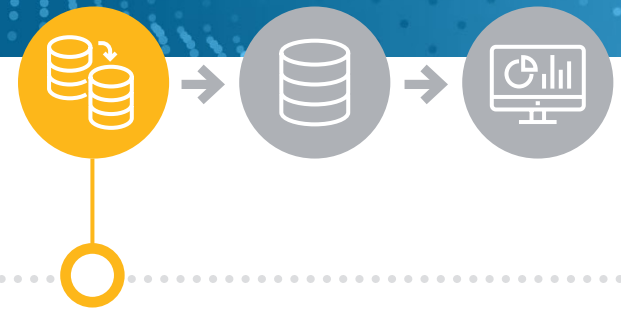
Technical information about HDIS

The Homeless Data Integration System (HDIS) is a statewide data warehouse that combines and processes data from the 44 local homelessness response systems in California.

HDIS enables the state to provide technical assistance for its local partners, to inform planning decisions, and to coordinate resources to more effectively prevent and end homelessness.

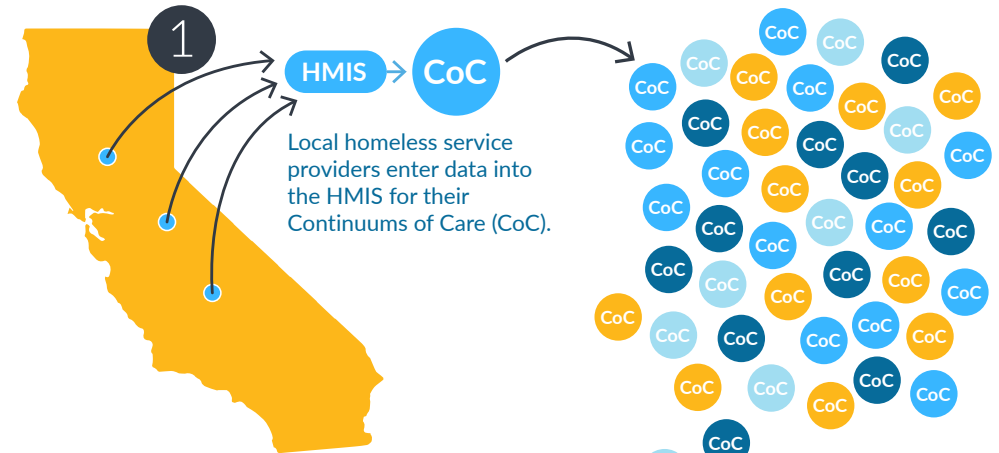
This document provides a behind-the-scenes view into the technology and processes that make HDIS.





Moving the data

1 California's homelessness response system is organized into 44 separate regional planning bodies called Continuums of Care (CoC). Each CoC maintains a separate local homeless management information system (HMIS). Local homeless services providers within the CoC enter data on clients who access services.



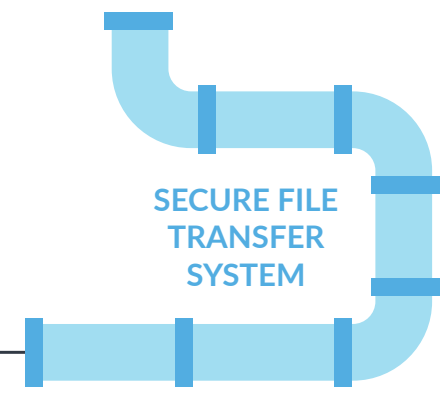
2 On a regular schedule, each CoC pulls data from their HMIS and sends them to HDIS via a secure file transfer system. The data are complex, with thousands of records organized into dozens of relational database tables.

2 44 separate CoCs send their data to the HDIS via a secure file transfer system.

3 HDIS then prepares the data for analysis and visualization.



3 **DATA PREPARATION**
The data are prepared for the statewide data warehouse.



Securing the data

The HDIS ensures data privacy in many ways. For example, multifactor authentication, secure socket layer encryption during data transfer, and encrypting data on servers.

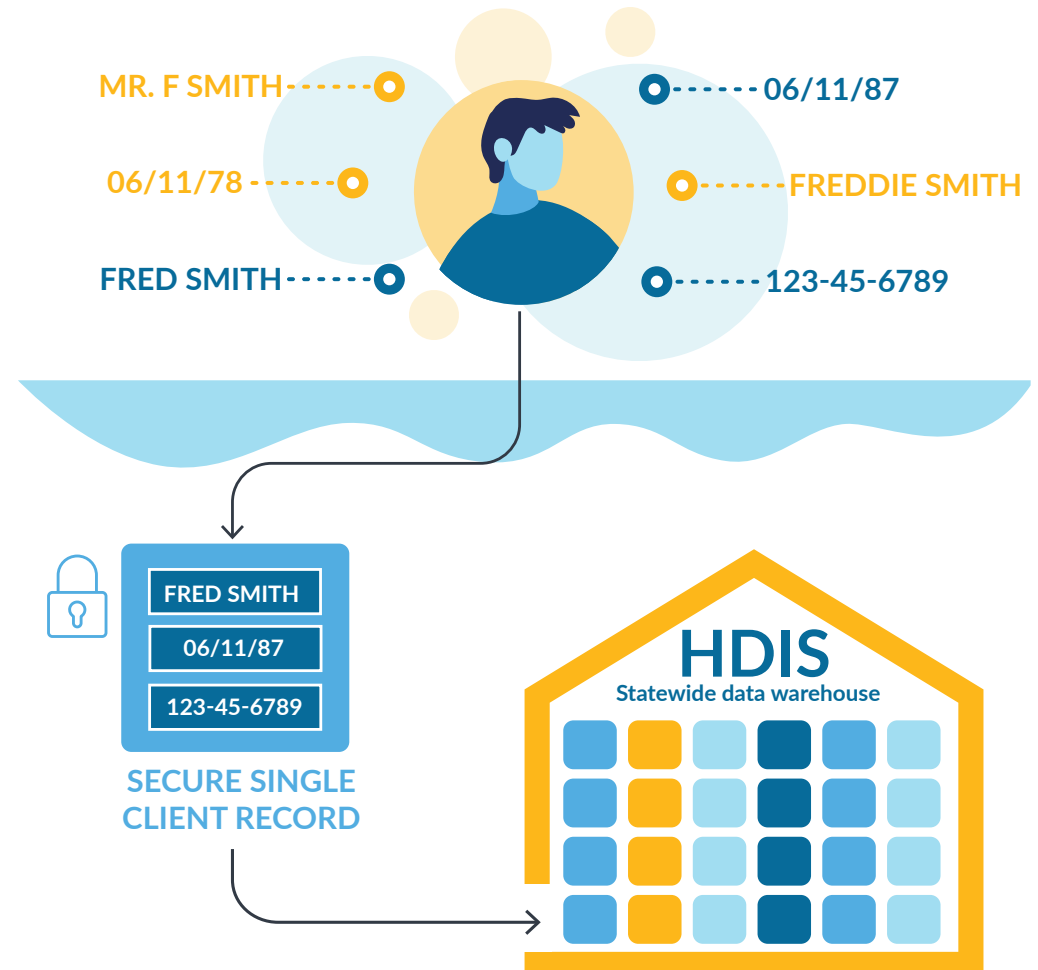


Preparing the data

Each CoC maintains its data using slightly different conventions. Before the data can be effectively used by HDIS, they need to be checked for accuracy and standardized. The process is called “data cleansing.”

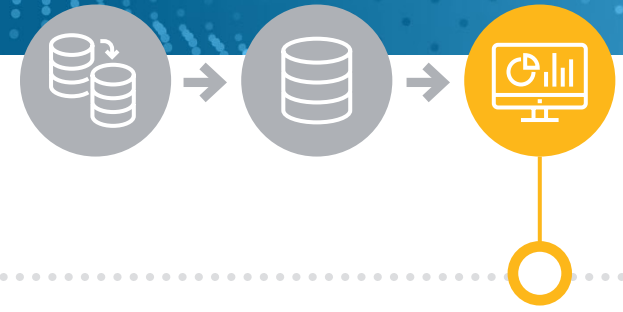
Part of data cleansing is identity resolution, using a method called master data management (MDM), to apply a set of rules that check and validate matches by cross-matching records, flagging errors, and combining duplicate entries. The end result is a single, standardized record for each client.

The cleaned and matched data are then placed into a statewide data warehouse – the central repository for HDIS – where the data are structured in a way to best answer questions by analysts, data scientists, and others.



Ensuring data accuracy

Data from CoCs may contain multiple records belonging to the same client. Drawing from Department of Housing and Urban Development standards and MDM best practices, the HDIS team set up a system of exact and probabilistic rules to compare pairs of records. Records with a high level of similarity are matched and merged together. This ensures that people are counted accurately in HDIS analysis.



Visualizing and analyzing the data

The final stage is seeing the data in action. HDIS data comes to life for policy makers, analysts and the public in two ways: through visualizations and analytics.

1 VISUALIZATIONS

Visualizations help people easily access the information in HDIS. Visualizing data reveals trends and insights that can improve the state's response to preventing and ending homelessness.



2 ANALYTICS

At the heart of HDIS is the data lake, which enables the state to employ robust data modeling, statistics, and analytics in a safe and secure manner.

HDIS is extensible and compatible with additional features such as automated connectors.



Ensuring privacy

All personal information in HDIS is protected against unauthorized use and disclosure in strict accordance with applicable law. Information used in visualizations has been de-identified, which ensures that individuals experiencing homelessness can't be singled out.

Where can I go to learn more?

Additional details and data visualizations are available on the [HDIS website](#).