



TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

# Data Quality

Managing, Normalizing & Maintaining



TEXAS  
Health and Human  
Services

---

Texas Department of State  
Health Services

# Data Quality

---

**Managing, Normalizing and Maintaining**

# Background



TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

The Texas Immunization Registry Quality Assurance (QA) team analyzes data for its integrity.

Data that is incorrectly recorded in the system:

- ✓ causes potential issues with searching and finding client/immunization data
- ✓ triggers an increase in questionable matches that registry staff must resolve manually





# Approach

The QA team focuses on one data field at a time, unless it is observed that multiple data fields can be normalized all at once.

Keeping the scope small helps prevent overcomplicating remediation techniques that could extend time, resources, effort and costs.



# Tools

---

## Software Used

- ImmTrac2 Application
- SQL Developer
- Tableau
- Excel



# Data Quality Characteristics

- Accuracy and Precision
- Legitimacy and Validity
- Reliability and Consistency
- Timeliness and Relevance
- Completeness & Comprehensiveness
- Availability and Accessibility
- Granularity and Uniqueness



TEXAS

Health and Human Services

Texas Department of State  
Health Services



# Data Quality Elements

1. Find
2. Capture
3. Secure
4. Manage
5. Normalize
6. Maintain
7. Aggregate
8. Understand
9. Report
10. Act Upon



# Data Quality Elements



TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

## Focus

Managing

Normalizing

Maintaining





TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

# Managing Data

Data management is an administrative process that includes acquiring, validating, storing, protecting, and processing key data to ensure the reliability, integrity, accessibility, security and timeliness of data for its clients and users.



# Data Analysis

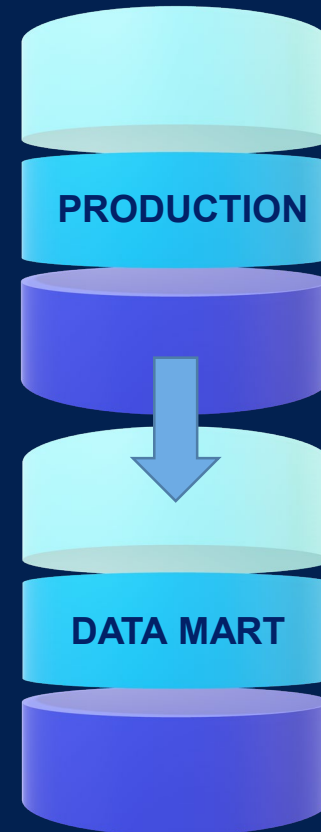
## Schemas

Production (transactional database):

- Registered Organizations and users
- Prescribing Authorities and Clinicians
- Data Exchange Logs

Data Mart (data subset):

- Clients
- Immunizations
- Client Immunization Schedules
- Audit Tables



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# Remediation



TEXAS

Health and Human Services

Texas Department of State  
Health Services

## System and Data Remediation can be handled by:

- Application Developers
- Database Administrators

## Training Staff and End Users:

- Business Analysts
- Interoperability Team
- Quality Assurance Team
- ImmTrac2 Trainers
- Operations Coordinator
- Customer Support Team



**FIX IT!**



# Inappropriate Data

Inappropriate data entry can be damaging to the patient/client, the healthcare provider and the IIS.

- Misleading data
- Duplicate data
- Incorrect data
- Inaccurate data
- Data that violates business rules
- Data without generalized formatting
- Incorrectly punctuated or misspelled data





# Data Business Rules

Recording TB test results in ImmTrac2 is considered an inappropriate data entry.

Allowing TB test results in the IIS could be misconstrued as a diagnosis and a privacy violation matter.

TB Test Results DO NOT indicate a diagnosis.

## TB Tests

- Negative
- Positive

# TB Test Results

---

## The options to prevent this were:

- Disallow TB tests to be entered in the IIS by removing the option to select them in the vaccine drop-down list
- Lock down (gray out) the Lot # field when TB tests are selected
- Create logic that does not allow words associated to TB test results to be entered. Like “NEG” and “POS”
- Disallow certain free text fields, like Client Notes field



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# TB Test Results

---

## The options to prevent this were:

- Disallow TB tests to be entered in the IIS by removing the option to select them in the vaccine drop-down list
- Lock down (gray out) the Lot # field when TB tests are selected
- Create logic that does not allow words associated to TB test results to be entered. Like “NEG” and “POS”
- Disallow certain free text fields, like Client Notes field



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# Normalizing Data

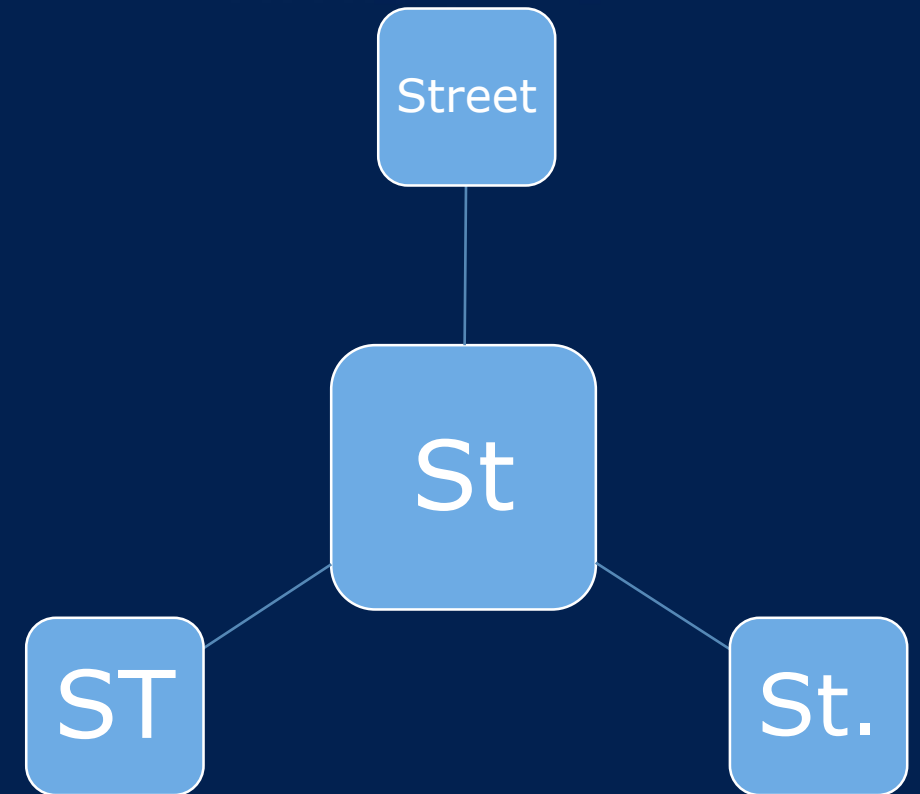


TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

## Normalization:

- is the process of restructuring data values in a relational database in accordance with a series of “normal” forms to improve data integrity.
- ensures that all your data looks and reads the same way across all records.
- can standardize fields including organization names, street addresses, mother's maiden names and client last names.





# Last Names and Suffixes

---

For example, here are three common types of last name anomalies that we often see:

- Escobar Jr.
- Garza Jr
- San Gabriel, JR.
- Smith, JR



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# Cleanup

This is how the data should look after cleanup:

<b>LAST_NAME</b>	<b>NAME_SUFFIX</b>
<b>ESCOBAR</b>	<b>JR</b>
<b>GARZA</b>	<b>JR</b>
<b>SAN GABRIEL</b>	<b>JR</b>
<b>SMITH</b>	<b>JR</b>



**TEXAS**  
Health and Human Services

Texas Department of State  
Health Services

# Other Data Anomalies

- Mother's maiden name unknown
- Facility address
- Numbers, special characters in the client name



**Unknown, Unk, Ukn, N/A, Not Known**



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# Maiden Name Types



TEXAS

Health and Human Services

Texas Department of State  
Health Services

MAIDEN _NAME	COUNT
NA	16003
SWK	3443
UNK	3359
SAME	1110
UNKOWN	642
UNKNOWN	489

MAIDEN _NAME	COUNT
UKN	232
SHELTER	224
NOT STATED	126
NOT AVAILABLE	123
MOTHER	67
MOM	37

# Using SQL Regular Expressions



TEXAS

Health and Human Services

Texas Department of State  
Health Services

```
WHERE (REGEXP_LIKE (C.MOTHERS_MAIDEN, '^U(N|K) (K|N)$')
OR REGEXP_LIKE (C.MOTHERS_MAIDEN_LAST, '^U(N|K) (K|N)OWN$'))
OR REGEXP_LIKE (C.MOTHERS_MAIDEN_LAST, '^MOTHER$|^MOM$|^NA$|N/A|^NONE$|^SAME$|CHILDREN|
SHELTER|NOT USE|ADOPT|FOSTER CARE|FOSTER CHILD|FOSTER PARENT|
^IES|DECEASED|NOT AVAIL|NOT APPL|NOT GIVE|NOT KNOWN|NOT LIST|
NOT PRO|NOT SP|NOT STATED|CPS|/','i')
```



# Maintaining Data

- **Systematic**
  - During capturing - as it enters the system
  - Scheduled Jobs - Systematic Cleanup
- **Manual**
  - Cleanup Tasks - assigned to staff
  - Training ImmTrac2 Users - Outreach Specialists

# Outreach Specialists

---

Responsible Entities:

- local health departments
- health service regions

Contain Outreach Specialists who reach out to providers in order to help improve our data quality.

They help with:

- training providers on how to consent clients into the system
- Perform assigned activities to help overall immunization rates



TEXAS  
Health and Human Services

Texas Department of State  
Health Services

# Summary



TEXAS  
Health and Human  
Services

Texas Department of State  
Health Services

It is truly a state-wide team effort in improving data quality for the Texas Immunization Registry.

# TEAM WORK





**TEXAS**  
Health and Human  
Services

**Texas Department of State  
Health Services**

# Thank you

---

**Arthur Lara – Quality Assurance Coordinator**  
**1100 West 49th Street**  
**Austin, Texas 78756**  
**(512) 776-3056**  
**Email: [arthur.lara@dshs.texas.gov](mailto:arthur.lara@dshs.texas.gov)**