Office of the National Coordinator for Health Information Technology

Data Quality Symposium

Standards and Implementation

ONC Tech Forum: Data Quality Symposium

High data quality is essential to every aspect of healthcare, research and epidemiology, public health reporting, and realizing the potential of health IT, and is a key ingredient to data aggregation, linkage and interoperability. It is accessible, usable, and actionable, whereas low data quality, such as duplicate patient records or obsolete information creates barriers to care delivery, research insights, and leads to higher costs and inefficiencies.

This symposium will review actionable data quality standards, assessment frameworks, metrics, and real-life challenges and success stories. Symposium attendees will gain knowledge and access to best practices and available solutions that they may apply in their everyday work.

12:25 – 12:45 PM EST

ISO 8000

ISO 8000 is the international standard for data quality that is concerned with principles of data quality, characteristics of data that determine its quality, frameworks for data quality, and processes to ensure data quality.

Peter Benson, Founding and Executive Director of the Electronic Commerce Code Management Association (ECCMA), ISO-8000 Project Leader.

Benefits of Quality Data

❖ Faster, more accurate and more reliable analysis, decision-making, and operational efficiency

❖ Reduces risks associated with incorrect or misleading information

❖ Compliance with regulatory requirements is easier, reducing the risk of costly penalties

Benefits of Portable Data

❖ Allows users to switch between different applications and platforms without losing access to their data

❖ Promotes collaboration by allowing multiple applications access to data

❖ Allows users to decide, how data is stored, and with whom it is shared, enhancing privacy and security

Code Management Association

[US-DE.BER:3031657]

dba

ECCMA

(Electronic Commerce Code Management Association)

www.eccma.org

If you are a US organization (company, university or government agency), you should consider joining the USTAG to ISO

❖ 501(c)(6) not-for-profit standards development organization formed in 1999 dedicated to improving the quality of data and digital transformation through the application of international standards

❖ ANSI appointed administrator of US Technical Advisory Group (USTAG) to ISO for:

  o ISO TC 184 - Automation systems and integration
  o ISO TC 184 SC4 - Industrial data
  o ISO TC 184 SC5 - Interoperability, integration, and architectures for enterprise systems and automation applications

❖ Project leader for:

  o ISO 22745 – Open technical dictionaries
  o **ISO 8000 – Data quality**
  o ISO 25500 – Supply chain interoperability and integration

Data Quality Standards
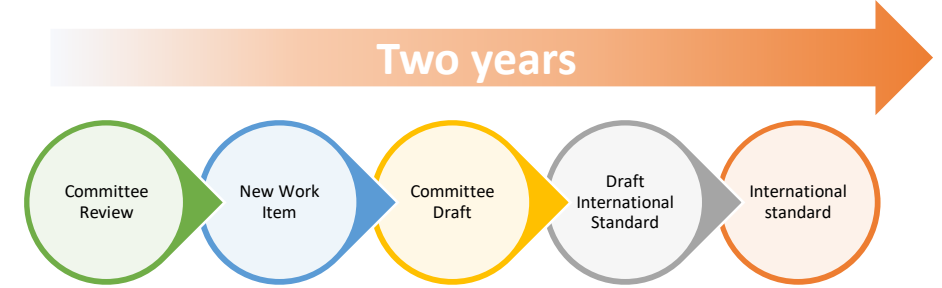
Standard	Scope	Status
ISO 8601	**Formatting date and time**	**International Standard**
ISO 22745	Open technical dictionaries and concept encoding	International Standard
ISO 8000-120	Formatting master data with provenance	International Standard
ISO 8000-115	**Formatting quality identifiers (PO numbers, part numbers, asset numbers)**	**International Standard**
ISO 8000-116	**Formatting Authoritative Legal Entity Identifiers (ALEI)**	**International Standard**
ISO 8000-117	Formatting identifiers in blockchains and secure links to off chain data	International Standard
ISO 8000-051	Formatting data governance policies	International Standard
ISO 8000-114	Interoperable Data Format (.idf)	Draft International Standard
ISO 8000-118	Natural Location Identifiers (NLI)	Committee Draft
ISO 8000-119	Transport Unit Identifiers (TUID)	New Work Item
ISO 25500-1	Supply Chain Interoperability and Integration - Overview	Committee Review
ISO 25500-2	Supply Chain Interoperability and Integration - Vocabulary	Committee Review
ISO 25500-110	Supply Chain Interoperability and Integration – Verification of Certificates	Committee Review

What is ISO 8000 Quality Data?

quality

degree to which a set of inherent characteristics of an object fulfils requirements
ISO 9000:2015

data quality

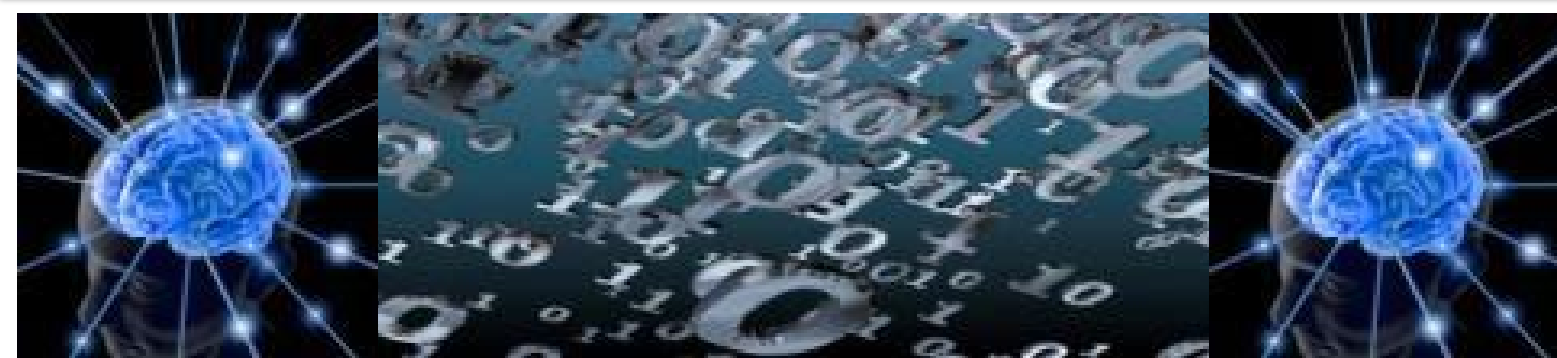degree to which a set of inherent characteristics of data fulfils requirements
ISO 8000-2:2022

ISO 8000 quality data goes beyond data profiling, the data must be syntactically and semantically explicit **and** it must meet <u>stated requirements</u>.

ISO 8000 quality data must be "fit for purpose"
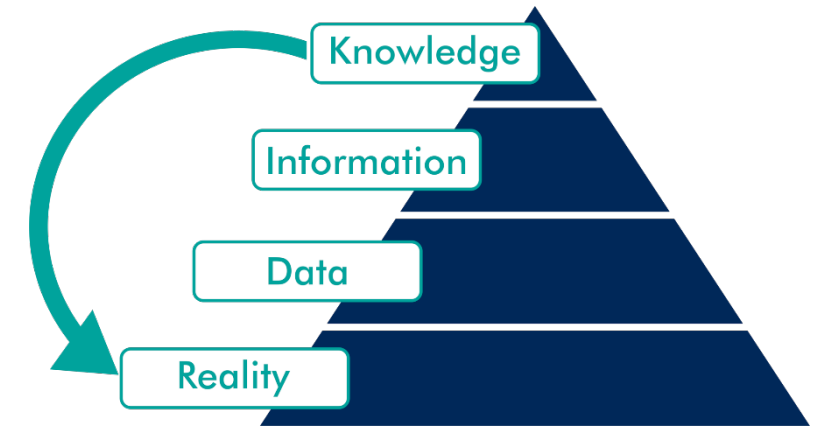
Information → **Data** → **Information**

Copyright only covers "fixed form" (*you can only copyright data not information*)

ISO definitions (ISO/IEC 2382-1:1993)

Information

knowledge concerning objects, such as facts, events, things, processes or ideas, including concepts, that within a certain context has a particular meaning
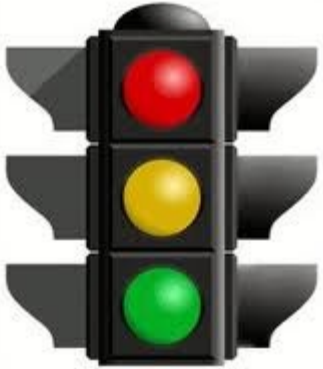
Data

Re-interpretable representation of information in a formalized manner suitable for communication, interpretation, or processing
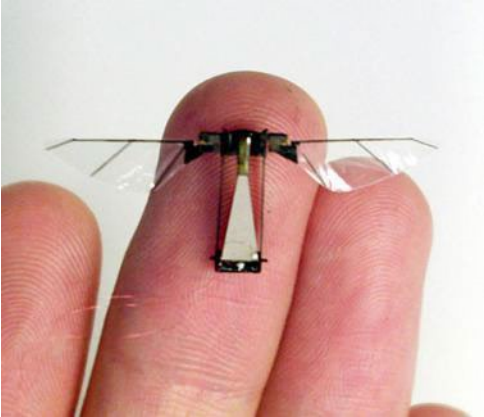
> data: fixed form into which information is transformed so that it can be stored or moved.
>
> Peter R. Benson

In South Africa a traffic light is called a robot!

Robot

Dictionary

A dictionary is an index of terms with provenance, pronunciation and other terminology used to explain <u>concepts</u>

Cap

1. a soft flat hat without a brim sometimes having a visor.

2. a protective lid or cover for an object such as a bottle, the point of a pen, or a camera lens.
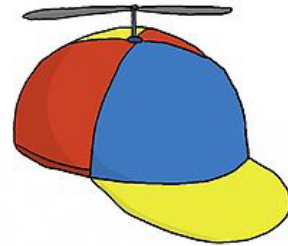
Terms   Abbreviations   **Terminology**   Definitions   Images

Two different approaches to building a dictionary?

British Model
Crowd Sourcing
It is what <u>you</u> say it is

French Model
Committee
It is what <u>we</u> say it is

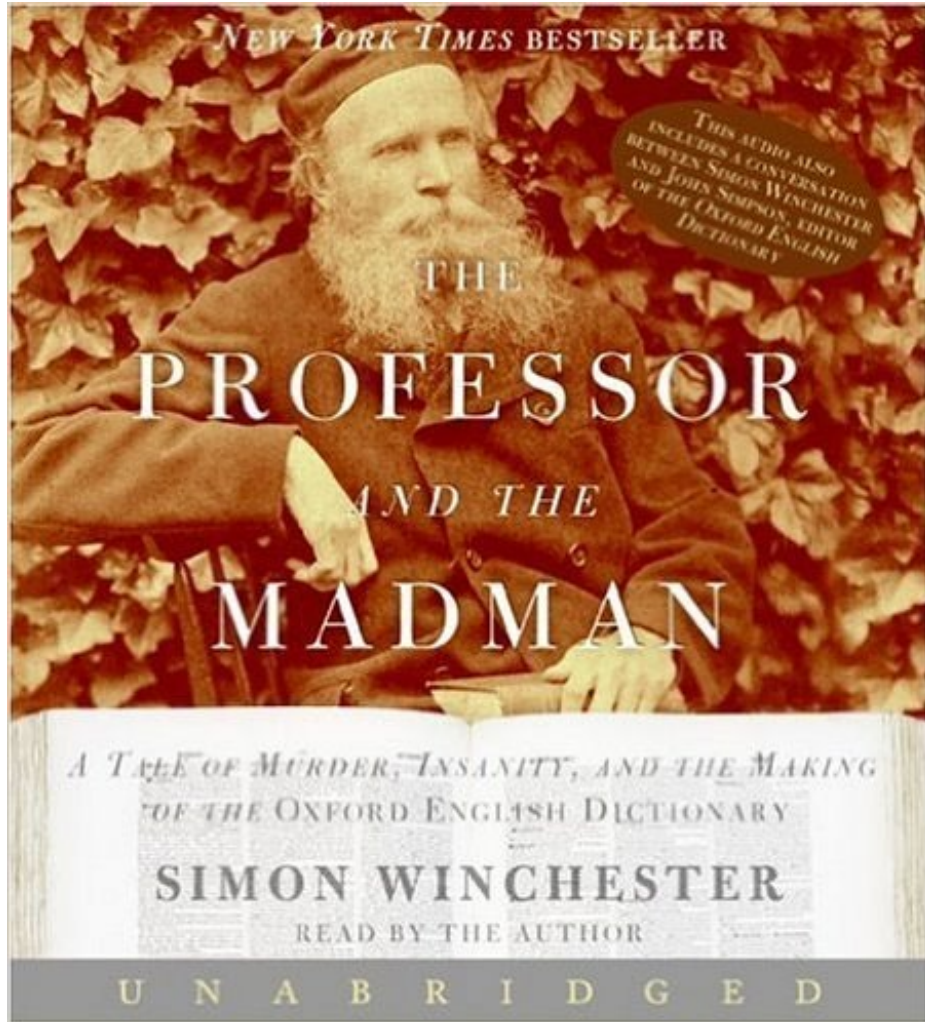The eOTD (ECCMA Open Technical Dictionary) is an ISO 22745-10 compliant central registry of terminology. Each concept and terminological component in the eOTD is assigned a **unique and permanent public domain identifier**.

Users create their corporate dictionaries as subsets of the eOTD and use the eOTD concept identifiers to manage **concept equivalence mapping.**

ISO 22745 - ECCMA Open Technical Dictionary (eOTD)

Terminology

| Terms | Abbreviations | Definitions | Images |

Public Domain Concept Identifier
ECCMA.eOTD:0161-1#xx-xxxxxx#1
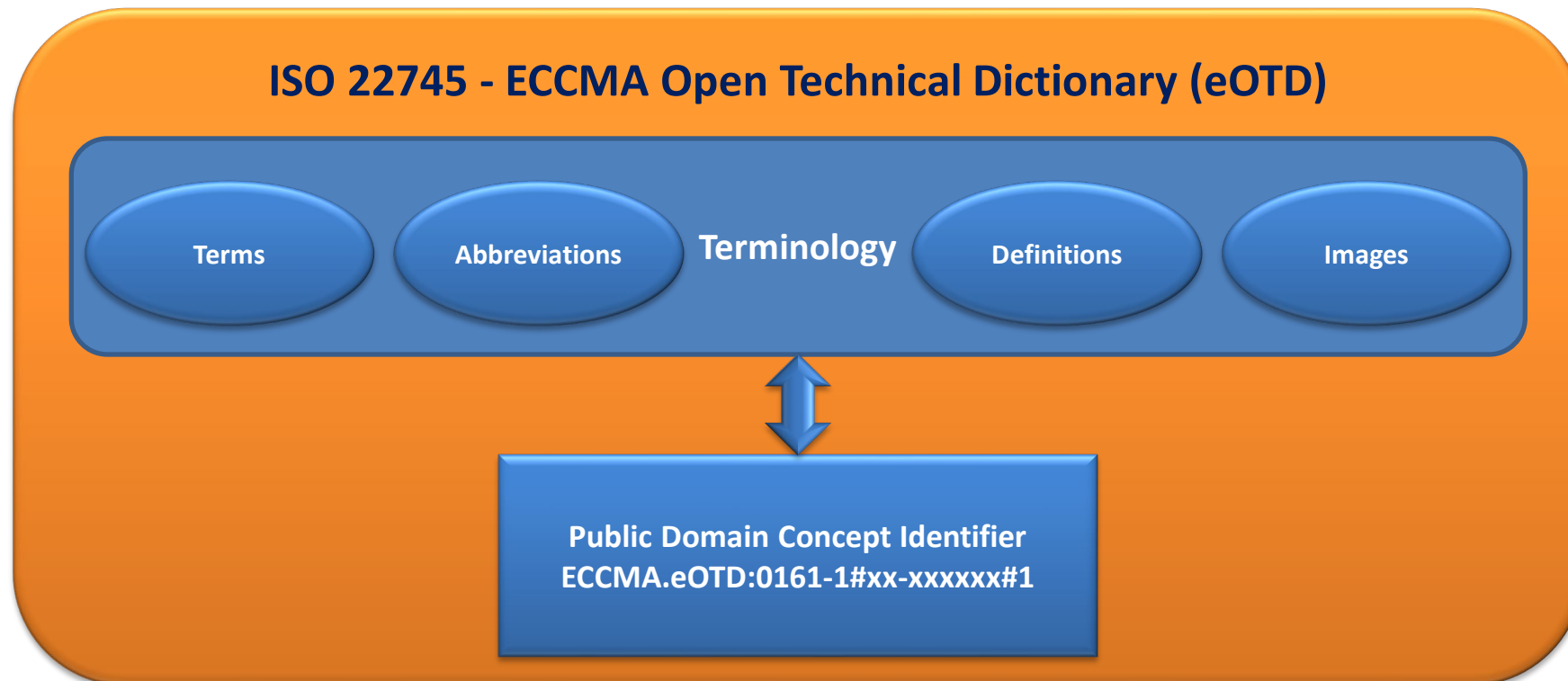
Just as with music notation and engineering symbols, concept identifiers are simply used to communicate more accurately in a language independent environment

Music

Engineering

- Diode
- Capacitor
- Inductor
- Resistor
- DC voltage source
- AC voltage source
- And gate
- Nand gate
- Or gate
- Nor gate
- Xor gate
- Inverter (Not gate)
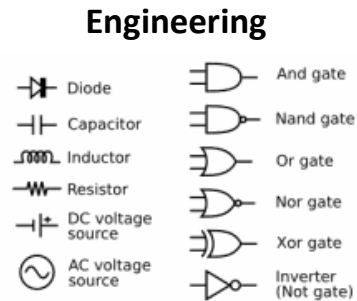
eOTD

A unique public domain identifier is assigned to a concept.

ECCMA.eOTD:0161-1#01-089388#1   **table**

ECCMA.eOTD:0161-1#01-086445#1   **chair**

ECCMA.eOTD:0161-1#02-018635#1   **weight**

ECCMA.eOTD:0161-1#02-005808#1   **length**

ECCMA.eOTD:0161-1#07-277660#1   **Monday**

ECCMA.eOTD:0161-1#05-001122#1   **kilogram**

Concept encoding using the eOTD

Spanner

Wrench

0161-1#01-89846#1

0161-1#01-89846#1

eOTD

0161-1#01-89846#1
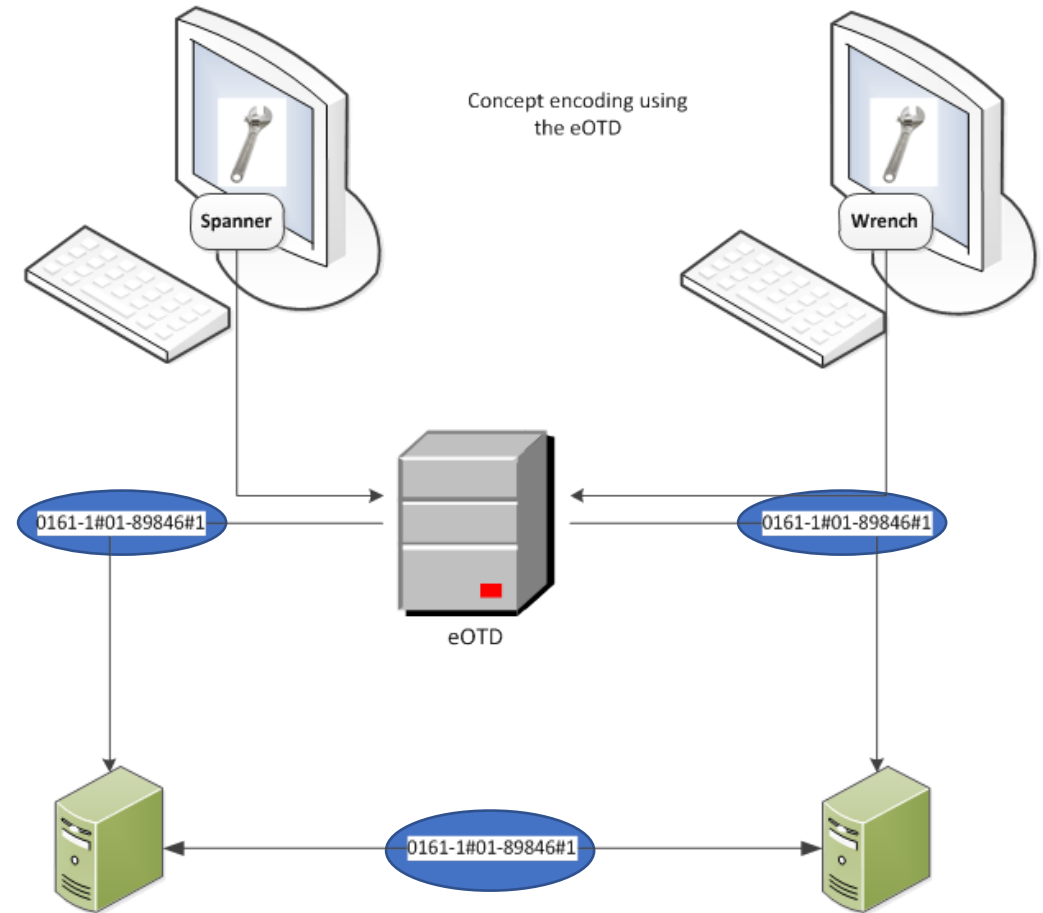
Terminology referenced in a standard model
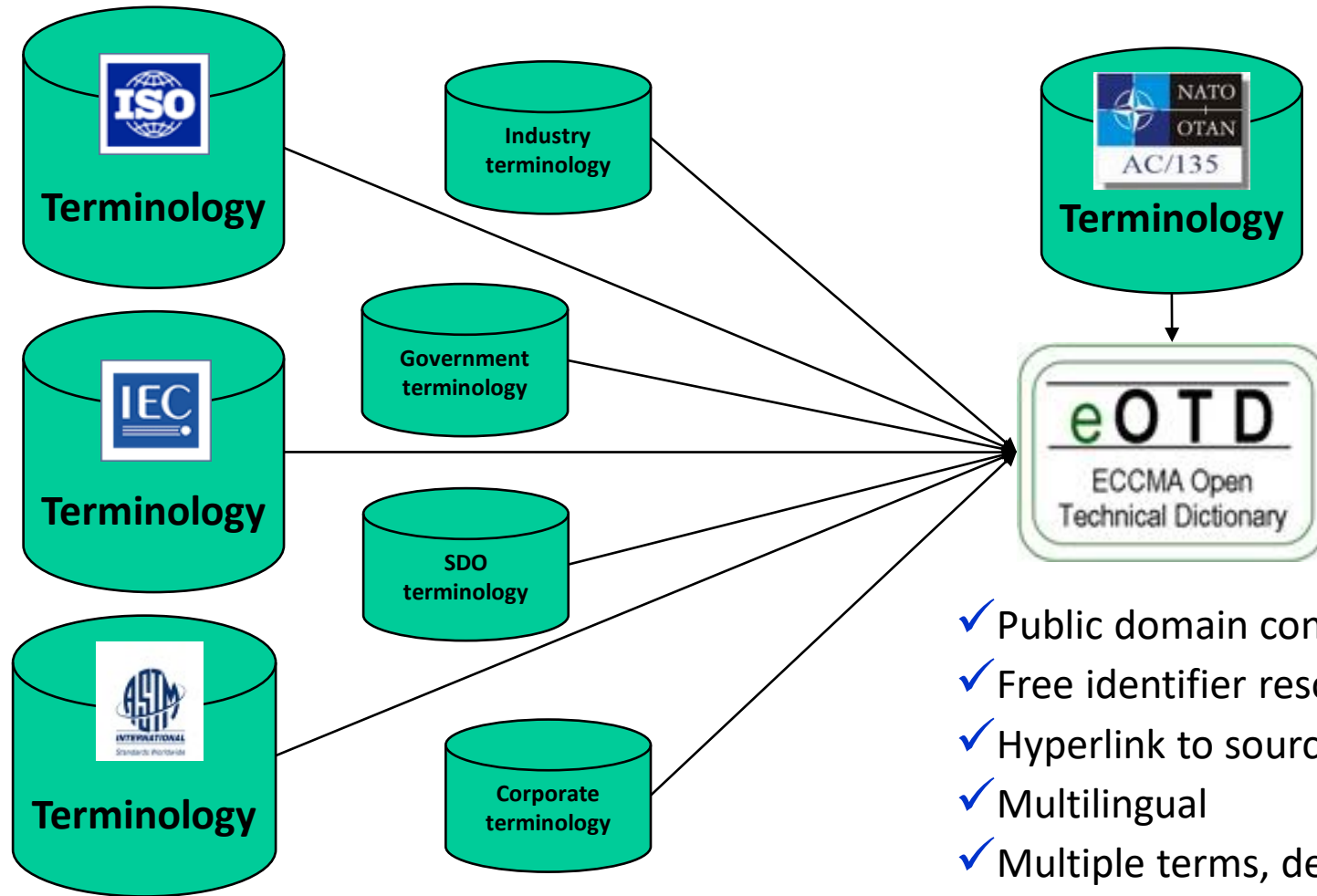
- ✓ Public domain concept identifiers
- ✓ Free identifier resolution to underlying terminology (web services)
- ✓ Hyperlink to source standards
- ✓ Multilingual
- ✓ Multiple terms, definitions and images linked to single concept identifier

Applying Concept Equivalence to facilitate semantic interoperability

SABIC

concept equivalence table

01-1073082	=	01-086142
01-1073082	=	01-068756

Sabic Bolt = Rockwell Bolt

Sabic Bolt = ASTM bolt

| eOTD Concept ID	Term	Ex ref	Originating Organization	Definition	Status
ECCMA.eOTD:0161-1#01-1073082#1	BOLT	IR237	SABIC	A fastener that is externally threaded on one end and generally with some style of head on the other end and is normally intended to be tightened or released by torquing a nut and designed to fasten objects together.	Active
Equivalent concepts					
ECCMA.eOTD:0161-1#01-086142#1	BOLT	-	Rockwell Automation, Inc	A fastener consisting of a threaded pin or rod with a head at one end and designed to be inserted through holes in assembled parts and secured by a mated nut that is tightened by applying torque. Крепежная принадлежность, представляющая собой стержень с нарезанной резьбой и головкой на одной стороне, предназначенный для помещения в отверстия собираемых вместе деталей с последующей их фиксацией с помощью гайки, затягиваемой до определенного крутящего момента.	Active
ECCMA.eOTD:0161-1#01-068756#1	bolt	F 1789 - F16	ASTM	headed and externally threaded fastener designed to be assembled with a nut	Active

ECCMA Master Data Conceptual Model
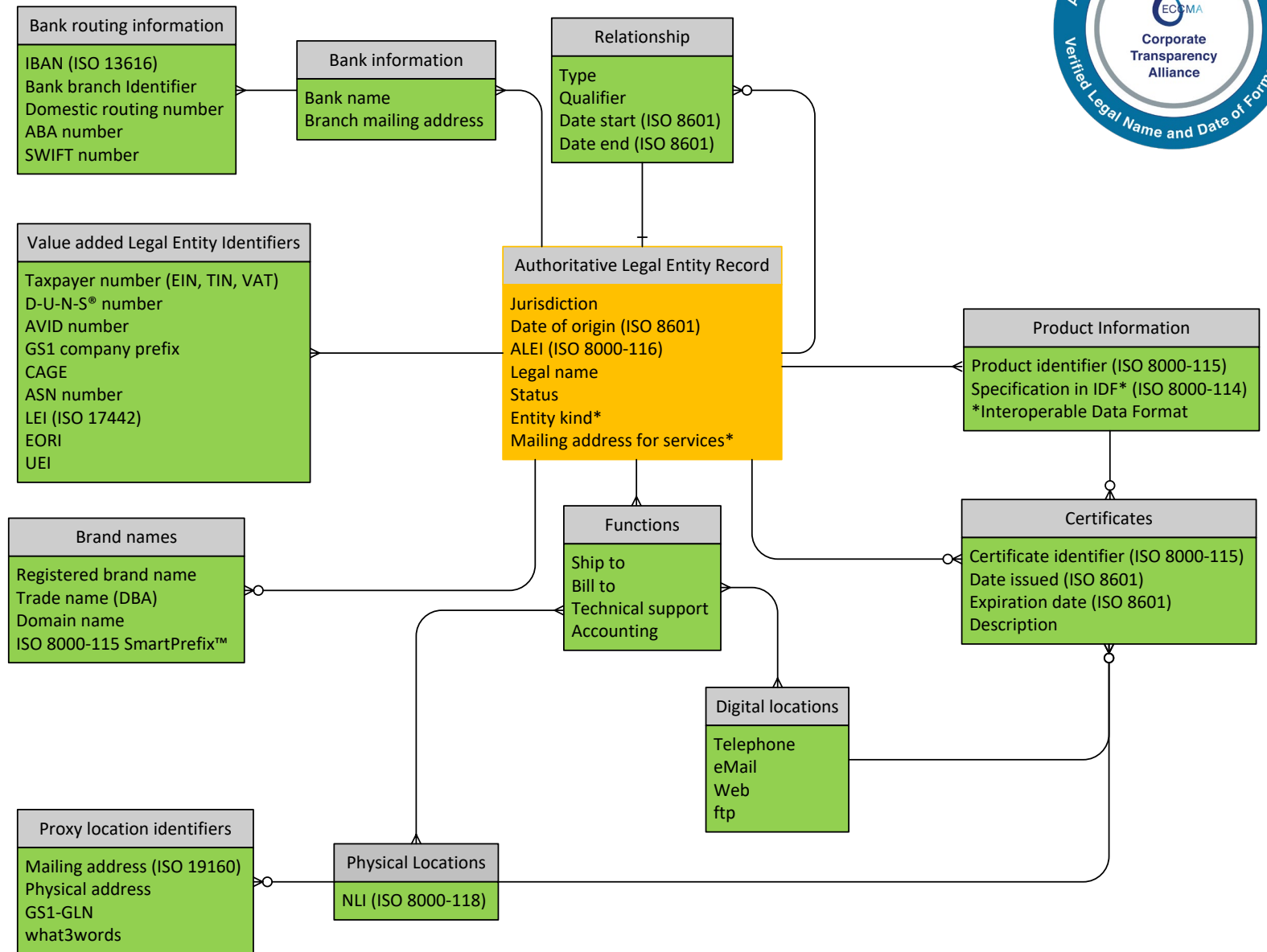
ISO 8000 Data Quality Assessment

eDQA

Reviews the quality of data and recommends improvements to maintain ISO 8000 quality data.

www.eccma.org

ISO 8000 defines data quality as **portable** data that meets **stated requirements**

Objectives
Provide an analysis of current data quality and recommend changes to data governance policies to maintain ISO 8000 quality data.

Includes:

- Development of an initial corporate dictionary as a subset of the ECCMA Open Technical Dictionary (eOTD),

- Development of an initial library of data requirements as a subset of the ECCMA Data Requirement Registry (eDRR) including initial rules for creating consistent item names and descriptions

- Data validation plan (data cleansing and data enriching)

- Data governance policies required to maintain ISO 8000 quality data.

Thank You

Peter R Benson | **Executive Director**
ECCMA | **www.eccma.org** | **ISO 8000 ALEI: [US-DE.BER:3031657]**
Data Driven Strategic Sourcing - Digital Transformation Powered by International Standards
O:**+1.610.851.4290** | M:**+1.610.462.5923** |
peter.benson@eccma.net
EST/EDT (-05:00 UTC/-04:00 UTC)
https://www.linkedin.com/in/peterrichardbenson/

Data Usability Taking Root

OVERVIEW

An initiative co-sponsored by AHIMA®

the sequoia® project

The Sequoia Project's Role

The Sequoia Project is a trusted, independent advocate for nationwide health information exchange.

Supports multiple independent initiatives, each with their own mission, governance, membership and structure.

SECURE	INTEROPERABLE	NATIONWIDE

Evolution of The Sequoia Project

2012

Sequoia Formed
April 2012

2014

Rebranded to
The Sequoia Project

Launched
Carequality

2018

Corporate Restructure

Transition to
Three Operating Entities

Interoperability Matters
Launches

2019

Launched
Info Blocking Compliance Work Group

Selected as
ONC TEFCA RCE

2020-22

Launched Three New Work Groups
Public Health

Data Usability

Stakeholder Engagement

Info Blocking Compliance Boot Camp

2023

Launched New Work Group

Consumer Voices

Launched
Data Usability Taking Root

Renewed as
ONC TEFCA RCE

It's one thing to get health data to the right place at the right time; it's quite another to make sure that data is complete and useful.

Cross-industry Guidance for Data Quality & Usability

3 years in the making

2 years of public input

266 engaged organizations

378 subject matter experts

It's time for this guidance to **take root**.

What is the difference between the **Data Usability Taking Root Movement** and the **Data Usability Workgroup**?

V3.0

V2.0

V1.0

Guidance for Data Provenance & Traceability of Change

Guidance for Data Integrity, Format and Trust

Guidance for Effective Use of Codes

Guidance for Data Tagging/Searchability

Guidance for Reducing the Impact of Duplicates

Guidance for Effective Use of Narrative for Usability

Data Usability **Workgroup**
Develops Guidance

Deployment **Guidance**

Taking Root **Movement**
Implements Guidance
An initiative co-sponsored by **AHIMA**

Development **Feedback**

268 Organizations **443 Participants**

Healthcare Providers 20%

Health IT Developers 18%

Other 15%

HIN/HIEs 13%

Federal, State, Local Government 13%

Consumer/Patient 5%

Standards Developer 4%

Health Plan/Payer 10%

Public Health 2%

Community of Practice

Roundtables

Technical Assistance

Testing Platform

In-person Convenings

Participation Levels

IMPLEMENTER   SUPPORTER   SPONSOR

25
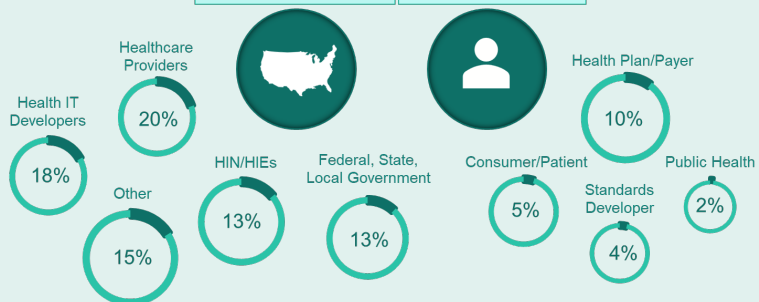
Why Join Our Data Usability Initiative?

Because…

- More complete quality data improves outcomes.

- Better data leads to better and timely decisions.

- High quality and usable data are more actionable

- Reduces clinician burden.

- Guidance promotes consistency across technologies.

- Practical, incremental improvements simplify implementation.

- Addresses a common challenge across all actors.

- It's the right thing to do.

Pragmatic Guidance

the sequoia® project

V1.0 Implementation guidance on clinical content for information exchange

- provider-to-provider
- provider-to-public health
- healthcare entity-to-consumer

Guidance for
Data Provenance & Traceability of Change

Guidance for
Data Integrity, Format and Trust

Guidance for
Effective Use of Codes

Guidance for
Data Tagging / Searchability

Guidance for
Reducing the Impact of Duplicates

Guidance for
Effective Use of Narrative for Usability

The Value of Useful Data

Public Health

- **Provenance & Traceability of Changes:**
  - A public health can leverage provenance to distinguish administered vaccines from a later recording of an externally sourced vaccine to identify duplicate data.

- **Effective Use of Codes:**
  - Allows patient history in the Individual Medical Management System (IMMS) or Vaccine Action Command and Coordination System (VACCS) to be exchanged consistently with codes.
  - Provides guidance for SARS-CoV-2 LOINC terms for COVID results, immunizations dose unit and lot number to improve semantic interoperability of data exchanged.
  - Makes immunization information distinct from patient or other party reports.

- **Data Tagging / Searchability:**
  - Public health organization can search data that pertains to certain criteria such as diagnosis code.
  - Enables inclusion of Clinical Notes and Document Narrative Linking.

The Value of Useful Data

- **Provenance & Traceability of Changes:**
  - Allows implementers to focus on consistency and presentation of provenance metadata in document sections.
- **Effective Use of Codes:**
  - Enables clinical decision support, graphing and trending of lab data.
- **Reducing the Impact of Duplicates:**
  - Known duplicates are identifiable between documents exchanged using consistent identifiers.
- **Data Integrity, Format and Trust:**
  - Improves patient matching with use of consistent patient demographics.
  - Enables a complete picture of a patient's history with use of patient summary and encounter documents to convey the complete patient story.
- **Data Tagging / Searchability:**
  - Enables a provider to search for document titles pertaining to certain criteria, i.e., diagnosis code.
- **Effective Use of Narrative for Usability:**
  - Enables consistent inclusion and linking of clinical narratives and notes with discrete data to provide mechanisms for clinicians to view and support better clinical decision making.

Provider

The Value of Useful Data

Vendor

- **Provenance & Traceability of Changes:**
  - Allows vendors/developers to focus on consistency and presentation of provenance metadata in document sections.
- **Effective Use of Codes:**
  - Enables clinical decision support, concept-based search, graphing and trending of lab data within platforms.
- **Reducing the Impact of Duplicates:**
  - Known duplicates are identifiable between documents exchanged using consistent identifiers within platforms.
- **Data Integrity, Format and Trust:**
  - Improves patient matching with use of consistent patient demographics for data sent and received.
  - Supports a complete picture of a patient's history with use of patient summary and encounter documents to convey the complete patient story.
- **Data Tagging / Searchability:**
  - Enables a vendor platform to search for document titles which pertain to certain criteria, i.e., diagnosis code.
- **Effective Use of Narrative for Usability:**
  - Enables vendor platforms to have consistent inclusion and linking of clinical narratives and notes with discrete data to provide mechanisms for clinicians to view and support better clinical decision making.

An initiative co-sponsored by AHIMA

The Value of Useful Data

Consumer

- **Provenance & Traceability of Changes:**
  - Allows consumers to understand and filter on organization and date/time of data captured during their journey.
- **Effective Use of Codes:**
  - Enables clinical decision support, graphing and trending of lab data.
  - Enables patient reported vaccines to be included in immunizations shared among providers.
- **Data Integrity, Format and Trust:**
  - Enables the consumer to search and find data more easily with use of consistent patient demographics.
  - Enables the consumer to find and share their complete patient story with new providers.
- **Data Tagging / Searchability:**
  - Enables a consumer to search for data related to certain criteria, such as diagnosis code.
- **Effective Use of Narrative for Usability:**
  - Provides value to the consumer by including them in the clinical reasoning and thoughts of the authoring provider.

The Value of Useful Data

- **Provenance & Traceability of Changes:**
  - Allows HIEs and HINs to focus on consistency and presentation of provenance metadata in on-demand created document sections.

- **Reducing the Impact of Duplicates:**
  - Known duplicates are identifiable between documents exchanged using consistent identifiers within platforms.

- **Data Integrity, Format and Trust:**
  - Improves patient matching with use of consistent patient demographics for data sent and received.
  - Supports a complete picture of a patient's history with use of patient summary and encounter documents to convey the complete patient story.

- **Data Tagging / Searchability:**
  - Enables an HIE/HIN platforms to search for document titles which pertain to certain criteria, i.e., diagnosis code.

- **Effective Use of Narrative for Usability:**
  - Enables HIE/HIN platforms to have consistent inclusion and linking of clinical narratives and notes with discrete data to provide mechanisms for clinicians to view and support better clinical decision making.

HIEs and HINs

The Value of Useful Data

- **Provenance & Traceability of Changes:**
  - Allows operational staff to leverage provenance metadata to improve data searchability and audits.
- **Effective Use of Codes:**
  - Enables clinical decision support, graphing and trending of lab data to make IT system more useful.
- **Reducing the Impact of Duplicates:**
  - Known duplicates are identifiable between documents exchanged using consistent identifiers.
- **Data Integrity, Format and Trust:**
  - Improves patient matching with use of consistent patient demographics.
  - Enables a complete picture of a patient's history with use of patient summary and encounter documents to convey the complete patient story.
- **Effective Use of Narrative for Usability:**
  - Enables consistent inclusion and linking of clinical narratives and notes with discrete data to provide mechanisms for clinicians to view and support better clinical decision making.

Operational
&
HIM Staff

Putting Guidance Into Practice

- **Identify where to start**
  - Which V1.0 sections are priorities?
  - Which can be done quickly?
  - What is the timeframe?
- **Track progress**
  - Potential self-reported score card promotes transparency and healthy competition
  - # elements supported
  - % of customers supporting
- **Incremental approach**
  - Enables rollout in conjunction with other IT projects
  - Elevates data usability for all IT projects - UAP
- **Other Considerations**
  - Leverage for governmental programs (e.g., EHR certification, USCDI, TEFCA, etc)
  - Address as part of Data Usability Round Table

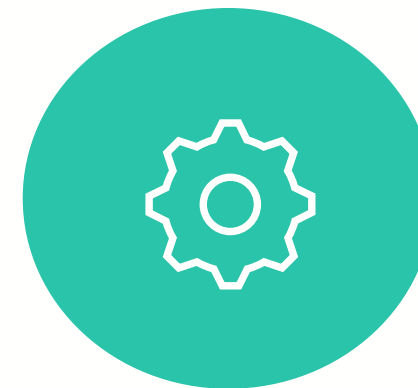Participants choose their own implementation pathway and pace…

Implementation Enablers

Technical Assistance + **Community of Practice** + **Testing Platform Services**

What makes this distinctive

Data Usability Guidance leverages existing standards to address pain points from end users on the frontline.

- The universal benefit of this work cannot be achieved in isolation.

- This work empowers diverse actors to affect change.

- The industry is entwined in interdependencies.

- When there is strength in numbers, momentum will accelerate.

- Collective action will solve a shared pain point.

- Those that adopt early will have first mover advantage.

- Practical focus can inform future versions of USCDI.

- Model of continuous improvement of data quality.

Data Usability Taking Root

Supporter

Pledges to support the data usability movement as a member of the data usability community of practice. Grants right to Sequoia to include logo in its Taking Root member directory. Participates in Data Usability Roundtables. Supporters that are also Sequoia members are invited to Taking Root Summits.

Implementer

Pledges to implement V1.0 data usability guidance across one or more topics within a defined timeline. Invited to participate in the data usability community of practice, the Data Usability Taking Root Planning Committee, and the Taking Root Summits. Grants right to Sequoia to include logo in its Taking Root member directory.

Sponsor

Pledges to provide sponsorship of the Taking Root Summit(s). Socializes and evangelizes the purpose and power of this work. Co-hosts Taking Root Summits and participates in The Data Usability Taking Root Planning Committee, Roundtables, and Summits.

Levels of Engagement

Roadmap

2023
- Early Supporters for V1
- Round Table
- Taking Root Summit

2024
- Expand participation; develop V2 to include FHIR
- Community of Practice
- Technical Assistance
- Implementation begins
- Movement grows

2025
- Community of Practice expands
- Technical Assistance expands
- Conformance Testing
- Movement grows

Join the Movement Now!

https://sequoiaproject.org/data-usability-taking-root-movement/
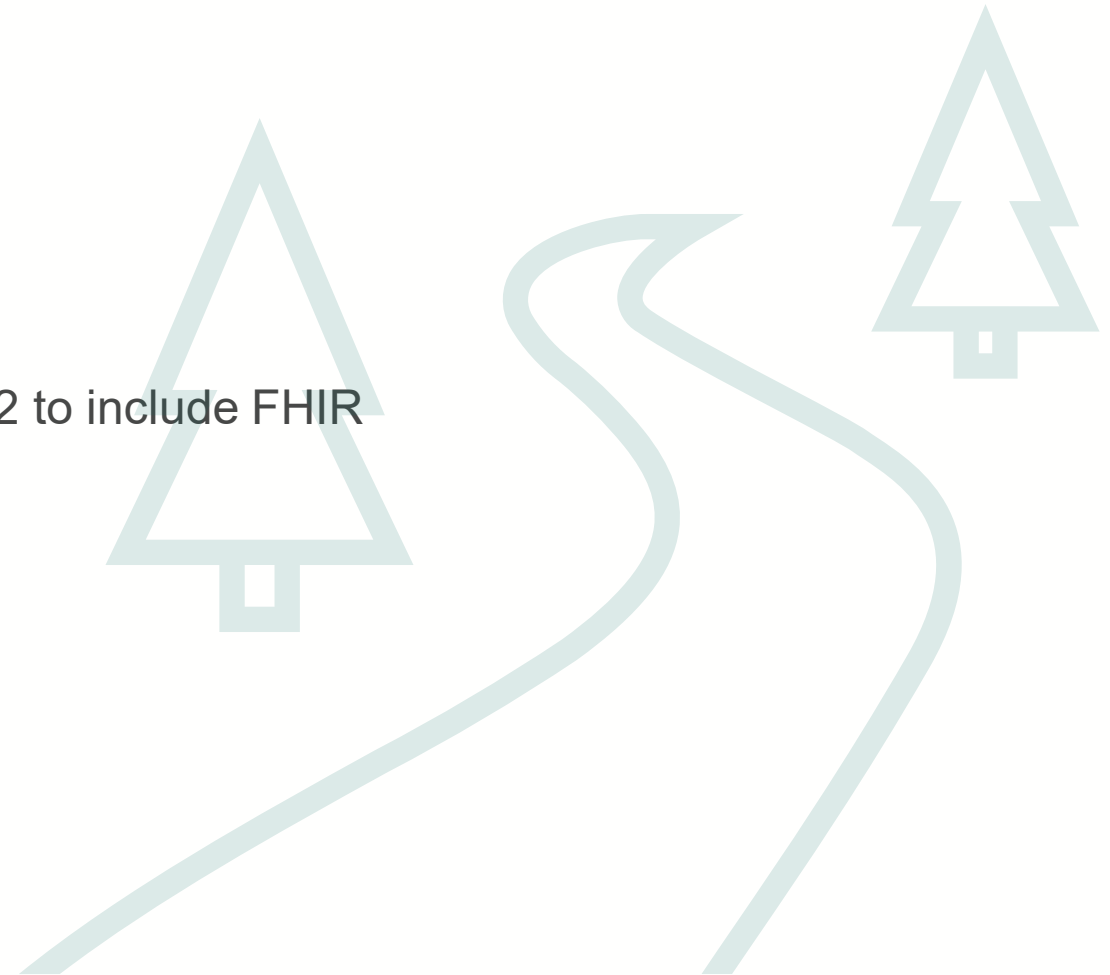
Contact Us

Thank you for your interest in The Sequoia Project's new **Data Usability Taking Root** Initiative for V1.

[Follow the development of V2.0 guidance](#)

If you would like to get in touch you can reach us at:

takingroot@sequoiaproject.org

Clinician & Health Information Network Perspective

Steven Lane, MD, MPH
Chief Medical Officer
Health Gorilla

Perspective

- Primary care physician, clinical informaticist and EHR user x > 30 years

- Early implementer of EHR (1990s), patient portal (2000), interoperability (2008-)

- Interoperability Champion:
  California HIEs, regional > national vendor network implementation, Direct messaging, eHealth Exchange, Sequoia Project, Carequality framework, FHIR accelerators
  ONC HITAC: USCDI, Information Sharing, TEFCA

- Health Gorilla:
  Joined as CMO one year ago to support the development of QHIN, implementation of new TEFCA exchange purposes not consistently supported by current exchange, and engagement of new interoperability participants.
  - Nationwide health information network and data sharing platform
  - One of seven HINs approved for onboarding as a TEFCA QHIN
  - Committed to improving the quality of exchanged data to optimize outcomes and value

Clinical Challenges

- **Data Explosion:** Growing data exchanges, trillions of transactions yearly, and petabytes of clinical data.

- **Integration Complexities:** Seamlessly integrating external data into workflows.

- **Data Quality Issues:** Inaccuracies, missing elements, and duplicates hinder decision-making.

- **Structured vs. Unstructured:** Balancing unstructured text data with structured entry.

- **Normalization:** Aligning data with evolving standards and terminologies.

- **Diverse Users:** Meeting varied data needs across user groups.

Goals of Interoperability

Five Rights	Healthcare Quintuple Aim

Five Rights

1. Right data
2. Right format
3. Right time
4. Right user
5. Right supporting workflows

Data Quality !!!

Healthcare Quintuple Aim

1. Improve health of the population
2. Lower per capita cost of healthcare
3. Improve patient experience
4. Improve provider/care team experience
5. Advance health equity

The need for high quality health data

- While many participants in the health care ecosystem remain frustrated by a lack of data access, others struggle to make optimal use of the large volume of data received in multiple formats, by multiple channels, from multiple sources.

- The quality of interoperable health data is widely variable.

- Different motivation for senders and receivers. Data sources may be driven primarily by data sharing **compliance** requirements. Recipients and users need to be able to **trust** the data.

- Quality deficiencies limit the value and utility of data for all healthcare participants and can contribute to medical errors, poor outcomes, burden, costs, and inefficiencies for individuals and populations.

- There are no established standards for defining, evaluating, quantifying, or managing the quality of data created, accessed, exchanged and used across the healthcare ecosystem.

Definition of Data Quality

- "Fitness for use/purpose"
  - Whose use, in what context?
    - Individuals / patients
    - Clinicians / providers
    - Public health / registries
    - Healthcare payers
    - Other insurers
    - Researchers
    - Health IT developers
    - Government agencies / policy makers

Data Science: Many dimensions of data quality

- **Accessibility** – *how readily the data can be located and accessed*

- **Accuracy** – *the degree to which data represent the things they were designed to measure*

- **Completeness** – *the percentage of data populated – records, attributes, reference values, truncation*

- **Coherence** – the degree to which the data are logically connected and mutually consistent

- **Conformity** – *the degree to which data values of the same attributes are represented uniformly*

- **Consistency** – *how well data aligns with another data set*

- **Credibility** – *The confidence the user places in the data*

- **Currency** – *how well the data reflects the true current state*

- **Integrity** – *the degree to which a defined relational constraint is implemented between two data sets*

- **Interpretability** – *the ease with which the user may understand and properly use and analyze the data*

- **Precision** – *the degree to which the data has been rounded or aggregated*

- **Relevance** – *the value contributed by the data*

- **Timeliness** – *the time lag between the described event and the appearance of relevant data in the data set*

- **Uniqueness** – *the degree to which the same data is represented multiple times*

- **Usability** – *the degree to data meets the needs of the user*

- **Validity** – *the closeness of data value to predetermined/expected values*

What Healthcare Needs

- Standard set of health data quality categories, dimensions and metrics
  - ➤ Describe and manage the data we use every day

- Efficient data correction
  - **Standardized Feedback:** A consistent process for informing data sources about problems
  - **Source-Level Correction:** Fix data quality issues at the source
  - ➤ Ensure cleaner data at the origin, enhancing efficiency and effectiveness.

Healthcare Data Quality Taxonomy
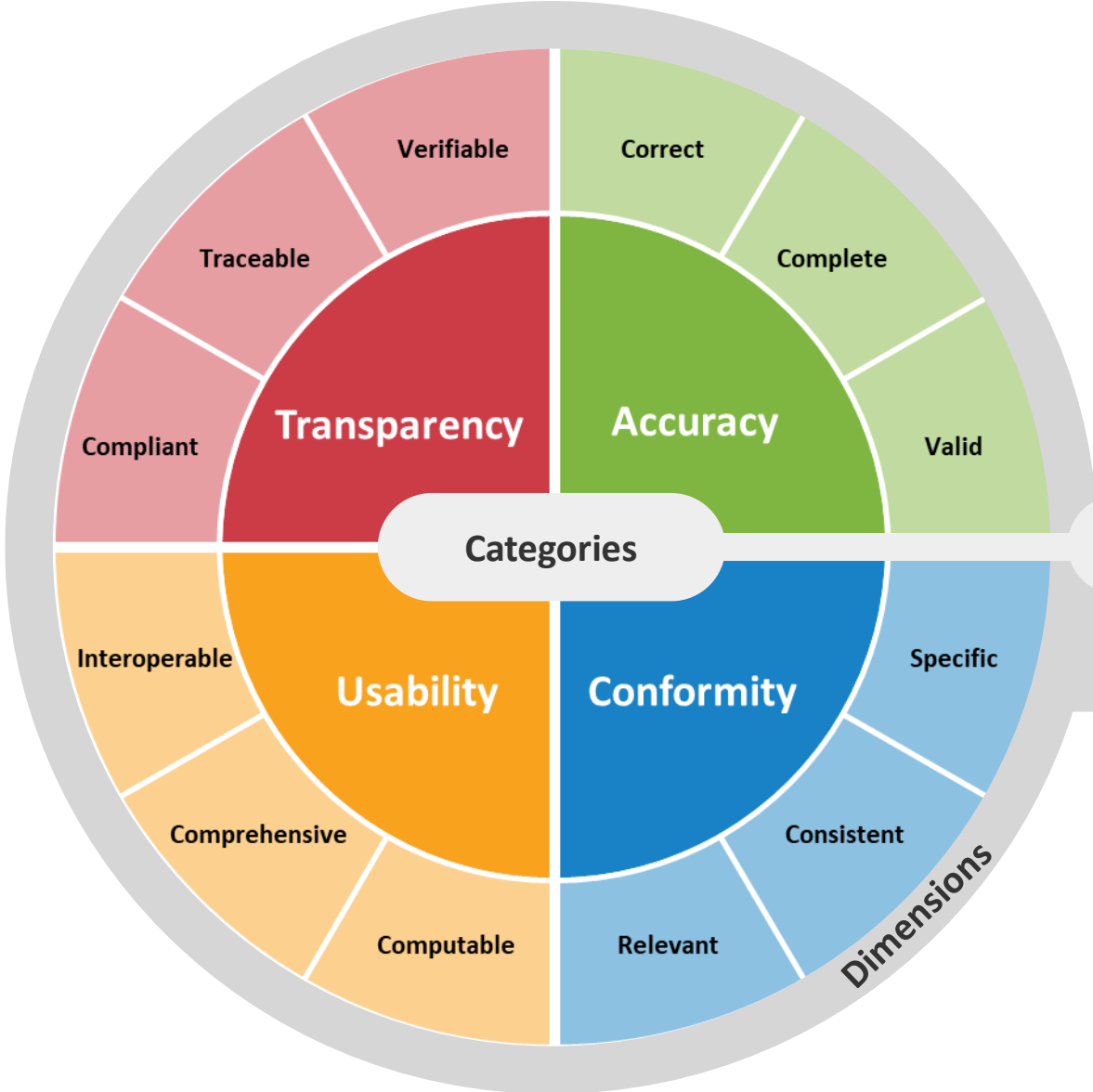
The **HDQT** starts with a simple 2-level taxonomy

The top level represents measure categories

The second level represents measure dimensions

Taxonomy of Healthcare Data Quality

Data Quality Categories

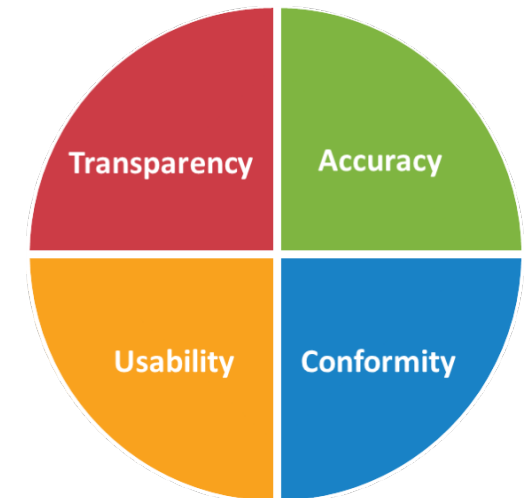**Usability**   Focuses on the ability of the data to be meaningfully leveraged by software applications

Accuracy   Focuses on the fundamental quality of the data itself.

Conformity   Focuses on assessment of the data within the context of the class, entity and use case.

Transparency   Focuses on whether the data can be trusted and ethically accessed.

Taxonomy of Healthcare Data Quality

Usability

Computable	The ability of healthcare data to be inherently usable by software applications.
Interoperable	The information can be exchanged and used across different systems and organizations in accordance with an interoperability profile
Comprehensive	The information has necessary, current data classes and elements for a given use case

Accuracy

Correct	The correctness and precision of healthcare data
Complete	The information has all necessary data elements and fields in the proper location
Valid	The information is in the correct data class and represented in a way that is valid according to the nature of the information

Conformity

Specific	The level of detail at which data is captured be fit for clinical purpose
Relevant	The information is appropriate and necessary to support the current use case
Consistent	The information does not create contradictions with respect to other elements or attributes in the current data

Transparency

Verifiable	The information includes meta-data that identifies its time, place, system and author of origin
Traceable	The path the information took to get from the origin to the receiver and any additions or remediations that were made in the process
Compliant	The information can be ethically and legally used without violating standards, including patient privacy regulations like HIPAA

Opportunities for HHS to advance Health Data Quality

- Support for industry initiatives:
  - VHA
  - Sequoia Project
  - Leading Edge Acceleration Projects (LEAP) in Health IT award: Boston Children's *CumulusQ Project*:
    - ➢ Iterative process to comprehend and assess the quality of structured and unstructured USCDI elements
    - ➢ Benchmark snapshots of data quality with root cause analysis of data anomalies

- ONC Health IT Certification

- Information Blocking rules

- TEFCA QHIN Technical Framework

- CMS Promoting Interoperability incentives

Health Gorilla's Data Quality Initiative – Goals

- **Initial Focus:** Assess Lab and Medication data classes

- **Completeness & Conformance:** Check for USCDI data element presence and vocabulary adherence

- **Deficiency Analysis:** Identify and quantify data quality issues

- **Automation vs. Human Intervention:** Determine fixable issues – automatically or with human help

- **Efficiency Methodology:** Create processes for fixing data both "at rest" and "in motion"

- **Sustainable Model:** Develop a long-term plan for expanding data quality efforts
  - ➢ All / evolving USCDI (USCDI+) data classes/elements/standards
  - ➢ Evolving clinical and data science

Priorities for USCDI data class evaluation

Foundational

- Patient Demographics / Information
- Data Provenance

High Priority

- *Laboratory*
- *Medications*
- Problems
- Allergies and Intolerances
- Immunizations

Medium Priority

- Procedures
- Vital Signs
- Clinical Tests
- Diagnostic Imaging
- Clinical Notes

Lower Priority

- Care Team Members
- Encounter Information
- Facility Information
- Goals and Preferences
- Health Insurance Information
- Health Status Assessments
- Medical Devices
- Patient Summary and Plan

Data Elements and Vocabulary Standards (V3, *V4*)

- **Laboratory Data Class**

1. Tests (LOINC)
2. Values/Results (SNOMED-CT®)
3. Specimen Type (SNOMED-CT®)
4. Result Status
5. *Result Unit of Measure (UCUM)*
6. *Result Reference Range (UCUM)*
7. *Result Interpretation (SNOMED-CT®, HL7)*
8. *Specimen Source Site (SNOMED-CT®)*
9. *Specimen Identifier*
10. *Specimen Condition Acceptability (SNOMED-CT®)*

- **Medications Data Class**

1. Medications (RxNorm +/- NDC)
2. *Dose*
3. *Dose Unit of Measure (UCUM)*
4. *Indication (SNOMED CT®, ICD-10-CM)*
5. *Fill Status*
6. *Medication Instructions*
7. *Medication Adherence*
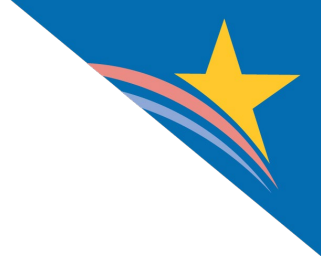
Evaluation data set

- 3 cohorts derived from Health Gorilla's ~25M patient database
  - Two geographically-based, one representing a specialty provider organization
  - All data de-identified

- Laboratory results data:
  - 10M laboratory result records from 92,000 patients
  - Derived from both laboratory result interface messages and treatment-based queries of EHR data via the Carequality, CommonWell and eHealth Exchange networks

- Medication data:
  - 600K medication records derived from 55K patients
  - Derived from network queries

After the break Charlie Harp from Clinical Architecture will review the findings of an initial quality assessment of Health Gorilla data
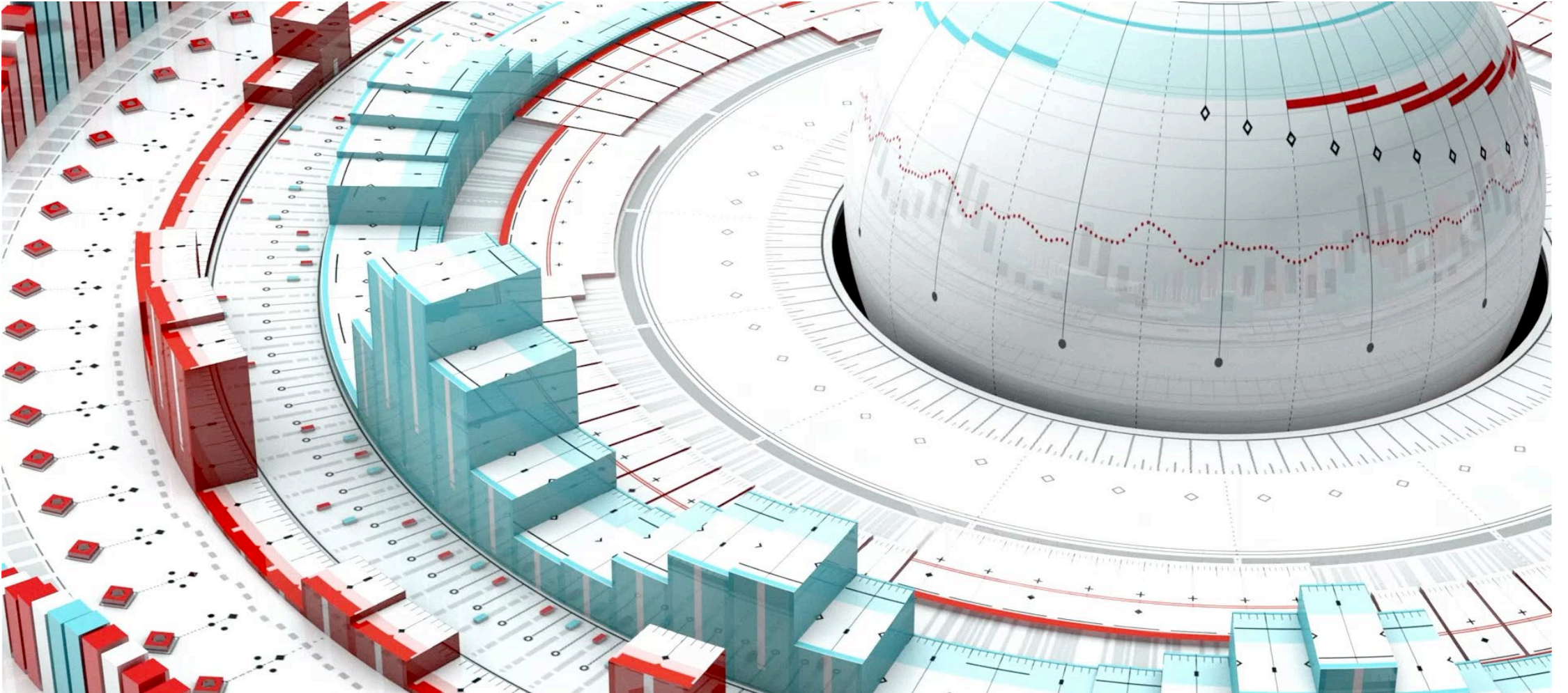
Break

Real World Observations and Best Practices

Measuring Data Quality in Healthcare

Charlie Harp
CEO

Perspective

- Software engineer working across Healthcare verticals > 35 years

- Worked in multiple solution spaces in Healthcare
  - Clinical Lab Systems, Interface Engines, Clinical Trials Platforms, Terminology Management, Clinical Decision Support Engines, Evidence-based Medicine Tools, Reference Data Distribution Networks, Interoperability Engines, Inference Engines and Data Quality Assessment platforms.

- Terminology Management, Interoperability and Data Quality Enthusiast:
  - IDNs (HCA, Common Spirit Health, Rochester Regional Health, MGB, LifePoint Health)
  - Government (DHA, VA, Israeli MOH)
  - Public Health (CDC, APHL)
  - Data Platform Vendors (FlatIron Health, Health Catalyst, dbMotion, others)

- Clinical Architecture:
  - Founded the company in 2007 – Offices in US and UK

Why is Measuring Data Quality Important?
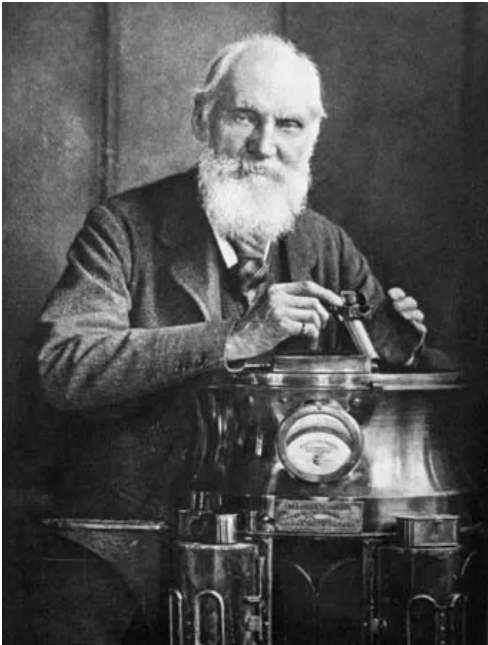
When you can measure what you are speaking about, and express it in numbers, you know something about it;

but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.
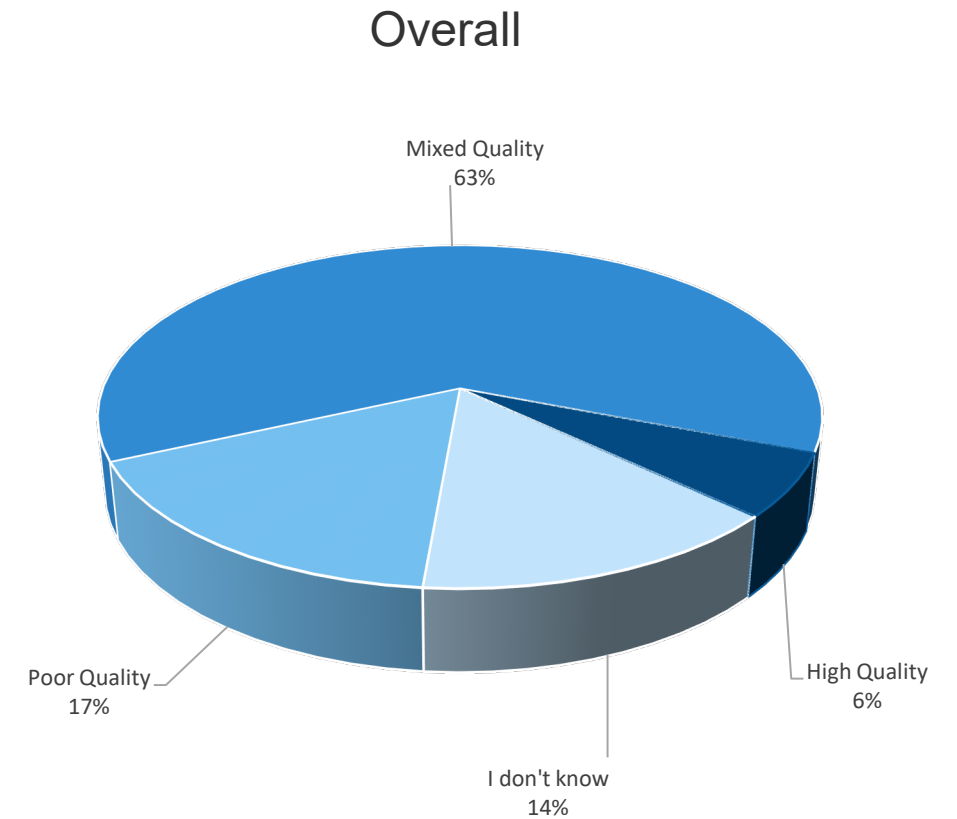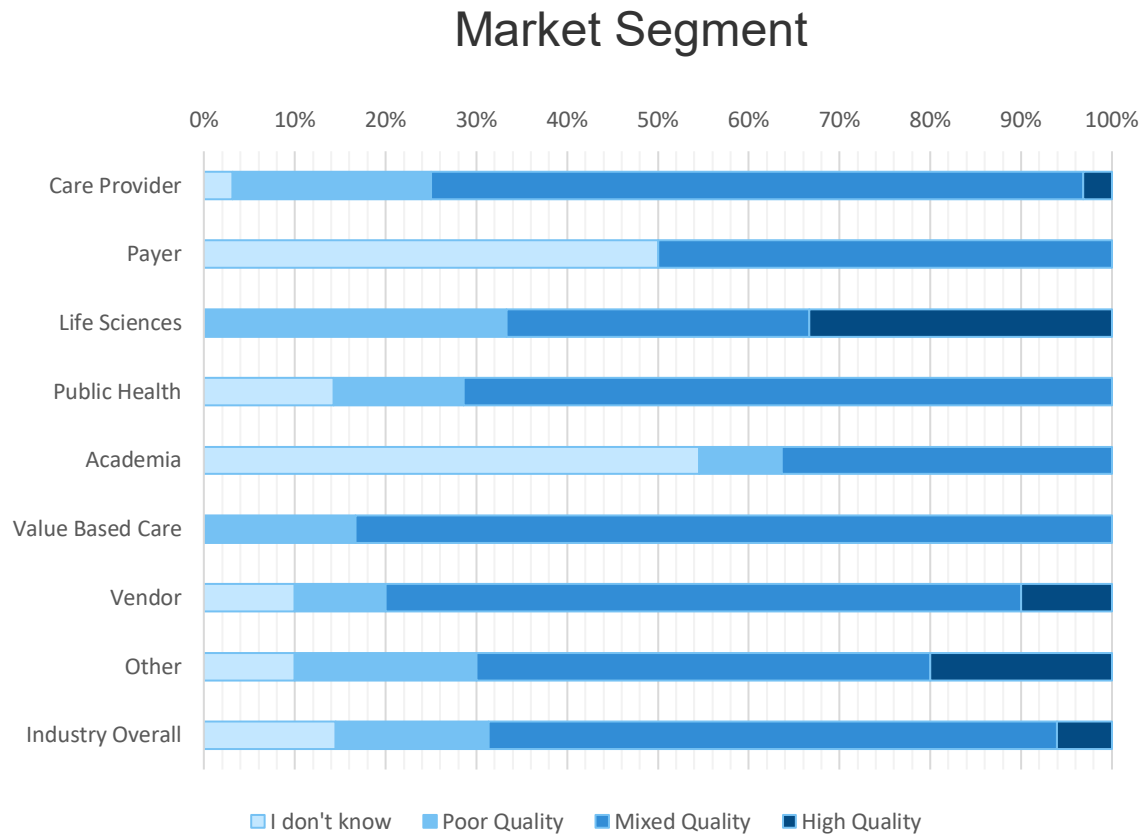
Lord Kelvin,
1883

The Perceived Quality of Shared Patient Data

The Perceived Quality of Shared Patient Data

How would you rate the **quality of the patient data** you get from sources **outside your enterprise**?
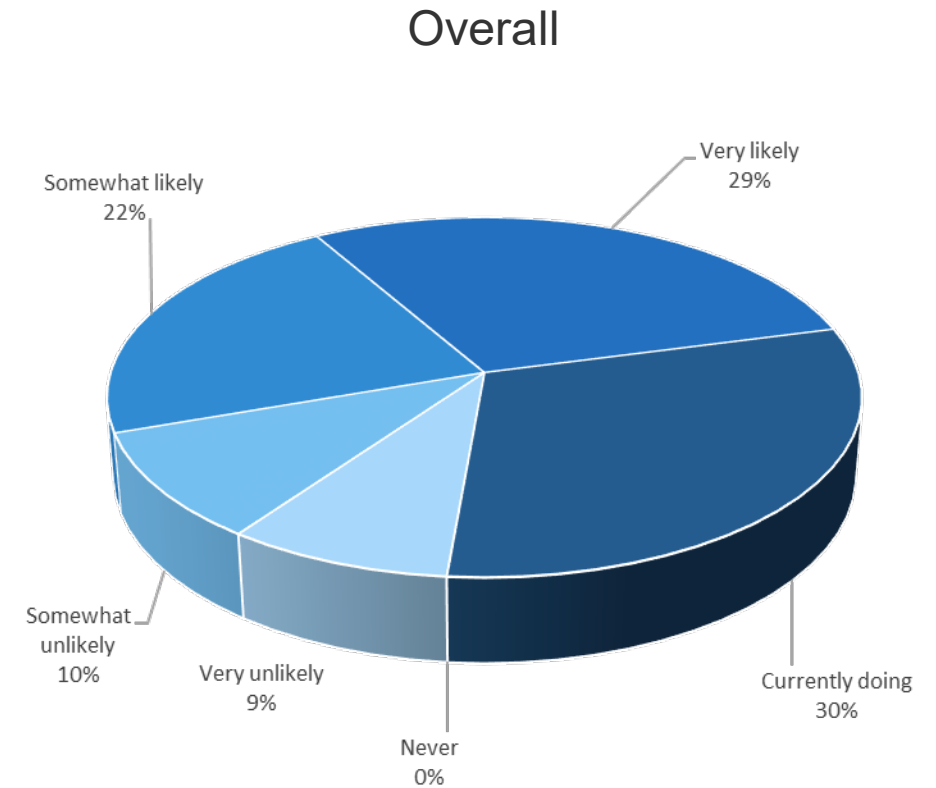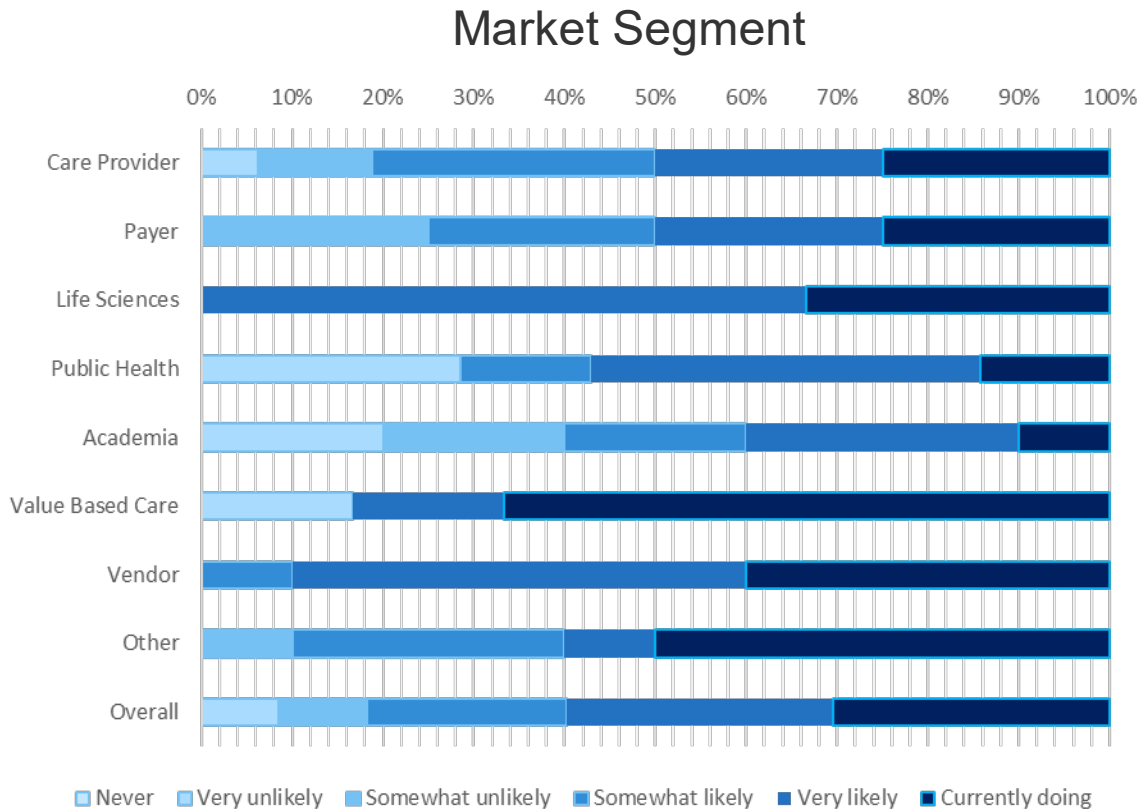
Market Segment

Overall

- Mixed Quality 63%
- Poor Quality 17%
- I don't know 14%
- High Quality 6%

The quality of interoperable health data is widely variable. Different motivation for senders and receivers.

Steven Lane, MD, MPH

The Perceived Quality of Shared Patient Data

Likelihood that your enterprise **would integrate external patient data** into your enterprise

Market Segment

Overall

Data sources may be driven primarily by data sharing **compliance** requirements. Recipients and users need to be able to **trust** the data.

Steven Lane, MD, MPH

Focus here today…

- Further introduce the Healthcare Data Quality Taxonomy (HDQT)

- Secure significant samples of real-world patient data for high volume, mature data classes

- Identify USCDI v3 as a pragmatic evaluation criteria for each data class

- Review how the USCDI evaluation criteria and HDQT could be applied to the samples

- Using the HDQT and USCDI v3 execute an evaluation of these samples

- Share the findings

- Discuss any insights that may be gleaned from this evaluation

Remember – Focus on Data Quality, not message syntax conformance

Evaluation Data Set

- 3 cohorts derived from Health Gorilla's ~25M patient database
  - Two geographically-based, one representing a specialty provider organization
  - All data **de-identified**

- Medication data:
  - 600K medication records derived from 55K patients
  - Derived from network queries

- Laboratory results data:
  - 10M laboratory result records from 92,000 patients
  - Derived from both laboratory result interface messages and treatment-based queries of EHR data via the Carequality, CommonWell and eHealth Exchange networks

Evaluation Criteria
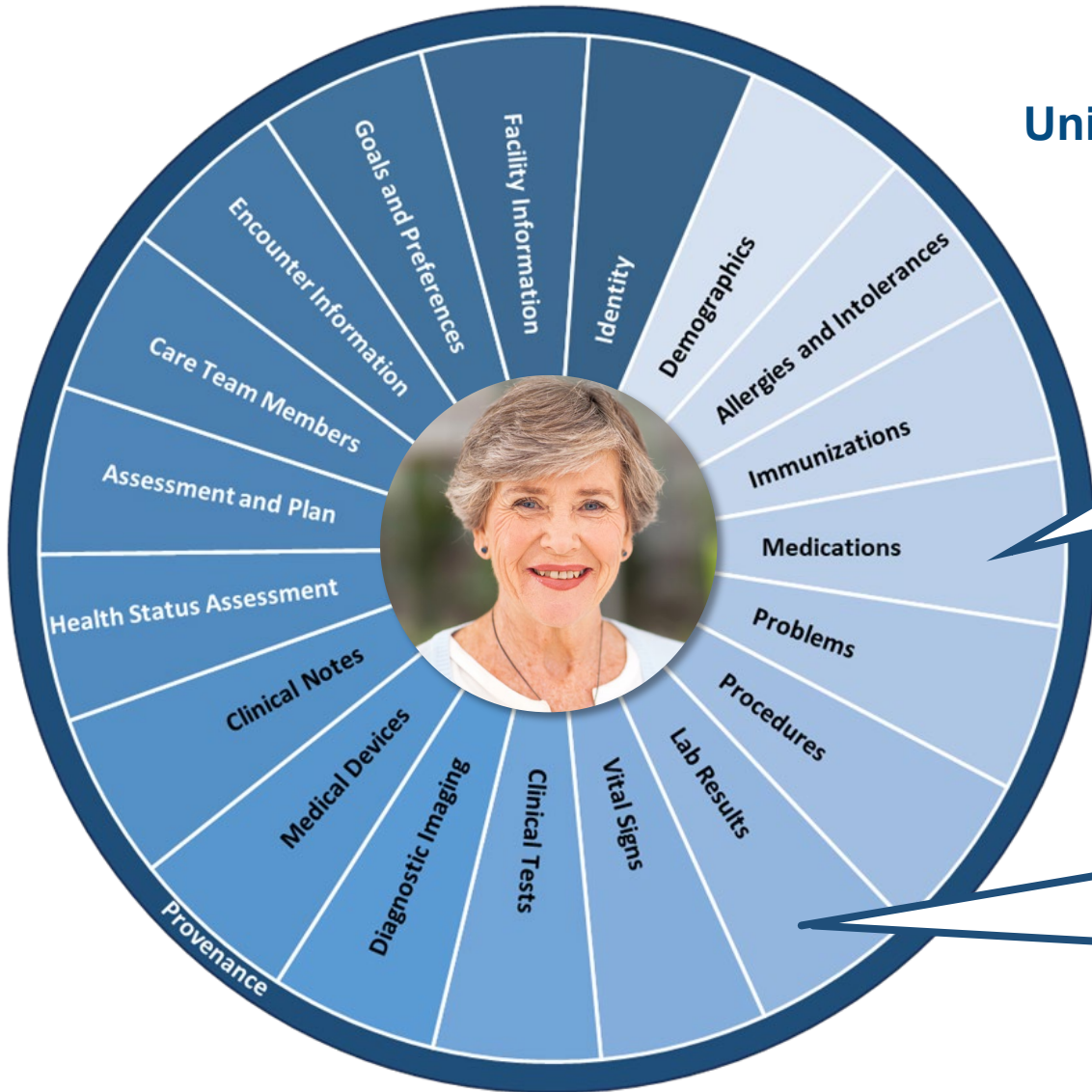
United States Core Data for Interoperability (USCDI) v3

Medications

Pharmacologic agents used in the diagnosis, cure, mitigation, treatment, or prevention of disease.
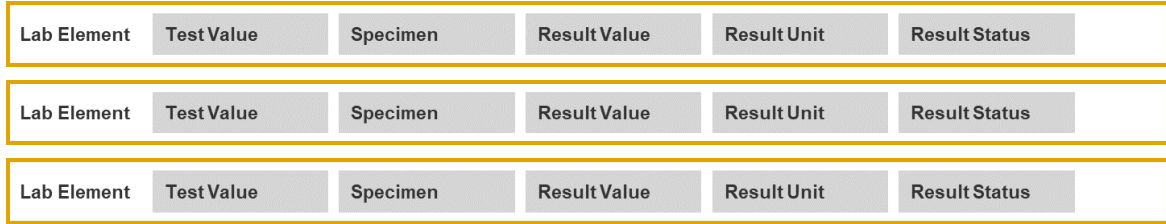
Medications
Dose
Dose Units of Measure
Indication
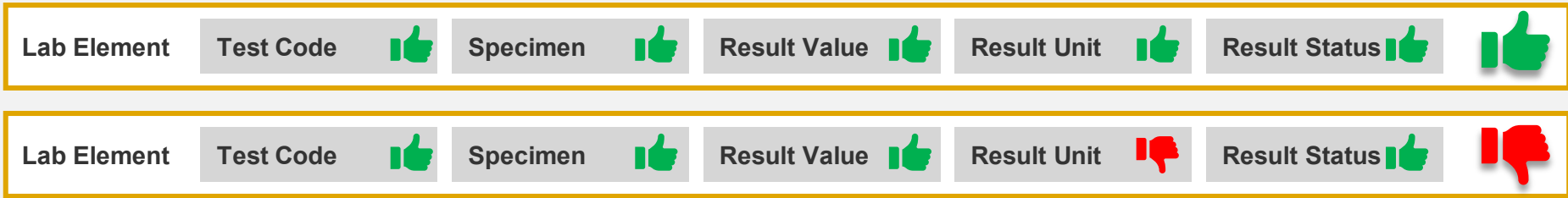Fill Status

Laboratory

Analysis of clinical specimens to obtain information about the health of a patient.

Tests
Values/Results
Specimen Type
Result Status

Evaluation Criteria acts as a Data Quality Gate
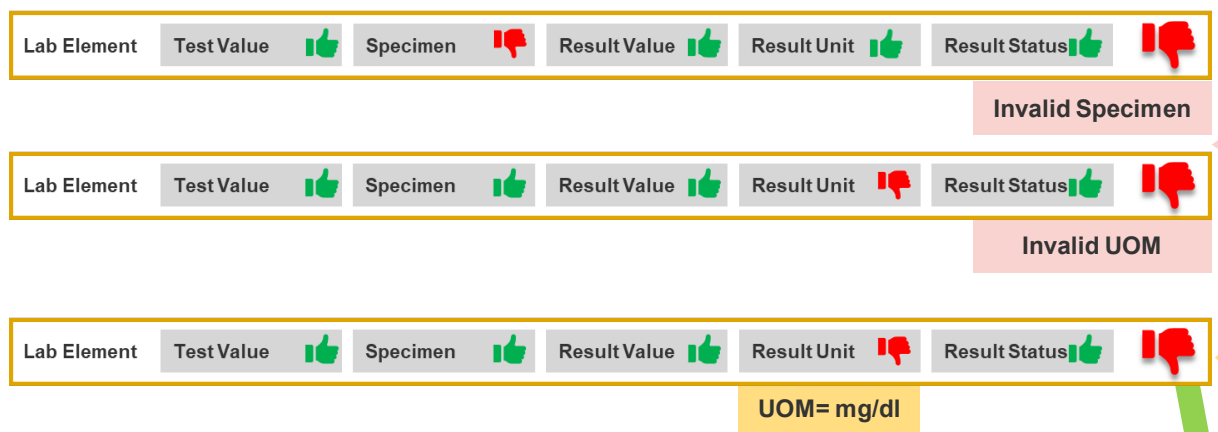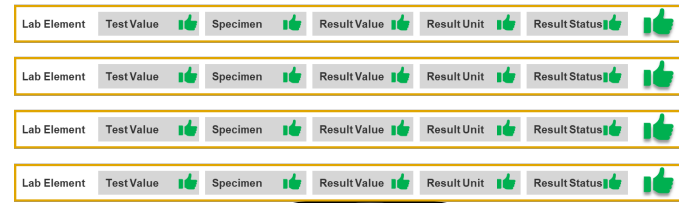
Evaluation Taxonomy

Healthcare Data Quality Taxonomy

HDQT

The top level represents measure categories

The second level represents measure dimensions

Intended to organize the measures necessary to perform the evaluation criteria to provide insight into the nature and potential root causes of the quality issues
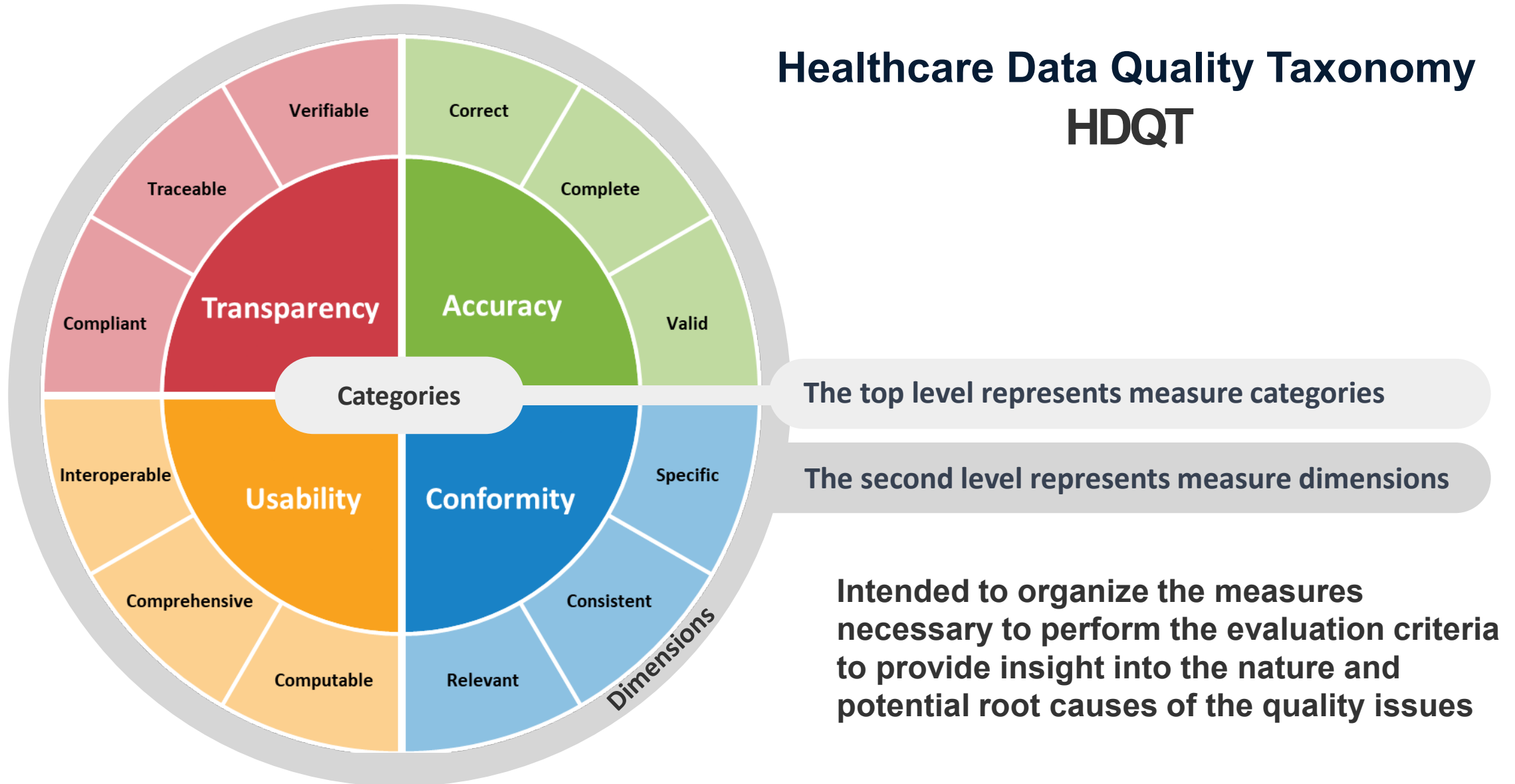
Healthcare Data Quality Taxonomy (HDQT)

Usability

Computable	The ability of healthcare data to be inherently usable by software applications.
Interoperable	The information can be exchanged and used across different systems and organizations in accordance with an interoperability profile
Comprehensive	The information has necessary, current data classes and elements for a given use case

Accuracy

Correct	The correctness and precision of healthcare data
Complete	The information has all necessary data elements and fields in the proper location
Valid	The information is in the correct data class and represented in a way that is valid according to the nature of the information

Conformity

Specific	The level of detail at which data is captured be fit for clinical purpose
Relevant	The information is appropriate and necessary to support the current use case
Consistent	The information does not create contradictions with respect to other elements or attributes in the current data

Transparency

Verifiable	The information includes meta-data that identifies its time, place, system and author of origin
Traceable	The path the information took to get from the origin to the receiver and any additions or remediations that were made in the process
Compliant	The information can be ethically and legally used without violating standards, including patient privacy regulations like HIPAA

Medication Data Quality Assessment
Using USCDI and HDQT

Medication Evaluation Data Set

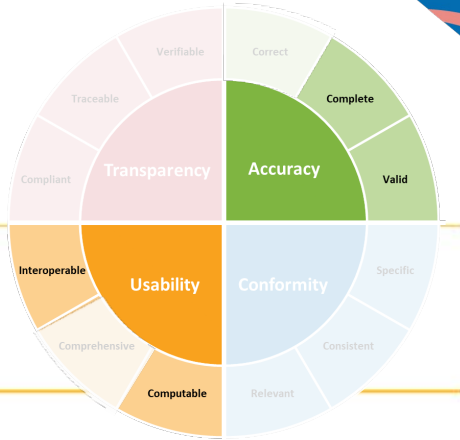| Med Concept					Fill Status	Dose Amount		Dose Unit	Instructions	Indication						
Medication Text	Medication Term	Medication C...	Medication CS	Medication CSN	Fill Status	Dose Low	Dose High	Dose UOM	Instructions	Indication Text	Indication CS	Indication Code	Indication Term	Note	Effective Date	Stop Date
albuterol (PROVENTIL...	albuterol (PROVENTIL) (2.5 ...	630208	2.16.840.1.113883.6.88	RxNorm	completed	2.5		mg						"2.5 mg, Nebuli...		2019-05-02
Clarinex 5 MG	desloratadine 5 MG Oral Ta...	352086	2.16.840.1.113883.6.88	RxNorm	active									Clarinex 5 MG		
Cialis 20 MG	tadalafil 20 MG Oral Tablet [...	402097	2.16.840.1.113883.6.88	RxNorm	active									Cialis 20 MG		
Ipratropium Bromide 0...	ipratropium bromide 0.042 ...	1797844	2.16.840.1.113883.6.88	RxNorm	active									Ipratropium Bro...		
meloxicam (MOBIC) 1...	meloxicam (MOBIC) 15 MG t...	152695	2.16.840.1.113883.6.88	RxNorm	aborted				"meloxicam 15 mg t...							2023-04-25
triamcinolone acetoni...	triamcinolone acetonide (KE...	1086275	2.16.840.1.113883.6.88	RxNorm	completed	40.0		mg		Right elbow pain	2.16.840.1.1...	74323005		"40 mg, Intra-a...		2023-03-07
methylPREDNISolone ...	methylPREDNISolone (MED...	762675	2.16.840.1.113883.6.88	RxNorm	active				Follow package dire...	Neck pain	2.16.840.1.1...	81680005		Follow package ...		
celecoxib (CELEBREX)...	celecoxib (CELEBREX) 200 ...	205323	2.16.840.1.113883.6.88	RxNorm	active	200.0		mg	Take 1 Capsule (20...	Neck pain	2.16.840.1.1...	81680005		Take 1 Capsule ...		
cyclobenzaprine (FLE...	cyclobenzaprine (FLEXERIL)...	828320	2.16.840.1.113883.6.88	RxNorm	active	0.0		mg	Take 1-2 Tablets (5...	Neck pain	2.16.840.1.1...	81680005		Take 1-2 Tablet...		
diclofenac Sodium (VO...	diclofenac Sodium (VOLTAR...	855633	2.16.840.1.113883.6.88	RxNorm	active	2.0		g	Apply 2 g every 8 h...	Neck pain	2.16.840.1.1...	81680005		Apply 2 g every...		
ergocalciferol (VITAMI...	ergocalciferol (VITAMIN D) ...	1367410	2.16.840.1.113883.6.88	RxNorm	active				"Vitamin D2 1,250 ...					"Vitamin D2 1,2...		
lidocaine (LIDODERM)...	lidocaine (LIDODERM) 5 %	1745091	2.16.840.1.113883.6.88	RxNorm	active	1.0		{patch}	Place 1 Patch ever...	Neck pain	2.16.840.1.1...	81680005		Place 1 Patch e...		
QUEtiapine (SEROque...	QUEtiapine (SEROquel) 50 ...	616487	2.16.840.1.113883.6.88	RxNorm	active	150.0		mg	Take 150 mg by mo...					Take 150 mg by...		
amphetamine-dextroa...	amphetamine-dextroamphe...	861232	2.16.840.1.113883.6.88	RxNorm	active	30.0		mg	Take 1 capsule (30 ...	"Attention deficit h...	2.16.840.1.1...	35253001		Take 1 capsule ...		
	Methotrexate	105585	2.16.840.1.113883.6.88	RxNorm	active	0.0			6 by mouth every ...							
	Folic Acid	310410	2.16.840.1.113883.6.88	RxNorm	active	0.0			1 every day							
	Enbrel Sureclick	1653225	2.16.840.1.113883.6.88	RxNorm	active	0.0			inject 50mg subcut...		2.16.840.1.1...	714.0	Rheumatoid Arthr...			
	Lidoderm	1011705	2.16.840.1.113883.6.88	RxNorm	active	0.0			apply 1 patch to ski...		2.16.840.1.1...	719.46	Pain Joint Lower L...			
	Tramadol HCL	835603	2.16.840.1.113883.6.88	RxNorm	active	0.0			1/2 - 1 by mouth fo...		2.16.840.1.1...	714.0	Rheumatoid Arthr...			
	Alendronate Sodium	904431	2.16.840.1.113883.6.88	RxNorm	active	0.0			1 every week							
	Hydroxychloroquine Sulfate	979092	2.16.840.1.113883.6.88	RxNorm	active	0.0			2 by mouth every d...		2.16.840.1.1...	714.0	Rheumatoid Arthr...			
LIPITOR (unknown str...	atorvastatin 10 MG Oral Ta...	617314	2.16.840.1.113883.6.88	RxNorm	active				take 1 tablet by or...					take 1 tablet by...		
	buprenorphine HCl 8 MG / n...	351267	2.16.840.1.113883.6.88	RxNorm	active											
Advair Diskus 250 mc...	60 ACTUAT fluticasone prop...	896212	2.16.840.1.113883.6.88	RxNorm	completed	1.0		inhalatio...	Inhale 1 inhalati...							2022-03-08
Plant Sterols and Stan...	Drug or medicament (substa...	410942007	2.16.840.1.113883.6.96	SNOMED Clinical Terms	active	950.0		mg	Take 950 mg by mo...					Take 950 mg by...		
Psyllium (METAMUCIL ...	Drug or medicament (substa...	410942007	2.16.840.1.113883.6.96	SNOMED Clinical Terms	active				Take by mouth					Take by mouth		
Ibuprofen (ADVIL PO)	Drug or medicament (substa...	410942007	2.16.840.1.113883.6.96	SNOMED Clinical Terms	active				by mouth					by mouth		

Medication Evaluation Criteria

Medications

Pharmacologic agents used in the diagnosis, cure, mitigation, treatment, or prevention of disease.

USCDI V1	USCDI V2	**USCDI V3**	USCDI V4	Level 2	Level 1	Level 0

Data Element				Applicable Vocabulary Standard(s)		

Medications — Computable — Complete — >
Pharmacologic agent used in the diagnosis, cure, mitigation, treatment, or prevention of disease.
- RxNorm Full Monthly Release, July 5, 2022 — Valid — Interoperable
Optional:
- National Drug Code (NDC), July 19, 2022

Dose — Computable — Complete — >
Amount of a medication for each administration.

Dose Units of Measure — Computable — Complete — >
Units of measure of a medication. (e.g., milligrams, milliliters)
- The Unified Code for Units of Measure, Revision 2.1 — Valid — Interoperable

Indication — Computable — Complete — >
Sign, symptom, or medical condition that leads to the recommendation of a treatment, test, or procedure.

Fill Status — Computable — Complete — > — Valid
State of a medication with regards to dispensing or other activity. (e.g., dispensed, partially dispensed, not dispensed)

Medication Evaluation Results
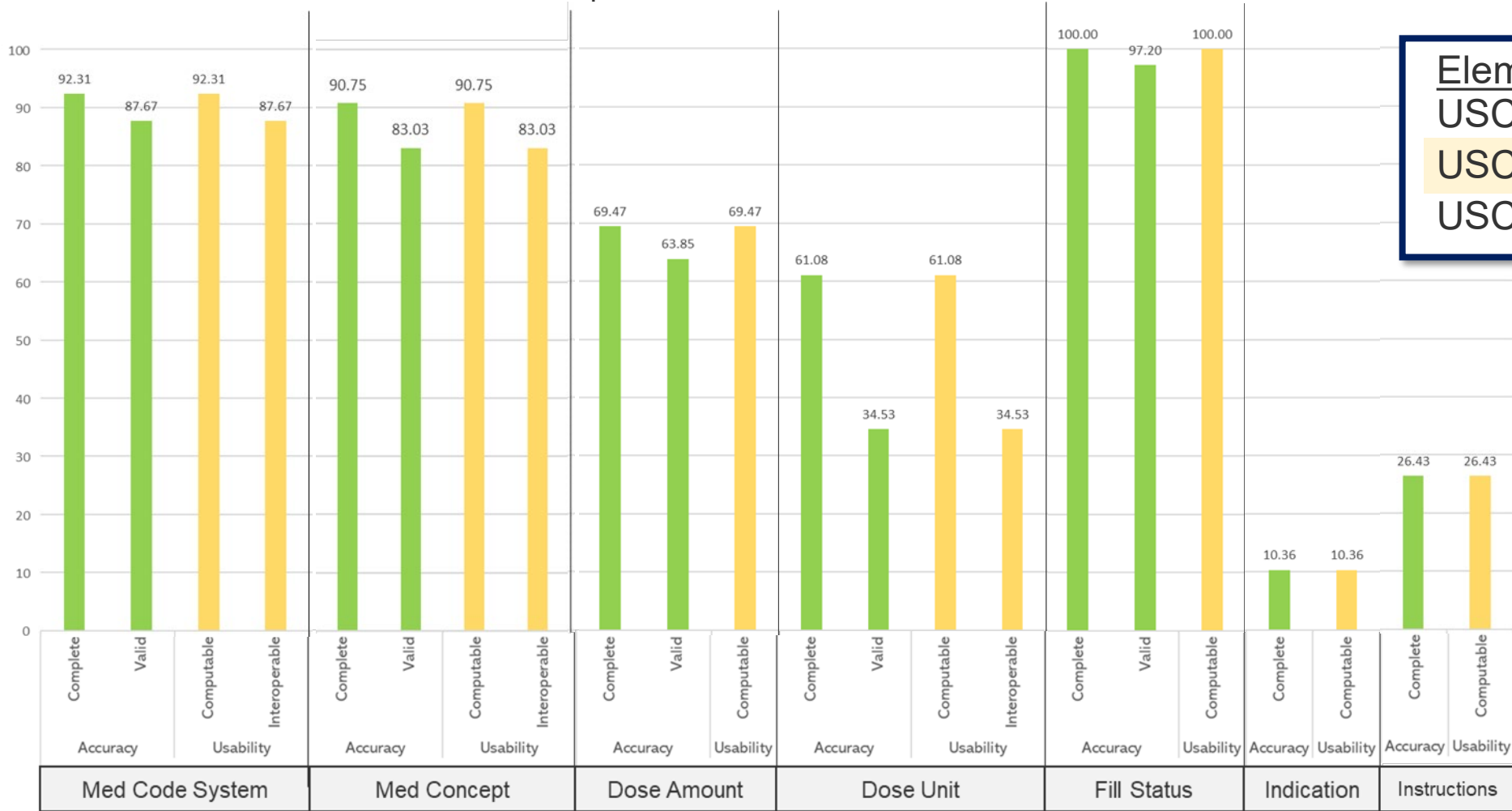
Medication data:

600K medication records derived from 55K patients

Element Results	
USCDI v1:	83.03 %
USCDI v2:	83.03 %
USCDI v3:	2.92 %

Lab Data Quality Assessment
Using USCDI and HDQT

Lab Evaluation Data Set

	Relevant	Verified			Test Concept				Status			Result Value	Result Unit	Specimen		

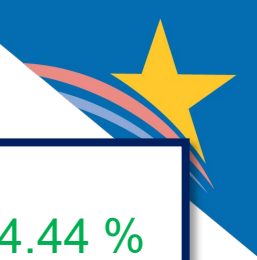| OBR4_1_panel C...	OBR4_2_panel Name	OBX14_date Ti...	OBX15_place ... ▲	OBX3_1_Observation ID	OBX3_2_observation NAME	OBX3_4_loi...	OBX_loinc ...	OBX11_Result STATUS	OBX2_value Type	OBX5_Observation VALUE	OBX6_Observation UNIT	OBR15_Specimen	OBX8_abnormal T...	OBX7_observation_Range
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015065	MCV	787-2		F	NM	90	fL			79-97
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	005025	WBC	6690-2		F	NM	4.1	x10E3/uL			3.4-10.8
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015123	Lymphs	736-9		F	NM	36	%			Not Estab.
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015933	Eos (Absolute)	711-2		F	NM	0.2	x10E3/uL			0.0-0.4
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	005033	RBC	789-8		F	NM	4.23	x10E6/uL			4.14-5.80
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015911	Immature Grans (Abs)	53115-2		F	NM	0.0	x10E3/uL			0.0-0.1
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015156	Basos	706-2		F	NM	1	%			Not Estab.
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015108	Immature Granulocytes	71695-1		F	NM	0	%			Not Estab.
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	105007	RDW	788-0		F	NM	12.4	%			11.6-15.4
005009	CBC With Differential/Platelet	2021-04-06T10:...	01	015909	Neutrophils (Absolute)	751-8		F	NM	1.8	x10E3/uL			1.4-7.0
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001040	BUN	3094-0		F	NM	16	mg/dL			8-27
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001099	Bilirubin, Total	1975-2		F	NM	1.2	mg/dL			0.0-1.2
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001016	Calcium	17861-6		F	NM	9.1	mg/dL			8.6-10.2
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001107	Alkaline Phosphatase	6768-6		F	NM	54	IU/L			39-117
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	100797	eGFR If Africn Am	88293-6		F	NM	105	mL/min/1.73			>59
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	100791	eGFR If NonAfricn Am	88294-4		F	NM	91	mL/min/1.73			>59
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001370	Creatinine	2160-0		F	NM	0.87	mg/dL			0.76-1.27
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001206	Chloride	2075-0		F	NM	101	mmol/L			96-106
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001198	Sodium	2951-2		F	NM	137	mmol/L			134-144
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	012039	Globulin, Total	10834-0		F	NM	2.8	g/dL			1.5-4.5
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001073	Protein, Total	2885-2		F	NM	6.9	g/dL			6.0-8.5
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001123	AST (SGOT)	1920-8		F	NM	46	IU/L		H	0-40
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	012047	A/G Ratio	1759-0		F	NM	1.5				1.2-2.2
322000	Comp. Metabolic Panel (14)	2021-04-06T11:...	01	001180	Potassium	2823-3		F	NM	4.4	mmol/L			3.5-5.2

Lab Evaluation Criteria

Laboratory

Analysis of clinical specimens to obtain information about the health of a patient.

USCDI V1	USCDI V2	**USCDI V3**	USCDI V4	Level 2	Level 1	Level 0

Data Element	Relevant	Verifiable		Applicable Vocabulary Standard(s)

Tests

	Computable	Complete	>	Valid	Interoperable

Analysis of specimens derived from humans which provide information for the diagnosis, prevention, treatment of disease, or assessment of health.

- Logical Observation Identifiers Names and Codes (LOINC®) version 2.72

Values/Results — Result Value

	Computable	Complete	>	Both standards are required: Valid

Documented findings of a tested specimen including structured and unstructured components.

- SNOMED International, Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) U.S. Edition,

Result Unit

	Computable	Complete		Valid	Interoperable

- The Unified Code of Units for Measure, Revision 2.1

Specimen Type

	Computable	Complete	>	Valid	Interoperable

Substance being sampled or tested. (e.g., nasopharyngeal swab, whole blood, serum, urine, wound swab).

- SNOMED International, Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) U.S. Edition, March 2022 Release

Result Status

	Computable	Complete	>	Valid

State or condition of a laboratory test.

Lab Evaluation Results

Laboratory results data:
10M laboratory result records from 92,000 patients

Element Results	
USCDI v1:	44.44 %
USCDI v2:	44.44 %
USCDI v3(ish):	2.21 %
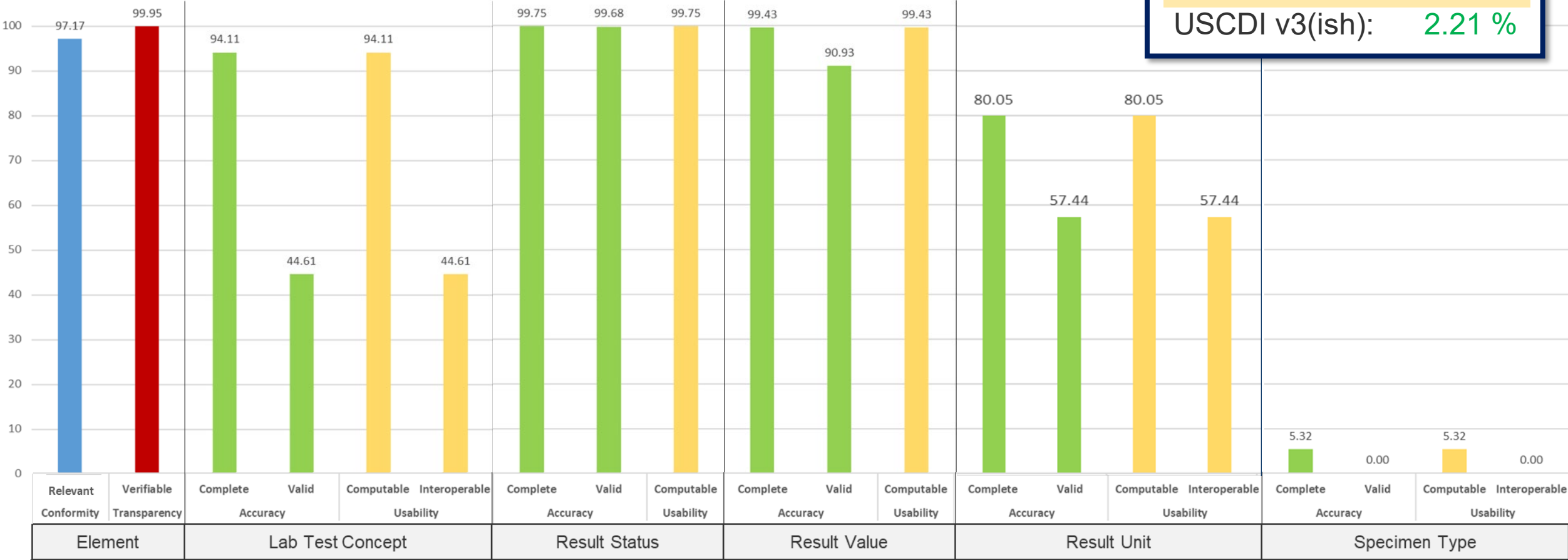
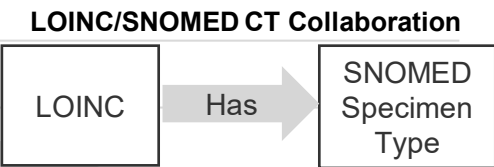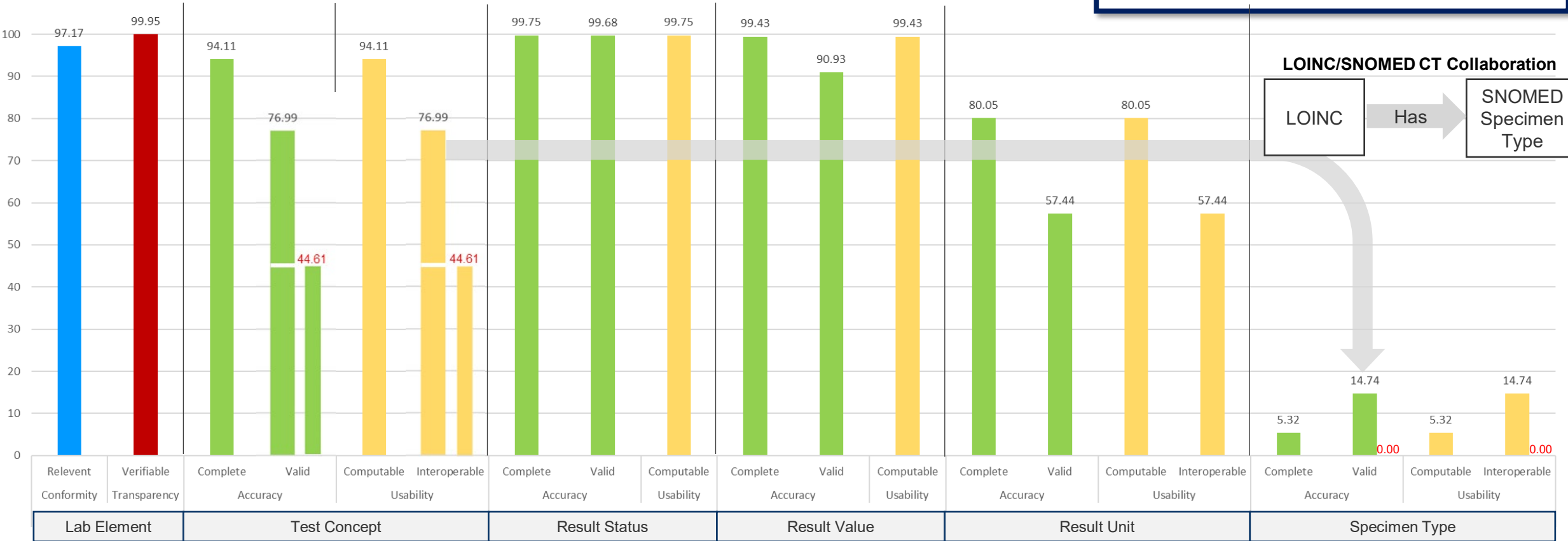# Lab Evaluation Results - Remediated

Laboratory results data:

10M laboratory result records from 92,000 patients

Element Results
USCDI v1:   71.63 % (44.44%)
USCDI v2:   71.63 % (44.44%)
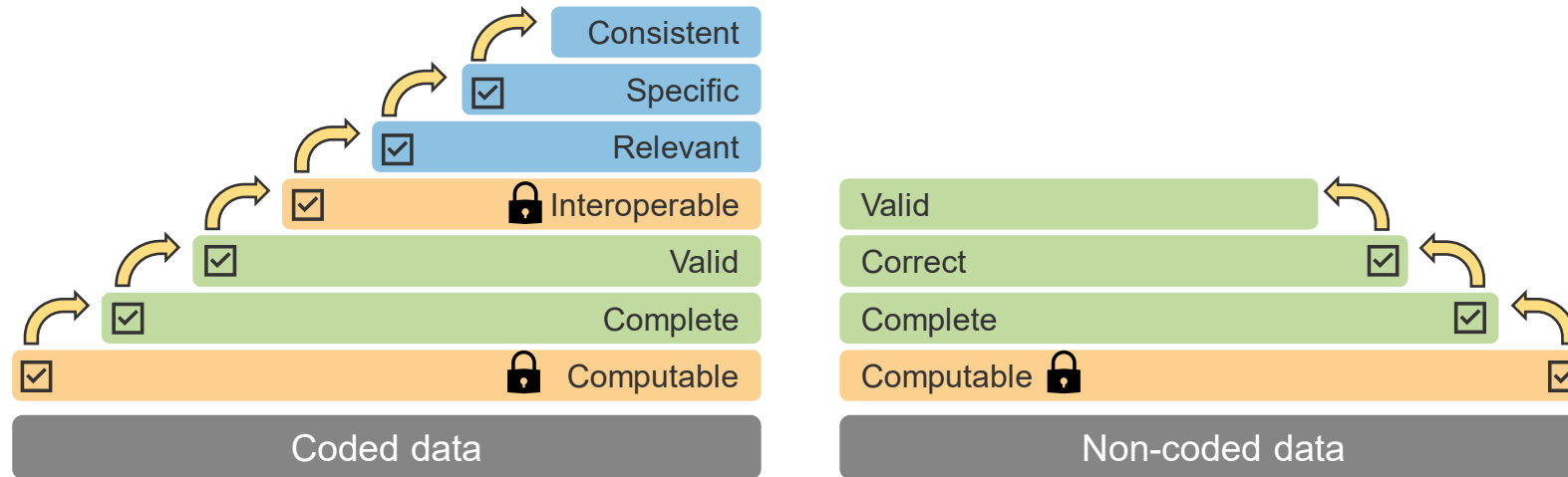USCDI v3:   14.74 %   (2.21%)

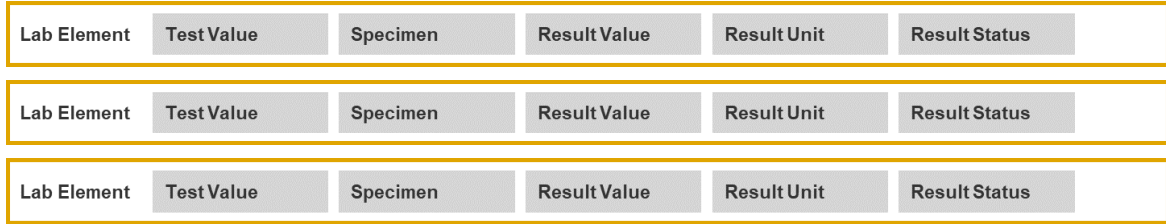# Observations

Dependency Pattern

Each dimension check acts like a gate for a more sophisticated check
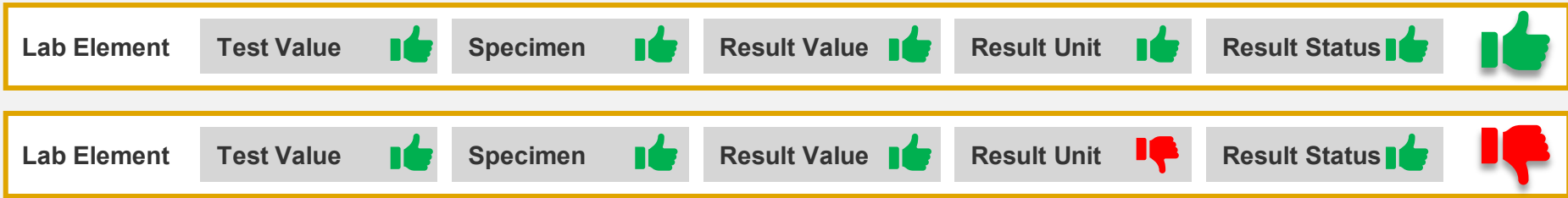

The application of a taxonomy revealed a pattern (albeit an obvious one)
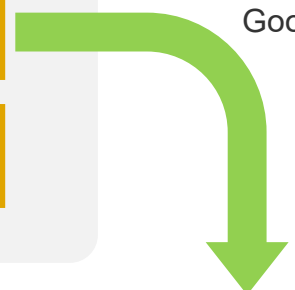
Evaluation Criteria acts as a Data Quality Gate
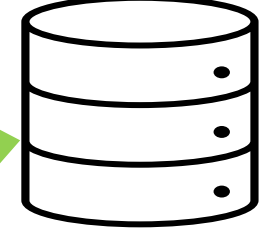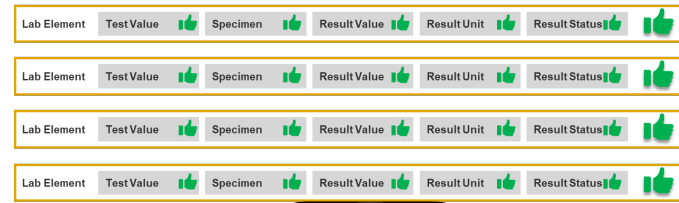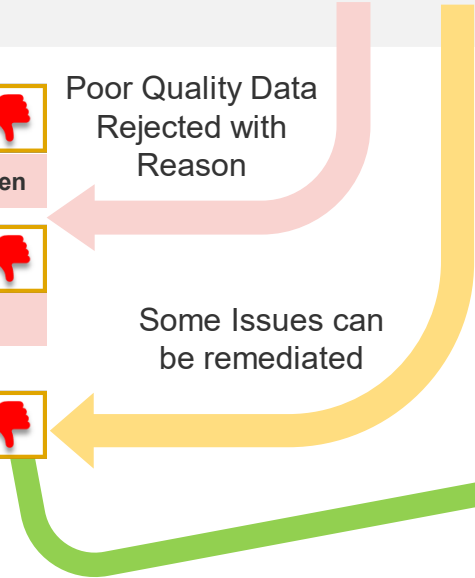
Medication Evaluation Results

Medication data:

600K medication records derived from 55K patients
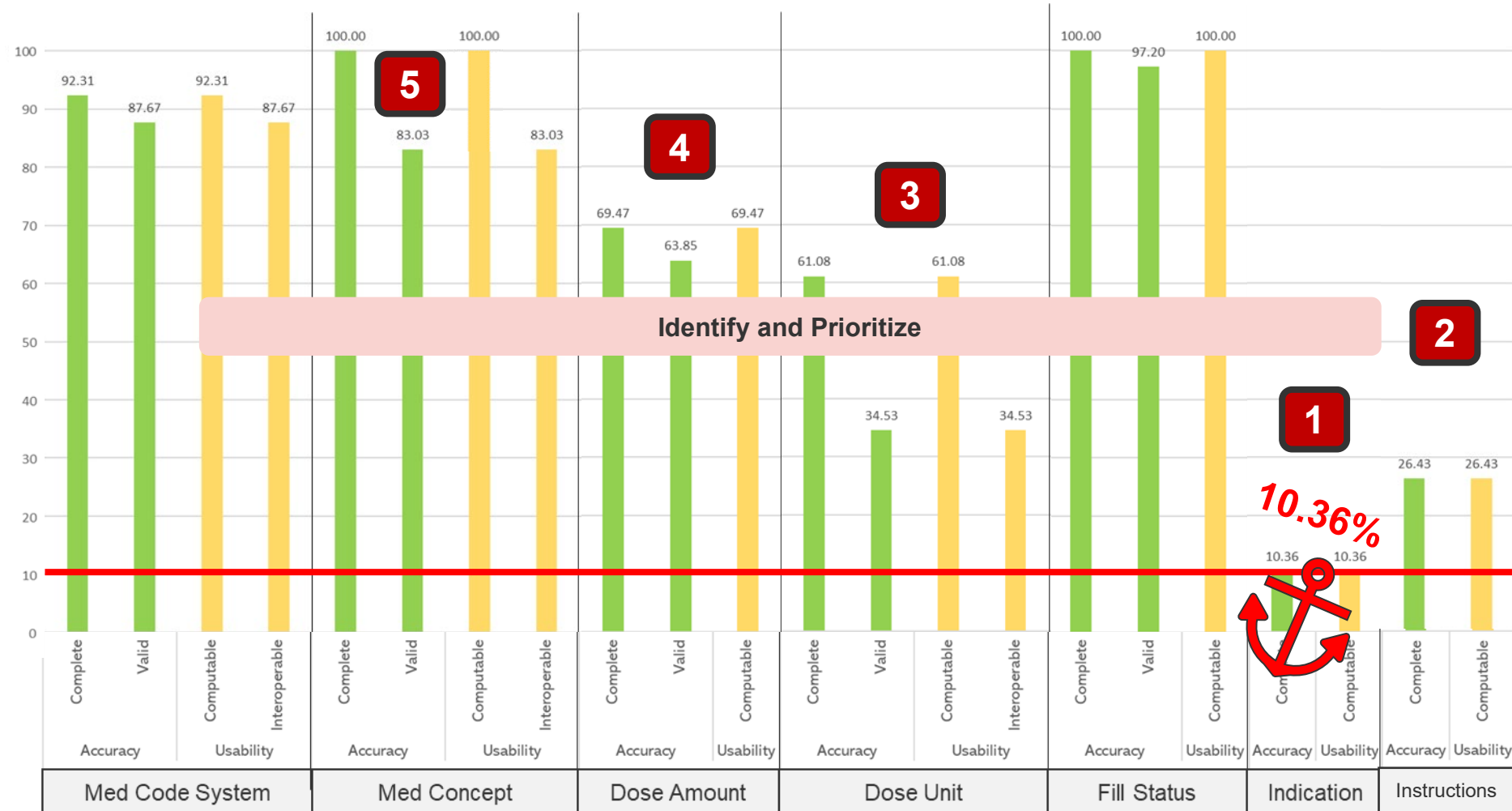
Conclusions

- USCDI v3 – a pragmatic evaluation criteria for meds and labs, if realistically applied

- HDQT – Useful start, but needs a bit of real-world evolution
  - Taxonomies reveal patterns and provide useful insights

- Coded Entities for Medications and Lab is key
  - If it doesn't have a code system, it's not coded

- Code remediation (mapping) can rescue uncoded or miscoded data

- Need more focus on collecting secondary attributes (specimen type, indication)*

- More progress on provenance, traceability, and compliance is needed

When you can measure what you are speaking about, and express it in numbers, you know something about it
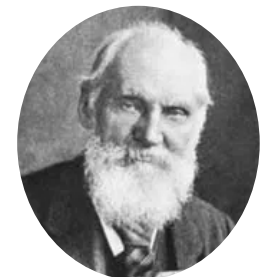
Thank You!

charlie_harp@clinicalarchitecture.com

AHIMA Data Quality and Integrity

AHIMA's Position Data Quality and Integrity

Healthcare data should be complete, accurate, consistent and timely throughout its life cycle. Health data is human data.

- Consistent standards across all HIPAA covered entities including mechanisms for accountability. Accuracy of patient health data hinges on being able to accurately **match patients with their data.**

- Support **content,** data mapping and documentation with consistent application of standards. A complete representation of the patient's clinical status.

- Adoption and use of technologies that support high-quality data. Available for access, exchange or use.

[Data Quality and Integrity | Advocacy (ahima.org)](ahima.org)

Data Quality in Matching Patients to Their Data

- Patient ID Now Research:
  - Spending 110 hours per week resolving patient identity issues

  - Spending $1 million annually (1/3 orgs)

  - 70% of respondents agreed or strongly agreed patients undergo or receive duplicative or unnecessary testing or services

- Health Disparities: Record overlays and duplication, combined with clinical information, have a disproportionate impact on underserved communities.

- Patient matching could be improved through the development and widespread adoption of standards for the demographic data elements that are relied on every day for matching patient records.

http://patientidnow.org/wp-content/uploads/2022/11/PIDN-Research-Findings-Final.pdf

Health Disparity Example

Potential Name Entries		
First	**Middle**	**Last**
Rogerio	Miguel	Turrietta
Rogelio	Michael	Turietta
Roger	Micheal	Turrieta
R.	M.	Turieta

- **Duplicate error potential**:
  - 1 health record number assigned to Rogerio Miguel Turrietta and 1 health record number assigned to Roger Michael Turrietta (same person with 2 versions of his name)

- **Overlay error potential**:
  - Rogerio Miguel Turrietta is mistakenly assigned the health record number which belongs to Rogerio Lino Turrietta (2 separate people mistaken as the same)

AHIMA Naming Policy Framework

AHIMA calls for the adoption and use of a person(s) demographic data element framework.

1. **Complete Legal Name** (FN, MN, LN – recording multiple names w/n a single field)
2. Newborns – distinct naming convention
3. Single births, Twin births, Multiple births, Safe Haven Baby & Adoption at birth, Birth pronouns, fetal care
4. Prefixes / Suffixes
5. Hyphens, Apostrophe's, Punctuation
6. Sex/Gender
7. Single legal name, mononyms
8. Unidentified Individuals
9. Complete address as captured for mailing purposes
10. Nicknames / Preferred Names / Aliases

Naming Policy | AHIMA

Texas Health Services Authority Data Quality and Content

Problem: Inconsistent Discharge Summary C-CDA content is impacting transitions of care.

- Too much and too little information

- Inconsistent across vendors and organizations

- Cognitive overload with electronic health records

- Concerned missing key components

- Lack of trust in data leads to falling back on phone calls, faxes or mail that delay care

Surveyed the Community

- "I often find C-CDAs not helpful, the hospital course section does not convey clinically relevant info. Fields are auto-populated. There is no narrative for the discharging physician to intentionally communicate with the physician they are handing off to. The discharge information is only as good as the relevant information that is provided by the physician"

- "Standardizing "how" things are displayed is as important as "what" the C-CDA contains. For example, if an organization doesn't use the Problem List, that's a real problem when others expect Problems to be listed."

- "Please consider the ORDER of items on the document to be Date of Admission-->Date of Discharge-->Discharge Location (Home, Skilled Nursing Facility, Rehabilitation)--> Chief Complaint-->Discharge Summary-->Consultations-->Laboratory tests/Radiology examinations/etc."

- "Needs to be lean and concise."

THSA Discharge Summary C-CDA Recommendation

1. Discharge Summary Narrative (aka Hospital Course)

2. Discharge Medications

3. Allergies

4. Admission Diagnosis

5. Discharge Diagnosis

6. Procedures: including Interventional Radiology, Cardiac Cath, operative procedures

7. Diagnostic Imaging – Advanced imaging for example: MRI, CT, PET, Nuclear Imaging, Ultrasound, Echo, & Venous Doppler

8. Laboratory – Recommend 1st and last laboratory result for every test. On those rare tests – they are only done once so would be included (ANA Rheumatoid)

9. Consultations

10. Assessment & Plan (includes future orders for follow-up with PCP and diagnostic tests)

11. Problem List

Discharge C-CDA Minimum Data-Set Content - Texas Health Services Authority (thsa.org)

Discharge Summary C-CDA Recommendation

- Shared with Sequoia Project Data Usability Implementation Work-Group; included in Data Usability Implementation Guide with recommendation Appendix B – 12/14/22

- Commented on USCDI V3
  - Recommendation from Office of National Coordinator to contact HL7 Structured Documents & Sequoia Project Data Usability Work Group
  - HL7 Structured Documents – **"you raised the bar on C-CDAs"**

- Participated with HL7 Implementation-A-Thon, 4/27/2023 as "Unhappy Documentation" – Supportive and moving recommendations forward

- Participated with HL7 Implementation-A-Thon, 8/24/2023, continued support with recommendation to include in technical standard for January 2024 vote

2022-12-14-Sequoia-DUWG-IG-Version-1-1.pdf (sequoiaproject.org)

Data Quality and Integrity

Healthcare data should be complete, accurate, consistent and timely throughout its life cycle. Health data is human data.

Connection of patient records supports completeness. Data Quality requires normalizing across vendor platforms and organizations for electronic sharing.

- AHIMA's Data Quality Policies are designed to evolve over time so that the community can support technological advancements and operational policies in this ever-changing healthcare environment.

- AHIMA Naming Policy Framework 2023: Essential Person Demographic Data Elements for Electronic Health Information is designed to evolve over time.

- Texas Health Services Authority Discharge Summary C-CDA Recommendation is following AHIMA's Data Quality and Integrity Policy and working with HL7 and Sequoia to implement.

Health Information Professionals

Health Information Professionals have extensive knowledge and expertise to contribute to developing policies around data quality and integrity as it relates to health information.

Leveraging the AHIMA Policies improves data quality, supports electronic transfer of clinical data, improves patient outcomes, create efficiencies and decreases cost.
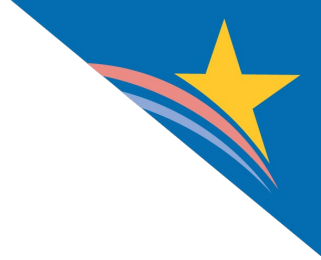
Veterans Health Information Exchange (VHIE) Clinical Data Quality (CDQ) Program Overview

Sandi Mitchell, RPh, MSIS, FASHP, FAMIA

VHIE Clinical Data Quality Innovation Team

Veterans Health Information Exchange (VHIE) Program

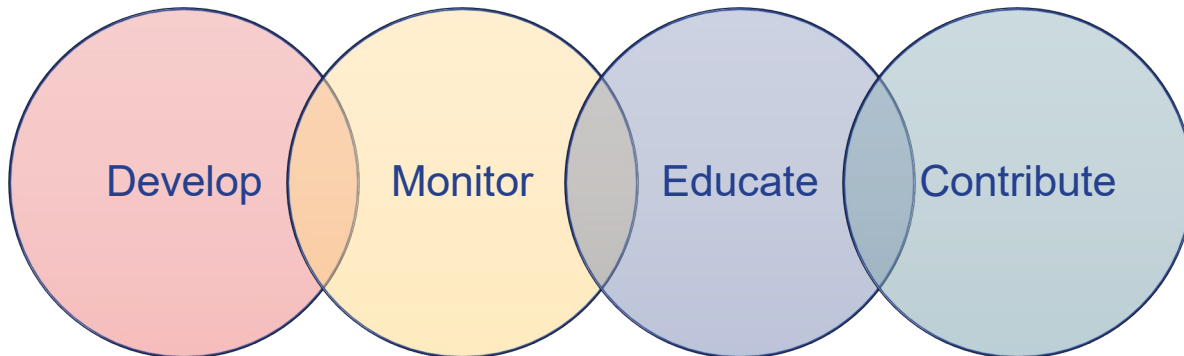Office of Health Informatics (OHI)

September 7th, 2023

VHIE Clinical Data Quality Program Overview

VHIE Clinical Data Quality Program Overview

Mission

VHIE Clinical Data Quality (CDQ) supports VA's journey to becoming a High Reliability Organization (HRO) by improving the quality and reliability of clinical data exchanged with participating external partners. CDQ engages the entire health care ecosystem to ensure clinicians have high quality data to support patient health care decision making.

Objectives

Develop   Monitor   Educate   Contribute

Process

Control
Continuous monitoring using sampling strategies to track clinical content improvement

Definition
Align Data Quality initiatives with VHA VHIE mission and objectives

Implementation
Support source implementation of solutions presented

What We Do

Assessment
Explore the clinical content behind the quality metrics

Improvement
Present insights based on analysis using simple deliverable tools

Analysis
Analyze the sources of clinical data quality issues based on the dataflow process

Impact, Value and Trust to Clinicians & Veterans

Impact to Clinicians and Veterans

VHIE Clinical Data Quality (CDQ) increases the reliability of data exchanged through VHIE by ensuring that non-VA organizations are aware of the importance of high-quality electronic health information exchange.
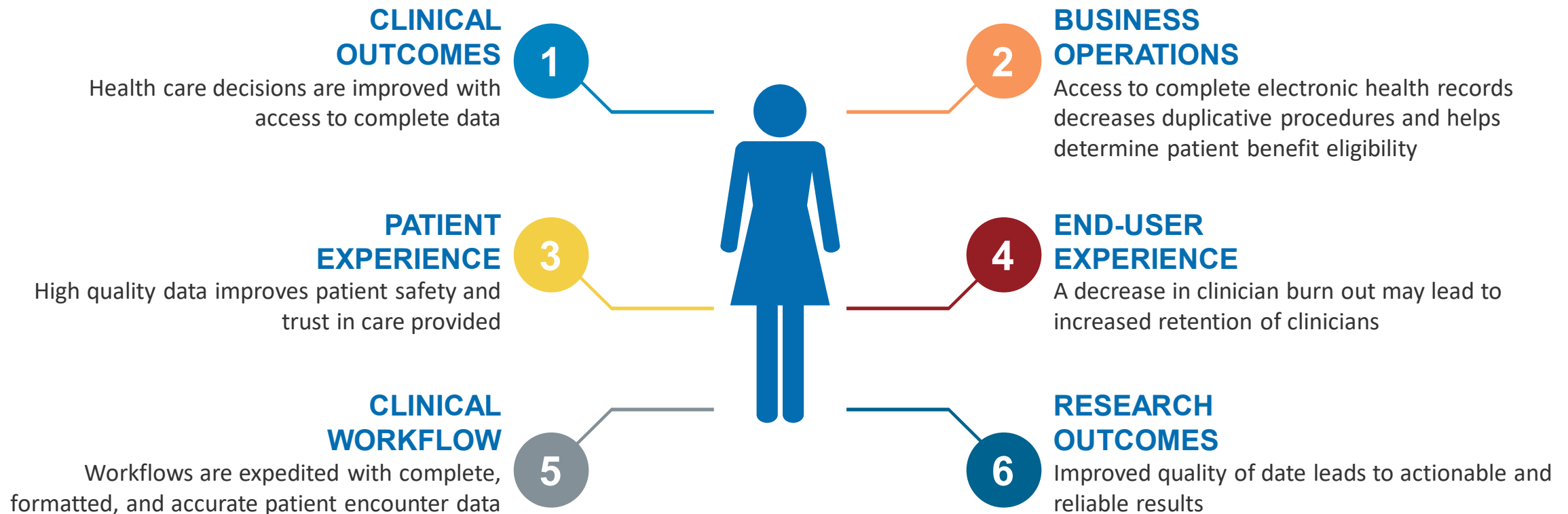
CLINICAL OUTCOMES

1

Health care decisions are improved with access to complete data

BUSINESS OPERATIONS

2

Access to complete electronic health records decreases duplicative procedures and helps determine patient benefit eligibility

PATIENT EXPERIENCE

3

High quality data improves patient safety and trust in care provided

END-USER EXPERIENCE

4

A decrease in clinician burn out may lead to increased retention of clinicians

CLINICAL WORKFLOW

5

Workflows are expedited with complete, formatted, and accurate patient encounter data

RESEARCH OUTCOMES

6

Improved quality of date leads to actionable and reliable results

These processes benefit VA Patients, and all patients, by impacting quality of care coordination throughout the entire health care ecosystem.

Clinical Data Quality Value to Clinicians & VA Patients

BETTER DATA, BETTER PATIENT CARE, BETTER OUTCOMES

Clinician

- **Clinical data usability leads to trust and adoption**
- **Data usability means fewer duplicate test procedures**
- **"Right data, right time" improves care coordination**

VA Patient

- **Enhanced clinical encounters**
- **Opportunity for improved outcomes**
- **Reduced time to diagnose and treat**
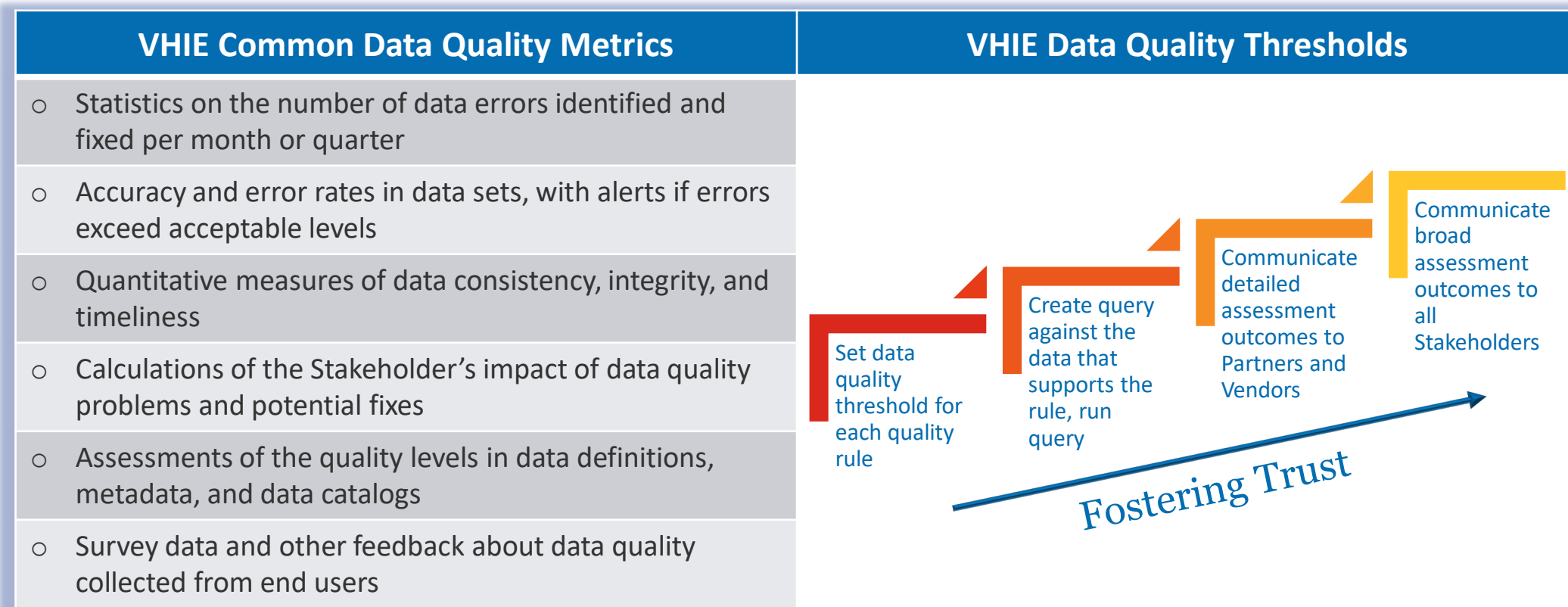- **Eliminate duplicate testing or the need to hand carry records**

VHIE CDQ Analyst

- **Awareness and Education**
- **Use Cases based on usability for the Clinician and Patient**
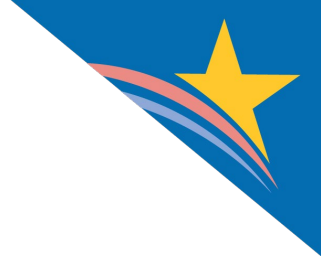- **Focus on Patient Safety**
- **Learning models**

Higher quality data exchange drives increased clinical adoption, reduces clinical burden,
and avoids duplication of services, leading to improved care coordination and patient satisfaction.

VHIE Methodology

VHIE Common Data Quality Metrics	VHIE Data Quality Thresholds
o Statistics on the number of data errors identified and fixed per month or quarter	
o Accuracy and error rates in data sets, with alerts if errors exceed acceptable levels	
o Quantitative measures of data consistency, integrity, and timeliness	
o Calculations of the Stakeholder's impact of data quality problems and potential fixes	
o Assessments of the quality levels in data definitions, metadata, and data catalogs	
o Survey data and other feedback about data quality collected from end users	

Set data quality threshold for each quality rule

Create query against the data that supports the rule, run query

Communicate detailed assessment outcomes to Partners and Vendors

Communicate broad assessment outcomes to all Stakeholders

Fostering Trust

Equip Clinicians with confidence in enhanced clinical data quality to improve outcomes and care coordination for patients.

VA Goals:
Alignment and Interoperability

Alignment with VA Goals

GOAL #1

VETERANS CHOOSE VA FOR EASY ACCESS, GREATER CHOICES, AND CLEAR INFORMATION TO MAKE INFORMED DECISIONS.

- Increased adoption across health care continuum
- Supports downstream systems, Artificial Intelligence, and Natural Language Processing

GOAL #2

VETERANS RECEIVE HIGHLY RELIABLE AND INTEGRATED CARE AND SUPPORT AND EXCELLENT CUSTOMER SERVICE THAT EMPHASIZES THEIR WELL-BEING AND INDEPENDENCE THROUGHOUT THEIR LIFE JOURNEY.

- Opportunity to improve patient outcomes with usable data
- Improve patient satisfaction with VA health care services

GOAL #3

VETERANS TRUST VA TO BE CONSISTENTLY ACCOUNTABLE AND TRANSPARENT.

- Non-VA Clinicians benefit from data across the health care continuum
- Retention of the original source document is required
- Improved VA business lines performance

GOAL #4

VA WILL TRANSFORM BUSINESS OPERATIONS BY MODERNIZING SYSTEMS AND FOCUSING RESOURCES MORE EFFICIENTLY TO BE COMPETITIVE AND TO PROVIDE WORLD-CLASS CUSTOMER SERVICE TO VETERANS AND THEIR EMPLOYEES.

- Empowers the VA to deliver high quality health care
- Clinical data quality auditing and scoring on each document
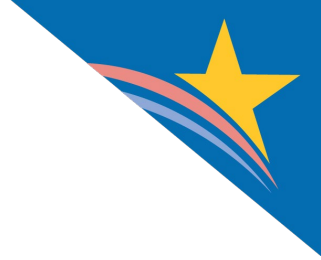- Identify potential wide-scale problems early in the timeline as a sentinel of national service

Clinical Data Quality Journey Towards Interoperability

- Interoperability of clinical data quality provides the environment to securely and reliably reach the right person at the right time, in the best manner to make an informed decision and take an action.

- Interoperable systems cross-populate data and allow authorized users to move between systems and seamlessly share data/information for consistent and easy access to secure health and benefits data.

- Interoperability creates opportunity to analyze results across use case assessments to identify opportunities to improve interoperability at the enterprise level.

VA PRINCIPLES
✓ Quality Service
✓ Management

Percentage Available by Clinical Domain

Viewable Partner Data Available

Percentage Range of Data Across Clinical Domains Provided by Partners/Vendors YTD*

(Estimated Percentage Range viewable by VA Clinicians)

90-100% Range Available **Problems**

13-93% Range Available **Allergy**

66-95% Range Available **Medications**
- Opportunities: Med Prescription at 80%

54-100% Range Available **Procedures**
- Opportunities: Epic.EAP.ID at 80%

21-99% Range Available **Immunizations**
- Opportunities: Lot Number at 40%

78-100% Range Available **Vital Signs**

40-100% Range Available **Results Domain**

- A high availability percentage suggests clinical and technical high-quality content consideration for display/ingestion.
- A low availability percentage warrants further examination by the source.

*In initial exploration of usability, ranges varied drastically
*Data derived from multiple participating external partners from 2022-23, randomized Consolidated Clinical Data Architecture (C-CDA) sample data sets
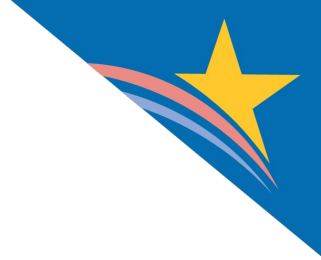
Clinical Data Quality Methodology

VHIE Clinical Data Quality Methodology

Clinical Data Quality's methodology includes an approach for scoring usability and is in the initial stages of development

✓ **Why develop a Clinical Data Quality Usability Methodology?**
- To articulate what clinical data quality means to the health care ecosystem
- To improve patient care and reduce clinician burden
- To augment conventional clinical data quality processes with measures aimed at assessing data usability
- To boost the efficacy to which evaluated clinical data can contribute to actionable, real-world value analytics

✓ **Thresholds Rationale**
- A percent minimum score must be attained for clinical reconciliation, which retains entire clinical intention
- Below percent minimum is deemed unusable and potentially detrimental to clinical decisions and patient well-being

CDQ Threshold Examples

An evolving Clinical Data Quality Usability Scale threshold implementation could be phased in over an x-year period.

Phase I: FY23- FYXX
- Attain 80% Usability

Phase II: FYXX - FYXX
- Attain 85% Usability

Phase III: FYXX - FYXX
- Attain 90% Usability

References

Use of a general purpose Healthcare Data Quality Framework (HDQF) provides a method to assess and visualize data quality to quickly identify areas for improvement.[1] VHIE Clinical Data Quality (CDQ) team is currently working with several HDQF frameworks, a JPSYS Usability framework and with an Elevance collaboration with DQFIT, an open source initiative.

There are many challenges with using EHR data for research including lack of data standardization, poor interoperability, privacy issues preventing aggregating complete healthcare records and poor data quality. The Office of the National Coordinator (ONC) is focusing attention on the first three issues, but the last issue, data quality, is not getting enough attention.[2]

However, as healthcare institutions across the country embark upon data governance initiatives, and as standard data elements become a reality for healthcare and health-related research, more and better machine-readable metadata are becoming available.[3]

Data-quality metrics falls into two main categories: (1) monitoring and (2) scorecards or dashboards. Monitors are used to detect violations that usually require immediate corrective action. Scorecards or dashboards allow numbers to be associated with the quality of the data and are more snapshot-in-time reports, as opposed to real-time triggers. Notice that results of monitor reports can be included in the overall calculation of scorecards and dashboards as well.[4]

Data-quality metrics need to be aligned with business KPIs throughout the company. Each LOB will have a list of KPIs for its particular needs, which must be collected by the data-quality forum and properly implemented into a set of monitors, scorecards, or both.[4]
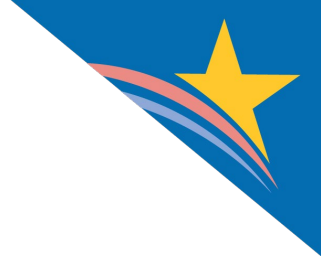
[1] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6568139/

[2] Butler M. Eight Predictions for ' 18 : Experts Prognosticate the Top HIM Topics for the Year Ahead and Advise on How to Prepare. J AHIMA. 2018;89(1):14–9. [Google Scholar]

[3] https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-6709-1

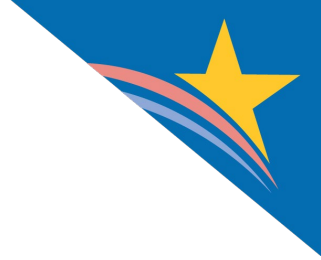[4] Data Quality Metric - an overview | ScienceDirect Topics

Q & A

Contact Us

To learn more about electronic health information exchange and clinical data quality, please contact us at vavhiecdq@va.gov or visit the VHIE website www.va.gov/vhie.

Thank you!

Veterans Health Information Exchange (VHIE) Program Office
Clinical Informatics Data and Management Office (CIDMO)
Office of Health Informatics (OHI)
Veterans Health Administration (VHA)

Break