CMS
CENTERS FOR MEDICARE & MEDICAID SERVICES

PUBLIC WEBINAR

Artificial Intelligence (AI) in Quality Measurement

Brian Anderson, MITRE

Jeffrey Geppert, Battelle

Agenda

Topics we'll cover include:

- Overview of the Coalition for Health AI (CHAI)
- A Nationwide Network of Health AI Assurance Laboratories
- AI Assurance in the Development and Evaluation of Clinical Quality Measures
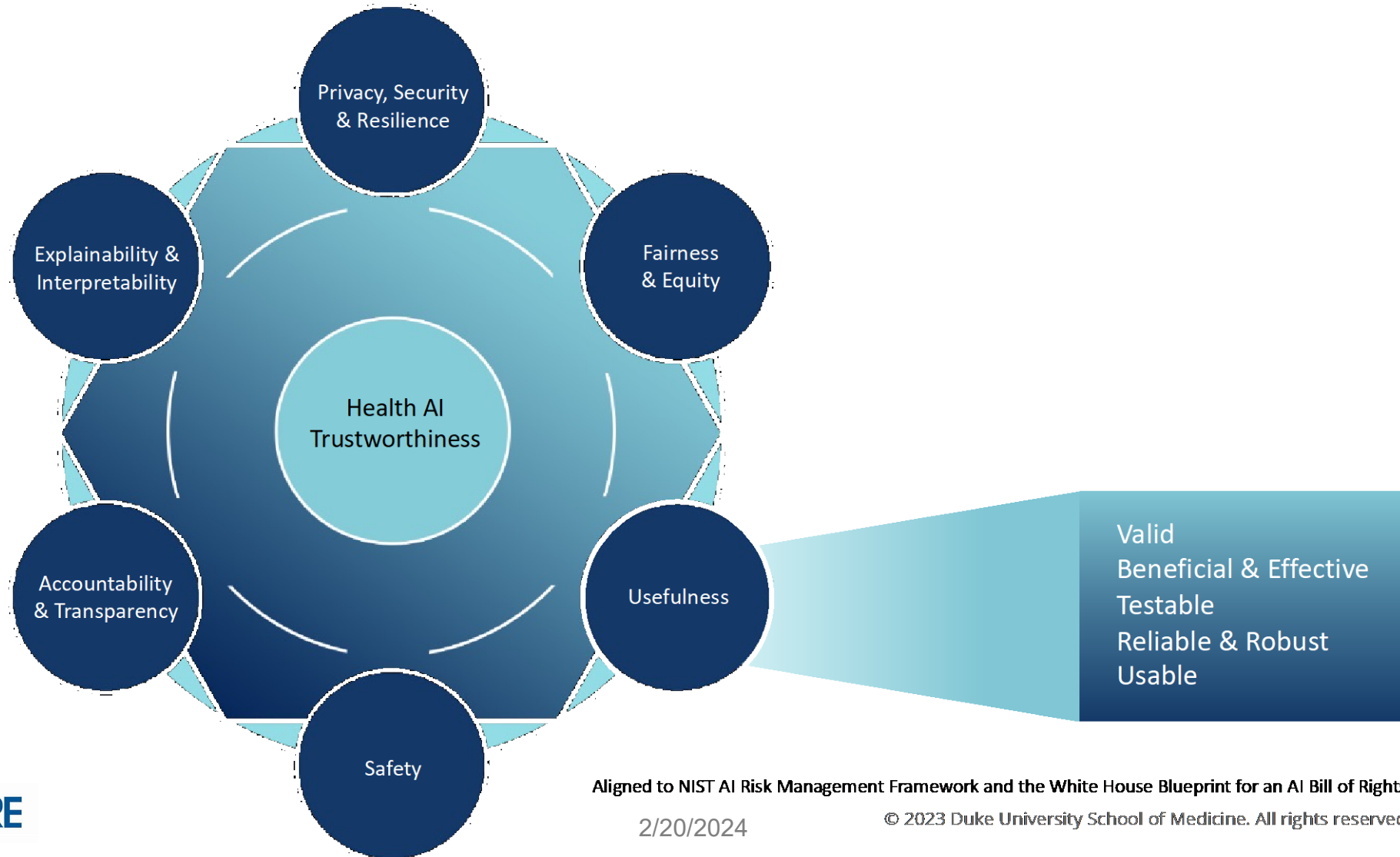
Who We Are

- **Over 1300+ Private Sector Organizations:** Health systems, payors, device manufacturers, technology companies, patient advocates

- **US Govt Partners:** HHS, FDA, ONC NIH, CMS, White House OSTP, AHRQ, VA, NIST, CDC, OCR
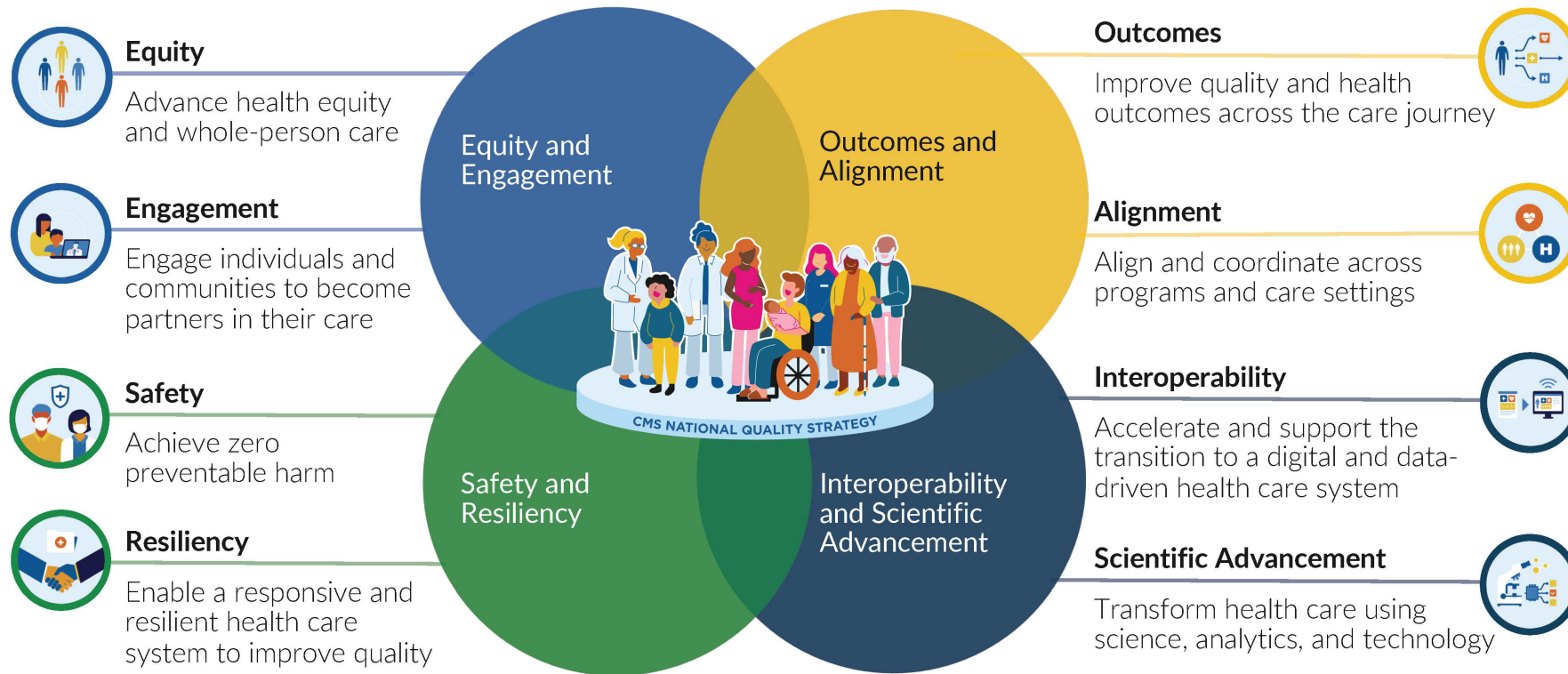
- Formally became **501c6 non-profit** in Jan 2024

MITRE

Core Principles for Responsible Health AI

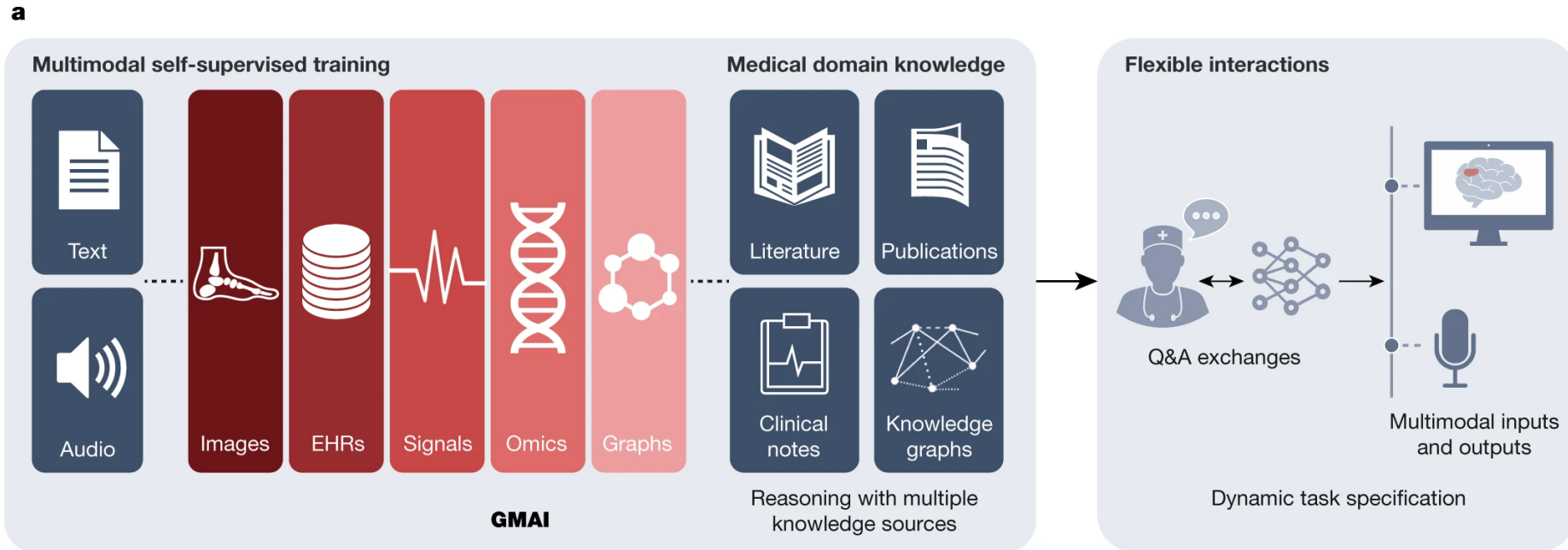Aligned to NIST AI Risk Management Framework and the White House Blueprint for an AI Bill of Rights

2/20/2024

4

CMS National Quality Strategy Goals

CMS National Quality Strategy Goals

Equity
Advance health equity and whole-person care

Engagement
Engage individuals and communities to become partners in their care

Safety
Achieve zero preventable harm

Resiliency
Enable a responsive and resilient health care system to improve quality

Equity and Engagement

Outcomes and Alignment

Safety and Resiliency

Interoperability and Scientific Advancement

CMS NATIONAL QUALITY STRATEGY

Outcomes
Improve quality and health outcomes across the care journey

Alignment
Align and coordinate across programs and care settings

Interoperability
Accelerate and support the transition to a digital and data-driven health care system

Scientific Advancement
Transform health care using science, analytics, and technology

MITRE

2/20/2024

a Multimodal self-supervised training: Text, Audio, Images, EHRs, Signals, Omics, Graphs (GMAI). Medical domain knowledge: Literature, Publications, Clinical notes, Knowledge graphs — Reasoning with multiple knowledge sources. Flexible interactions: Q&A exchanges, Multimodal inputs and outputs — Dynamic task specification.

b Applications: Chatbots for patients, Interactive note-taking, Augmented procedures, Grounded radiology reports, Text-to-protein generation, Bedside decision support.

Regulations: Application approval; validation; audits; community-based challenges; analyses of biases, fairness and diversity

Acronyms: Electronic health records (EHRs), Generalist medical artificial intelligence (GMAI)

Bridging the Gap between Risk Prediction & Quality Improvement

- Model Development for Outcome Risk Prediction & Stratification
  - CMS AI Outcome Challenge
    - Bias Mitigation
    - Transparency
    - Explainability & Interpretability
- Tying Quality Improvement Metrics to Clinical Outcomes
- Weighting & Instructional Tuning for Equity
- Causal Machine Learning & "Patients Like This"

MITRE

Reducing Burden – the power of LLMs

- Rethinking digital Quality Measures (dQMs) and Measurement Calculation Tools (MCTs)

2/20/2024

Acronyms: Electronic health records (EHRs), Fast Healthcare Interoperability Resources (FHIR)
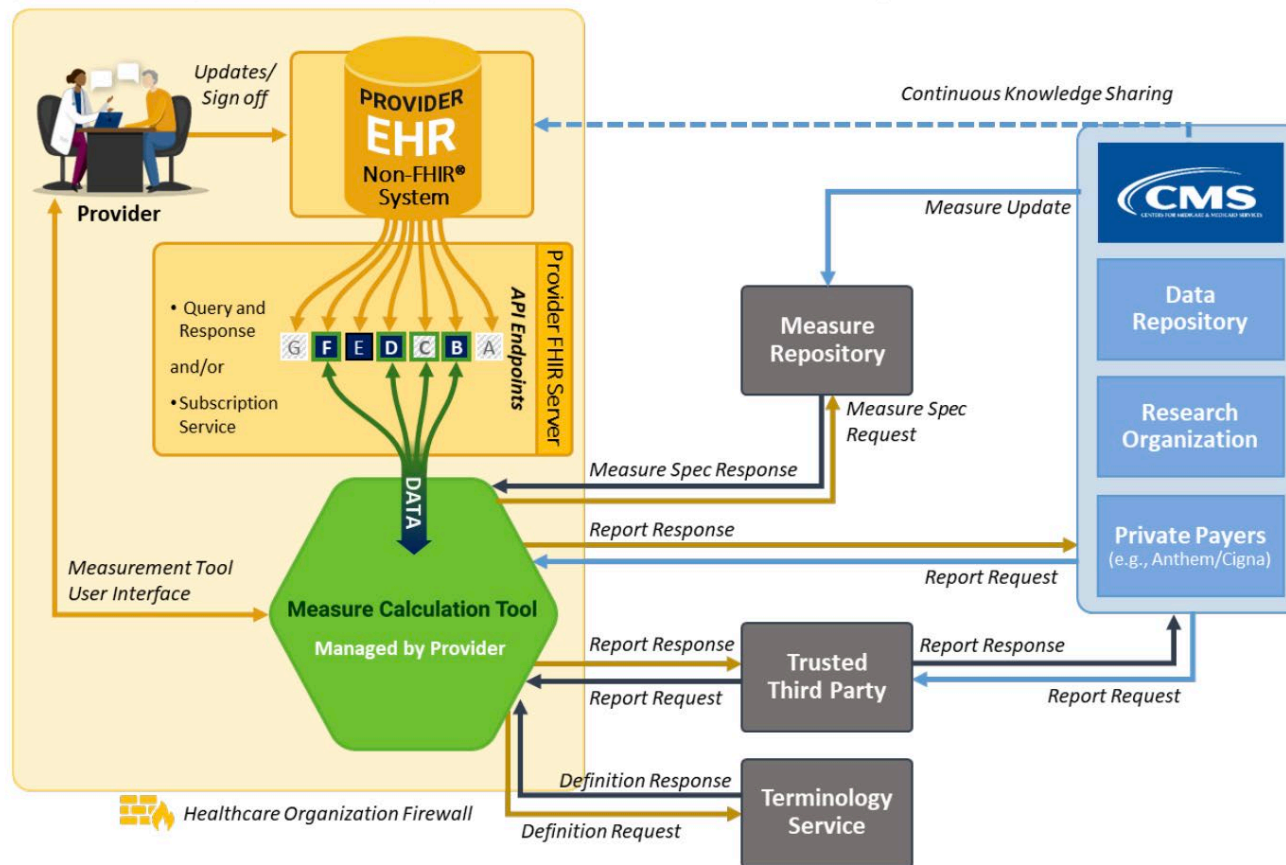
Generative AI and Summarization:
Automating Measurement Capture from Conversation to Orders

AI Assurance in the Development and Evaluation of Clinical Quality Measures

AI Assurance in the Development of CQM

An analogy with Patient Reported Outcome Measures-Performance Measures (PRO-PMs)

PRO-PMs use a Patient Reported Outcome Measure (PROM) in the construction of a measure

There are separate processes and criteria for evaluating the feasibility, reliability, and validity of the PROM separate from the PRO-PM

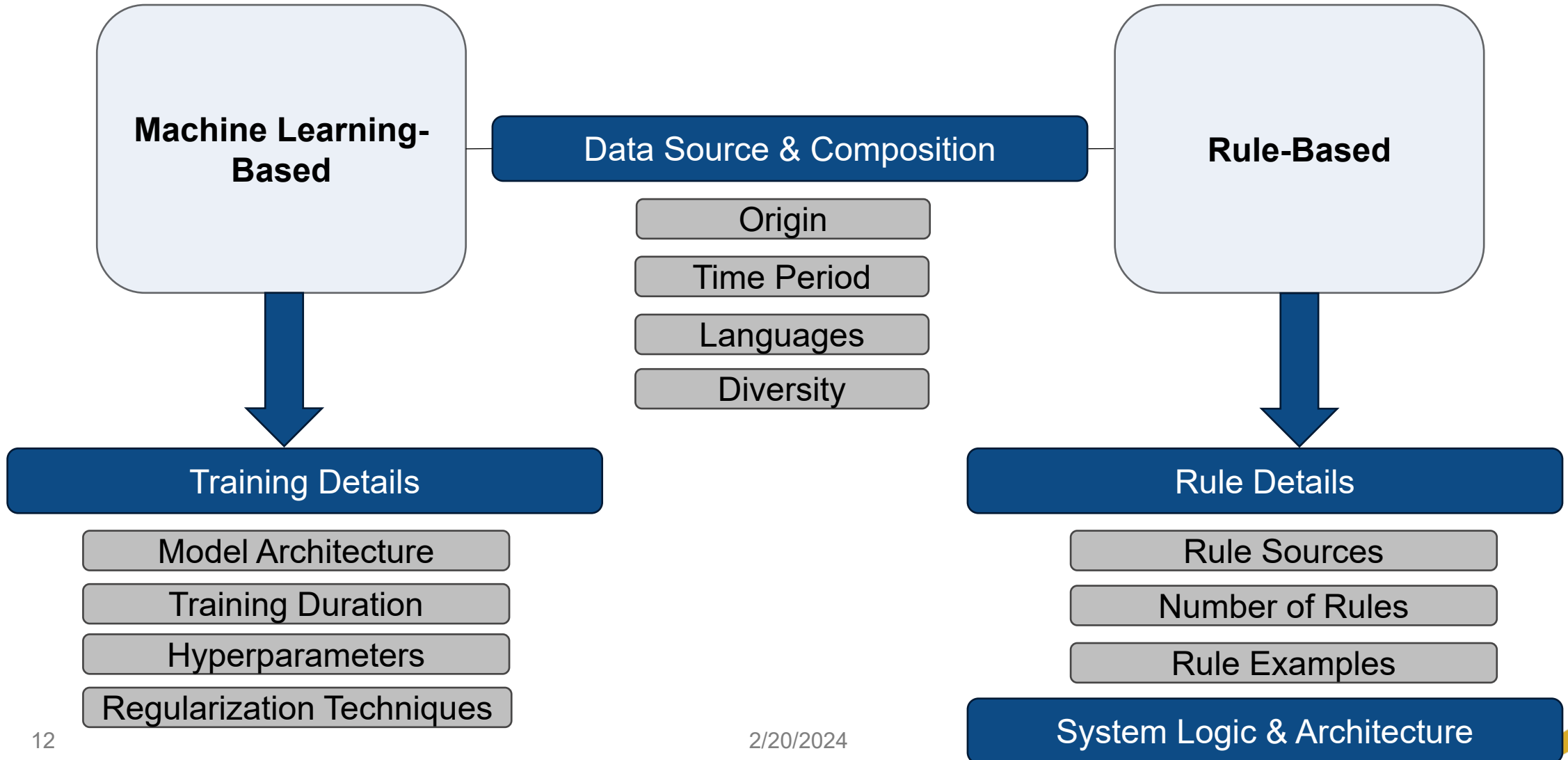**An analogy with electronic Clinical Quality Measures (eCQMs)**

eCQMs use specific standards for data elements and logic (e.g. Fast Healthcare Interoperability Resources)

There are separate feasibility, reliability, and validity considerations for eCQMs in addition to standard considerations that apply to all CQM
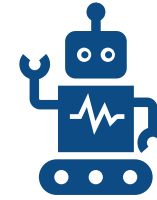
Clinical Quality Measures (CQMs) that use Artificial Intelligence (AI) in some way

A Potential AI "Model Card" for CQM

Machine Learning-Based

Data Source & Composition

Rule-Based

- Origin
- Time Period
- Languages
- Diversity

Training Details

- Model Architecture
- Training Duration
- Hyperparameters
- Regularization Techniques

Rule Details

- Rule Sources
- Number of Rules
- Rule Examples
- System Logic & Architecture

2/20/2024

A Potential AI "Model Card" for CQM Cont.

Rule-Based

Parsing Strategy

Rule Evaluation Order

Fallback Mechanisms

Dependencies & Resources

Dictionaries & Lexicons

Grammar Frameworks

Coverage & Limitations

Coverage

Known Gaps

Adaptability & Extension

Customizability

Integration Points

History & Evolution

Rule Revision History

Rationale

Collaborators & Contributors

Validation & Testing

Test Suites

Performance Metrics

Traceability

2/20/2024

A Potential AI "Model Card" for CQM Cont. 2

General

Performance Metrics

- Accuracy
- Precision, Recall, F1 Score
- Perplexity
- Benchmarks

Ethical & Bias Considerations

- Bias Mitigation
- Known Limitations
- Transparency Report

Usage Recommendations

- Intended Use
- Unsupported Uses
- Safety Measures

Versioning & Update Information

- Version Number
- Update Frequency
- Changelog

System Requirements

- Hardware Requirements
- Software Dependencies

License & Provenance

- Licensing Details

2/20/2024

AI Assurance in the Evaluation of CQM

Criteria/Assertions	Question	External Validity to Quality Program Population
Meaningfulness	Degree of uncertainty in how, why, for whom, and under what conditions a measure yields net benefit	
Feasibility	Minimize Burden	Does the explicit articulation of the people, processes, and technology required for data collection and reporting extrapolate to the quality program population? Do most entities in the quality program population have access to those people, processes, and technology
Appropriateness of scale	Degree of uncertainty in how, why, for whom, and under what conditions a measure's net benefit is distributed across persons and/or entities (net net benefit)	How are benefits and harms distributed across identifiable subpopulations of either persons or entities?
Time to value realization	Degree of uncertainty in how, why, for whom, and under what conditions a measure's net benefit and/or net net benefit will change over time	How might benefits and harms change over time?

AI Assurance in the Evaluation of CQM – Appropriateness of Scale – Lessons from Environmental Justice

Inequality

Refers to a difference in access across populations

Inequalities have many potential drivers and do not necessarily indicate a need for policy intervention (e.g., related to personal choice, cultural factors, disposable income)

Inequity

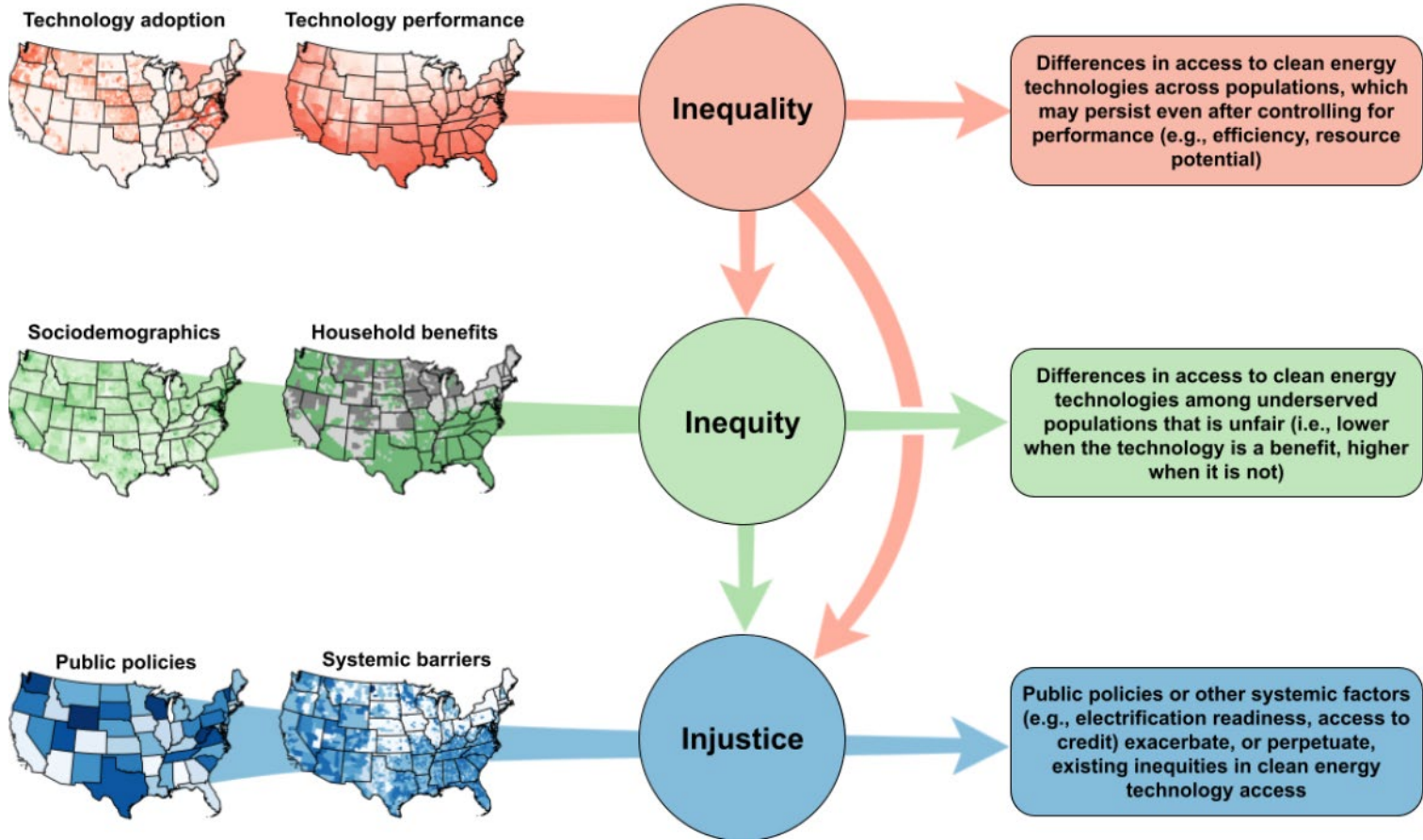Refers to a difference that persists even after controlling for these drivers/factors

Specifically, there may be lower access in underserved or under-resourced communities (e.g., rural communities, lower-income households)

Injustice

Refers to cases where policies or other persistent or systematic factors exacerbate, or perpetuate, existing inequities

2/20/2024

Differing levels of access to AI-enabled health care among various communities. For instance, high-income communities have more advanced AI technologies available in their health care systems compared to low-income communities.

Underserved or marginalized populations have less access to AI-enabled health care technologies due to factors like cost, lack of awareness, or limited AI-enabled health care infrastructure.

Systemic issues or policies that exacerbate or perpetuate inequities in access to AI-enabled health care. For example, regulations that do not ensure equal access or fail to address biases in AI algorithms

2/20/2024

Questions?

2/20/2024

Battelle
MMSSupport@battelle.org

CMS
Melissa Gross
melissa.gross@cms.hhs.gov

Angela Wright
angela.wright@cms.hhs.gov

Gequincia Polk
gequincia.polk@cms.hhs.gov