



Volunteer eHealth Initiative



The Challenges of Aggregating Patient Data from Multiple Sites

Janet King

Technical Project Manager

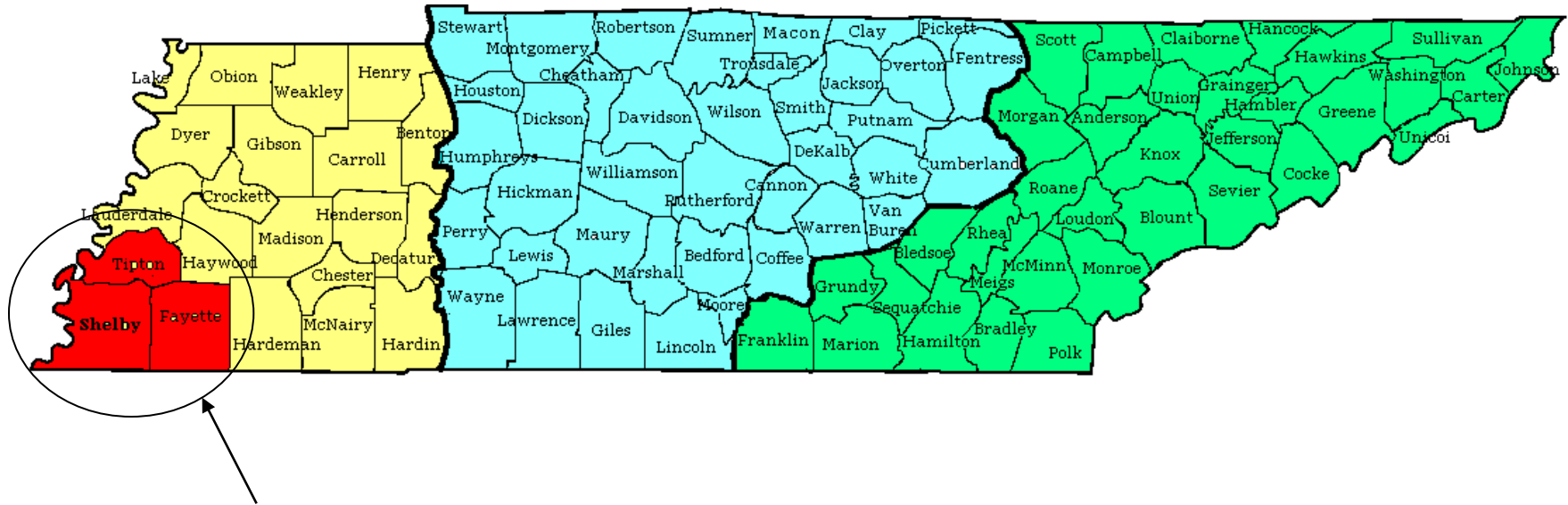
Regional Health Initiatives

Vanderbilt Center for Better Health

Funding: AHRQ Contract 290-04-0006; State of Tennessee; Vanderbilt University.

This presentation has not been approved by the Agency for Healthcare Research and Quality

Where are we talking about?



All parties recognize that health care is regional and that a significant number of individuals seeking care in Tennessee are residents of one of the 8 bordering states
Note – There are other regional initiatives and state-wide HIT initiatives funded by HHS, AHRQ and HRSA in the state



The Expectations from our Participants....

- Entities will participate as data contributors and end-users for a 5 year period and participate in planning efforts around sustaining the Clinical Data Exchange for the long-term.
- Governance will be a neutral organization at a neutral site determined by agreement of the board.
- **The proposed architecture imposes minimal data exchange requirements upon the entities publishing data to a vault.**
- **Participating entities will own and maintain the data stored in their individual data vault.**
- **Vanderbilt University will be responsible for aggregating data published by the participating entities and implementing the aggregation algorithms, data display, and reports.**
- Participating entities will take part in a testing effort to validate data aggregation and algorithms.
- Early deployment will be focused on select care settings where value is self-evident.
- The Clinical Data Exchange is not intended to replace existing clinical systems.
- General application training materials (e.g., login and basic navigation) will be provided by the Volunteer eHealth Initiative to ensure consistency across the deployed care settings.
- Entities will be responsible for incorporating process changes and delivering training to support adoption and utilization of the Clinical Data Exchange into the clinical workflow.
- Long-term operational support needs from participating entities are still being determined, therefore the ability to fill operational FTE needs is not being assessed.
- Long-term goal will be to incorporate data into participating entities' current systems.



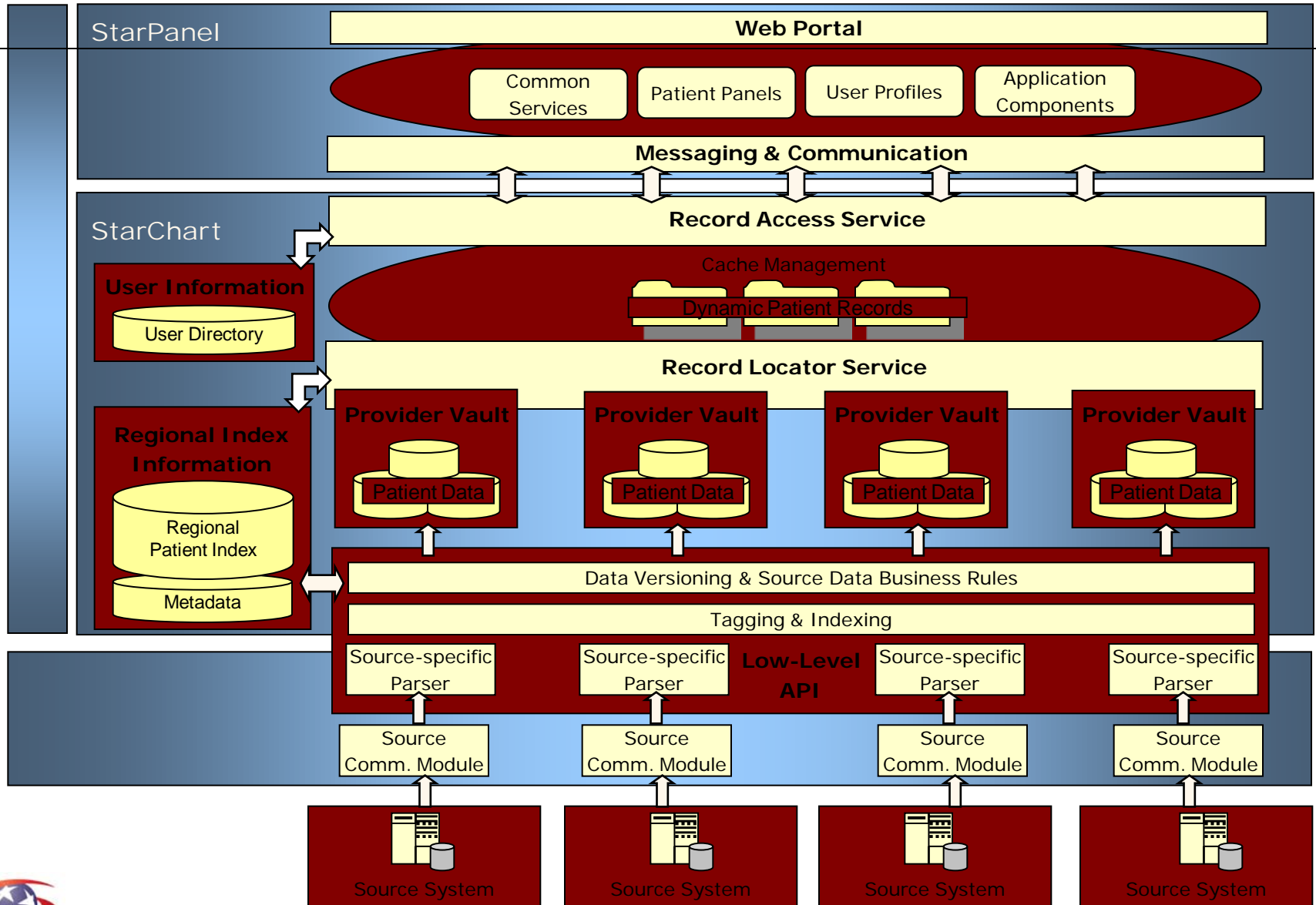
Core Entities and Data Elements currently received

Core Entity	Patient ID Data	Lab Results	Encounter Data	Diagnostic Codes	Medications	Allergies
BMHCC – 4 sites	✓	✓	✓			
Le Bonheur Children's Hospital	✓	✓	✓			
Methodist University Hospital + 5 sites	✓	✓	✓			
The Regional Medical Center (The MED)	✓	✓	✓			
Saint Francis Hospital & Bartlett	✓	✓	✓		✓	
St. Jude Children's Research Hospital	✓	✓	✓			
Shelby County/Health Loop*	✓	✓	✓			
UTMG	✓					
Memphis Managed Care-TLC	✓	✓	✓	✓	✓	

*Available patient population will grow as EMR roll-out continues through 2006



High-level Regional Architecture

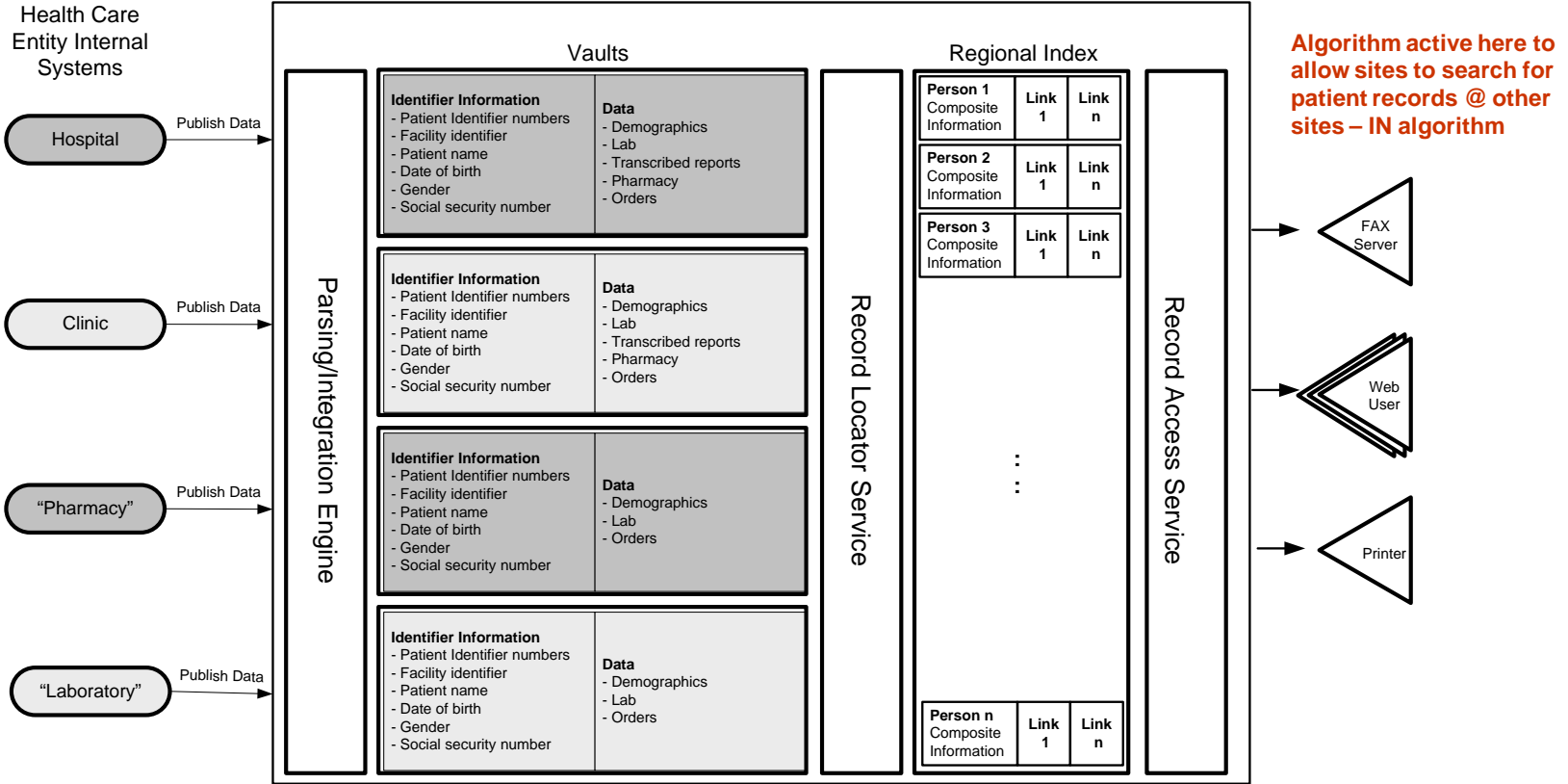


Regional Solution High-level Data Flow Schematic

Algorithm inherent to Oracle here to check if patient has previous records

Algorithm active here to match PID to clinical record

Algorithm active here to create linkages of site records to RHIO RPI



Data is published from data source to the exchange

- Participation Agreement
- Patient Data
- Secure Connection / Real-Time

Exchange receives data & manages data transformation

- Mapping of Data
- Parsing of Data
- Standardization of Data
- Queue Management

Organizations will have a level of responsibility for management of data

- Issue Resolution
- Data Integrity
- Entities are responsible for managing their Data

Data bank compiles and aggregates the patient Data at the regional level

- Compilation Algorithm
- Authentication
- Security
- User Access



Challenge #1 – Data Feeds

- Data feeds in varying formats
 - HL7 v 2.1 – 2.3
 - XML code
 - Flat files
 - Multiple laboratory systems with differing specs
 - Multiple parsing programs created for a site
 - Understanding all the statuses and their meanings
 - *Many parser changes came late in the process with site QA review*
- Data feeds received in varying intervals
 - Real-time and real-time cycles
 - Daily Batch
 - Weekly Batch
 - *Health Plan member file (update)*



Challenge #2 - Algorithm development & testing

- We reviewed articles written to evaluate of the effectiveness of various methods. We followed other published experiments in tuning the code.
- Our process began by doing simple deterministic matches of patient demographic information between 2 sites' data where we knew we would find overlap (matches.)
 - Our initial searches were based upon an exact match of SSN, name, DOB and Gender.
 - Feeling confident from our review of the match results and the sites confirmation of our matches, we boldly began to “tweak’ the algorithm.
 - One of our first changes was to check for female name changes. We tweaked the algorithm to check first name on females where there was an SSN and DOB match.



Challenge #2 - Algorithm development & testing

- What was most interesting was what we saw in the matched results that required us to go back to the sites to ask some questions.
 - Does the site use default SSNs such as 888-88-8888 or all 9's or some other equivalent to fill in a field when the patient cannot give the number?
 - Does the site use any default MRNs that is duplicated from one patient to another?
 - How are trauma patients or John /Jane Does registered in their systems?
 - How are single birth newborns named; how are multiple birth newborns named?
 - How are specimens named and assigned numbers in their systems?



Challenge #3 – QA Review by Sites

- Compared their data in our system to the same data in their system
- Some sites used samples of data from our database
- Others used their ADT reports
- Completed Issue tracker spreadsheet for inconsistencies of data
- General issues were missing results due to report statuses not displayed
- No matching issues found to date
- Some changes by vault to formatting of result comments



Challenge #4 - LOINC Mapping

- LOINC mapping of clinical lab results can improve the presentation of data.
- With mapped data, we can display results in a tabular format in columns or we can graphically represent the same data over time.
- In this project, we divided the LOINC mapping into multiple phases.
- We (the project) would be responsible to do the actual mapping to LOINC using site resources as little as possible.



Challenge #4 - LOINC Mapping

- VUMC clinicians (working on the project) identified a subset of laboratory test results to begin
- A list of ~40 codes including test panels
- Lab test results that would be useful in the treatment of a patient in the Emergency Department
- Prepared spreadsheets for sites to use to enter test result data & sent to the sites
- Completed spreadsheets used by VUMC project resource to assign appropriate LOINC code
- Used RELMA™ tool from Regenstrief



LOINC - Next Steps

- Create Oracle tables to load site interface code, result name, LOINC code
- Load tables
- Change parsers to pick-up LOINC code as a part of data storage
- Create tabular display screens in web browser to present lab results standardized to LOINC across sites
- Send LOINC mappings back to sites to load in their LIS
- Select next 50 codes to map
- Define process to do annual review and update of LOINC codes



Questions

