# Envisioning Data Liquidity: The DCRI-Pew Data Interoperability Project

NIH Collaboratory Grand Rounds

March 22, 2019

James E. Tcheng, MD

Professor of Medicine / Professor of Informatics

Duke Clinical Research Institute / Duke Center for Health Informatics

DukeHealth

# DCRI-Pew Data Interoperability Project

- Interoperability of what?
- Why not native data interoperability?
- The DCRI-Pew Project
- Envisioning data liquidity - next steps

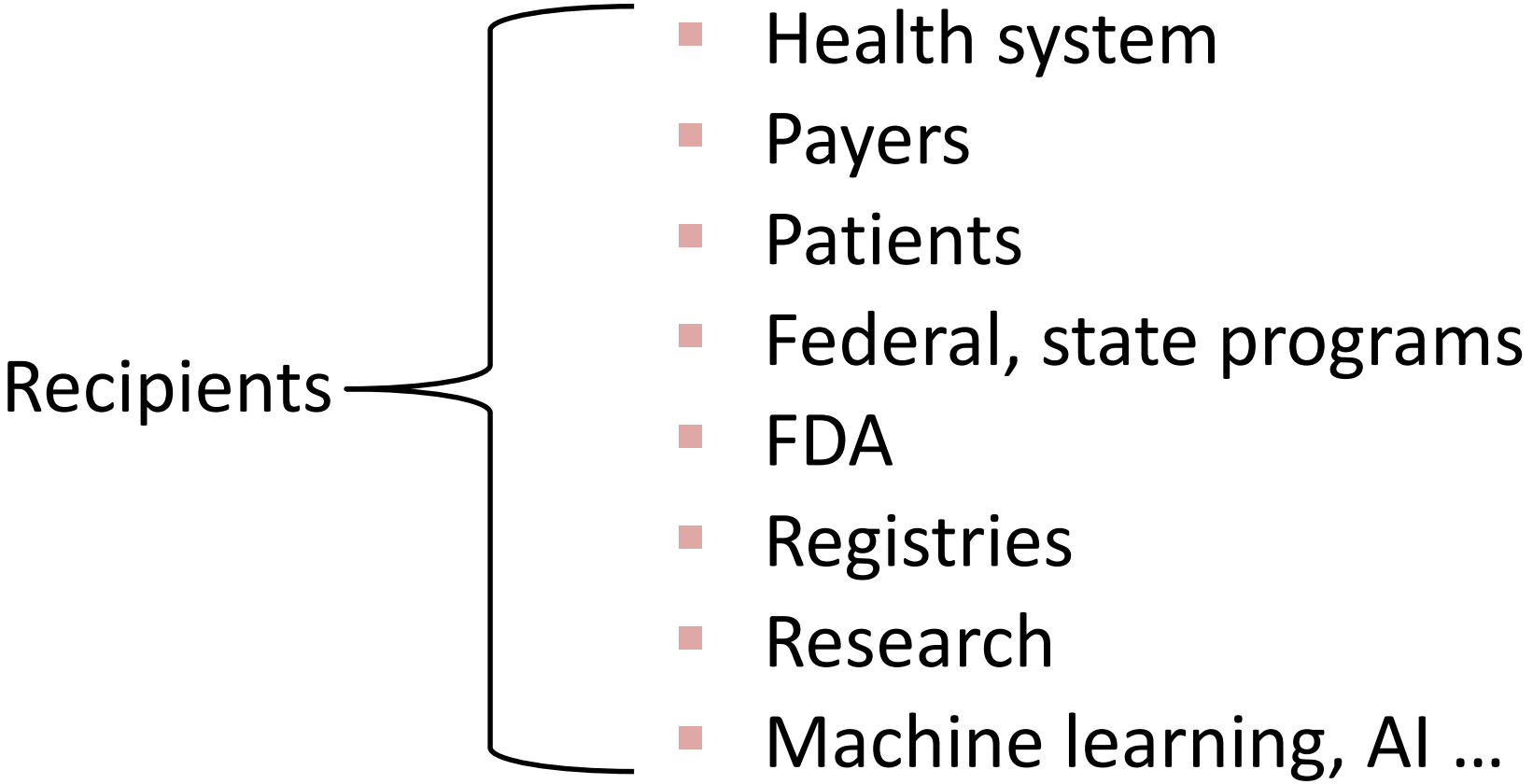
# The View from the President's Office

- 2004 - President Bush establishes a 10 year goal to develop the electronic health record (EHR)
- 2009 - President Obama signs ARRA, pushes EHR adoption through incentives, targets full implementation by 2016

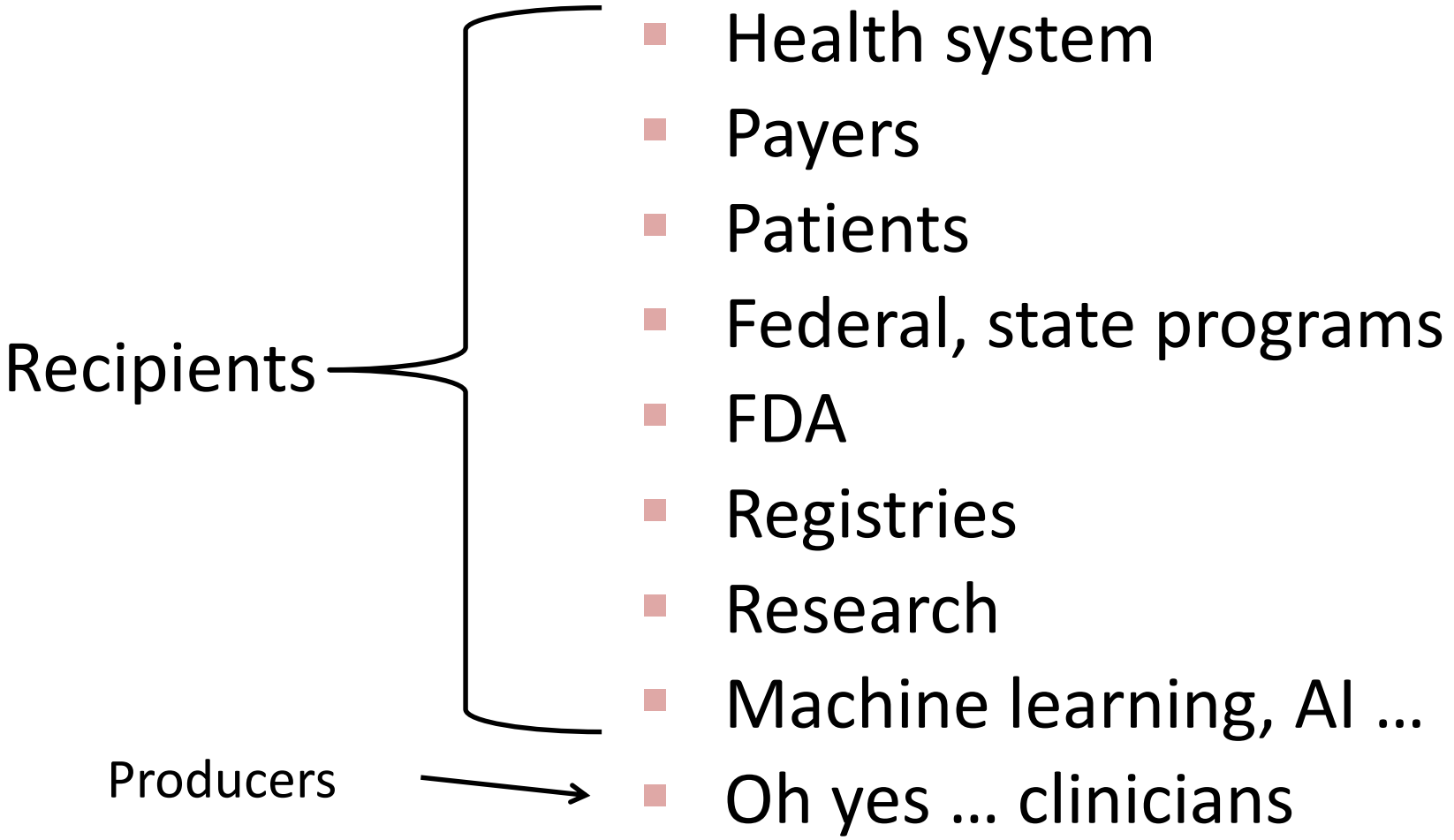
# 10 Years & $36 Billion Dollars Later …
# *Are We There Yet?*

| Envisioned | Reality |
| --- | --- |
| EHR "Meaningful Use" | EHR meaningless burden |
| Usability and productivity | Death by a thousand clicks |
| Patient engagement | AVS drivel |
| Effective clinical care | CDS trivial pursuit |
| Population health | Resource consumption focus |
| Bending healthcare cost curve | Cost control and penalties |
| Better provider work life | NOT! |
| Torrent of real-world data | Puddles of document exchange |
| Big (clinical) data analytics | Transactional (admin) data |
| Leveraged RCTs via registries | Electronic bridge to nowhere |

# Data Demand: Multiple Masters



# Data Demand: Multiple Masters




… who are time-challenged, short-staffed, overloaded with information and have increasing expectations placed upon them

# ARRA HITECH
# HIT Standards Committee

## Clinical Operations Workgroup Report

**Jamie Ferguson, Chair**
Kaiser Permanente

**John Halamka, Co-chair**
Harvard Medical School (& HITSP)

**20 August 2009**

# HIT Committee: Standards for Interoperability

- Clinical Operations is recommending standards for interoperability **between entities**, *not* within an entity
- Recommended standards should *not* apply to internal data capture, storage or uses – only to **external representation and data exchange** between entities
- Content should be able to be represented in the specified vocabularies and exchanged in the specified standards **at the boundary** between entities, regardless of how it is managed internally
  - Many methods may potentially be used to achieve interoperability standards, e.g., mapping, external services, or native data capture

# Edge-Based Interoperability

- SNOMED Clinical Terms (SNOMED CT)
  - International Health Terminology Standards Development Organization (IHTSDO)
- Logical Observation Identifiers, Names and Codes (LOINC)
  - Regenstrief Institute for Healthcare
- RxNorm
  - National Library of Medicine

Focus on recording clinical content

- International Classification of Diseases – Clinical Modification (ICD-9/10-CM)
  - World Health Organization
  - National Center for Health Statistics
- Current Procedural Therapy (CPT)
  - American Medical Association

Focus on reimbursement

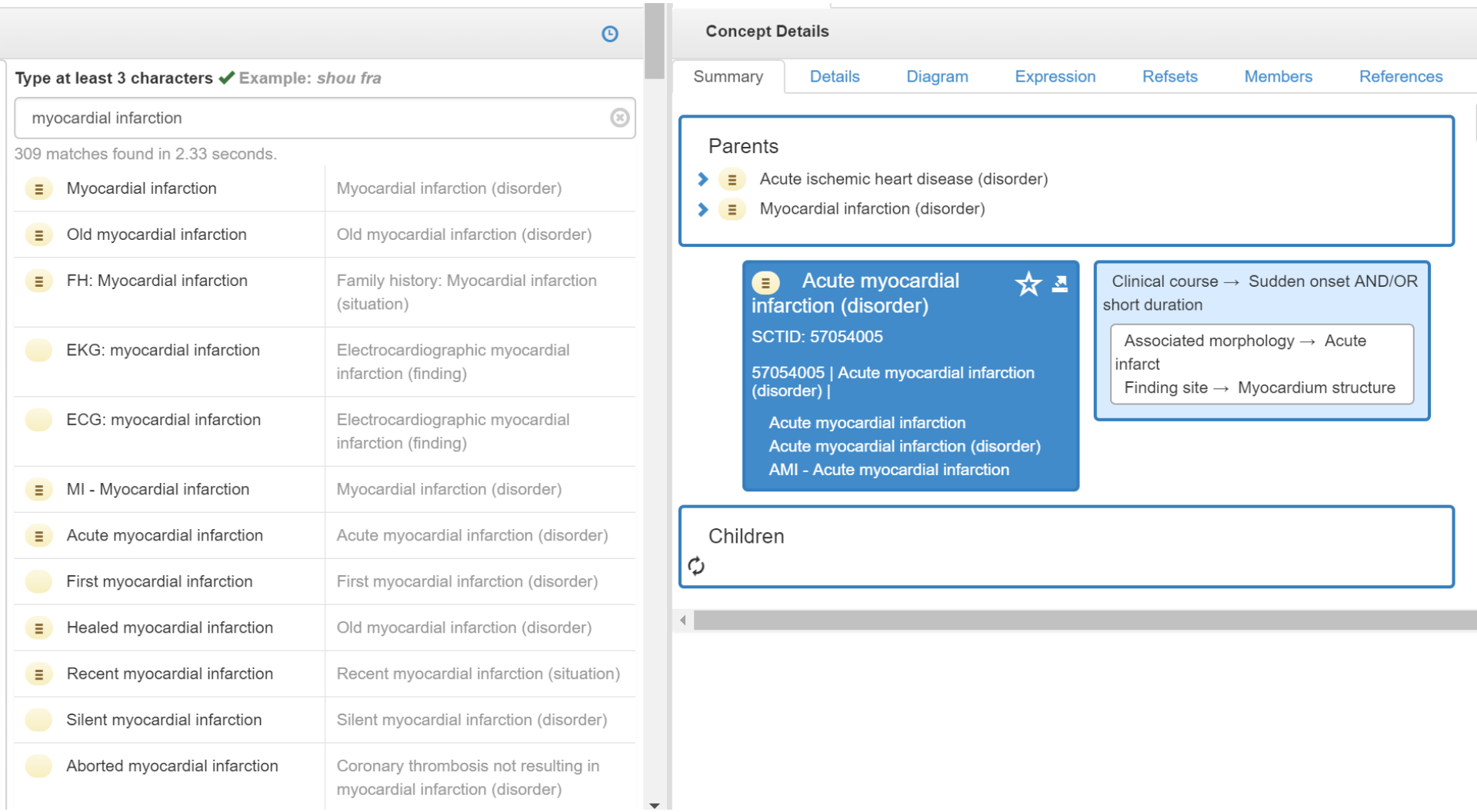
Search Term: myocardial infarction

SNOMED-CT

Returns 308 matches in 2.33 seconds

Term defined by pathologic, anatomic relationships

No clinical definition



TECHNOLOGY

The New York Times

# For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights

By STEVE LOHR AUG. 17, 2014

- ETL: extract, transform, load
- Mappings: syntactic & semantic
  - Map source data tables to destination data model
  - Map source terms → terminologies
  - Map of terminologies ← destination data model
  - Verification of preservation of semantics
- Repeat for every point to point connection
  - ETL not scalable

# How Registries Solve the Data Capture Problem

Home > NCDR > Registries > Hospital Registries > CathPCI Registry

NCDR CathPCI Registry®

## Standardized NCDR data elements and processes

The CathPCI Registry uses standardized data elements and definitions for:

- Patient demographics for diagnostic coronary angiography and percutaneous coronary intervention (PCI) procedures
- Patient history/risk factors, cath lab visit indications and coronary lesion information
- Provider and facility characteristics
- PCI Indications, lesion information, intracoronary device utilization and intra/post-procedure events
- 30-day and 1-year follow-up information on patients who had PCI

The registry supports a variety of data entry and submission options including certified third-party vendors and secure web-based entry. Data collection options

https://cvquality.acc.org/NCDR-Home/registries/hospital-registries/cathpci-registry

# How Registries Solve the Data Capture Problem

Home > NCDR > Registries > Hospital Registries > CathPCI Registry

NCDR CathPCI Registry®

Standardized NCDR data elements and processes

- Patient demographics for diagnostic coronary angiography and percutaneous coronary intervention (PCI) procedures
- Patient history/risk factors, cath lab visit indications and coronary lesion information
- Provider and facility characteristics
- PCI Indications, lesion information, intracoronary device utilization and intra/post-procedure events
- 30-day and 1-year follow-up information on patients who had PCI

The registry supports a variety of data entry and submission options including certified third-party vendors and secure web-based entry. Data collection options

https://cvquality.acc.org/NCDR-Home/registries/hospital-registries/cathpci-registry

# How Registries Solve the Data Capture Problem

Home > NCDR > Registries > Hospital Registries > CathPCI Registry

NCDR CathPCI Registry®

## Standardized NCDR data elements and processes

- Patient demographics for diagnostic coronary angiography and percutaneous coronary intervention (PCI) procedures
- Patient history/risk factors, cath lab visit indications and coronary lesion information
- Provider and facility characteristics
- PCI Indications, lesion information, intracoronary device utilization and intra/post-procedure events
- 30-day and 1-year follow-up information on patients who had PCI

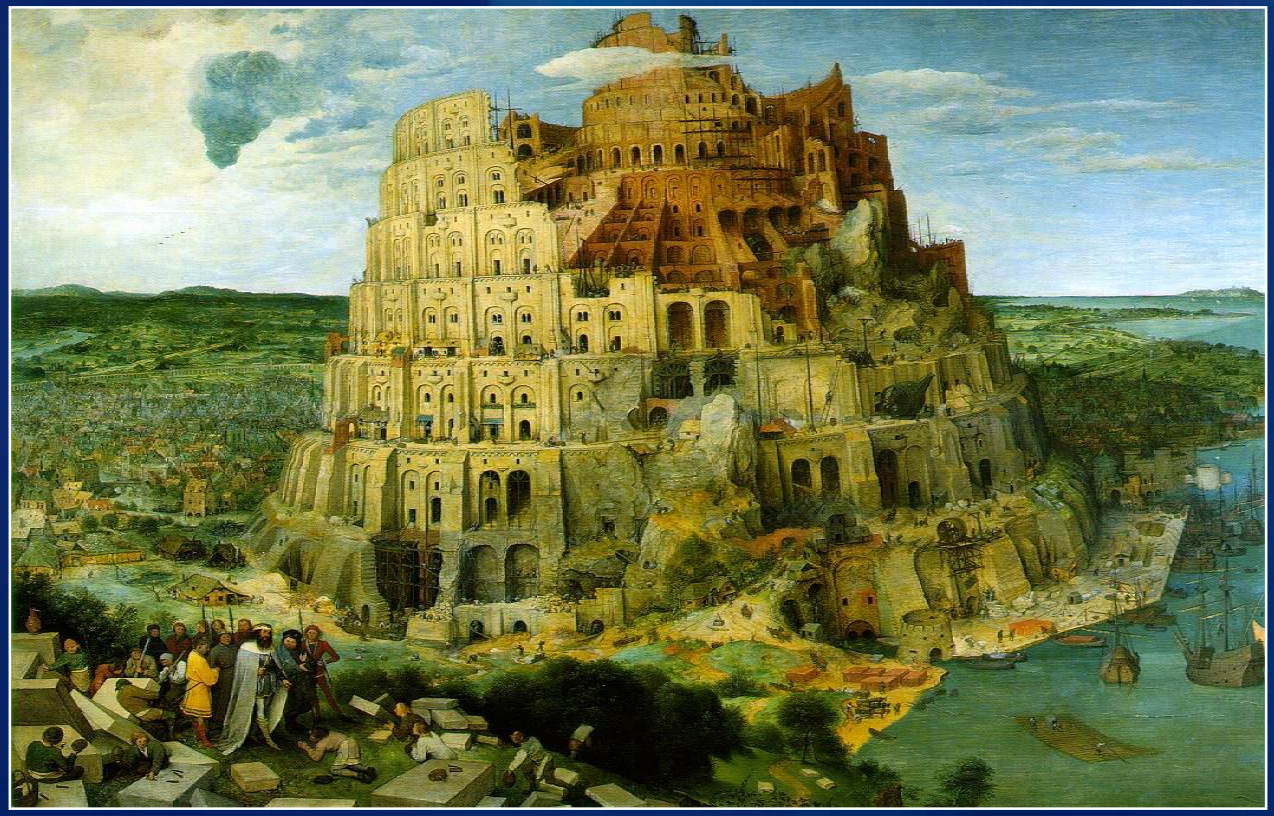
The registry supports a variety of data entry and submission options including certified third-party vendors and secure web-based entry. Data collection options

https://cvquality.acc.org/NCDR-Home/registries/hospital-registries/cathpci-registry

# Swivel Chair Interoperability *Wes Rishel*

@PaulLomax: The most unbelievable aspect of the Star Trek universe is that every ship they meet has compatible video conferencing facilities …

# THE Foundational Issue



Tower of Babel

*Pieter Bruegel the Elder and Pieter Bruegel the Younger, 1563*

# The Big Idea:
# *Native Data Interoperability, End to End*

- Defined (key) clinical concepts
- Key clinical concepts captured as data
- Specified representation of data in database systems
- Data capture integrated into workflow
- Capture once, use many times …
- And reduce / eliminate need for ETL!

# Project Goals

THE PEW CHARITABLE TRUSTS

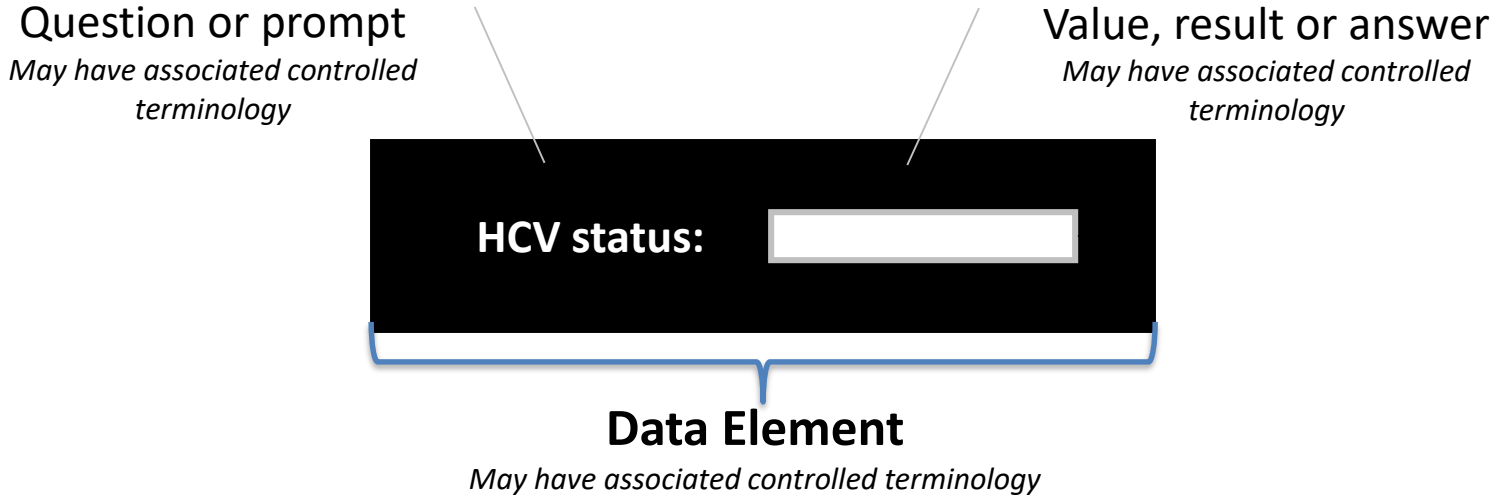
- Evaluate current state of registries
  - Identify common concepts shared across >20 registries
  - Assess use of data standards for those concepts
- Identify predicate work in CDE interoperability
  - Environmental scan
  - National common data models
- Create an implementation guide
  - All-in-one package of recommendations for database developers
  - Catalyze governance, structural, operational, and technical transformations

™Office Depot

# Methods

- Perform environmental scan
- Collect registry case report forms (CRFs), data dictionaries, data model representations
- Abstract common clinical concepts
- Determine concordance of data representations, use of data standards
  - Across registries
  - Across national common data models (OMOP, SENTINEL, PCORnet); FHIR representations
- Specify common data elements, key metadata
  - Clinicians
  - Database developers

# What is a Data Element?




- A data element is a question – value pair
- Considered the smallest meaningful unit of data exchange
- Formally defined in ISO/IEC 11179-1 and 11179-3
- Typically have a unique identifier, a definition, and valid values
- Interpretation requires context (e.g., date/time of collection, method of measurement, or person, place or thing to which the data pertains)

# Data standards are like toothbrushes:

Data standards are like toothbrushes:

*Everybody agrees we need them, but nobody wants to use anyone else's.*

Various attributions

# US Core Data for Interoperability (USCDI)

https://www.healthit.gov/sites/default/files/draft-uscdi.pdf

*Table 1: Draft USCDI Version 1 Data Classes*

| Draft USCDI Version 1 Data Classes | |
|---|---|
| 1. Patient name | 2. Sex (birth sex) |
| 3. Date of Birth | 4. Preferred Language |
| 5. Race | 6. Ethnicity |
| 7. Smoking Status | 8. Laboratory tests |
| 9. Laboratory values/results | 10. Vital signs |
| 11. Problems | 12. Medications |
| 13. Medication Allergies | 14. Health concerns |
| 15. Care Team members | 16. Assessment and plan of treatment |
| 17. Immunizations | 18. Procedures |
| 19. Unique device identifier(s) for a patient's implantable device(s) | 20. Goals |
| 21. Provenance | 22. Clinical Notes |

# USCDI – Relevant to Registries?

- Patient name
- Date of birth
- Race
- Smoking status
- Lab values / results
- ~~Problems~~
- ~~Medication allergies~~
- Care team members
- ~~Immunizations~~
- UDI
- ~~Provenance~~
- Sex
- ~~Preferred language~~
- Ethnicity
- Laboratory tests
- Vital signs
- Medications
- ~~Health concerns~~
- ~~Assessment / plan of rx~~
- Procedures
- ~~Goals~~
- ~~Clinical notes~~

# Ethnicity (Reg.CRF's)

| Data Element Name (CRF Label) | Permissible Values | Concordance |
|---|---|---|
| Ethnicity | Hispanic or Latino<br>Non Hispanic or Latino | 6 |
| Ethnicity | Hispanic of Latino<br>Not Hispanic or Latino Not Disclosed | 1 |
| Patient Ethnicity | Hispanic or Latino<br>Not Hispanic or Latino Patient declined to provide<br>Unknown | 1 |
| Ethnicity Type | Mexican<br>Mexican-Americano<br>Chicano<br>Puerto Rican<br>Cuban<br>Other Hispanic<br>Latino or Spanish Origin | 2 |
| Hispanic | No<br>Unknown<br>Yes | 1 |
| Hispanic or Latino Ethnicity | No<br>Yes | 2 |
| Hispanic Origin (maternal) | Mexican American<br>Chicano<br>Puerto Rican<br>Cuban<br>Other<br>Spanish/Hispanic/Latino<br>Hispanic, NOS | 1 |
| Is Patient of Hispanic Origin? | Yes<br>No<br>Unknown | 1 |
| Hispanic, Latino or Spanish Ethnicity | Yes<br>No<br>Not Documented | 1 |

# Example: Date of Birth (CDMs, FHIR)

| Date of Birth | | |
|---|---|---|
| **Data Element Field Name** | **Field Type** | **Concordance** |
| **Date of Birth** | **Date** | **2 (CCDS, CCRF)** |
| **Derived (year_ / month_ / day_of_birth) YEAR_OF_BIRTH, MONTH_OF_BIRTH, DAY_OF_BIRTH** | **Separate fields** | **1 (OHDSI)** |
| **Patient.birthDate** | **Date** | **1 (FHIR)** |
| **BIRTH_DATE** | **Date** | **2 (PCORnet, Sentinel)** |

# Key CDE Metadata (data about data)

1. Clinical concept label (human prompt – CRF, data entry screen)
2. Clinical definition
3. Clinical allowed values (human prompt – CRF, data entry screen)
4. Clinical allowed values definitions
5. Database field label
6. Database field data type / format (e.g., char, date, integer, values set)
7. Database field business rules (edit checks, range checks, etc.)
8. Database allowed values (as stored in db)
9. OID
10. Reference ontology concept binding
11. Reference ontology allowed values bindings
12. FHIR references (profiles, resources)
13. Other sources, references, notes

# Recommendation: Sex

1. Clinical concept label: Sex [Birth Sex, Sex (Birth Sex)]
2. Clinical definition: The biological sex of a patient, assigned at birth, not to be confused with the social construct of gender.
3. Clinical allowed values: F, M, UNK [Female, Male, Unknown]
4. Database field label: SEX, birthsex
5. Database field data type / format: Value Set – Char(3)
6. Database field business rules:
7. Database allowed values: F | M | UNK
8. Allowed values definitions: Female, Male, Unknown - a proper value is applicable, but not known. Includes ambiguous, variations of unknown, and variations of null.
9. Reference ontology concept: LOINC: LL3324-2, Sex assigned at birth
10. Reference ontology allowed values: LOINC: LA3-6, LOINC: LA2-8, LOINC: LA4489-6
11. FHIR references: https://www.hl7.org/fhir/us/core/StructureDefinition-us-core-patient.html; FHIR Resource: https://www.hl7.org/fhir/us/core/StructureDefinition-us-core-birthsex.html; Value Set: https://www.hl7.org/fhir/us/core/ValueSet-us-core-birthsex.html
12. Sources / references / notes: 2015 CCDS and USCDI, C-CDA Birth Sex observation

# Candidate Common Concepts → CDEs

7 As Is (more or less)

- Patient name
- Date of birth
- Sex
- Race
- Ethnicity
- Procedures
- UDI

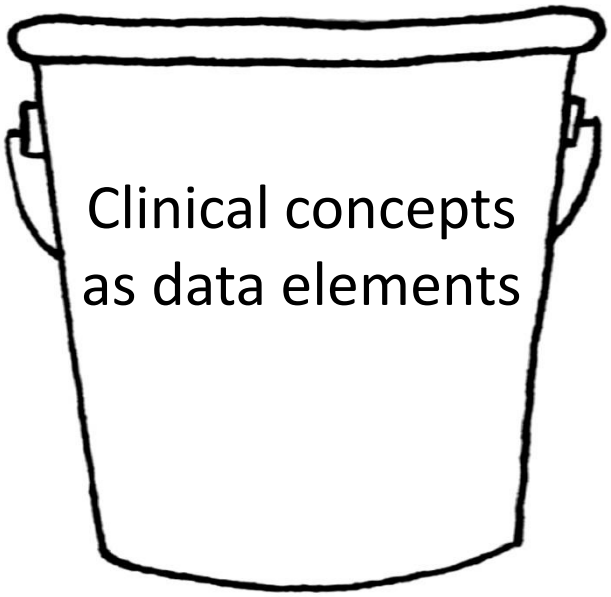
8 Adjusted (select modifications)

- Vital signs: height, weight, BP, pulse
- Lab results (via model)
- Medications (via model)
- Care team: only doctor
- Smoking status (via model)
- *EtOH use
- *Substance abuse
- *Vital status (death)

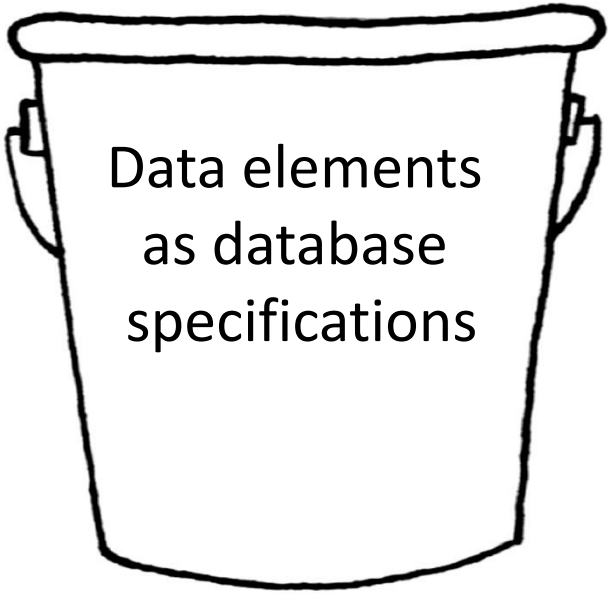
https://dcri.org/registry-data-standards

*not in USCDI

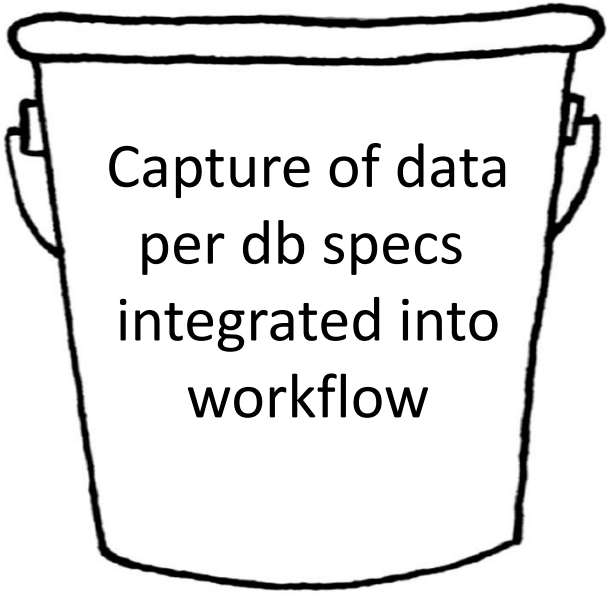
# Steps to Native Data Interoperability




Professional societies
Academic consortia
FDA




Informatics modeling
Regulation (ONC, ASC X12)
HIT vendors

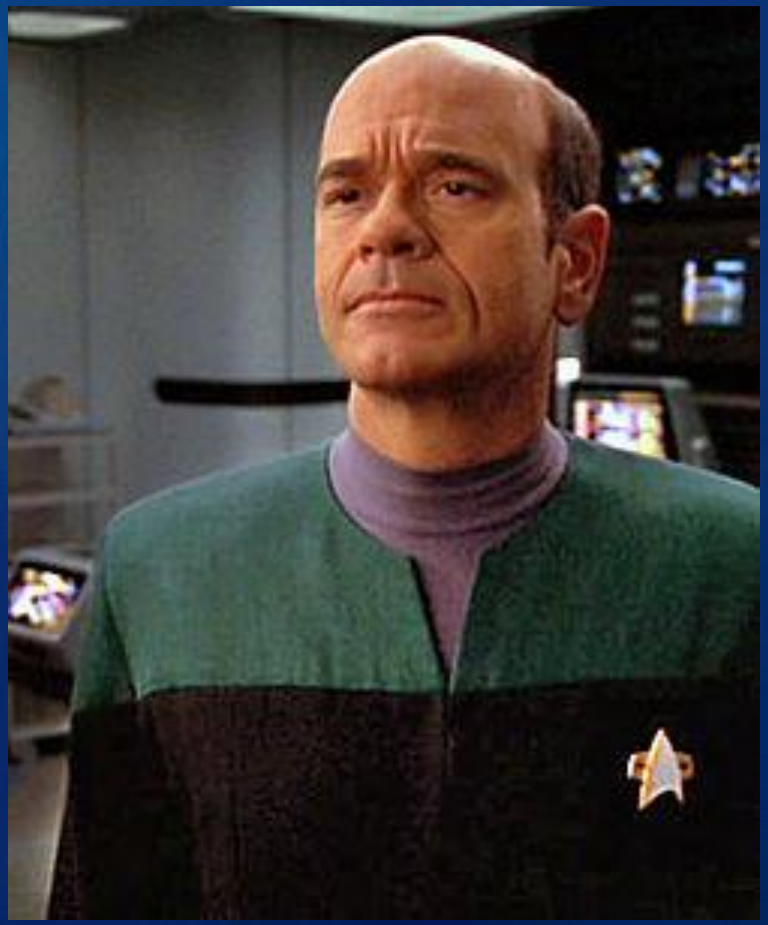



HIT vendors
Healthcare entities
Professional societies

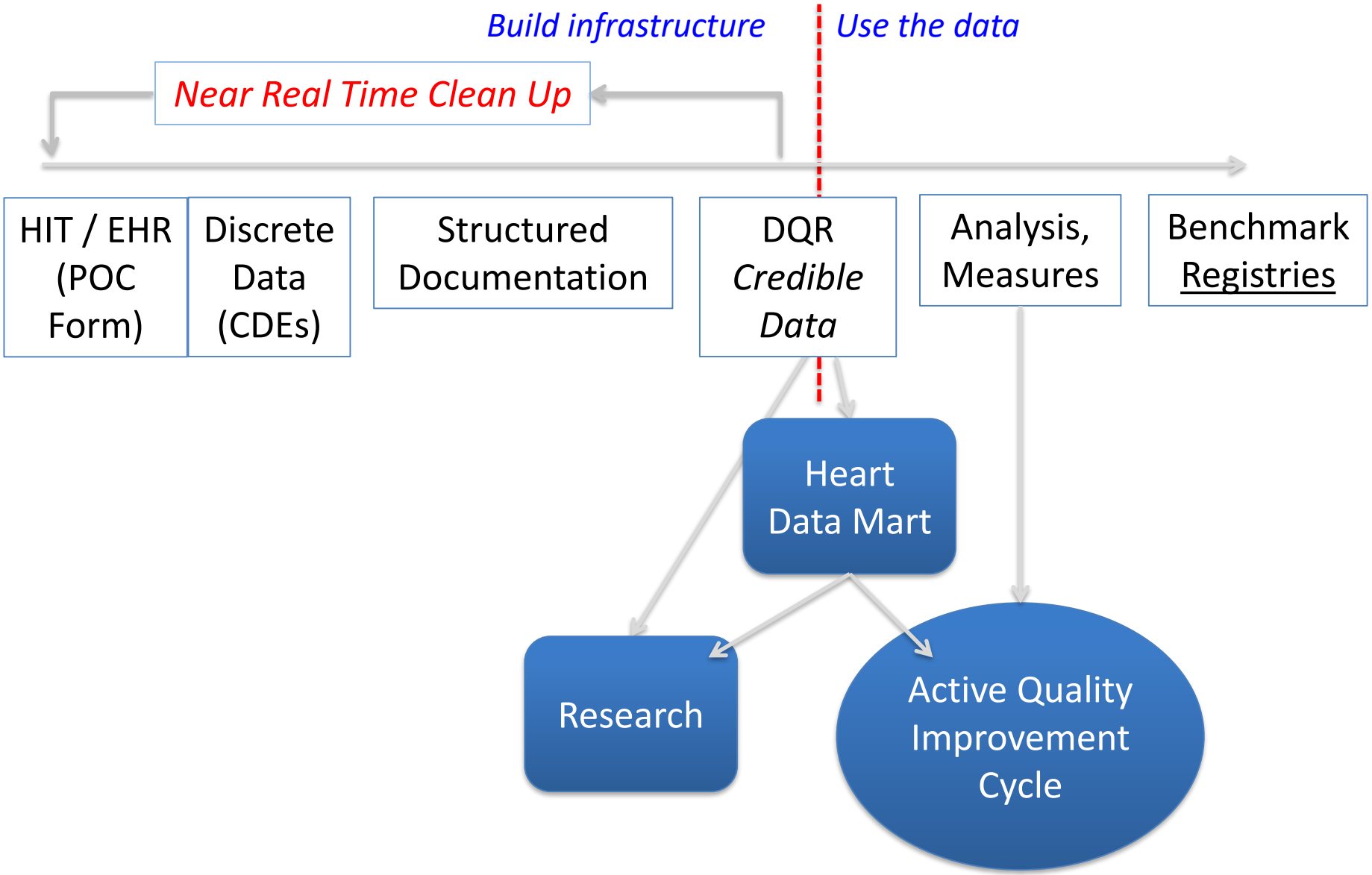
# FDA Coordinated Registry Networks

- Orthopedics (joint replacement) - ICOR
- Vascular intervention – VISION (RAPID)
- Cardiovascular disease – CDCRN (TAVR, etc.)
- CIEDs – EP PASSION
- Prostate ablation – SPARED
- Robotics
- Women's Health Technology
- Hernia repair
- Neurology (stroke intervention) – DAISI
- Breast implants – NBIR
- GI (bariatric devices) – CATNIP
- TMJ
- Venous infusion catheters – VANGUARD

# “Dammit, Jim, I’m a Doctor, Not a Computer!”



# Duke Heart Center - Dataflow End State



# Concurrent Data Capture: Key Concepts

- Capture data once, use many times
- Directed data capture, relevant (pertinent) charting, charting by exception
- Distributed data capture, integrated into workflow
- Team-based documentation
- Data persistence, data liquidity
- Data compilation into views (reports)
- Semantic interoperability
- = Structured reporting

# Interoperability Loci

- Clinical care ↔ Registries ↔ Research ↔ Reporting
  - Common, cross-registry / EHI data elements
  - Minimum core (domain-specific) data elements
  - Quality and outcome measures (typically summative)
  - UDI: reference data in GUDID, AUDI databases
- Data transfer, representation
  - HL7 v2+, FHIR
- Common data models (generic data aggregation)
  - SENTINEL, PCORNet, i2b2, OMOP OHDSI
- Analytics
  - Data aggregation and analysis
  - Distributed analysis

# Is Healthcare Changing for the Better …

## *The Common Denominator*

Clinical documentation
Administrative reporting
Clinical decision support
Quality and performance
Analytics, research
Device safety, surveillance
Machine learning, AI
Big Data
Etc., etc., etc.

# From Concepts to Action

*Creating the ecosystem …*

- Registry Community – core clinical CDEs
  - Technical (database) representation for implementation across registries
- FDA - Coordinated Registry Networks
- ONC - USCDI open comment period
- Informatics – terminology modeling
  - HL7 Common Clinical Registry Framework
  - Modeling – Clinical Information Modeling Initiative
- Clinical Community – structured reporting!

# *Thank You!*

james.tcheng@duke.edu

Visit the Project website:

https://dcri.org/registry-data-standards