Foundational Factors Critical to Responsible AI: Ethics, Equity and Data

Session 10 | March 12th, 2024

Susan Dwyer, PhD
Associate Professor of Philosophy
University of Maryland College Park

Stephen J. Konya III, BAA
Senior Advisor to the Deputy National Coordinator, and Innovation Portfolio Lead for the Office of the National Coordinator for Health IT (ONC), U.S. Department of Health and Human Services (HHS)

Denis Newman-Griffis, PhD
Data Scientist | National Institutes of Health Clinical Center
Lecturer (Assistant Professor) in Data Science | University of Sheffield

Jessica Skopac - PhD, JD, MA
Senior Principal , Chief Engineer
Division of Healthcare Payment, Innovation, and Quality
The MITRE Corporation

informa markets

HIMSS 24
March 11-15 | Orlando

Session Moderator:
Jessica Skopac, PhD, JD, MA

Senior Principal Health Policy Analyst
The MITRE Corporation

Susan Dwyer, PhD

Associate Professor of Philosophy
University of Maryland College Park

Stephen J. Konya III, BAA

Senior Advisor to the Deputy National Coordinator, and Innovation Portfolio Lead for the Office of the National Coordinator for Health IT (ONC), U.S. Department of Health and Human Services (HHS)

Denis Newman-Griffis, PhD

Lecturer (Assistant Professor) in Data Science, University of Sheffield
Data Scientist, National Institutes of Health Clinical Center

Conflicts of Interest

Susan Dwyer, PhD

- No conflicts of interest reported

Stephen J. Konya III, BAA

- No conflicts of interest reported

Denis Newman-Griffis, PhD

- No conflicts of interest reported

Jessica Skopac - PhD, JD, MA

- No conflicts of interest reported

Agenda

- Background
- Current State of AI Regulation
- Potential Impacts of AI on Vulnerable Populations
- Data Integrity Implications for Equity
- Ethical Considerations

Learning Objectives

By the end of this session, attendees will be able to:

- Describe factors that are foundational to the development of AI solutions, including ethics, biased data & data quality related factors
- Explain how foundational factors impact the development & effectiveness of AI solutions
- Apply knowledge about ethics, biased data & data related factors that impact real-world examples of AI
- Compare integrations of AI into clinical practice in terms of their adherence to foundational principles of ethical & effective AI

Current State of AI Regulation

Stephen J. Konya III, BAA

Senior Advisor to the Deputy National Coordinator, and Innovation Portfolio Lead for the Office of the National Coordinator for Health IT (ONC)

HIMSS 24

Definitions

Predictive AI (aka Predictive Analytics): Predictive analytics is the use of data to predict future trends & events. It uses historical data to forecast potential scenarios that can help drive strategic decisions. (https://online.hbs.edu/blog/post/predictive-analytics)

Algorithm: A clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result. (NIST SP 800-107 Rev. 1)

Artificial Intelligence (AI): a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. (15 U.S.C. 9401(3))

Machine Learning: a set of techniques that can be used to train AI algorithms to improve performance at a task based on data. (EO 14110 §75195 (3)(t))

Generative AI: class of AI models that emulate the structure & characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, & other digital content. (EO 14110 §75193 (3)(b))

Challenges of Regulating Technology: Lessons Learned from Health IT Standards

Specific enough to be meaningful

Broad enough to evolve with technology

Cross-Agency Health AI Regulatory Activities

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Office for Civil Rights

FDA U.S. FOOD & DRUG ADMINISTRATION

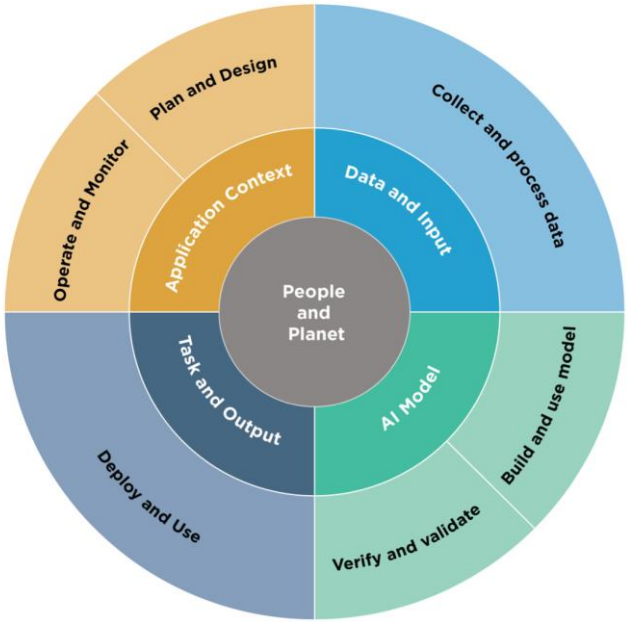
ONC
Office of the National Coordinator for Health Information Technology

Applicable Federal Policies		
Nondiscrimination in Health Programs & activities Proposed Rule (section 1557 of the Affordable Care Act)	CDS & Device Software Function-related Guidance Documents	ONC Health IT Certification Program (HTI-1 rulemaking)
Who Must Comply?		
Health care provider & health plan using AI to support decision-making in covered health programs & activities	Manufacturer of device software functions (e.g., AI-enabled software that meets the definition of a medical device)	Developers of certified IT that supply a predictive decision support interventions (DCI) as part of its Health IT Module
What Must Be Done?		
Not use clinical algorithms in any discriminatory ways (§ 92.210), covered entities may be held liable for decisions made in reliance on clinical algorithms	Require FDA-approval for demonstrating the device software function's safety & effectiveness	Provide transparency information about predictive DSIs to clinical customers & engage in risk management practices

https://www.federalregister.gov/documents/2022/08/04/2022-16217/nondiscrimination-in-health-programs-and-activities

Cross-Agency Health AI Regulatory Activities continued

National Institute of Standards & Technology AI Risk Management Framework (1/2023)



High Level AI Risk Management Activities
https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf page 10

CMS Medicare Advantage (MA) Rule (4/2023)

MA organizations must:

> ensure that they are making ***medical necessity determinations based on the circumstances of the specific individual…as opposed to using an algorithm*** (p.235)

> must comply with amended § 422.566(d)…which requires that a ***denial based on a medical necessity determination must be reviewed by a physician or other appropriate health care professional*** (p.235)

https://federalregister.gov/d/2023-07115

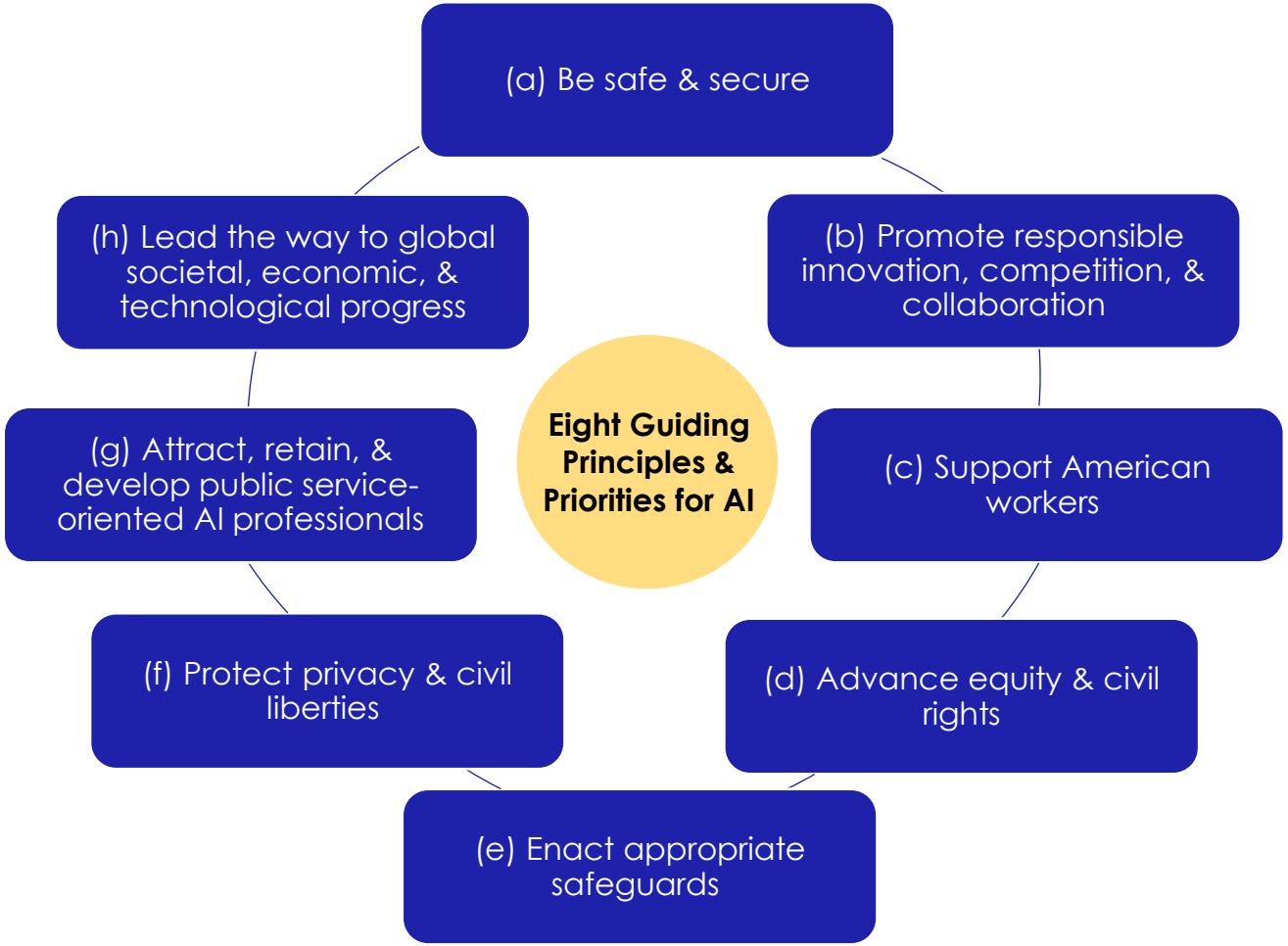
Executive Branch AI Activities

Executive Order 13960 (12/2020)

- Agencies shall b e guided by the following:
  - ***Lawful & Respectful of Our Nation's Values***
  - ***Purposeful & Performance-driven***
  - ***Accurate, reliable, & effective***
  - ***Safe, Secure, & Resilient***
  - ***Understandable***
  - ***Responsible & Traceable***
  - ***Regularly Monitored***
  - ***Transparent***
  - ***Accountable***

https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government

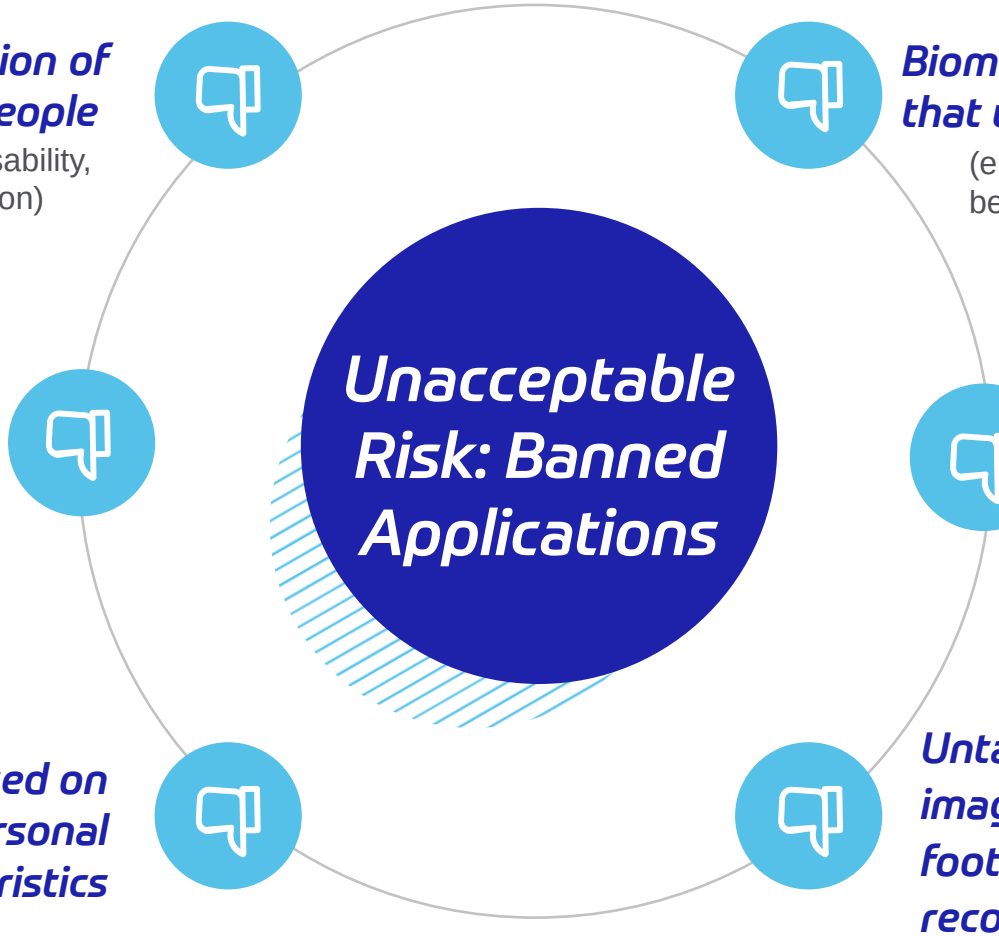
Executive Order 14110 (10/2023)



https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Overview of European Union Regulation (9/2023)

Artificial Intelligence Act: Deal on Comprehensive Rules for Trustworthy AI



Exploitation of vulnerabilities of people
(e.g., due to their age, disability, social or economic situation)

Biometric categorization systems that use sensitive characteristics
(e.g., political, religious, philosophical beliefs, sexual orientation, race)

Manipulation of human behavior to circumvent free will

Emotion recognition in the workplace & educational institution

Social scoring based on social behavior or personal characteristics

Untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases

Levels of Risk:
- **Unacceptable**
- **High**
- **Limited**
- **Minimal**

https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai

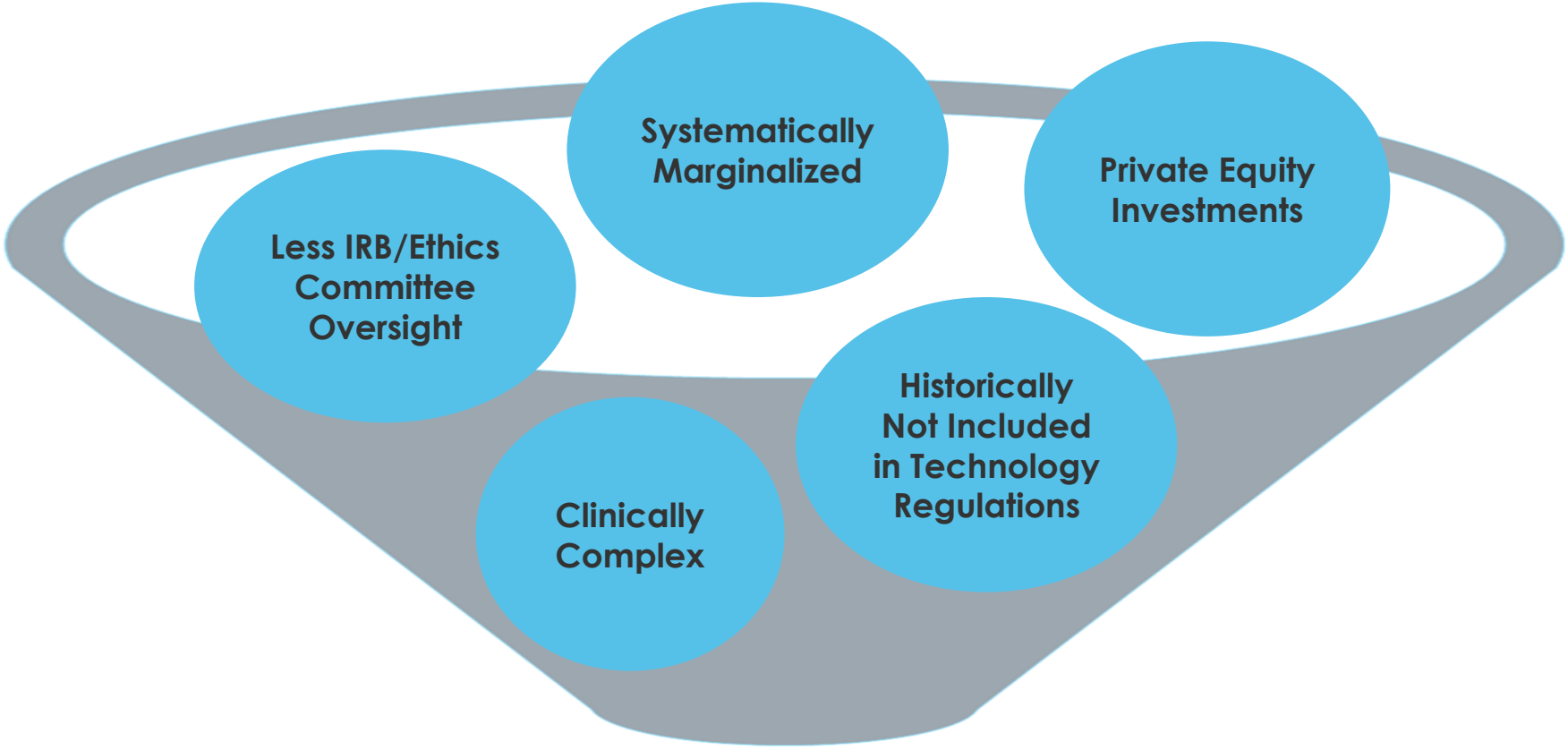
Potential Impacts of AI on Vulnerable Populations

Jessica Skopac, PhD, JD, MA

Senior Principal Health Policy Analyst
The MITRE Corporation

HIMSS 24

What Makes Post Acute Care (PAC) Consumers Particularly Vulnerable?



Sources available upon request

Untangling "Truthiness" and "Trustworthiness" in AI

"Truthiness"

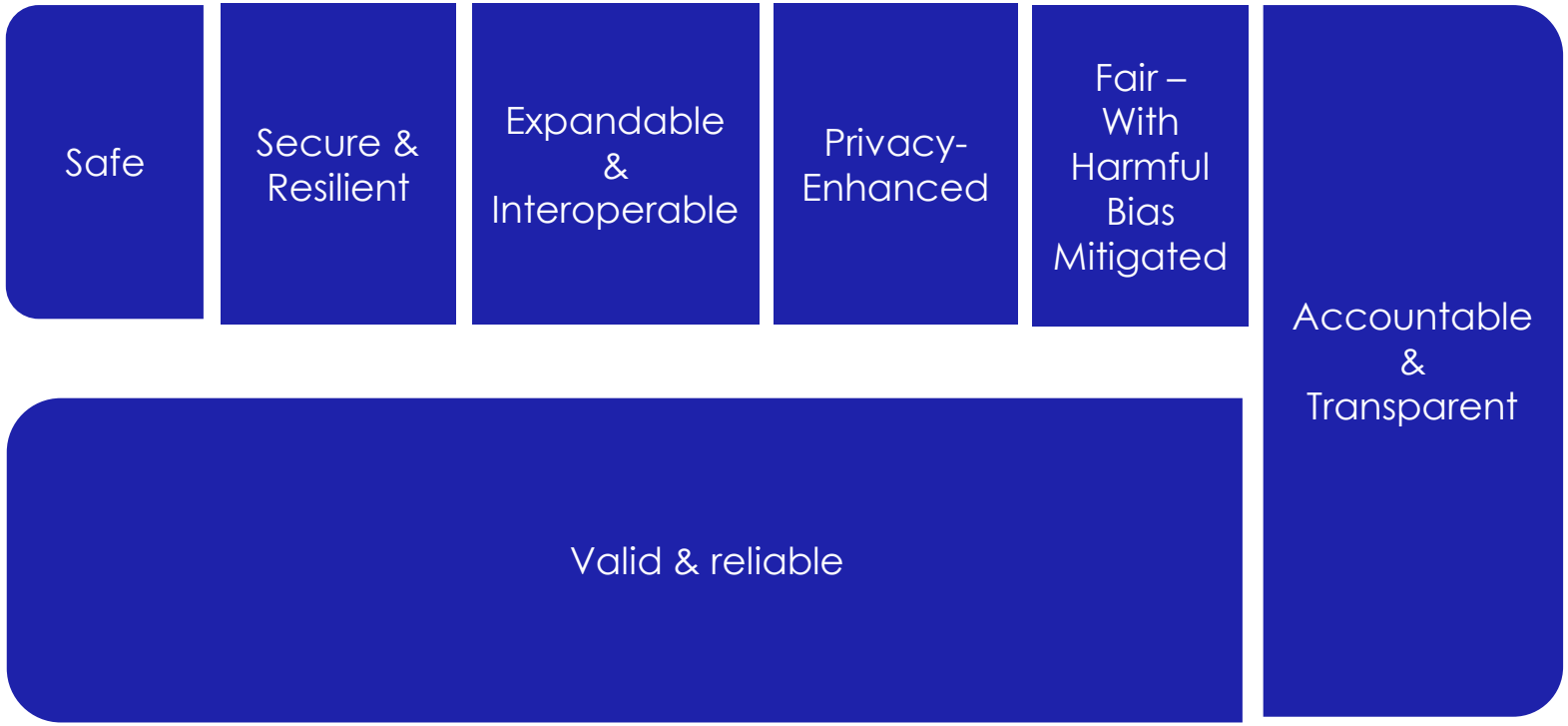
vs.

Trustworthiness

"the belief in what you feel to be true rather than what the facts will support"

- Stephen Colbert

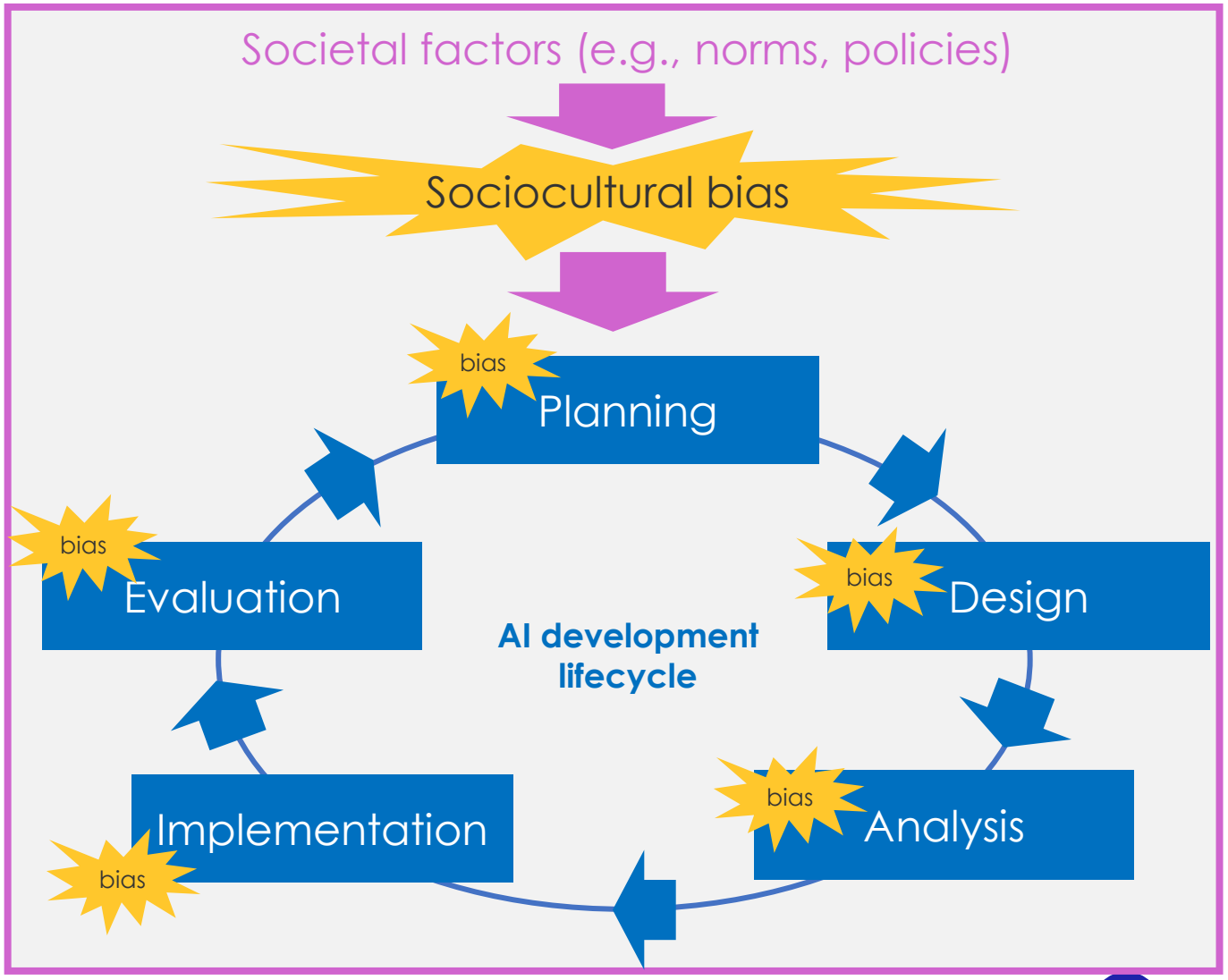
https://www.oed.com/search/dictionary/?scope=Entries&q=truthiness



https://airc.nist.gov/AI_RMF_Knowledge_Base/AI_RMF/Foundational_Information/3-sec-characteristics

Emerging Points of Failure in AI

- AI is flawed when trained with incomplete or problematic data
- AI itself is not flawed, but AI is deployed in a way that
  - Unintentionally results in disproportionate impacts to vulnerable populations
  - Intentionally & systematically disproportionately impacts vulnerable populations
  - Is inconsistently integrated into the workflow, resulting in inconsistent, inequitable outputs



Based on: Software Testing: Testing Across the Entire Software Development Life Cycle By Gerald D. Everett, Raymond McLeod, Jr.

Example: PaidLeave.ai to Help Consumers Identify Resources for Paid Family Leave in New York

- Paid leave is underutilized
  - Lack of awareness of benefits
  - struggle with administrative burden to access benefits
- PaidLeave.ai pilot to help parents in New York state access & apply for paid family leave benefits
  - Uses a human-like voice
  - helps parents maximize benefits & provides an action plan to complete & submit their claim

> *PaidLeave.ai is a powerful model for how generative AI can change the way we approach customer service for good. This is a big step forward in helping people more easily access the benefits they need to care for their families…*
>
> Craig Newmark, founder of Craigslist

> *PaidLeave.ai is a clear, high-impact example of how AI can help deliver public good.*
>
> Julie Samuels, President & Executive Director of Tech: NYC

Tools like PaidLeave.ai could easily be adapted to identify resources for PAC services rather than Paid Family Leave

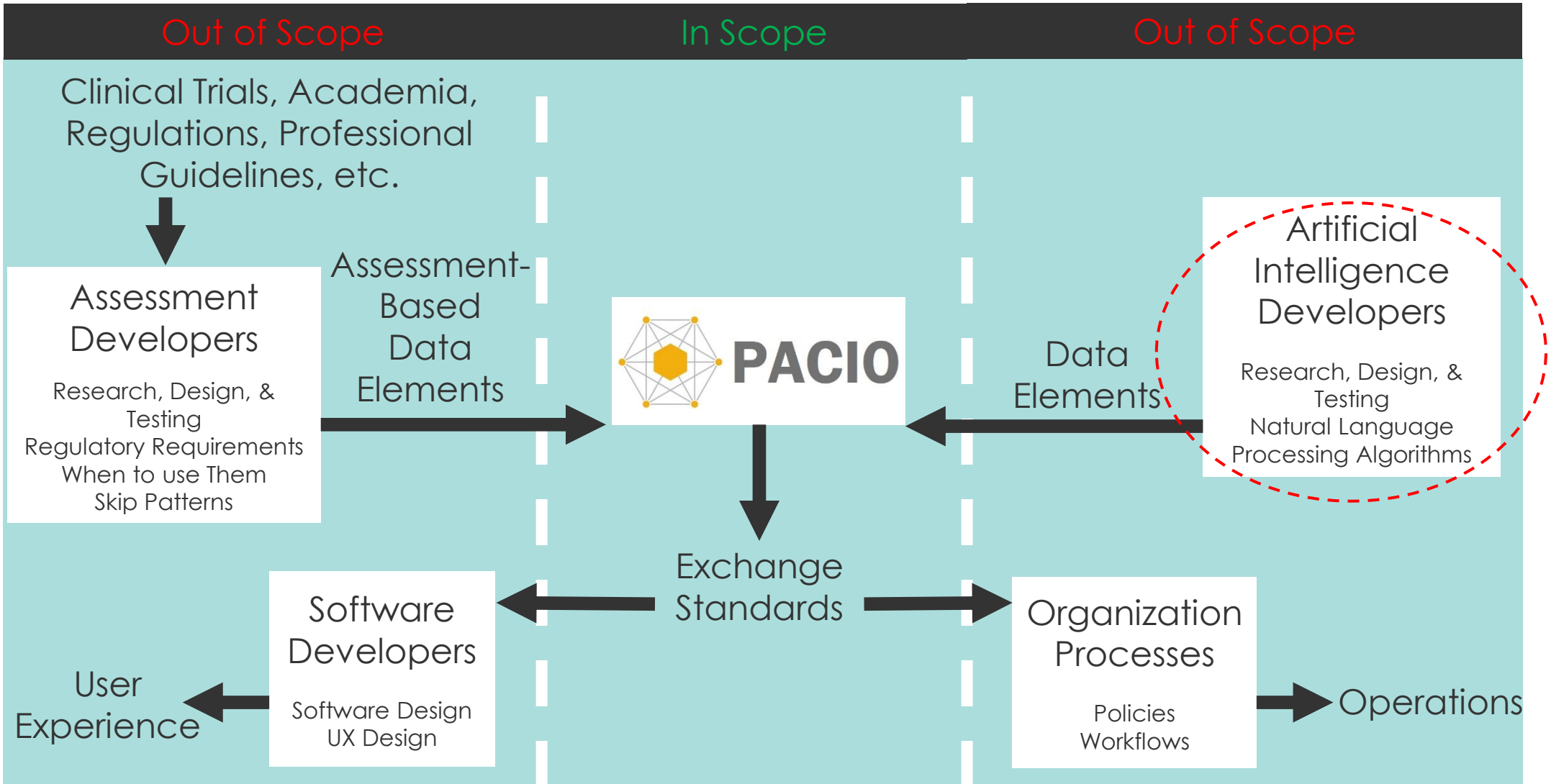
https://www.prnewswire.com/news-releases/moms-first-launches-paidleaveai-to-help-parents-in-new-york-access-paid-family-leave-302005318.html

Example of Laying a Foundation for Responsible AI with Standards: PACIO Project

Established February 2019, the PACIO Project is a collaborative effort between industry, government & other stakeholders, with the goal of establishing a framework for the development FHIR implementation guides to facilitate health information exchange.

http://pacioproject.org

PACIO Project: Scope



http://pacioproject.org

Data Integrity Implications for Equity

Denis Newman-Griffis, PhD

Lecturer (Assistant Professor) in Data Science

[1]University of Sheffield

[2]National Institutes of Health Clinical Center

Why Use AI?

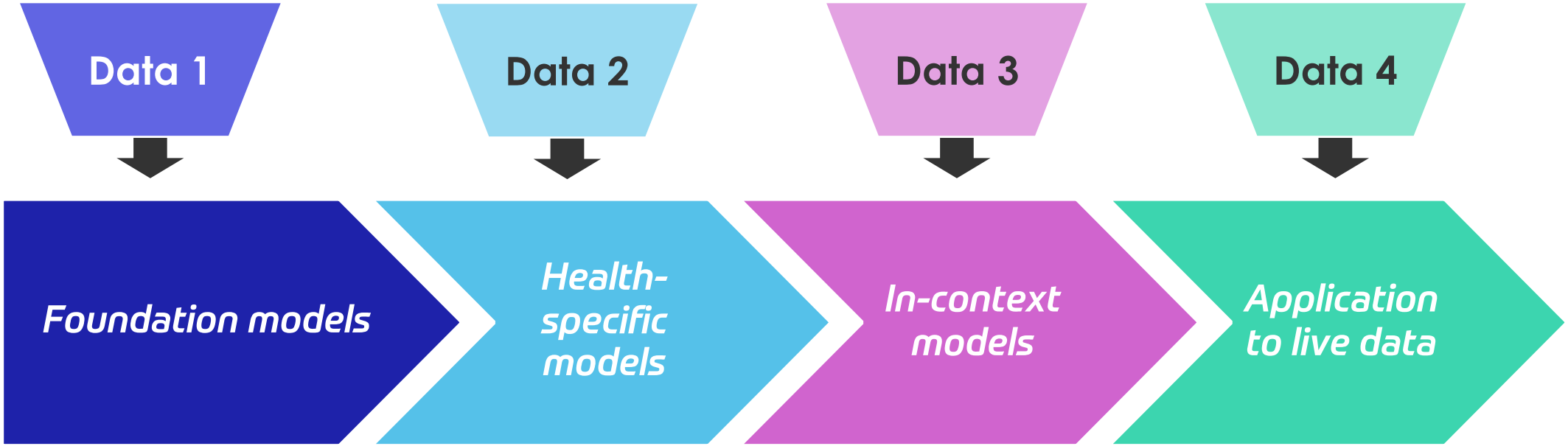
- **Learn** from patterns in data
- Deliver **better quality**, **more efficient** service

Our Data and Analytics Must Be:

- Does it mean what we think it does?
- Does it tell us what we want it to?

- Does it represent the people we are serving?
- Does it tell us what is important to them?

Today's AI Workflows: Many Entry Points for Data



- Non-health data
- Trained to do something else

- **Health** data
- Trained to do something else

- **Health** data
- Trained to do **target task**

- **Health** data
- Trained to do **target task**
- **Actually being used!**

Stages of Data Risk

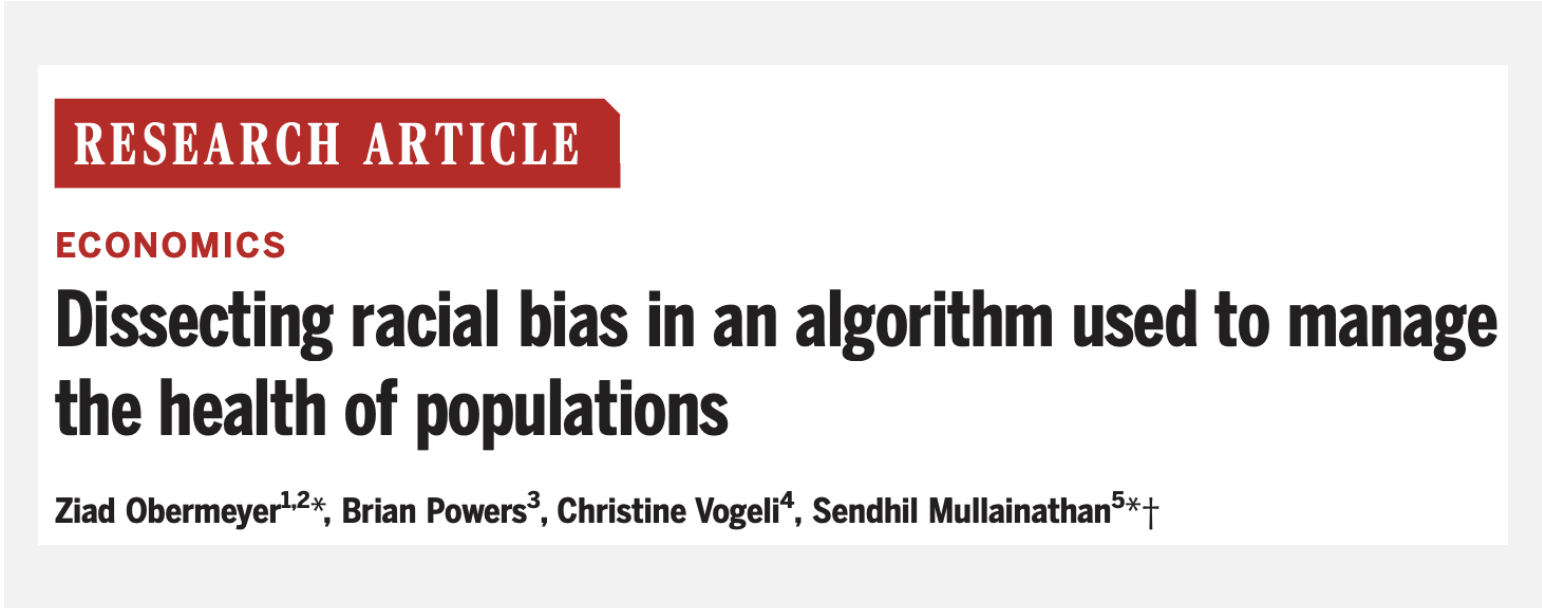
from ***Person***
to ***Information***

from ***Information***
to ***Data***

from ***Data***
to ***Insight***

Example: Risks in Measurement

- Race-blind algorithm for estimating health risk
- Based on estimated cost of care
- But less is spent on Black patients!
- Result: **racially-biased care**



RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2]*, Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5]*†

Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-453. doi:10.1126/science.aax2342

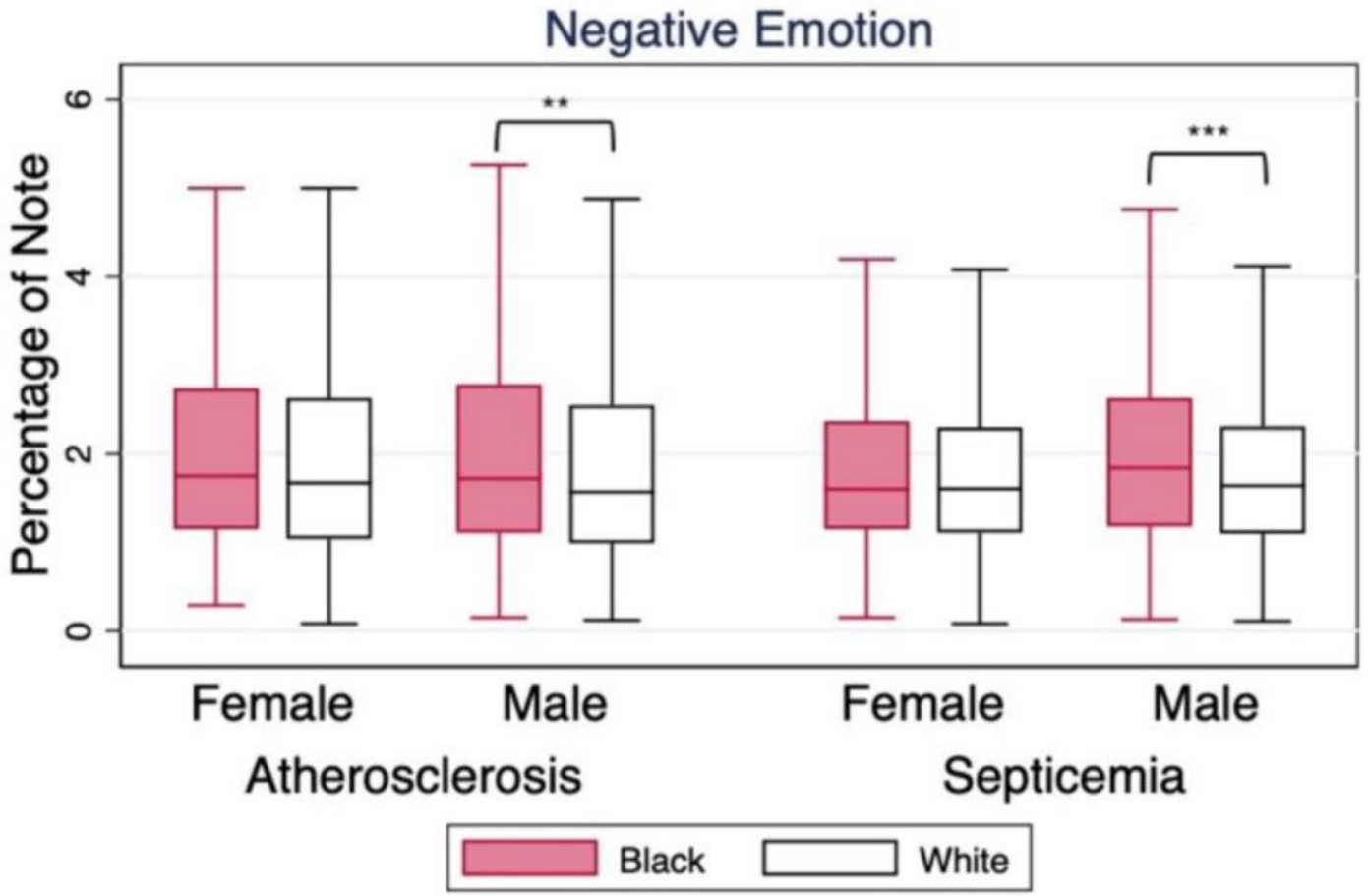
Key Questions: Measurement

Data are not neutral…
…and may not tell you what you think!

- How reliably is this data point captured?
- Am I confident it tells me what it should?

Example: Risks in Data Capture

- Provider biases can affect **quality** and **experience** of patient interaction
- They also affect **recorded data!**
- Different tone in EHR notes for White vs Black patients
- Result: **Biased readers and algorithms**



Penn, J, Newman-Griffis, D. Half the picture: Word frequencies reveal racial differences in clinical documentation, but not their causes. AMIA (2022).

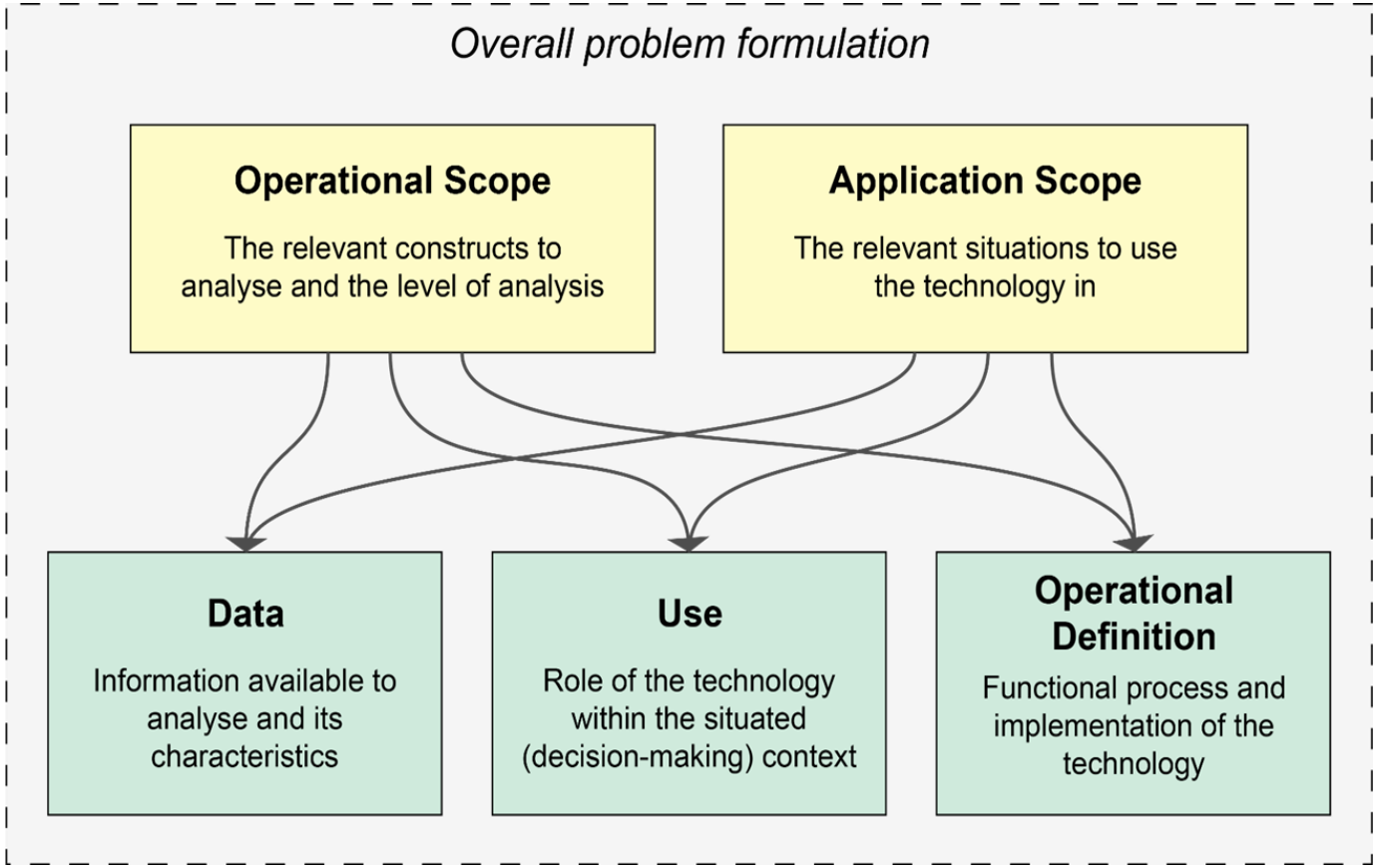
Key Questions: Data Capture

What we choose to record and how it affects the AI using our data

- Are data recorded the same for all my patient populations?
- Whose perspective does my data represent?

Example: Risks in Use of AI

- Many different ways to define and measure **disability**
- Many design decisions in building AI systems
- AI designers' definitions may **not match** providers OR patients!
- Result: **incompatible and harmful AI systems**



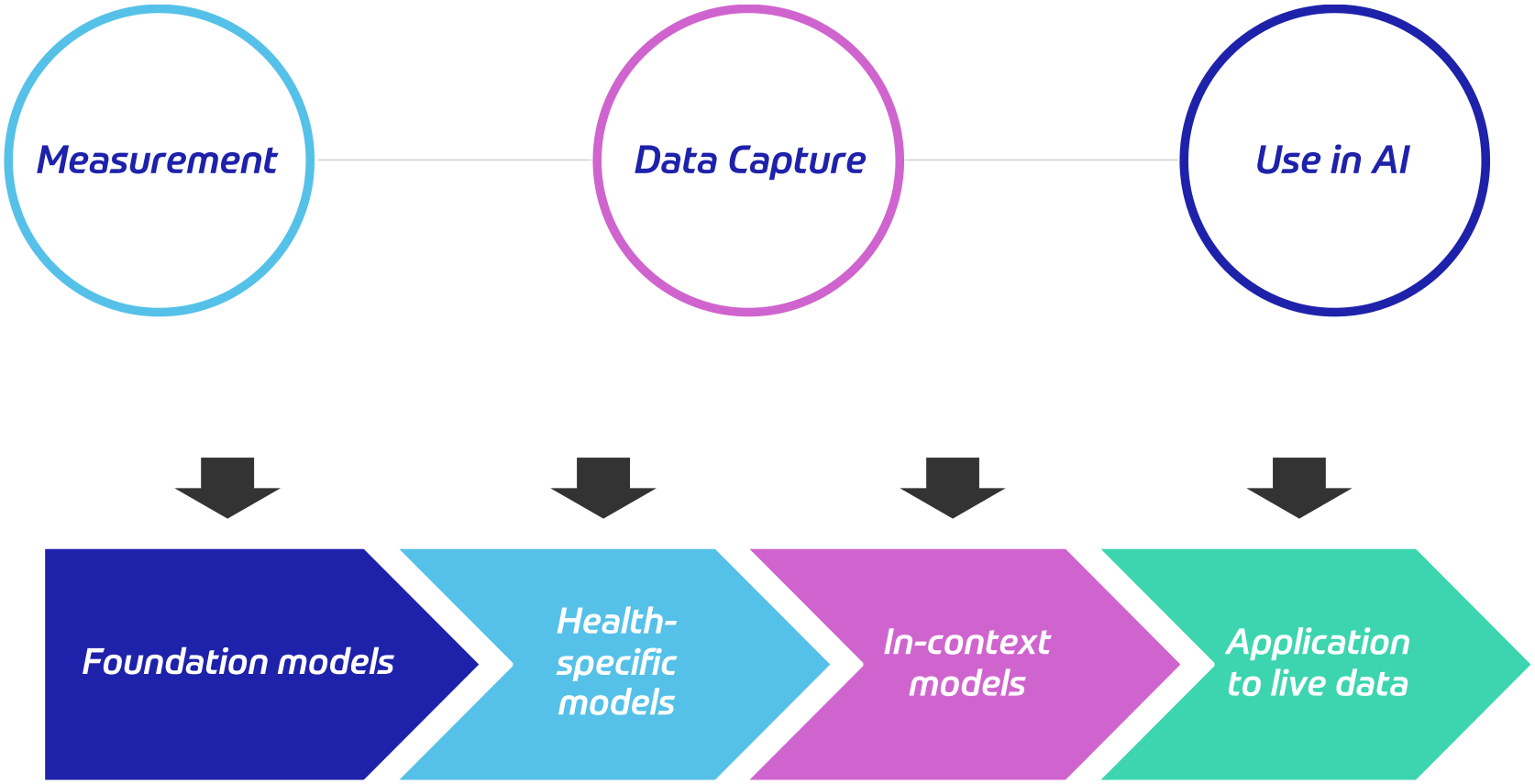
Newman-Griffis, D, et al. Definition drives design: Disability models and mechanisms of bias in AI technologies. *First Monday* (2023).

Key Questions: Use in AI

AI systems are full of invisible design decisions about data and patients

- How does my AI implementation operationalize health and care?
- What information am I emphasizing and why?

Data Risks Affect Every Stage of Today's AI Pipeline



Tools to Assess Data & AI Integrity

1

Data auditing

- Consistent quality across populations
- Data history in underlying models

2

No assumptions in AI design

- Clear agreement on definitions
- Feedback cycles and oversight throughout
- Involve patients!

3

Evaluate in context

- Effects on **your** process with **your** patients
- More than just AI accuracy!

4

Monitor & mitigate

- Continuous improvement
- Measure for bias, respond and adjust

Ethical Considerations: Use of AI in Post-Acute Care of the Elderly

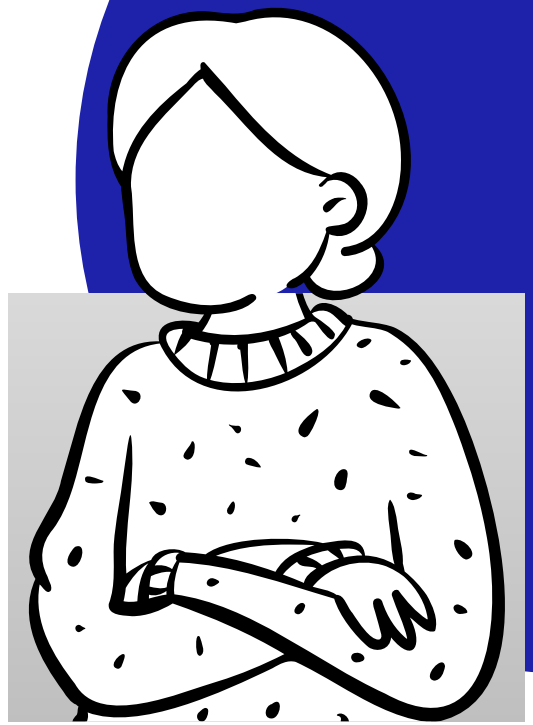
Susan Dwyer, PhD

Associate Professor of Philosophy
University of Maryland College Park

HIMSS 24

Case Study: B

- 81-year-old female lives alone
- Around 12am: Falls in bathroom
- Around 8am: Discovered by a friend
- Ambulance transports her to hospital
- Conditions: Bleeding duodenal ulcer, diverticulitis, Type II diabetes, anemia, severe dehydration, low blood pressure, confusion, disoriented
- Discharged after 6 days
- At home: weak, frightened, cannot drive; overwhelmed



Issues & Questions & Context

Issues

- Living & health circumstances can quickly change
- Medical, psychological, spiritual disruptions
- Absence of engaged family
- Lack of plans

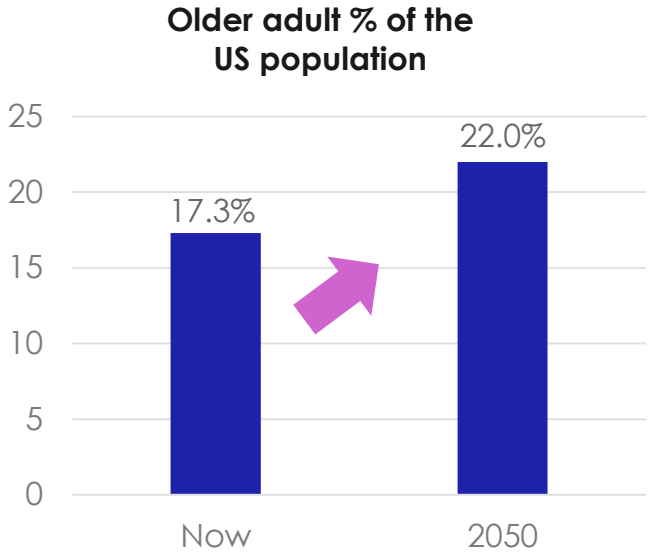
Pew Research Center. 2019/ "Religion and Living Arrangements Around the World". Associated Press-NORC Center for Public Affairs Research. 2021. "Long Term Care in America: Americans Want to Age At Home".

Questions

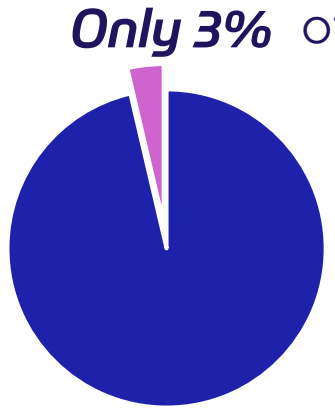
- How are such individuals to be assisted & supported?
- What role can AI-driven systems play here?
- What ethical considerations ought designers of such systems take into account?

Elements

- Rapid growth of older adults in post-acute care
- Current eldercare crisis in USA
- Increase in pressure to deploy AI to manage eldercare
- Few Americans have eldercare insurance



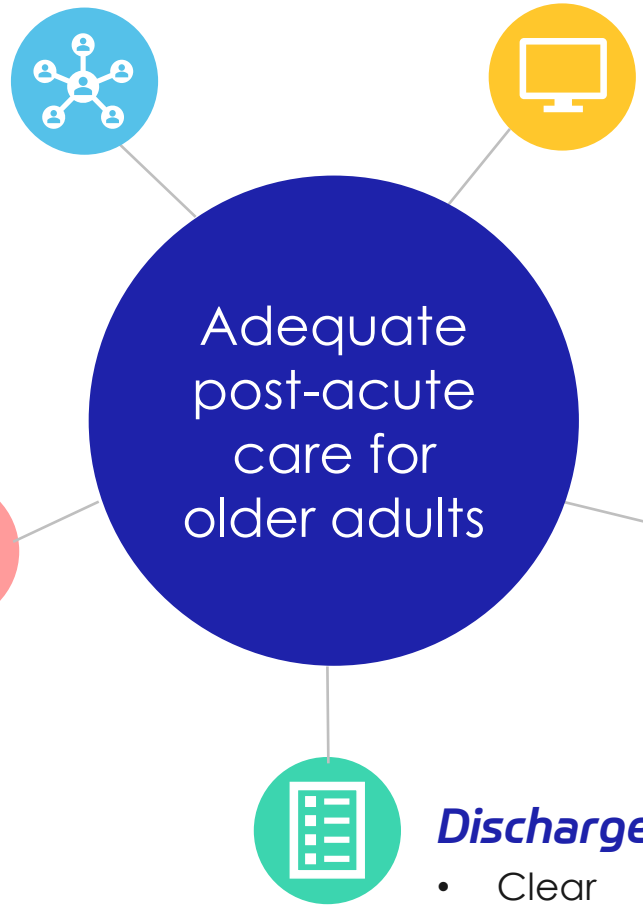
https://www.prb.org/resources/fact-sheet-aging-in-the-united-states



of Americans have long term care insurance

https://www.limra.com/en/newsroom/industry-trends/2022/do-consumers-really-understand-long-term-care-insurance/

Obvious Needs



Newly-needed specialists

- Referrals
- Appointments

Medical records

- Accurate
- Up-to-date
- Sharable with healthcare providers & caregivers

Therapy providers

- Referrals
- Appointments

Medications

- Dispensing
- Management

Discharge orders

- Clear
- Understandable
- Actionable

Family/friends may not be able/willing to support care needs

If Appropriately Designed, AI Can:

1. ***Support care coordination***
   AI-supported platforms can help share & analyze information

2. ***Facilitate self-care engagement***
   AI can support ongoing patient/caregiver education, recommend plans of action, deliver reminders/nudges

3. ***Support home-based care***
   AI assistive robots can monitor daily activities, report anomalous behaviors, provide alerts to caregivers, offer 'company' & cognitive engagement

No healthcare technology can be appropriately designed unless its designers' pay attention to substantive ethical issues

AI Ethical Principles

1. Transparency & accountability
2. Justice & fairness (non-discrimination)
3. Non-maleficence
4. Responsibility
5. Privacy
6. Safety & security
7. Freedom & autonomy

All can be deployed in the design of AI systems for post-acute care

Hongladarom, Soraj and Jerd Bandasak. "Non-western AI ethics guidelines: implications for intercultural ethics of technology." AI & Society, April 2023.

Familiar Ethical Considerations

Elder-centered AI

The principle-based approach to designing ethics into AI will need to be rethought from the perspective of the elderly & with attention to the varied dimensions late in life

Autonomy

Respecting the wishes, decisions, & bodily integrity of competent adults has been a central plank of bioethics from the beginning

- The desire "help" & "support" older adults must be tempered with respect for such individuals to refuse that help & support
- The risk is that AI systems will default to a particular conception of what is "best" for such patients

Privacy

Respecting patient privacy is also baked into bioethics & into law

- It's not unusual for conditions that a person has kept secret from family (if in the picture) &/or friends to be revealed in emergency situations, leaving the elder to contend with intense feelings of embarrassment or shame in addition their health & living needs
- The risk is that AI systems will default to revealing all the data it 'considers' pertinent.

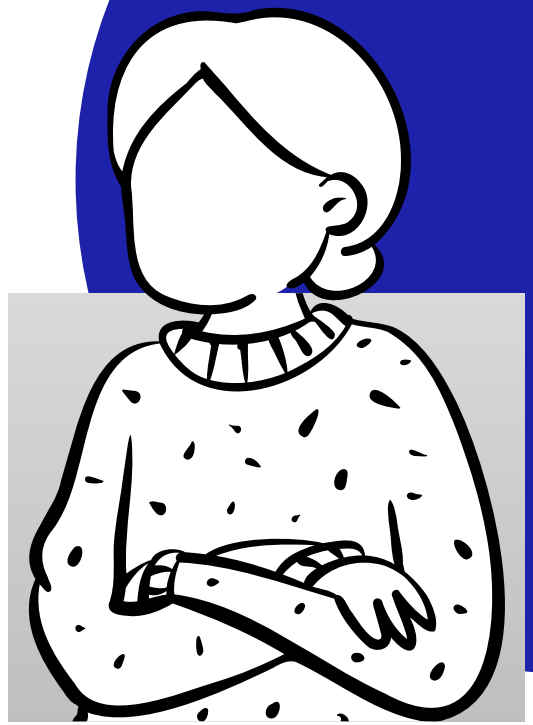
Something is Still Missing

Remember B

- Acute health crisis
- Little agency
- Does not understand what to do
- Multiple significant disruptions

These bring up thoughts of mortality for B

- We all die, each death is different



Implications

Confrontation With Mortality

- Post-acute care goes beyond practicalities
- Impacts patient, family, friends

Lack of Patterns

- Few patterns in dying
- Yet patterns fuel AI

Desire for Control over Dying

- Want control over dying process
- Do not want/need AI involved

Two Options to Pursue to Build Appropriate Datasets

- **Benefits**: access to ground truths about what is needed & wanted from all parties concerned; reduced reliance on assumptions; increased trustworthiness of systems
- **Challenges**: assuring equal epistemic authority on panels; translating qualitative data to computable data; protecting against vested interests

- Little known about ethical concerns of patients & families
- Cho et al. (2020):
  - tension & disagreement between patients & family
  - disagreement & lack of communication with healthcare providers
  - uncertainty about limited decision-making capacity
  - lack of knowledge of end-of-life care & planning
- Generalities emerge that might permit AI-supported scenario tools

Cho et al. 2020, "Patient and Family Descriptions of Ethical Concerns," *The American Journal of Bioethics* No. 6, 52-64.
Bennett V et al. 2023. "Development of a Lived Experience Panel to inform the design of embedded pragmatic trials of dementia care interventions." Journal of the American Geriatrics Society. 2023.

Discussion & questions

Thank you

Contact information

Dr. Jessica Skopac - jskopac@mitre.org
Dr. Denis Newman-Griffis - d.r.newman-griffis@sheffield.ac.uk
Dr. Susan Dwyer - dwyer@umd.edu
Stephen Konya III – stephen.konya@HHS.gov