# The ACT Network

Powered by NCATS CTSA Program

Understanding
the ACT Data Structure

The data in the ACT Network comes from raw electronic medical records or EMRs belonging to patients who have received services at participating institutions. All EMR data is de-identified prior to use in ACT.

EMR data has several limitations. EMRs are made up of both structured data like demographic information, diagnostic codes, medications and lab results and unstructured data like provider notes and radiology images.
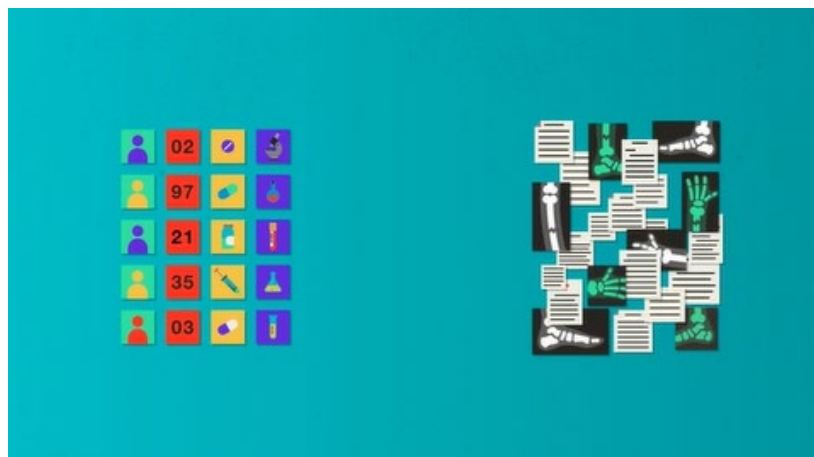
ACT does not include unstructured data. EMRs typically don't contain information about current or past clinical trial participation, medical adherence, costs incurred for services or health insurance plans.


Some providers may prefer to enter patient information using unstructured data fields rather than structured fields. Or may choose to enter information based on what is billable.

Additionally, the healthcare industry has not reached a point where every patient has one complete and accurate EMR. They may just have bits and pieces and often those bits and pieces are inconsistent with EMR data from another source whether or not their using the same software.

This can create significant barriers to sharing healthcare information between and even within institutions.

The ACT Network ontology harmonizes query terms across the network and local sites map that ontology to their local i2b2 repository.
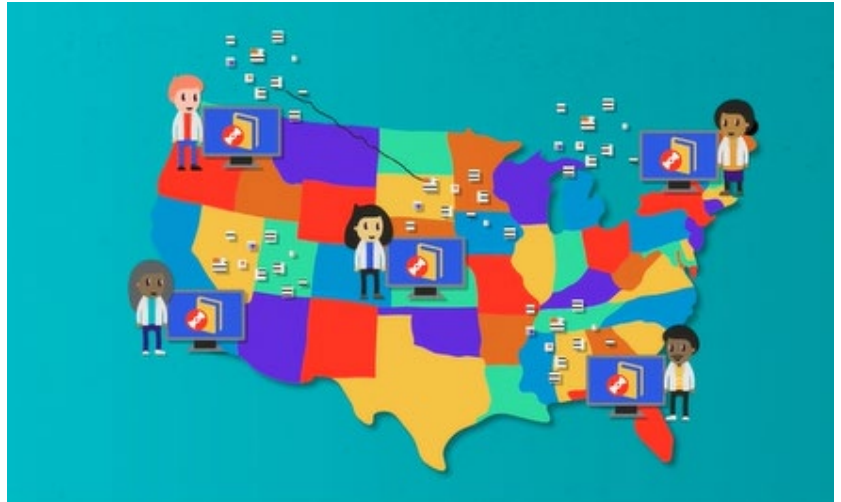
First, the raw data is collected from the institution's EMR system or clinical data warehouse. All protected health information as defined by HIPAA is removed with the exception of dates. The extracted data is then mapped to a common data structure using software called Informatics for Integrating Biology and the Bedside known as i2b2.

The ACT Network ontology and i2b2 ontology both use parent-child relationships which enables hierarchical queries.
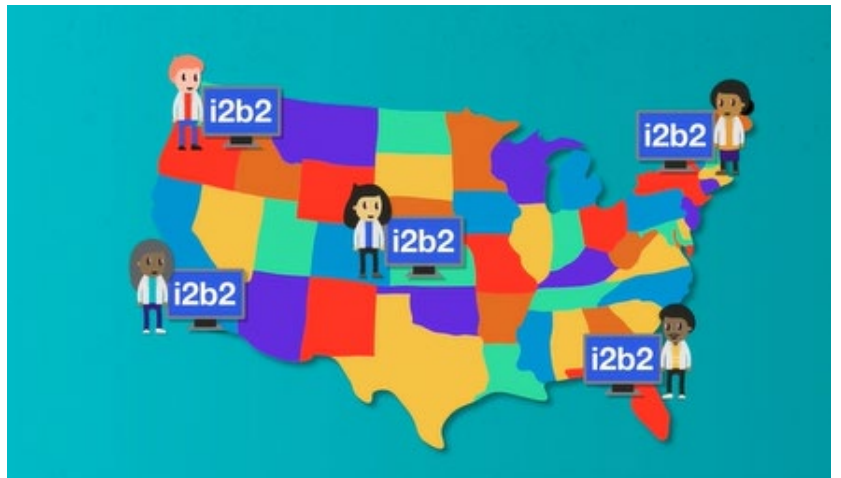

Using software called SHRINE (installed at each local institution), ACT users can query the de-identified data stored in each site's local i2b2 repository.

All of the procedures used to extract and convert the raw data adhere to recognized best practices, using secure computer servers, and under the supervision of medical center personnel.

While running queries, investigators may find discrepancies in results between ACT queries and queries conducted on local databases at their home institution.

This may be attributed to the data conversion process or could occur because counts obtained through ACT are obfuscated by plus or minus 10 and rounded to the nearest multiple of 5 to protect patient confidentiality.
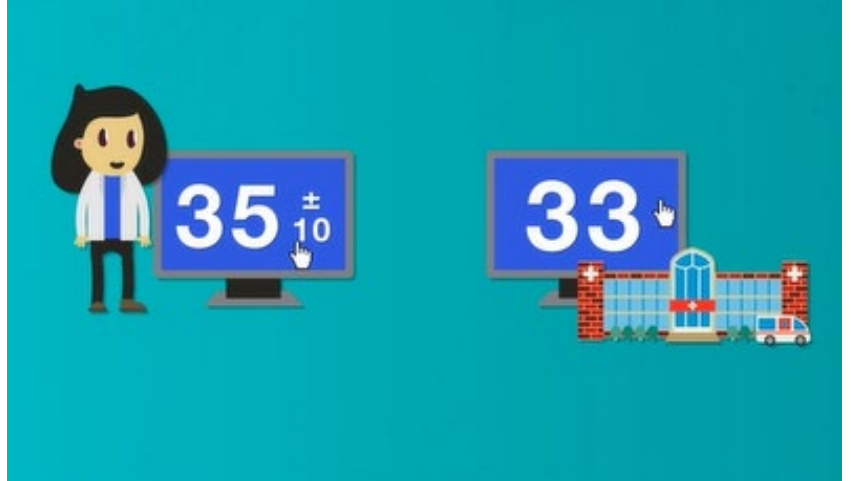
This is why ACT queries should be used in conjunction with Local database queries when doing research.

ACT

The ACT Network

Powered by NCATS CTSA Program

Explore. Discover. Connect