Vermont's Patient Deduplication Tool

Guidance • Support • Prevention • Protection

Kyle Mitofsky &
Bridget Ahrens
AIRA Conference 2013

VERMONT
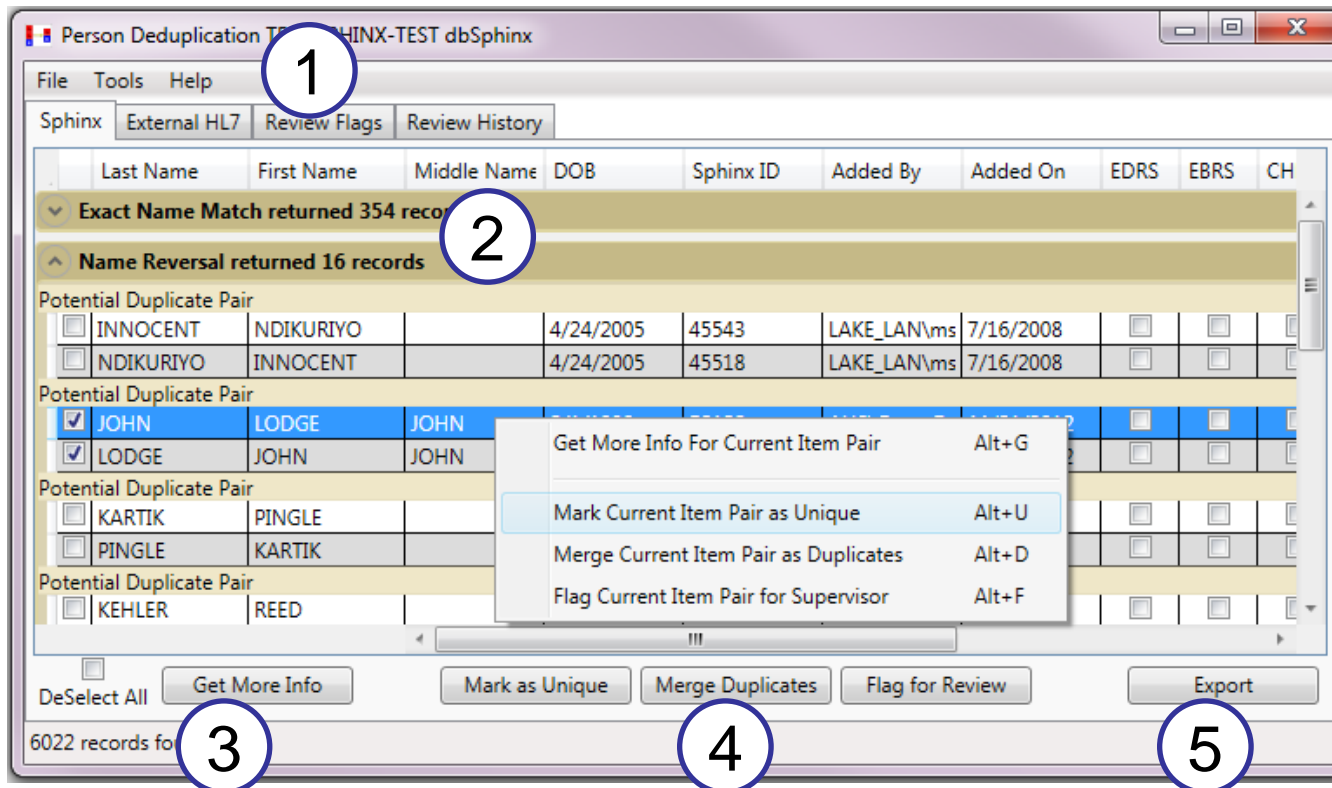**DEPARTMENT OF HEALTH**

Overview

- Project Goals

- Tool For Manual Review

- Deduplication Process

- Matching Algorithms

- Auditing Decisions

- Results

Project Vision

- Goals
  - Reduce manual processing time
  - Improve internal data quality
- Challenges
  - Allowing multiple users
- Key Innovations
  - Marking non-duplicates as **Unique**
  - Soundex and Tokens **Algorithms**

Tool For (Automated) Manual Review

1. Sources
2. Match Type
3. More Info
4. Actions
5. Export
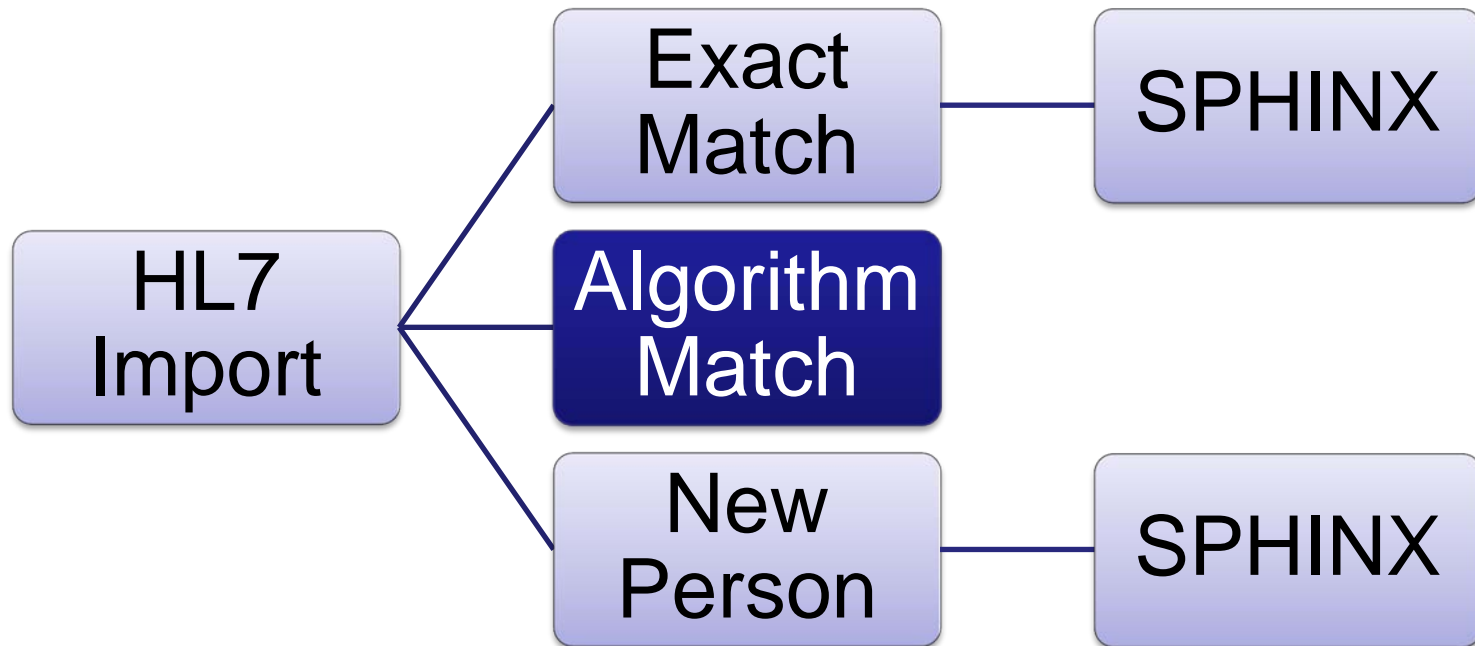
4

State Assets

- Relatively Small Size
- SPHINX - Combined Demographic Data
  - Immunization Registry
  - Electronic Birth Registry System
  - Electronic Death Registry System
  - Child Hearing Screening
  - Lead Registry
  - Newborn Bloodspot Registry

Process Outline

- Check all records in nightly batch

- Display potential pairs in application

- Reviewers investigate and take action:

  - Mark as Unique
  - Merge as Duplicates
  - Flag for Manager Review

Incoming Records

- Hold bad data for further review

Matching Algorithms

- Exact Name Match
- Name Reversal
- Name Parts Missing
- First Name Typo
- Name Typo Hyphen
- Name Change
- Date Of Birth Typo
- Soundex Match First Name
- Soundex Match Last Name
- Name Tokens

Algorithm - Soundex

- ## Phonetic Algorithm for indexing
  - First Letter + 3 Digit 'Sound'

SOUNDEX()	Matt	Mat	Matthew	Max
Keep first letter	Matt	Mat	Matthew	Max
Drop vowels and H/W	Mtt	Mt	Mtt	Mx
Replace consonants with digits	M33	M3	M33	M2
Remove pairs	M3	M3	M3	M2
Pad to 4 characters	M300	M300	M300	M200

Algorithm - Tokens

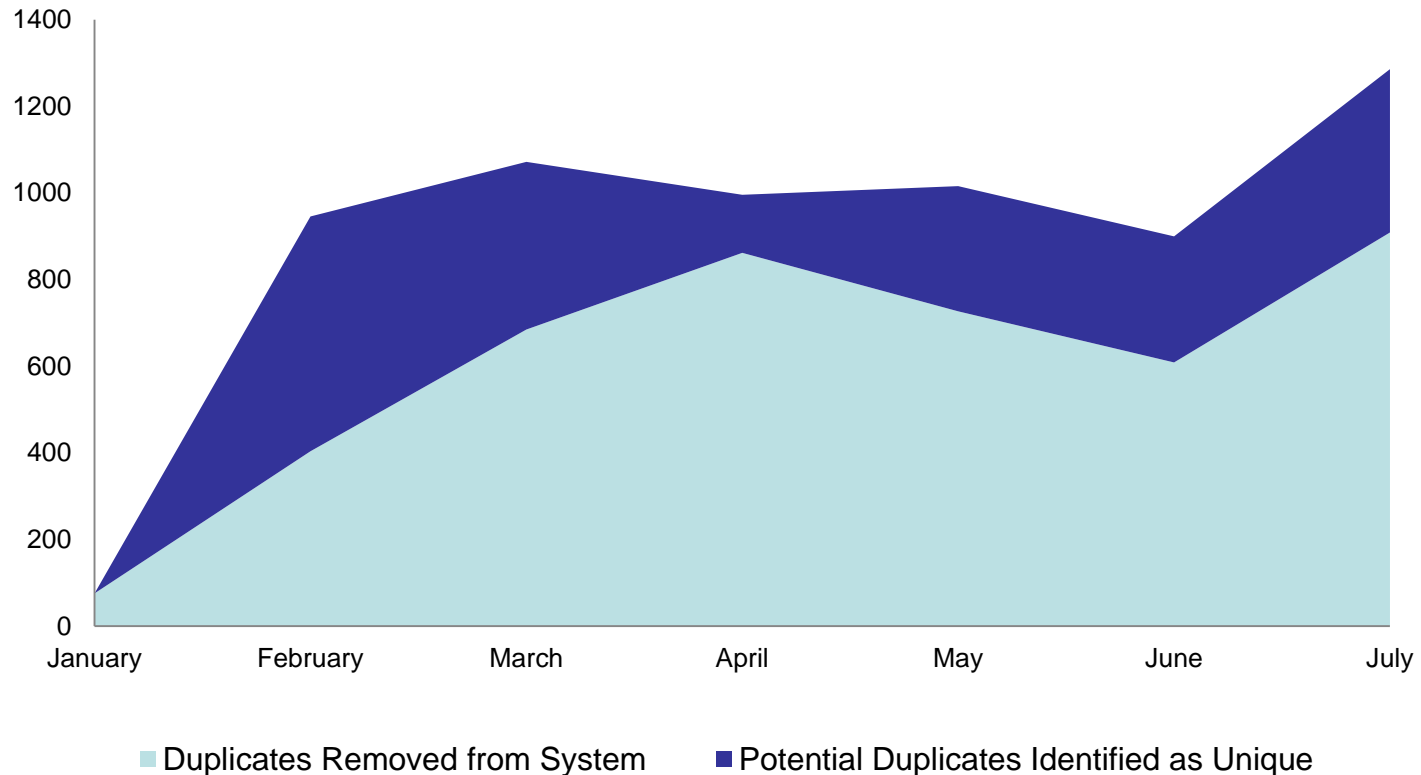- Puts all fields in large bin
- Must match on (n-1) fields

First Name	Last Name	Middle I.	Mother
Bob	Jones	A	Leanna
Jones	Leanna	C	Bob

Manual Audit

- Log all actions preformed
  - Failed and Successful
- Available for view and reports

Results

Total Review And Identified

Legend: ■ Duplicates Removed from System ■ Potential Duplicates Identified as Unique

12

Thank You!

Author	Title	Contact
Kyle Mitofsky, BA	Systems Developer II	Kyle.Mitofsky@state.vt.us
Bridget Ahrens, MPH	Immunization Registry Manager	Bridget.Ahrens@state.vt.us
Karen Clark, MS	IT Manager I	Karen.Clark@state.vt.us
Nicholas Pain, MS	Public Health Analyst	Nicholas.Pain@state.vt.us
Megan Barnes, BA	Health Data Administrator	*Contact Bridget Ahrens
Matthew Sontum, MS	Senior Systems Developer	Matthew.Sontum@state.vt.us
Grace Demler	Systems Developer II	Grace.Demler@state.vt.us