# Data driven policy for pandemic response: Data, knowledge, and action well outside of the ivory tower

**Jessie Tenenbaum, PhD**
Duke University School of Medicine

**Chief Data Officer,**
**NC Dept of Health and Human Services**

@jessiet1023

**November 3, 2022**

# Disclaimer

“Photo” credit DALL-E

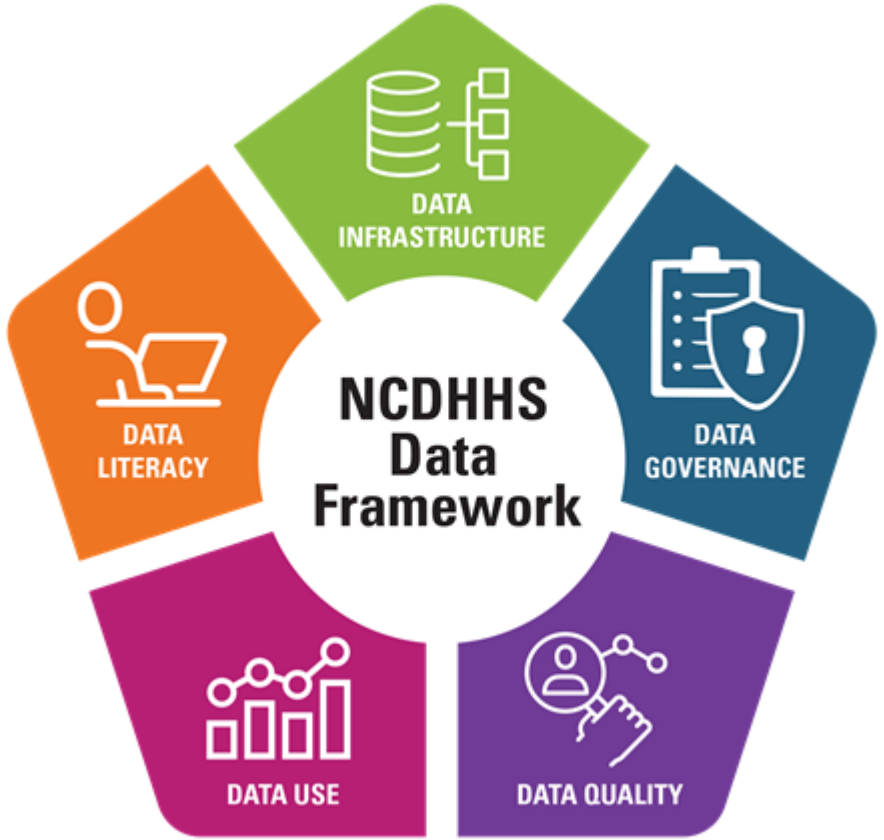
https://labs.openai.com/

NC DEPARTMENT OF
# HEALTH AND HUMAN SERVICES

**https://www.ncdhhs.gov/**

# NCDHHS Data Office

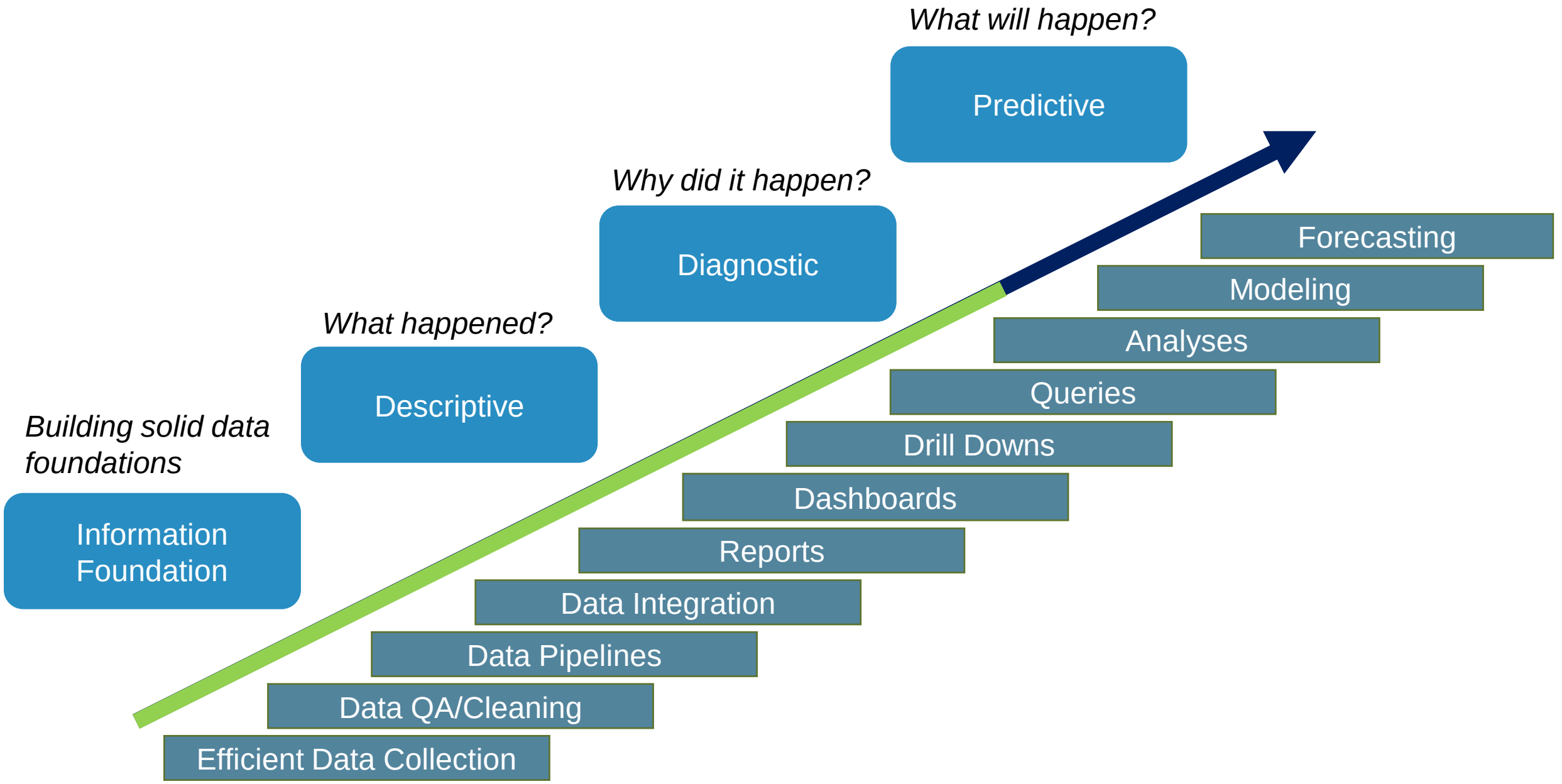
- Established 2019
- Led by Chief Data Officer
- Drive Department-wide cultural transformation
- Data as 1st class citizen, strategic business asset
- Data strategy and roadmap based on 5 data pillars



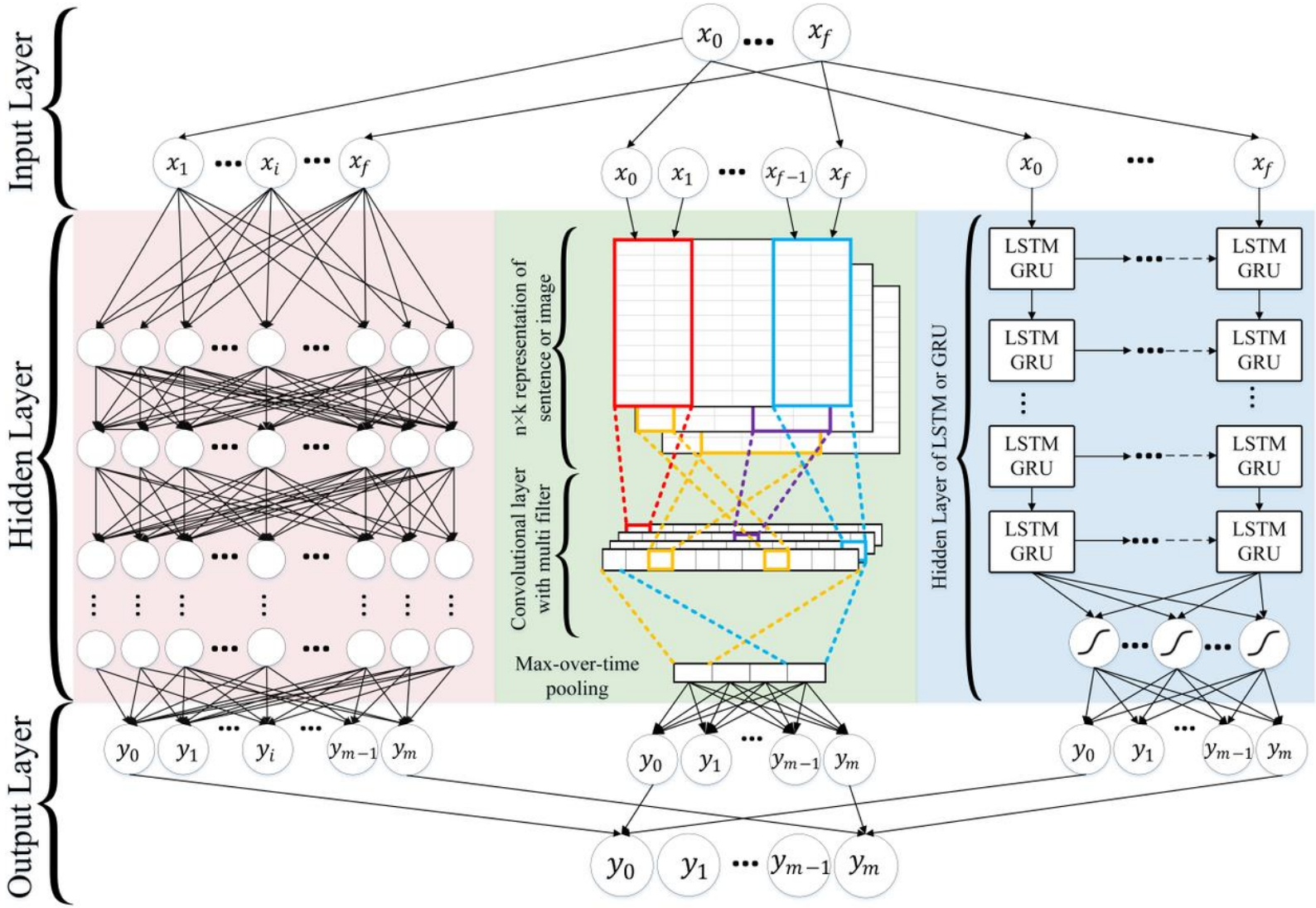
# Data Strategy Framework

- **Data Infrastructure:** Technology used to store, exchange, and access data
- **Data Governance:** People, processes, and technology for data security, management, and access
- **Data Quality:** Tools and processes to ensure data are valid, accurate, and trustworthy
- **Data Literacy:** Workforce training across all levels of baseline knowledge
- **Data Use:** Reports, visualization, and analysis

# Moving from descriptive to predictive



# What academics are focused on



# What most of the world needs

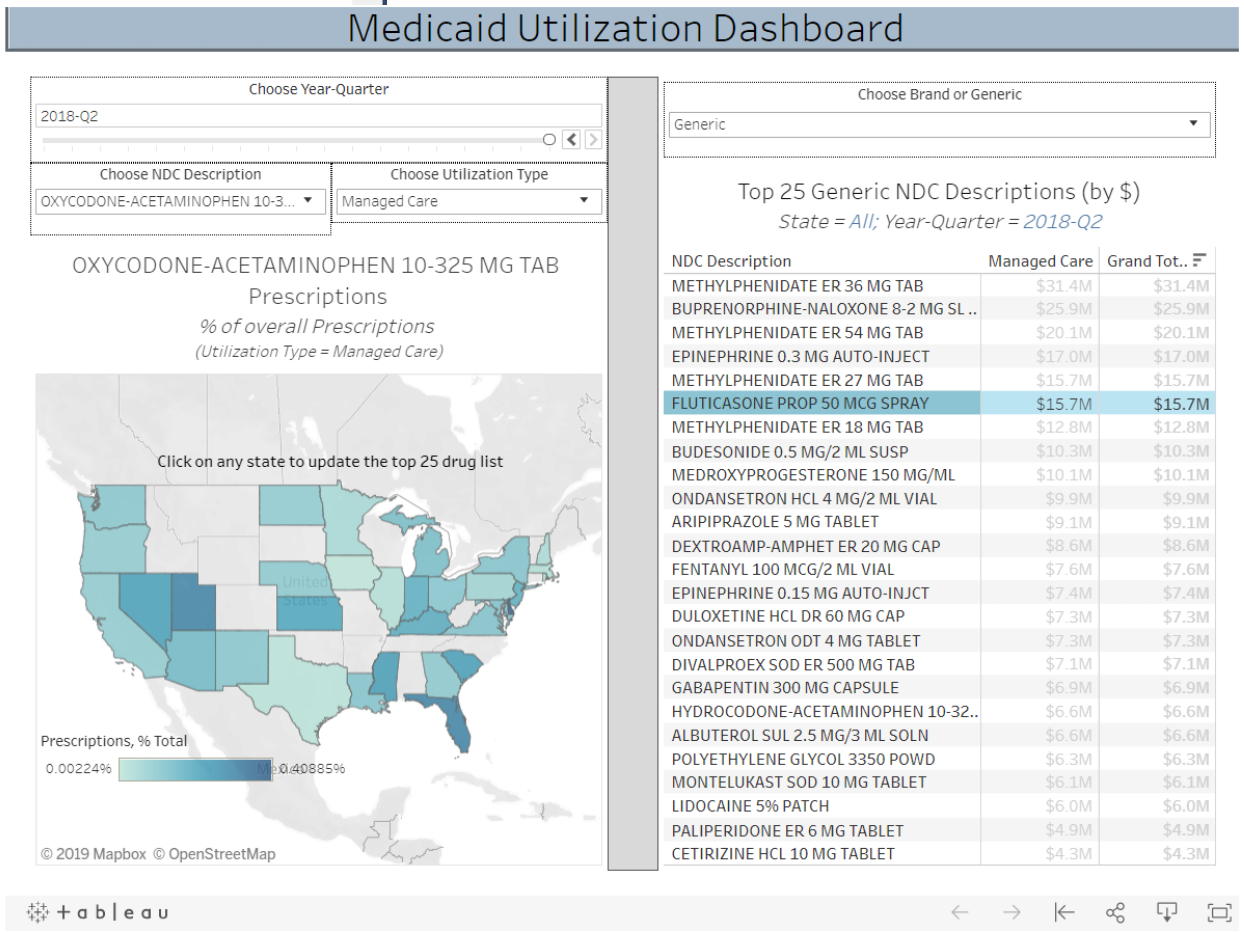
```sql
-- SQL Server Group By Example
USE [SQL Tutorial]
GO
SELECT Education,
    COUNT([EmpID]) AS [Total IDs]
    FROM [Customer]
    GROUP BY Education
```

100 %

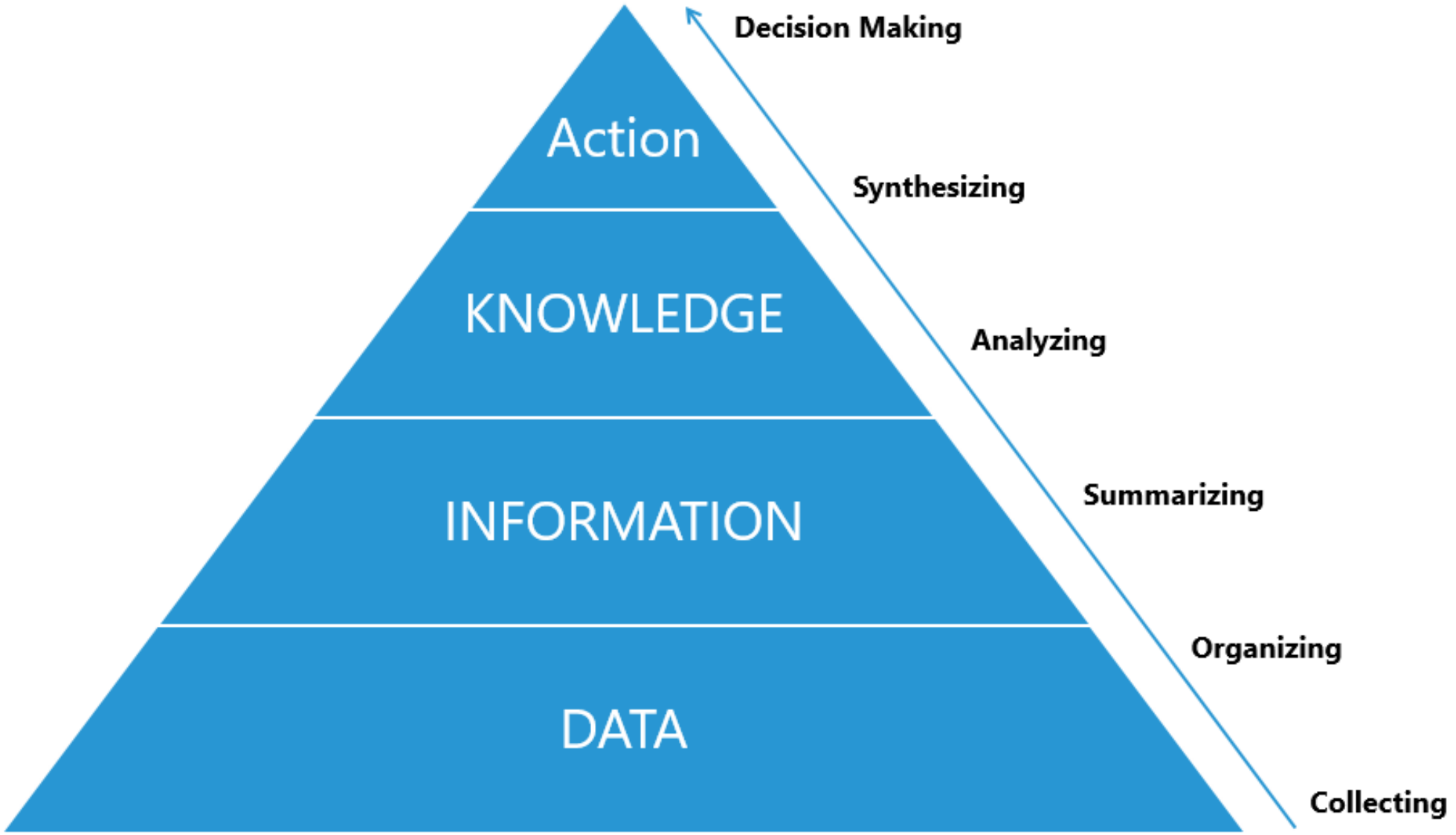
Results | Messages

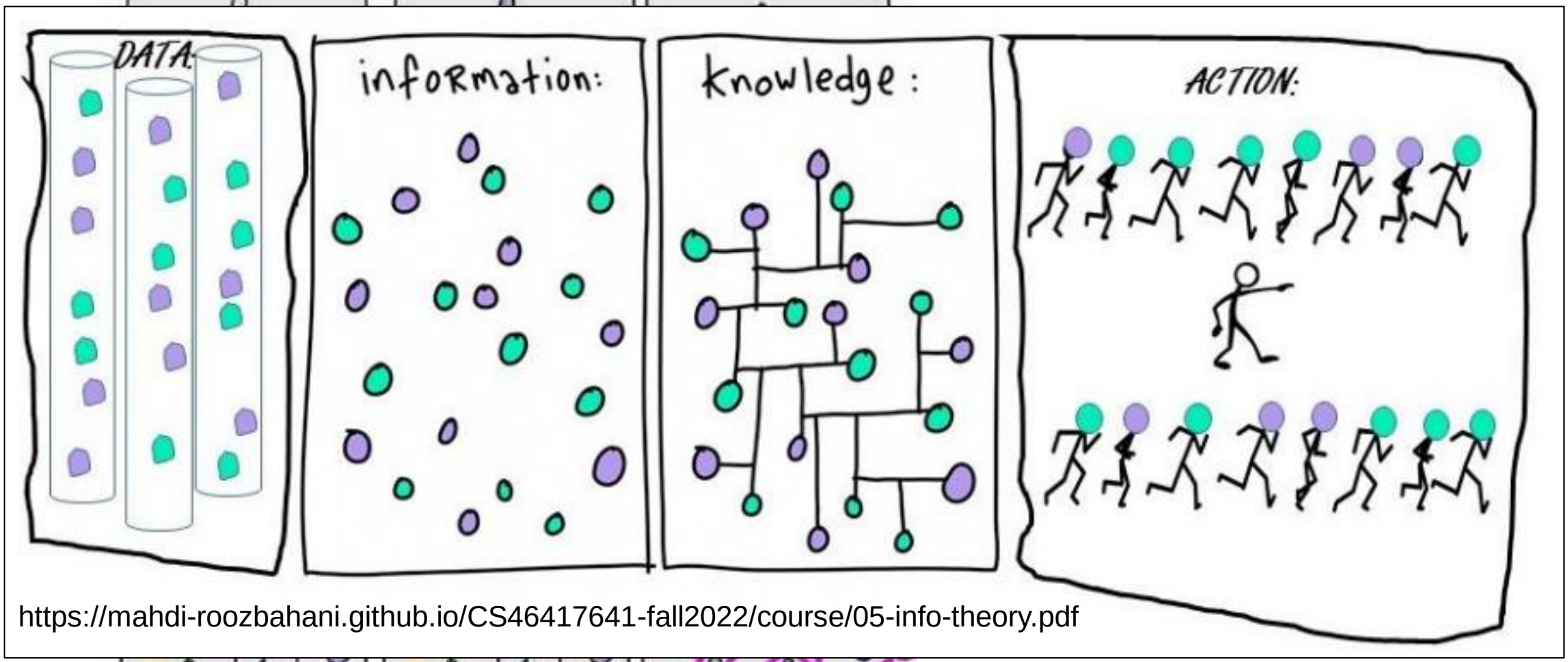
| | Education | Total IDs |
|---|---|---|
| 1 | Bachelors | 4 |
| 2 | Education | 2 |
| 3 | Graduate Degree | 2 |
| 4 | High School | 1 |
| 5 | Masters Degree | 2 |
| 6 | Partial College | 1 |
| 7 | Partial High School | 3 |

Query exec... | PRASAD (12.0 SP1) | PRASAD\sures



# The central dogma of informatics: DIKA





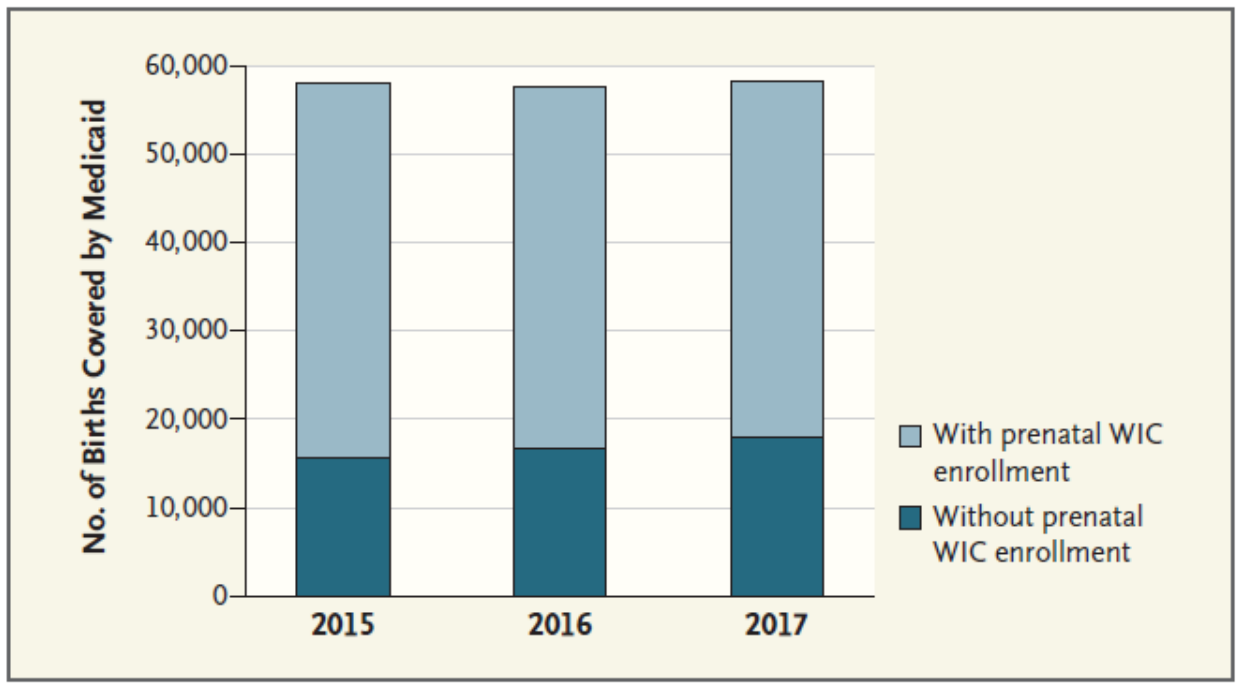

**Original source unknown**

The NEW ENGLAND JOURNAL of MEDICINE

Perspective FREE PREVIEW

# Focusing on Population Health at Scale — Joining Policy and Technology to Improve Health

Aaron McKethan, Ph.D., Seth A. Berkowitz, M.D., M.P.H., and Mandy Cohen, M.D., M.P.H.



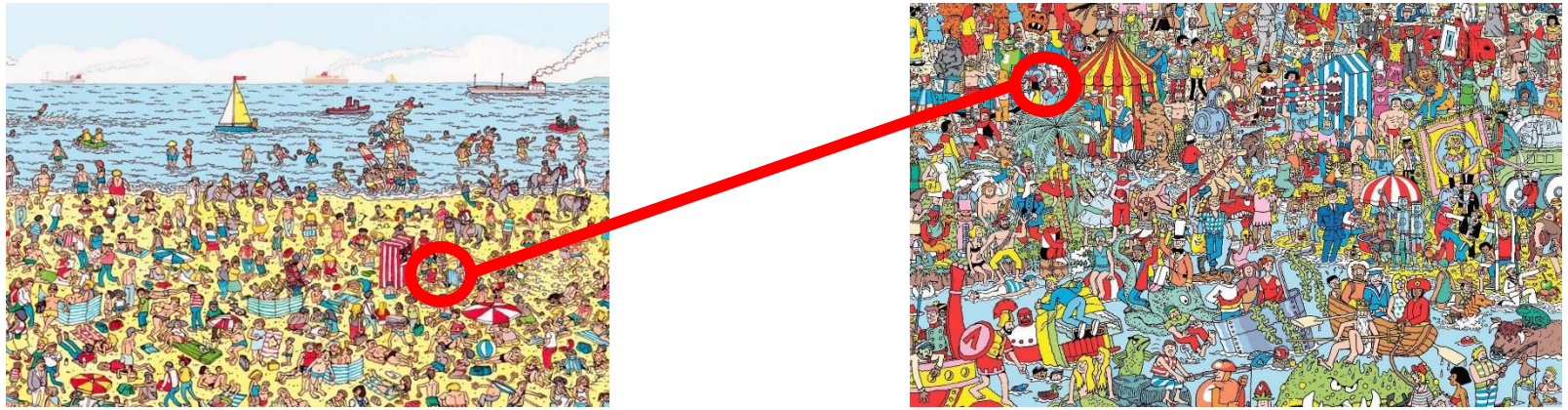

**Medicaid-Covered Births with and without Concurrent Prenatal WIC Enrollment, North Carolina.**

# In order to gain actionable knowledge, need 2 things:

1. **The ability to integrate data between divisional silos (“Data Integration”)**

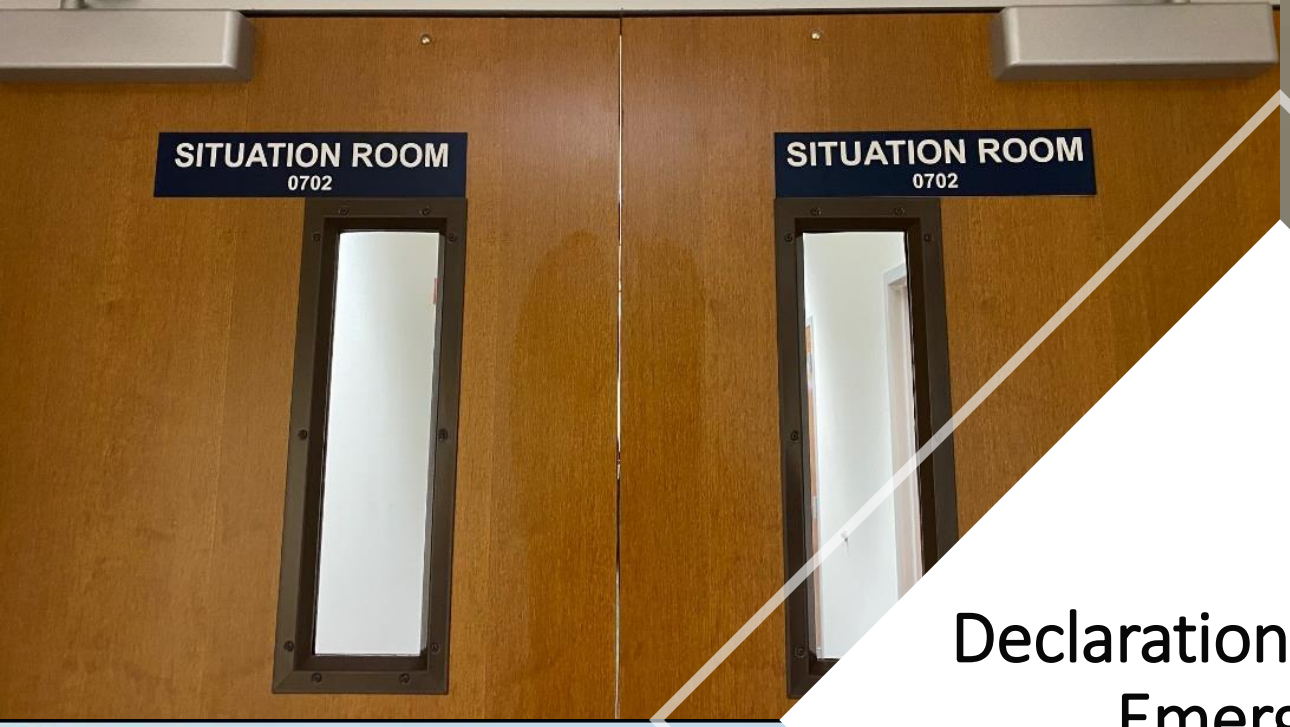
2. **The ability to identify and link the same individual from different datasets (“Entity Resolution”)**

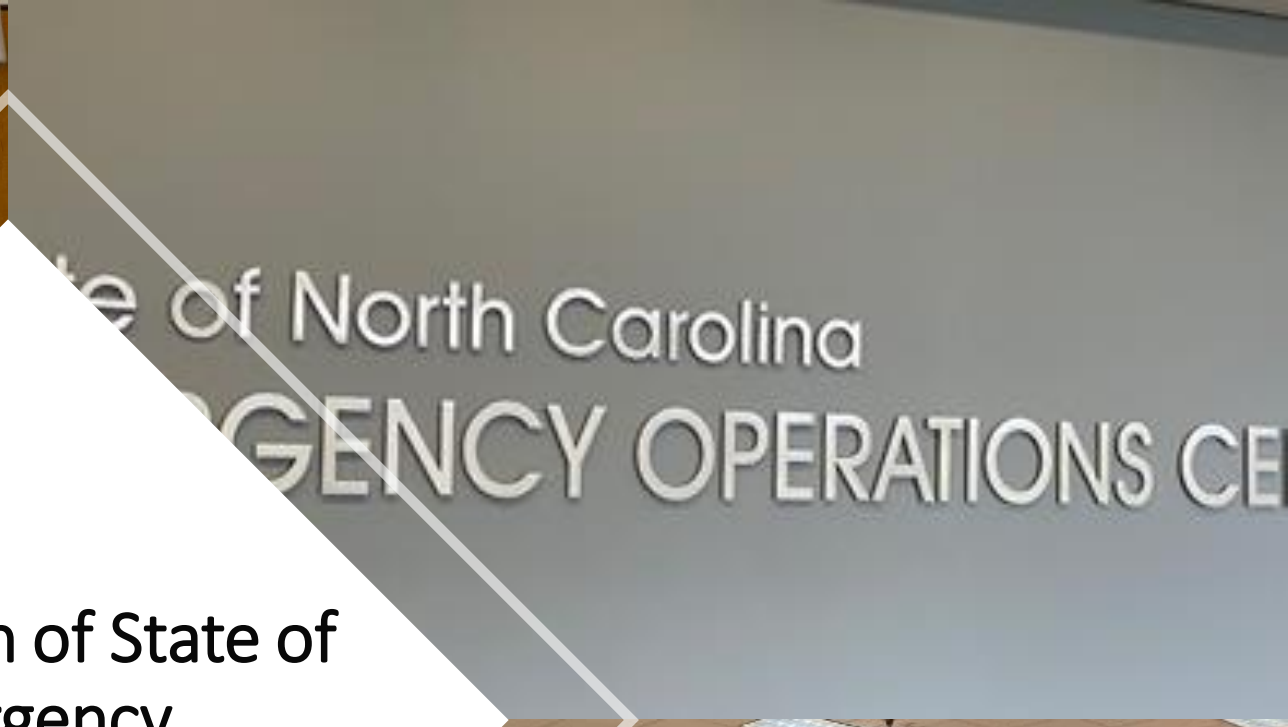
ID: 123456
Name: Waldo



ID: 123456
Name: Waldo

# We interrupt our regularly scheduled Data Office business to bring you…



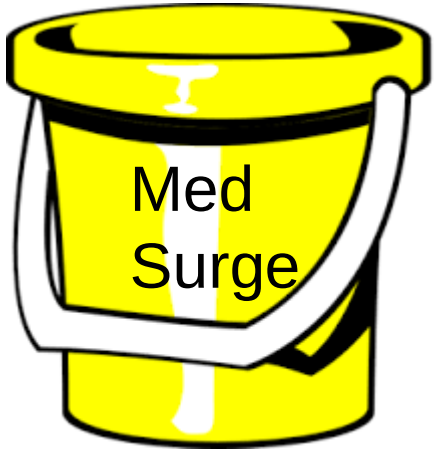

# Declaration of State of Emergency

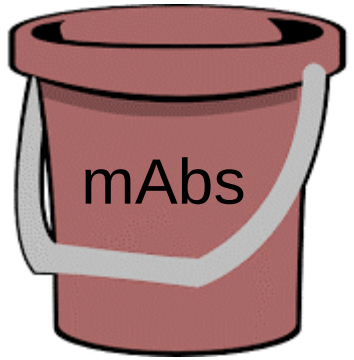
*March 10, 2020*

# Early Actionable Questions to Enable Data Driven Policy

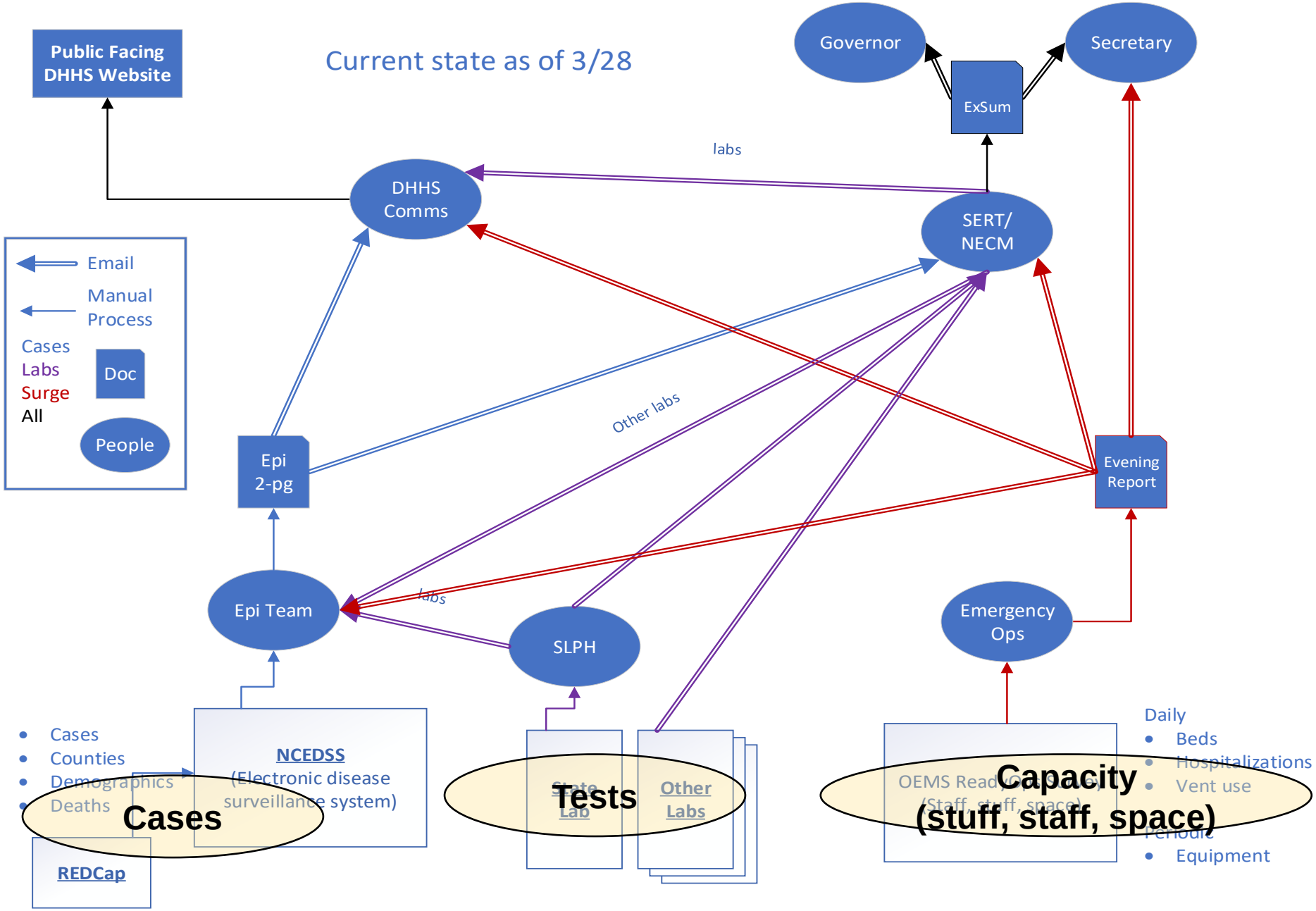
- How many cases will we see? When will our “epi curve” peak?
- How much (extra) PPE is needed, and where can we get it from?
- Will we run out of hospital beds? ICU beds? Ventilators?
- Should we shut down bars and restaurants? Schools? Businesses?
- How can we support families who are unable to work (either because workplace is shut down, or childcare is unavailable)?

# Buckets of NC DHHS COVID-19 Data

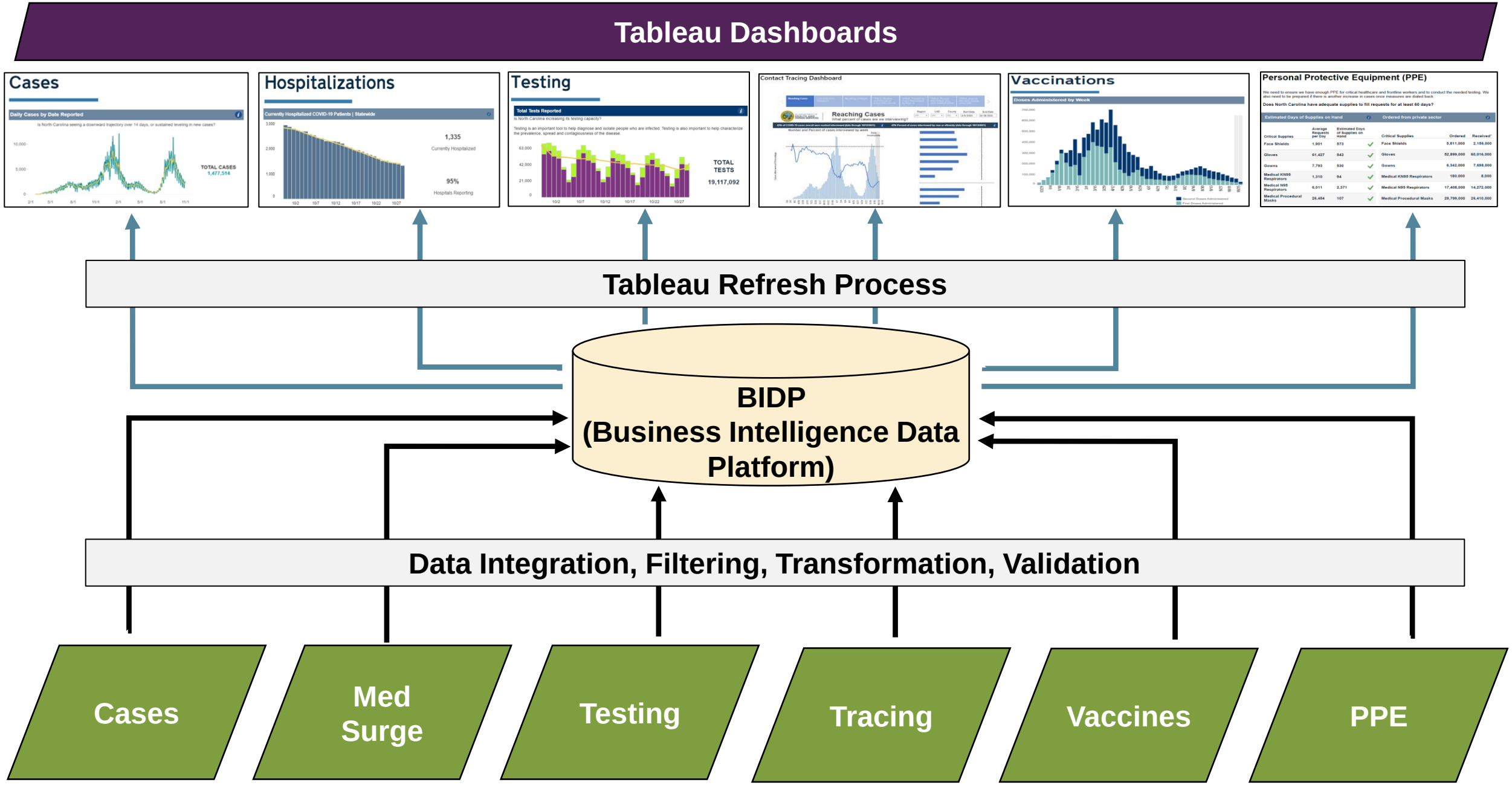


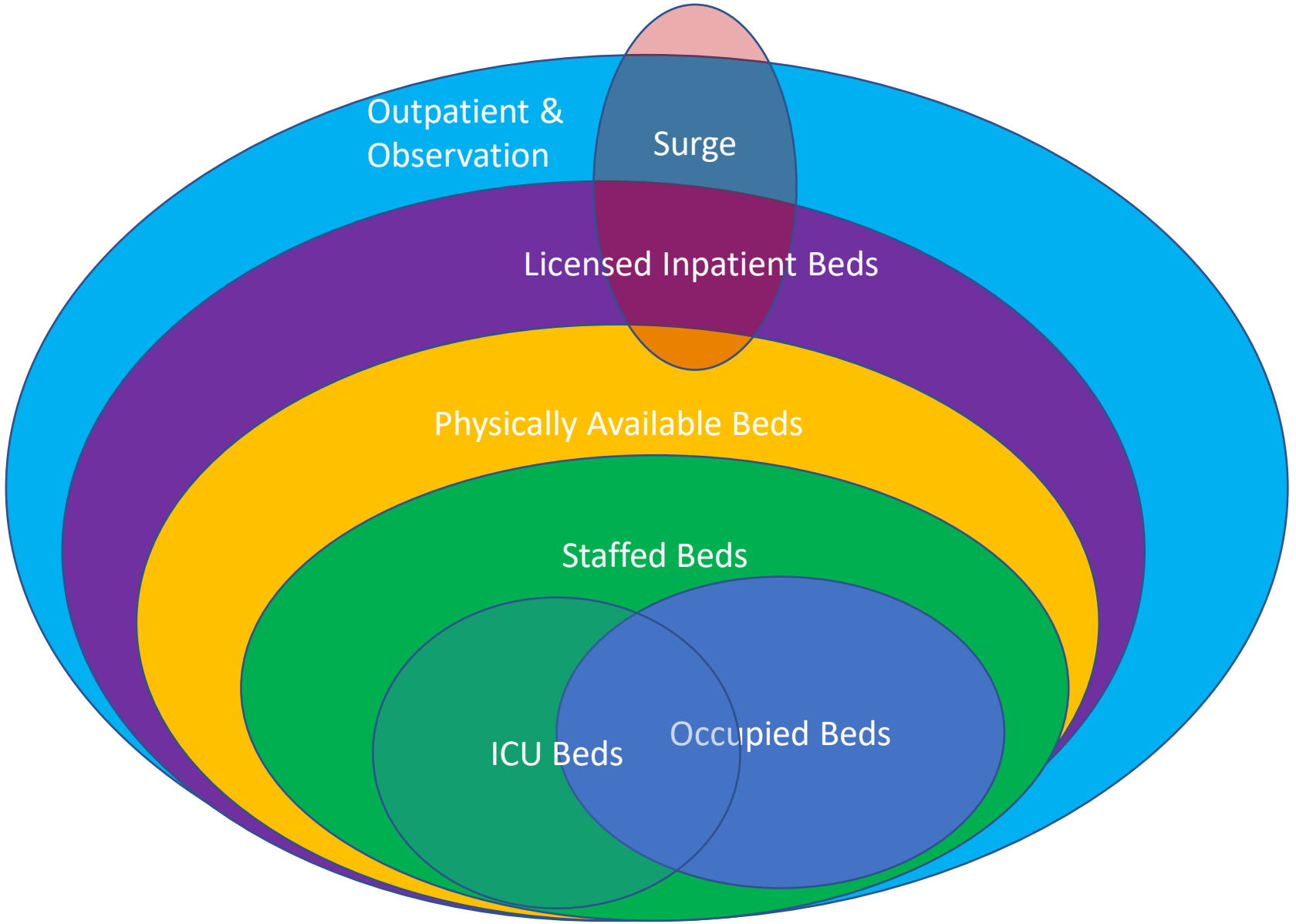


# Baseline Data Flow (as of March 2020)

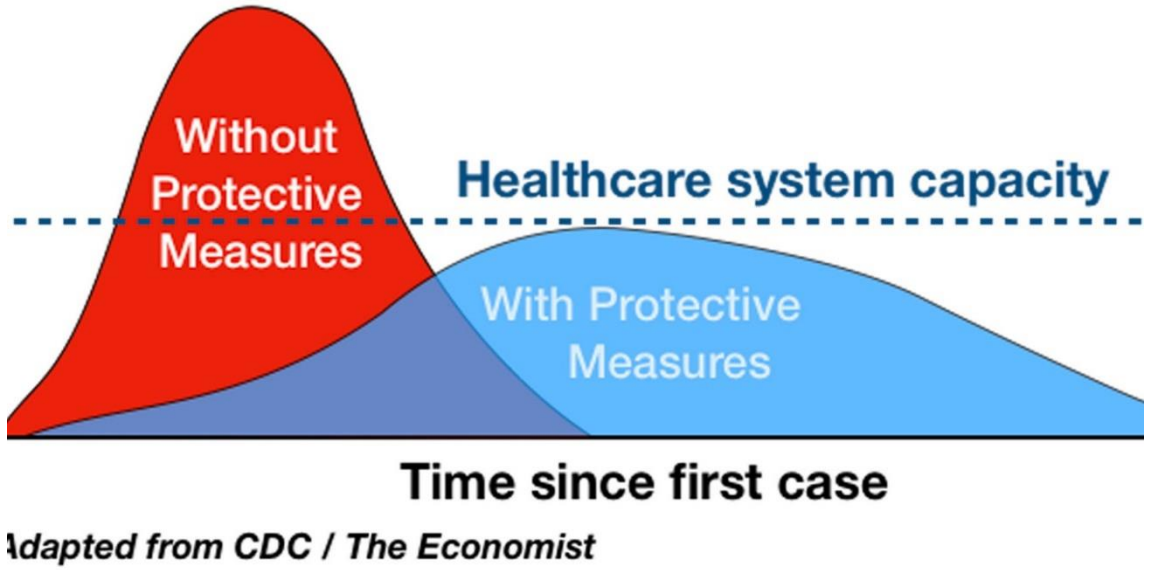


# Today's Data Flow

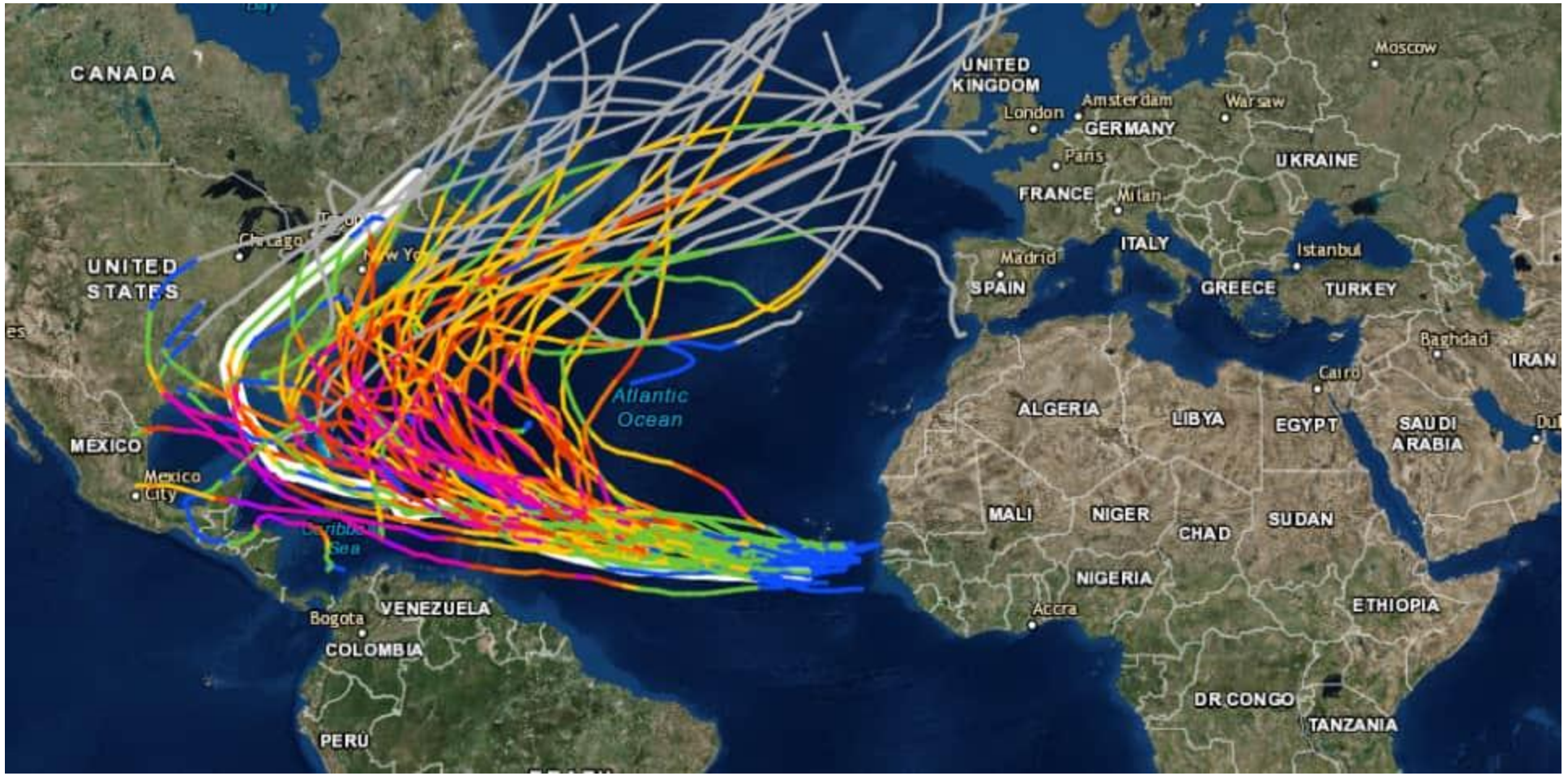


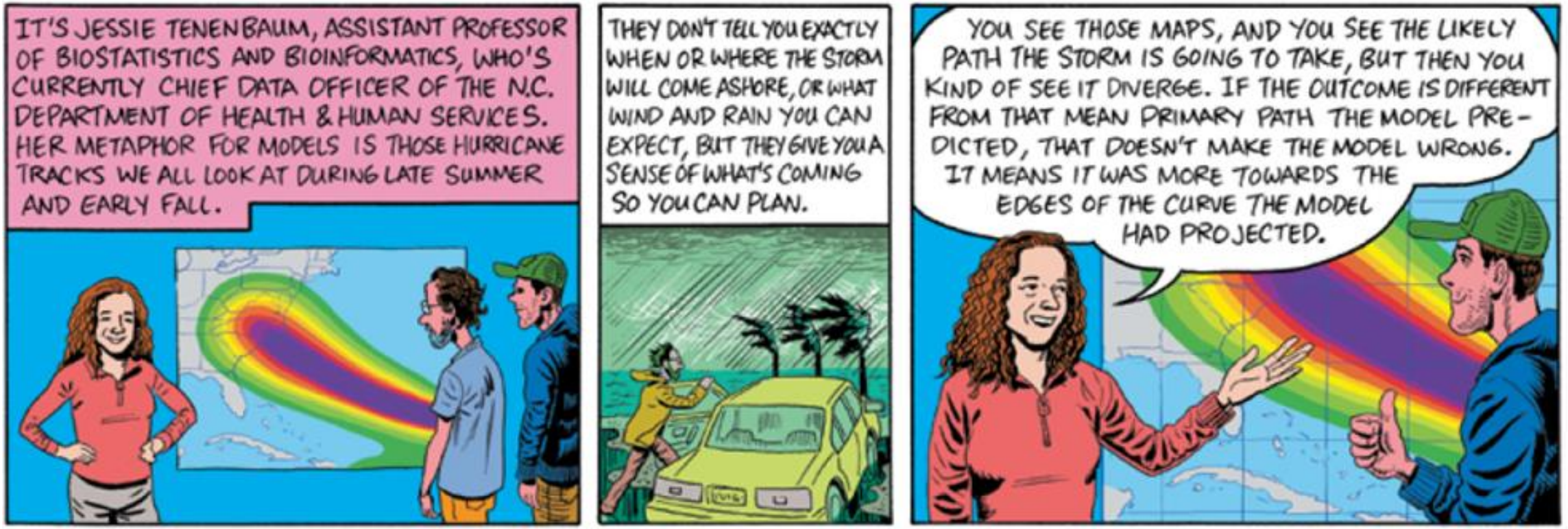

Outpatient &
Observation
Surge
Licensed Inpatient Beds
Physically Available Beds
Staffed Beds
ICU Beds
Occupied Beds

# Data Modeling: Team science at work




Courtesy of Aaron McKethan


CANADA
UNITED KINGDOM
London
Amsterdam
GERMANY
Warsaw
Moscow
Paris
UKRAINE
FRANCE
Milan
ITALY
Madrid
SPAIN
Istanbul
GREECE
TURKEY
UNITED STATES
Chicago
New York
Atlantic Ocean
Baghdad
IRAN
Cairo
ALGERIA
LIBYA
EGYPT
SAUDI ARABIA
MEXICO
Mexico City
MALI
NIGER
CHAD
SUDAN
NIGERIA
Accra
ETHIOPIA
Caribbean Sea
VENEZUELA
Bogota
COLOMBIA
DR CONGO
TANZANIA
PERU




Duke Magazine, Summer 2020

Story by Scott Huler, Illustrations by James Yamasaki

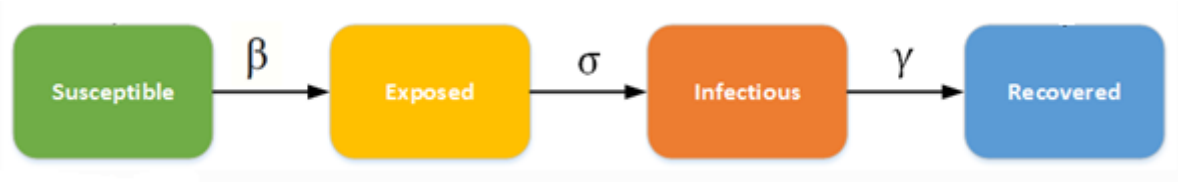
# Geospatially Explicit Agent Based Model (ABM) Framework

- ABM includes ~10.4 million agents
  - Based on the NC synthetic population: Dataset of anonymous synthetic persons and households, located geographically
  - Each agent represents a NC resident and has attributes (e.g., age)
- Each agent's physical location, health status, and life status are simulated daily

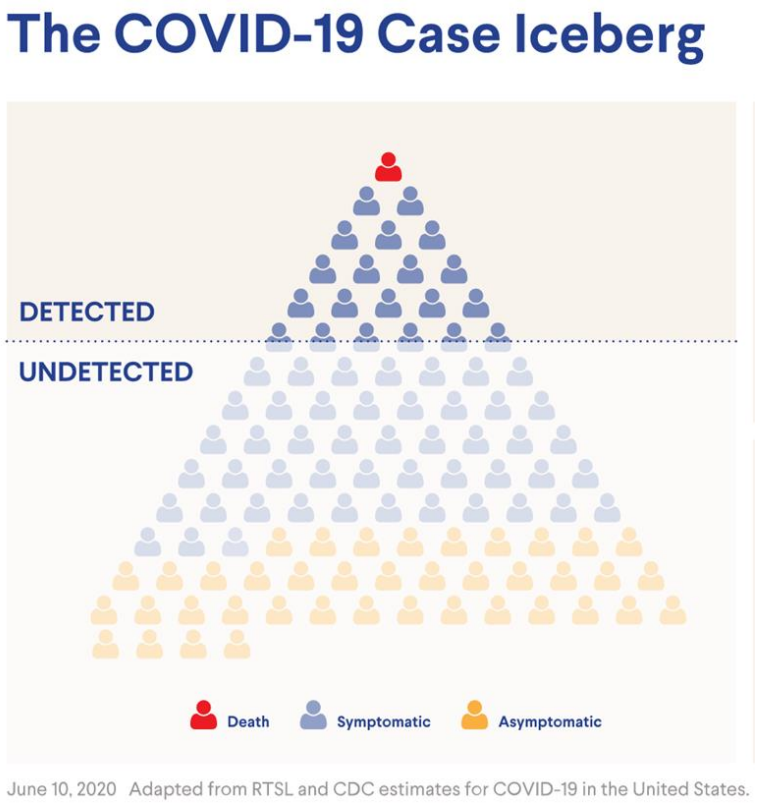
Courtesy of Dr. Sarah Rhea, NC State University & RTI International

# NC County Level SEIR* models

- Predict new SARS-CoV-2 infections for the next 30 days
- Leverage NC county-level reported case data while accounting for unreported infections




*Susceptible-Exposed-Infectious-Recovered Model Diagram
https://institutefordiseasemodeling.github.io/Documentation/malaria/model-seir.html




Source: Prevent Epidemics
https://preventepidemics.org/covid19/science/insights/covid-19-case-iceberg/

Courtesy of Dr. Sarah Rhea, NC State University & RTI International

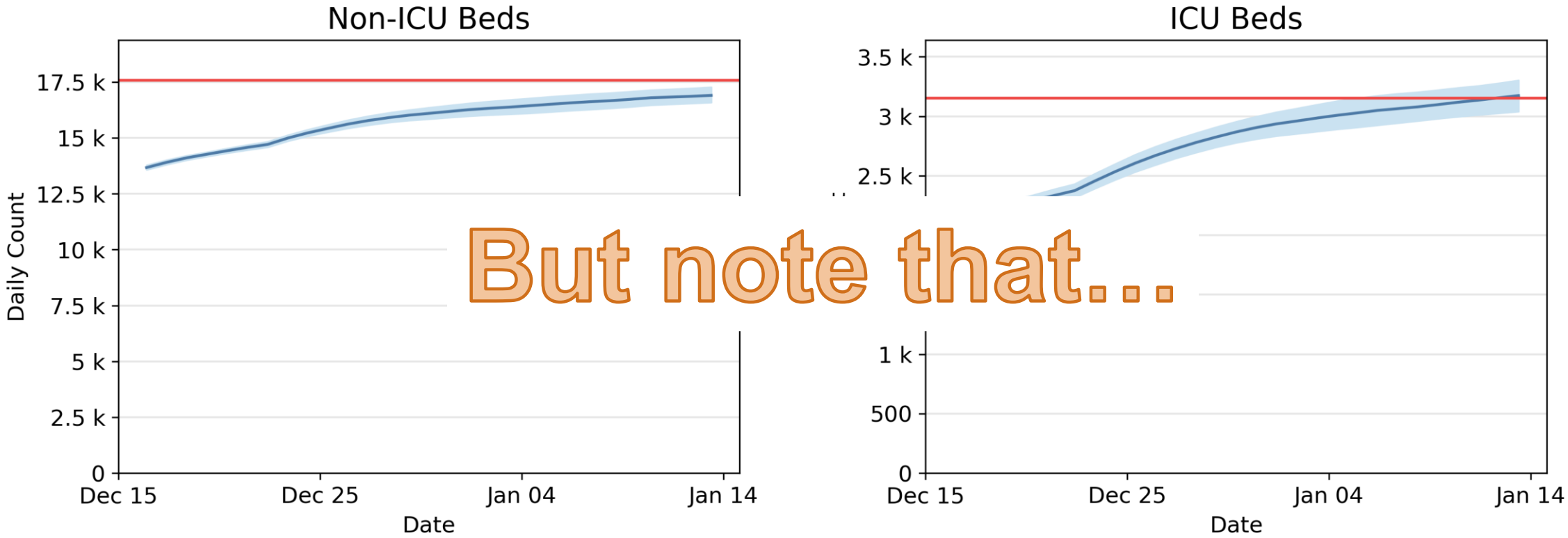
# Model-Projected Patient Demand for Non-ICU & ICU Beds by Day - STATEWIDE

**Statewide**

Model-projected mean number of patients needing a hospital bed on a given day* statewide and by North Carolina flu monitoring region. Modeled scenarios consider an effective reproductive number ($R_e$) range** and a reported cases multiplier to account for unreported/undocumented infections at model initialization and assume 25% of hospitalized patients requiring ICU-level care. Regional demand values are based on the modeled hospitals in each region. See model documentation for additional detail and references.

**$R_e$ = 1.2 - 1.4**

**Updated Dec 17, 2020**




**But note that…**

Light blue area: 25th and 75th percentiles of model output

Red line: Number of non-ICU operational beds or licensed ICU beds, as indicated (per NC DHHS, July 2020); scaled to account for rounding in forecasts

Parameter values and assumptions are subject to change as knowledge of COVID-19 evolves

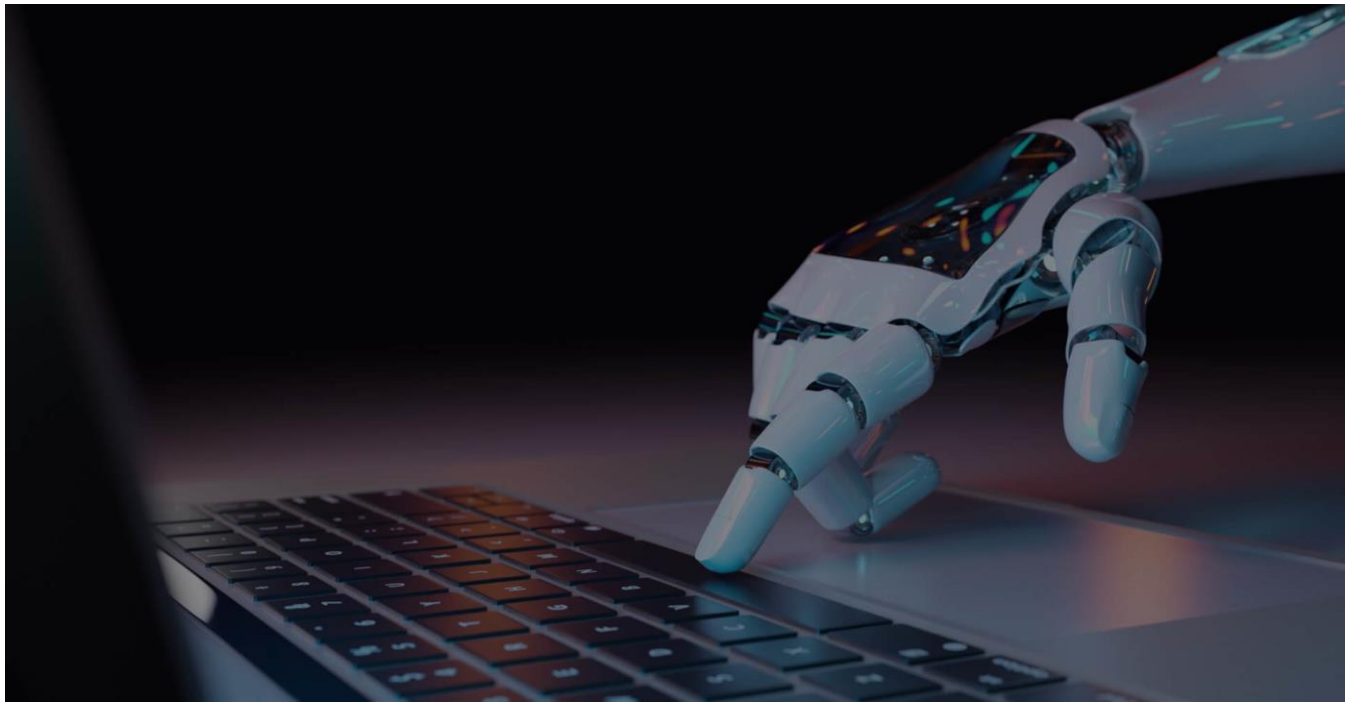
*Based on 100 model runs and using a sample of 2.5 million agents for each run, with results scaled to the 10.4 million agents representing the North Carolina population

**$R_e$ ranges are sampled from a uniform distribution

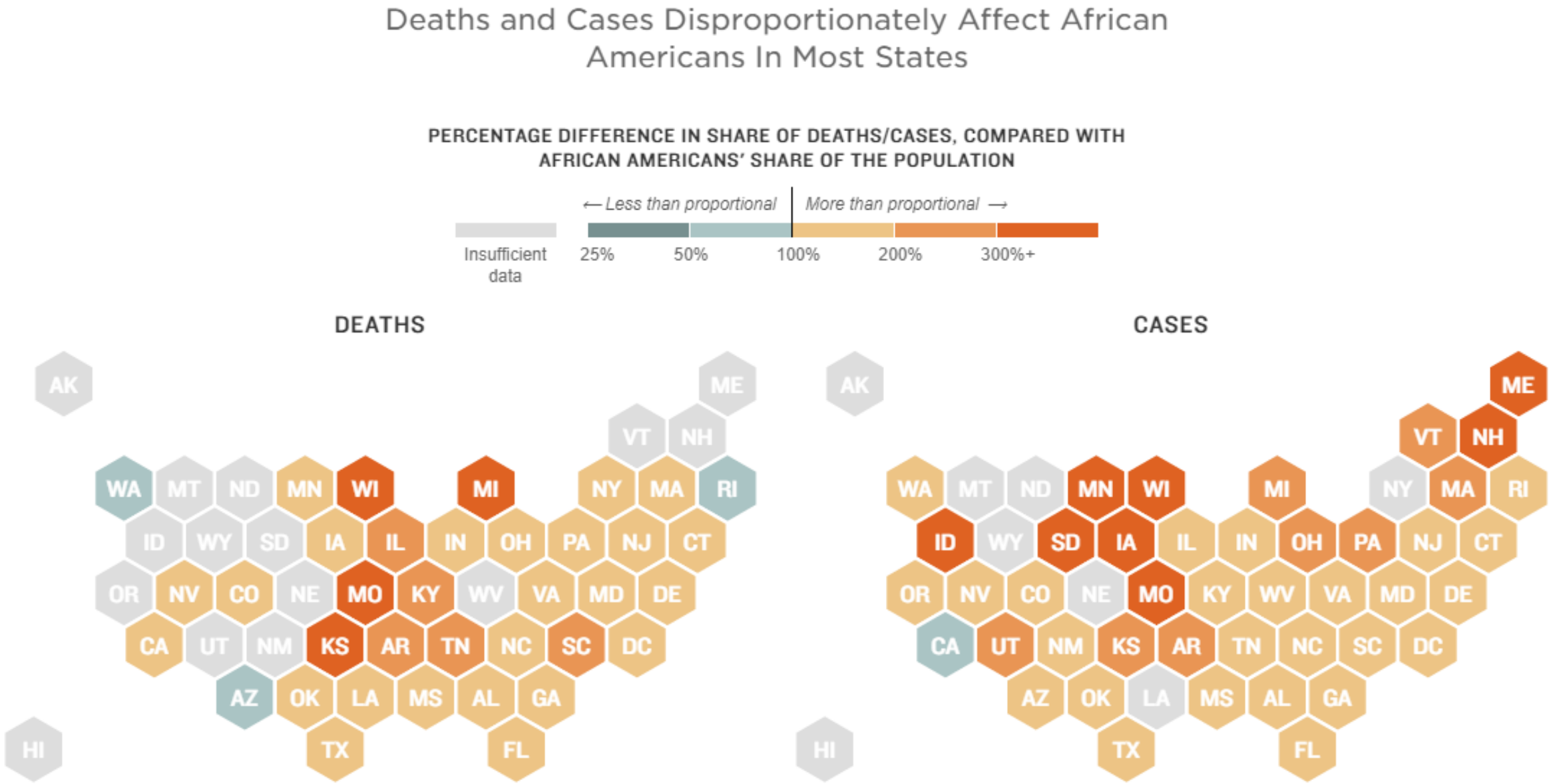
Dr. Sarah Rhea,
NC State University & RTI International

# Notable Exception where AI has proven useful in Public Health

- RPA: Robotic Process Automation
- Software to perform mindless human tasks, often to make up for poorly designed and/or non-interoperable systems



# COVID disproportionately impacted historically marginalized populations



# You can’t improve what you can’t measure

- What is the distribution of race/ethnicity among
  - COVID-19 cases?
  - COVID-19 deaths?
  - People tested for COVID-19?
  - People vaccinated against COVID-19?

Bloomberg

Prognosis

**One U.S. State's Laser Focus on Data Helps Shrink Racial Vaccine Gap**

North Carolina has racial data for 99.6% of people who got a shot in the state

By Angelica LaVito
March 6, 2021, 8:30 AM EST



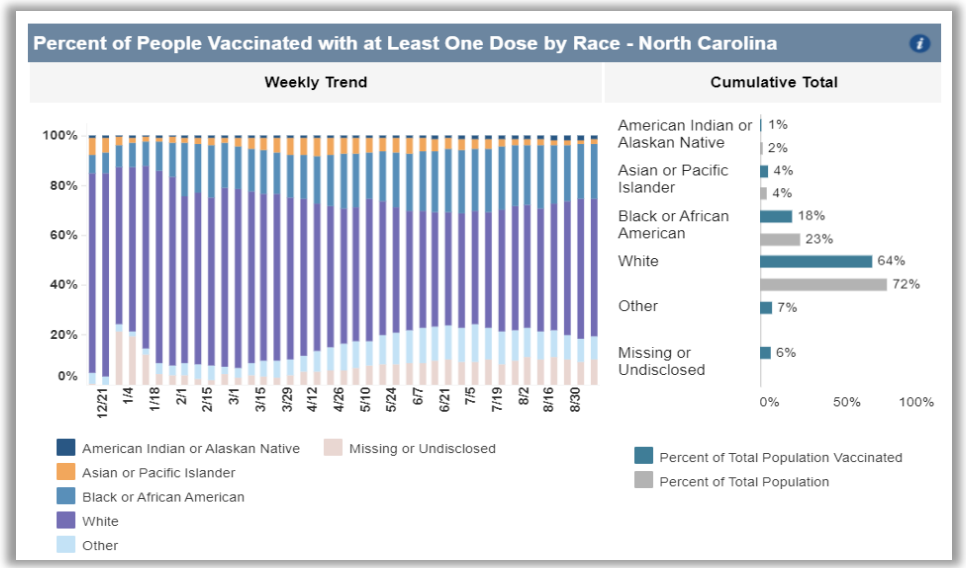
A health-care worker puts a bandage on man after he received the Moderna Covid-19 vaccine at the Durham County Department of Public Health in Durham, N.C., on Jan. 8. *Photographer: Julia Wall/The News & Observer via AP*

LIVE ON BLOOMBERG
Watch Live TV >
Listen to Live Radio >

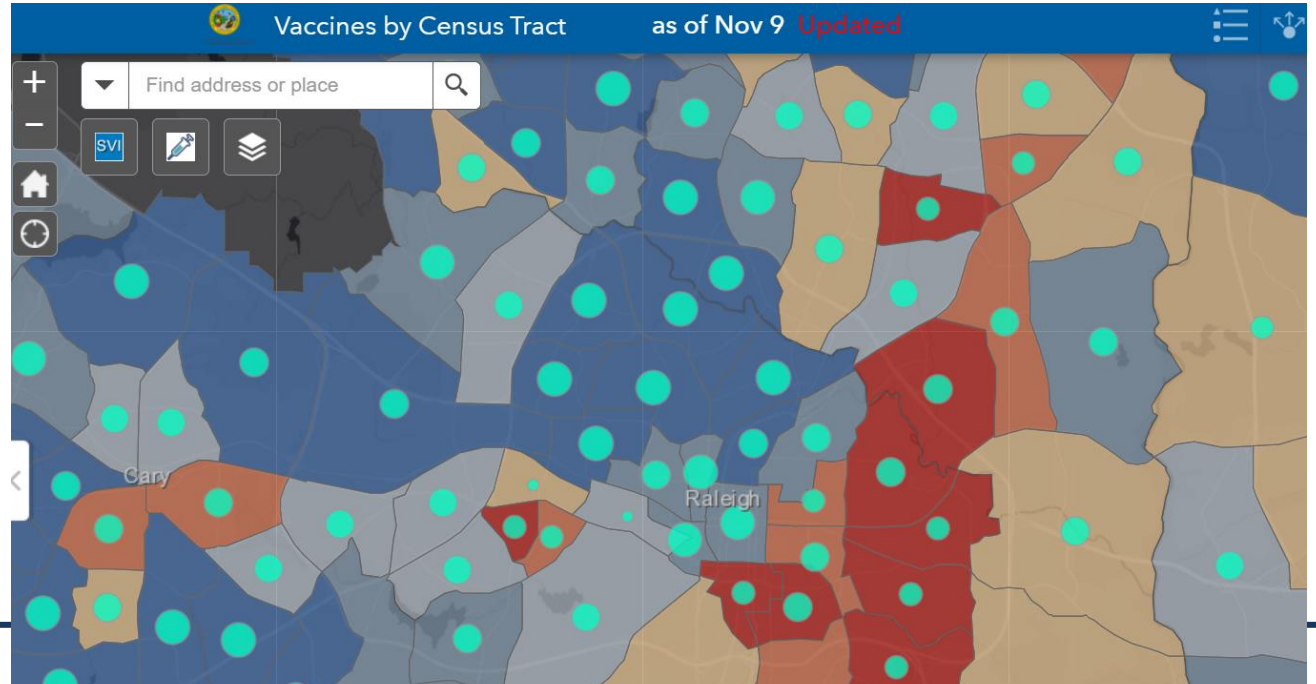
IG
Open an account and start trading in minutes
Trade now
Forex trading involves risk. Losses can exceed deposits.

# NC Leads the Country in Collecting Equity Data

- Data tools built to infuse equity data insights across entire vaccine operation

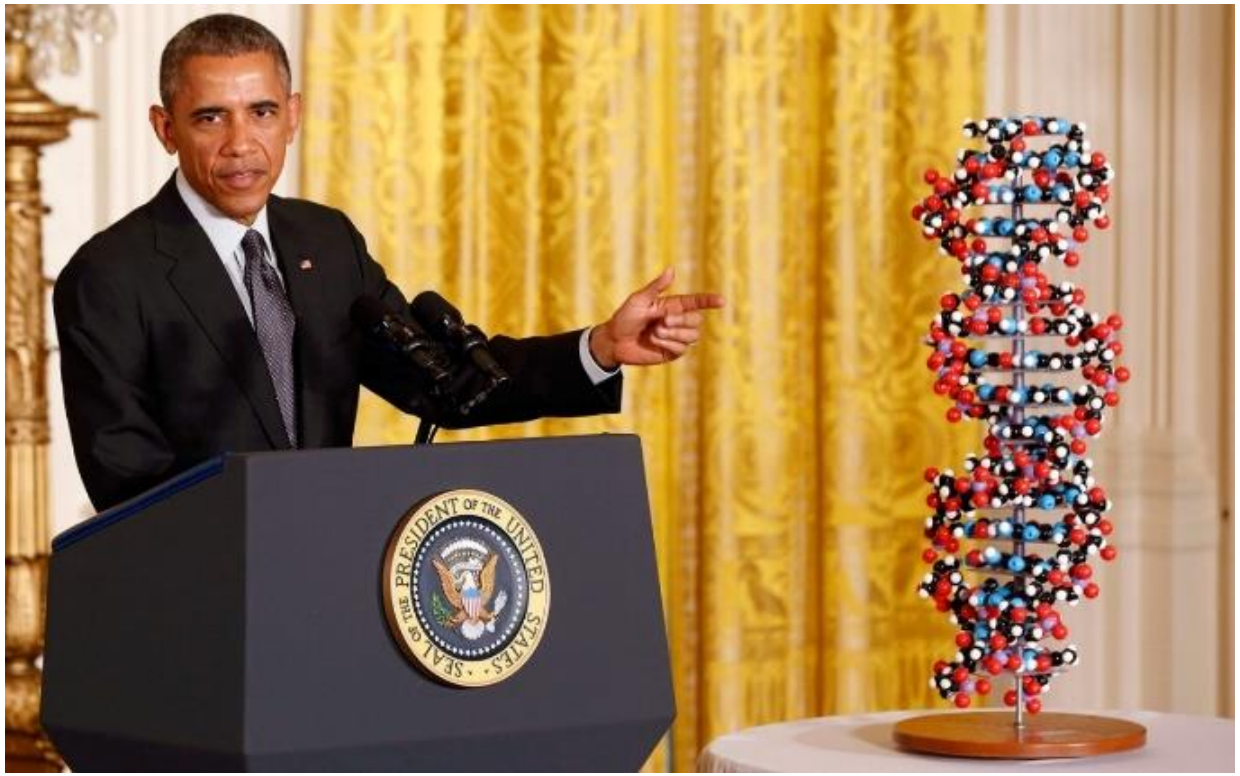



https://covid19.ncdhhs.gov/dashboard

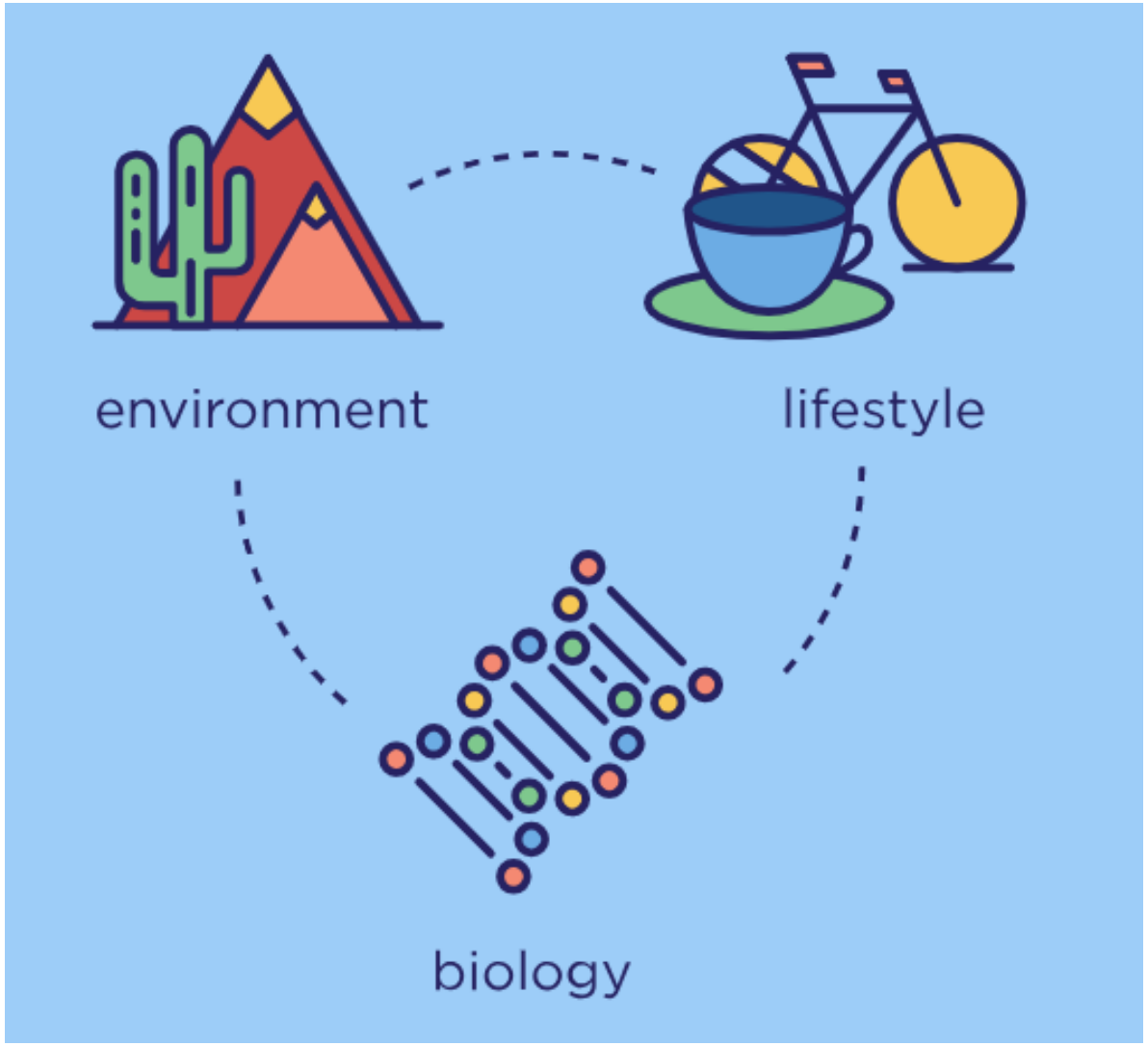


# We now return to our regularly scheduled Data Office business…

# Remember this pic?



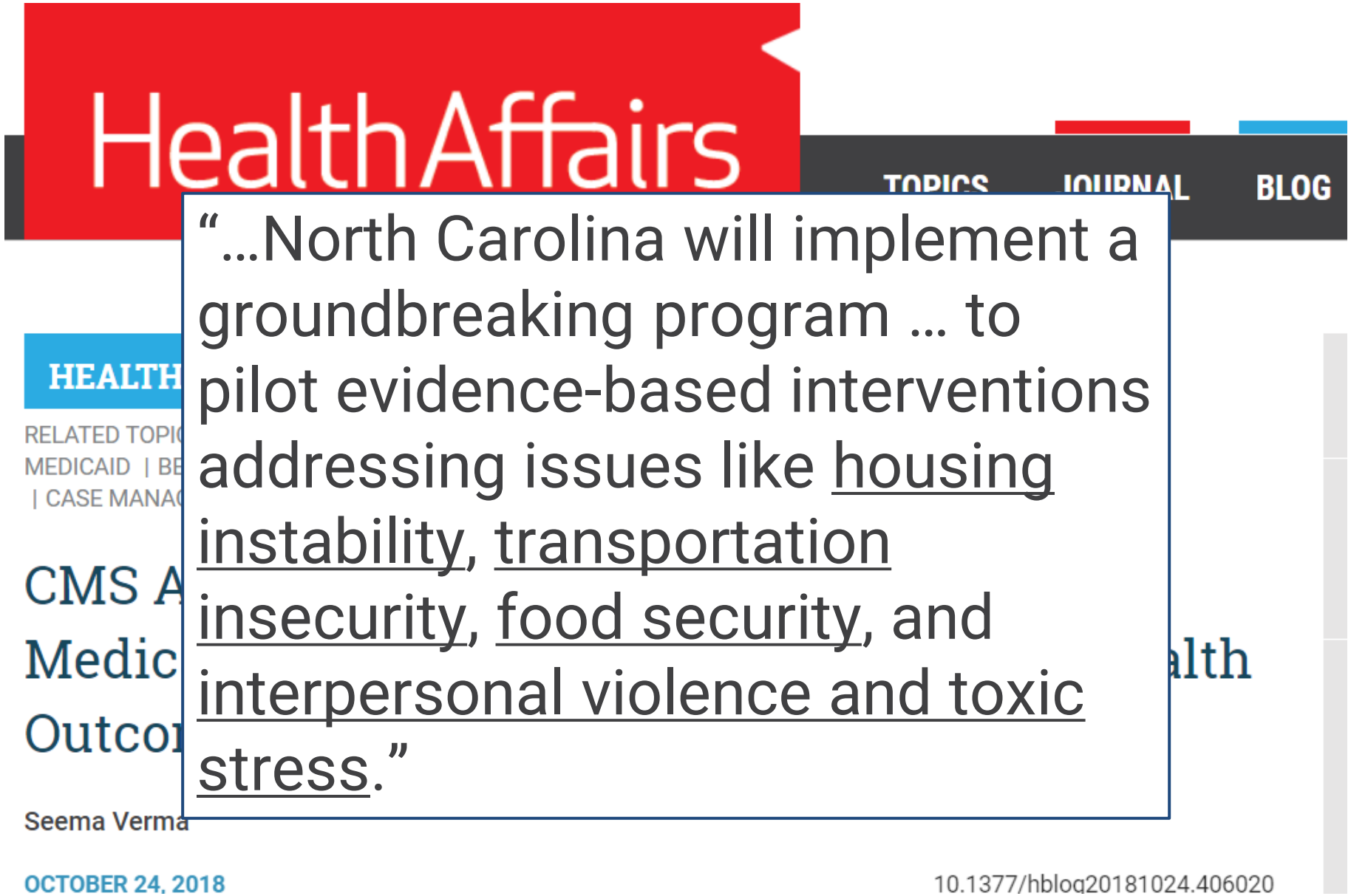
All of Us
RESEARCH PROGRAM




https://allofus.nih.gov/

# Beyond DNA

https://www.publichealthnotes.com/social-determinants-health-sdh/


HealthAffairs
TOPICS
JOURNAL
BLOG
HEALTH
RELATED TOPI
MEDICAID | BE
| CASE MANA
CMS A
Medic
Outcor
alth
Seema Verma
OCTOBER 24, 2018
10.1377/hblog20181024.406020
"…North Carolina will implement a groundbreaking program … to pilot evidence-based interventions addressing issues like housing instability, transportation insecurity, food security, and interpersonal violence and toxic stress."

# Homelessness is a medical condition

- **Hospital pays $1,000 per individual per month for housing.**
- **Significant drop in ER use among individuals given housing options.**

09:35

## To Reduce Chronic Homelessness, A Chicago Hospital Is Treating It As A Medical Condition

February 21, 2019 By Jeremy Hobson

HOUSTON CHRONICLE

# Texas lawmakers look to Uber, Lyft to transport Medicaid patients

By Allie Morris, Austin Bureau | March 12, 2019 | Updated: March 12, 2019 3:17 p.m.

# Screening Questions

**Food**

1. Within the past 12 months, did you worry that your food would run out before you got money to buy more? (Y/N)
2. Within the past 12 months, did the food you bought just not last and you didn't have money to get more? (Y/N)

**Housing/Utilities**

3. Within the past 12 months, have you ever stayed: outside, in a car, in a tent, in an overnight shelter, or temporarily in someone else's home (i.e. couch-surfing)? (Y/N)
4. Are you worried about losing your housing? (Y/N)
5. Within the past 12 months, have you been unable to get utilities (heat, electricity) when it was really needed? (Y/N)

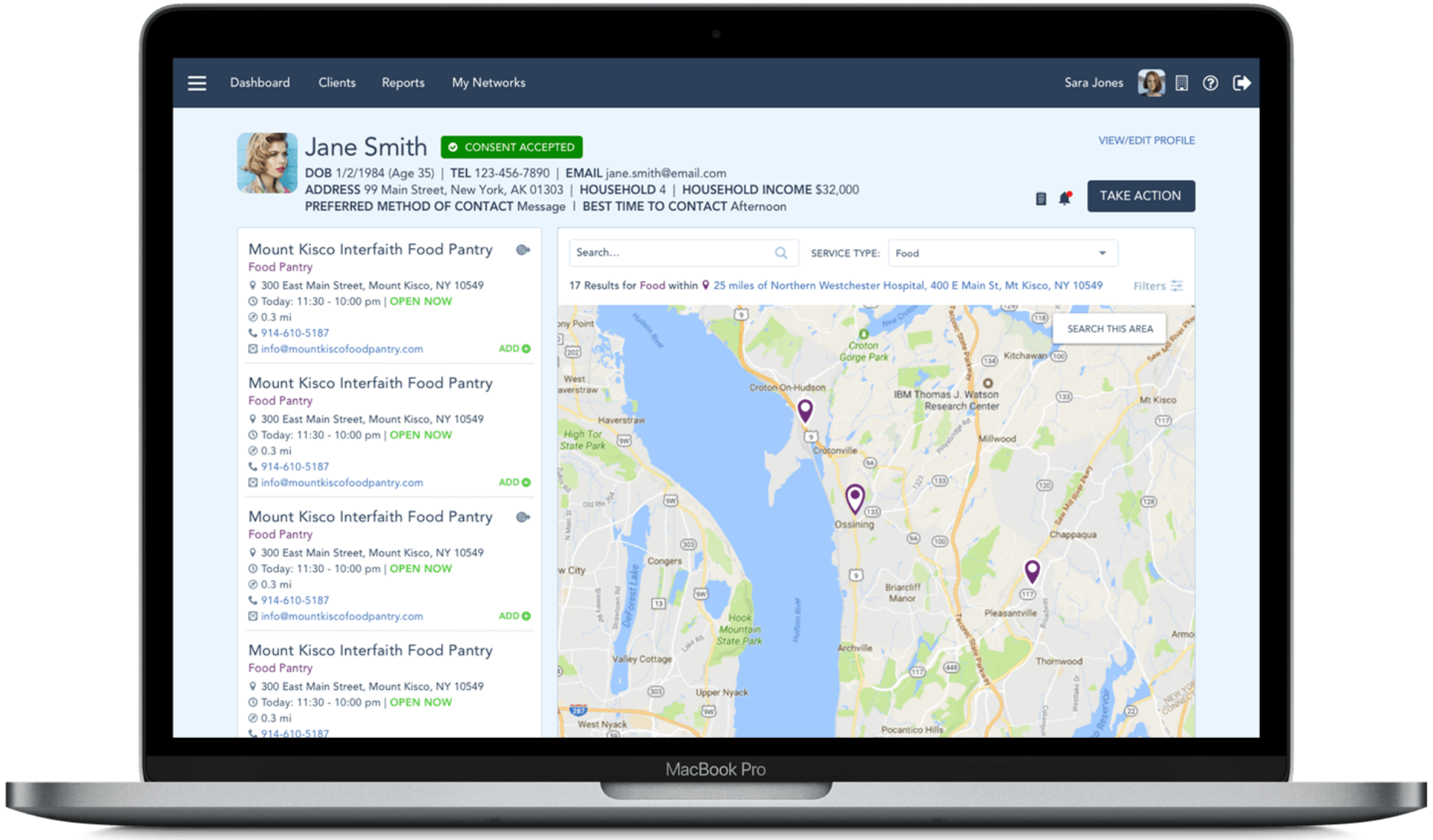
**Transportation**

6. Within the past 12 months, has a lack of transportation kept you from medical appointments or from doing things needed for daily living? (Y/N)

**Interpersonal Safety**

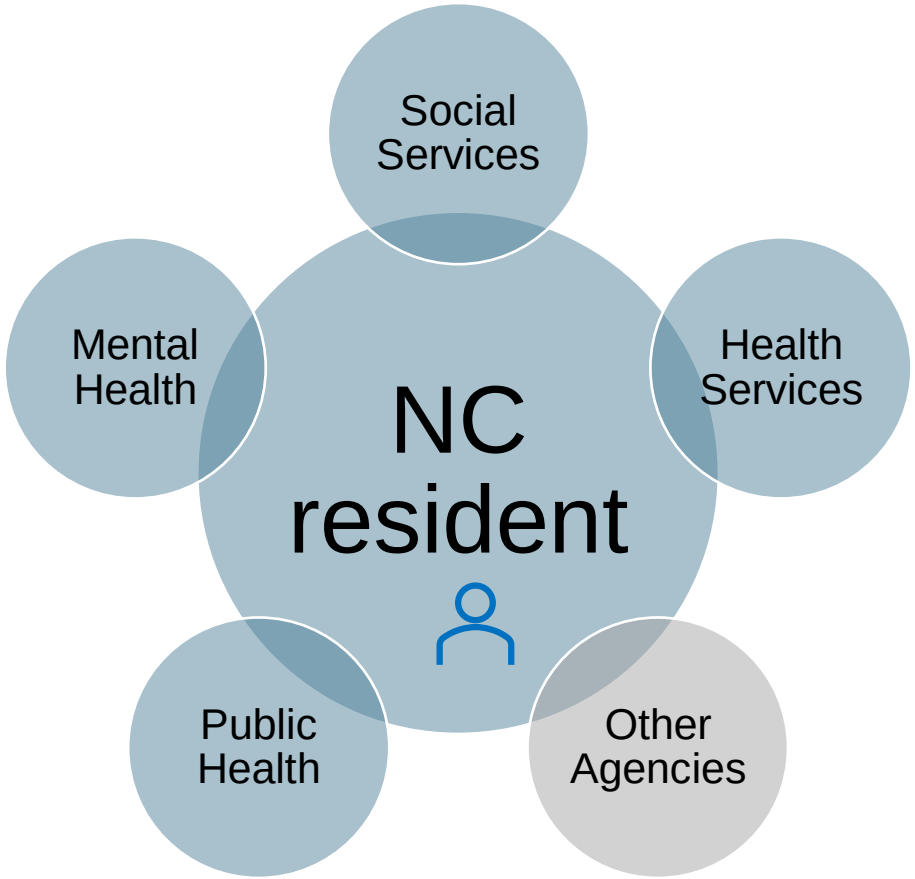
7. Do you feel physically and emotionally unsafe where you currently live? (Y/N)
8. Within the past 12 months, have you been hit, slapped, kicked or otherwise physically hurt by anyone? (Y/N)
9. Within the past 12 months, have you been humiliated or emotionally abused by anyone? (Y/N)

# NCCARE360 Closed Loop Referral Platform



# Whole Person Health

- Goal: link data to facilitate a “Whole Person Health” view of the people we serve.
  - Real-time individual level
  - Aggregate analysis to inform policy
- Requires the ability to integrate data across divisional silos
- Which requires ability to link records between systems



# Motivating Questions (Examples)

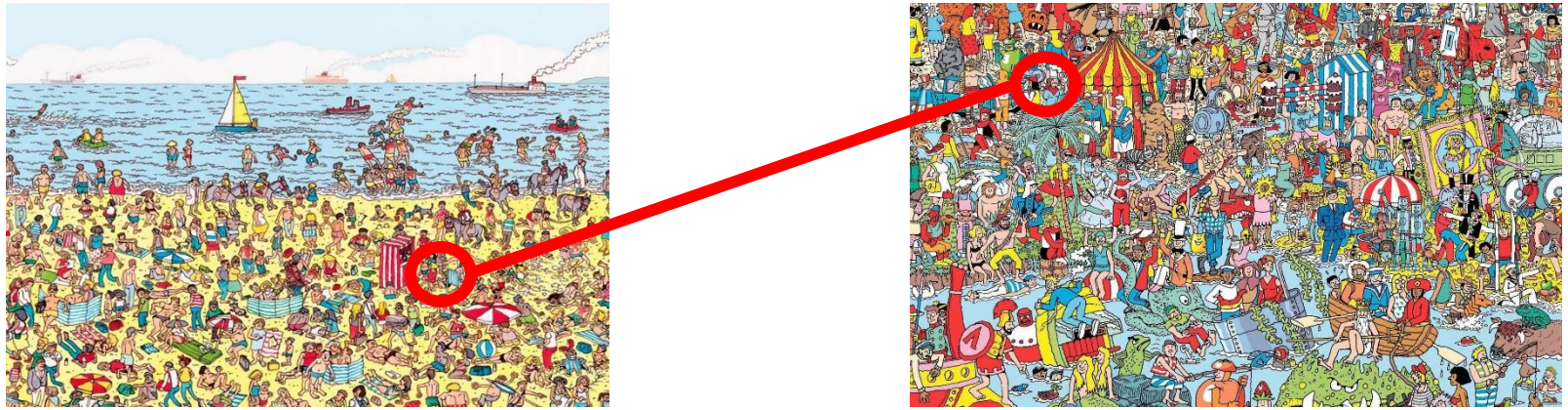
- What % of Medicaid beneficiaries have been vaccinated?
- Which children in foster care have prescriptions for >4 psychotropic medications?
- Who receives regular food assistance referrals and therefore may benefit from SNAP but is not enrolled?
- What % of people experiencing homelessness have been vaccinated?
- What is the relationship between early grade outcomes (e.g., third grade reading) and different early childhood conditions (e.g., early learning, health, housing, child welfare)?

# In order to achieve Whole Person Health, need 2 things:

1. **The ability to integrate data between divisional silos (“Data Integration”)**

2. **The ability to identify and link the same individual from different datasets (“Entity Resolution”)**

ID: 123456
Name: Waldo

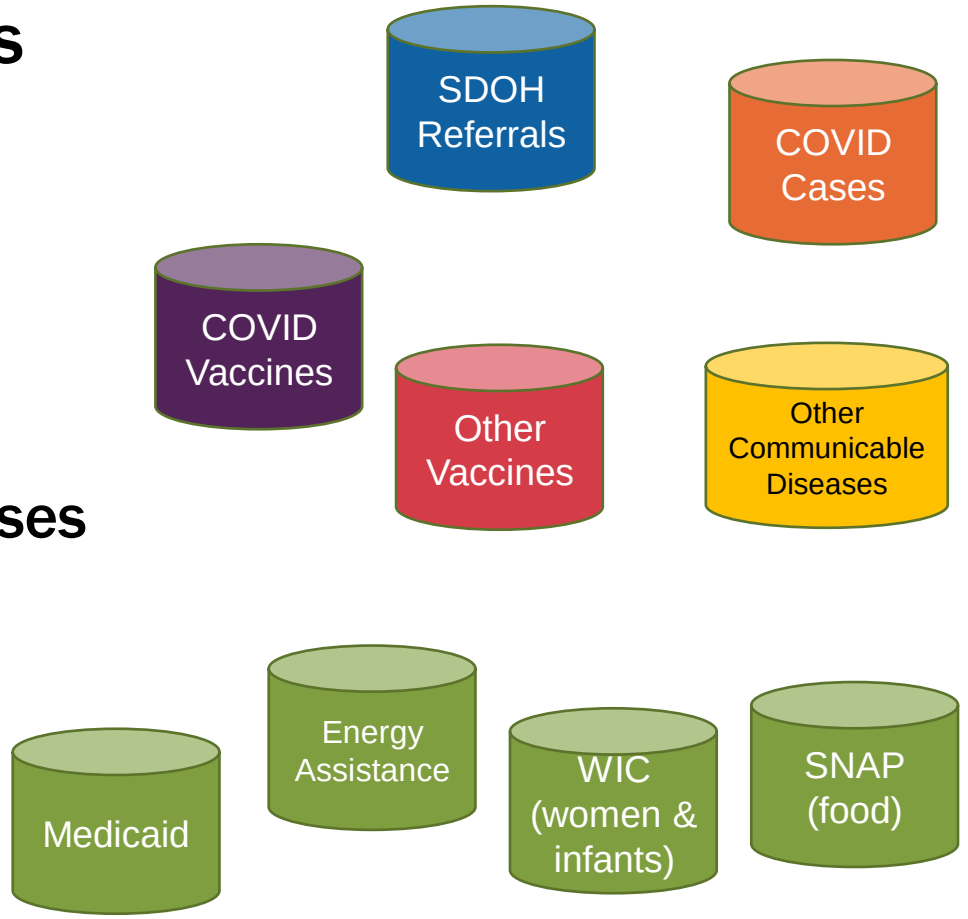


ID: 123456
Name: Waldo

# Answering those questions is currently difficult at best

- Data live in silos that do not share IDs
- Probabilistic match is possible, but
  - Labor-intensive
  - Prone to error

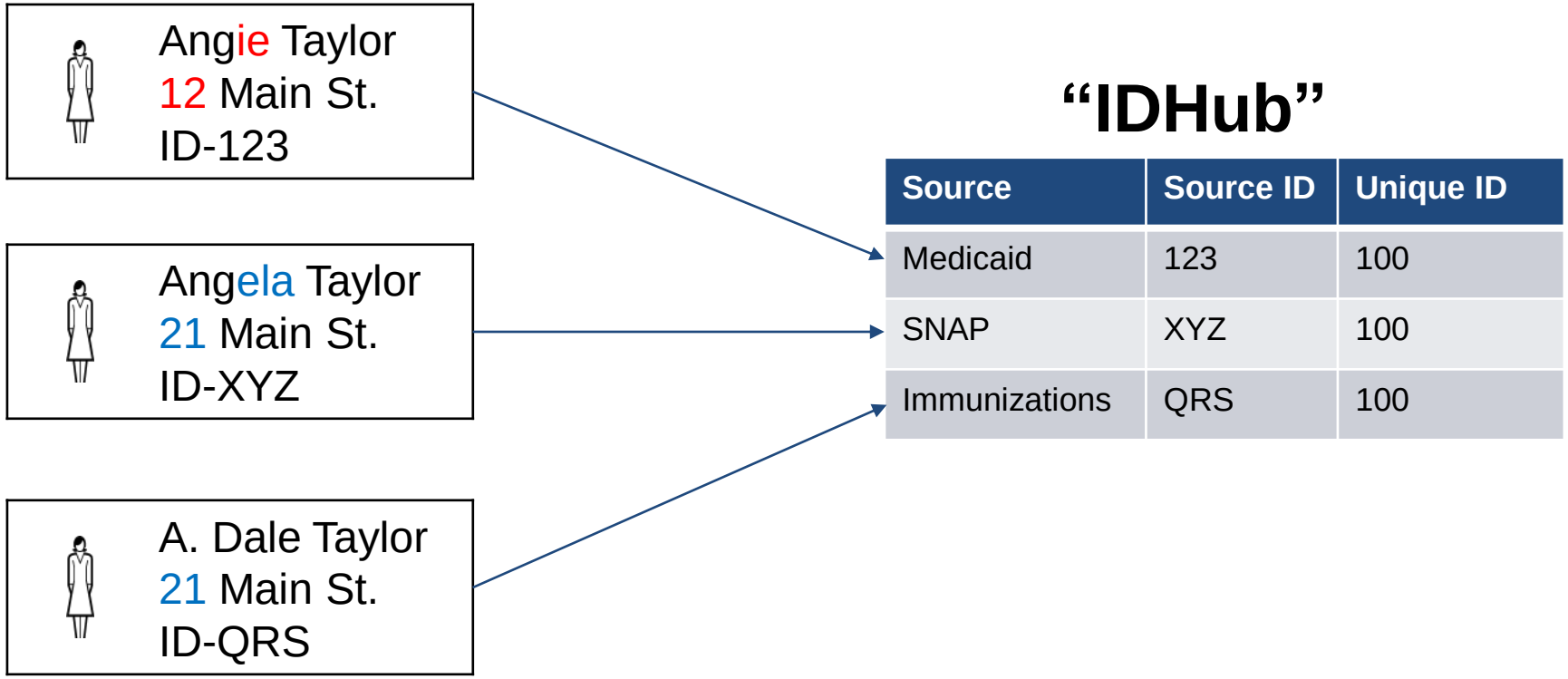
  - Example: initial approach for post-vax cases



# Options for “solving” entity resolution

1. Each system uses its own ID, probabilistic “fuzzy match” between systems as needed
2. One universal ID, e.g. state-wide or National Health Identifier
3. Somewhere in between
   a. Multiple “standard” identifiers
   b. Map them to each other!
   c. Refer to that mapping for efficient data integration

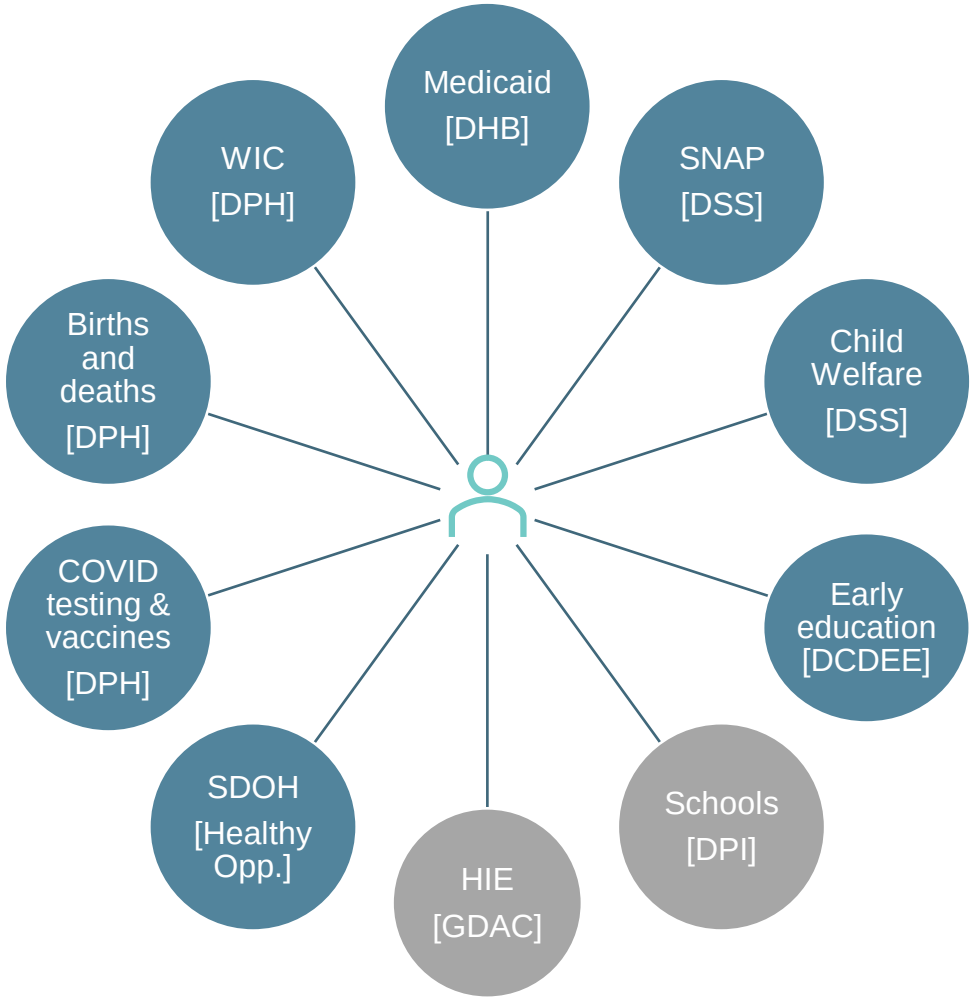
# Map each ID to a universal unique identifier…




## “IDHub”

| Source | Source ID | Unique ID |
|---|---|---|
| Medicaid | 123 | 100 |
| SNAP | XYZ | 100 |
| Immunizations | QRS | 100 |

# …enabling Whole Person Health

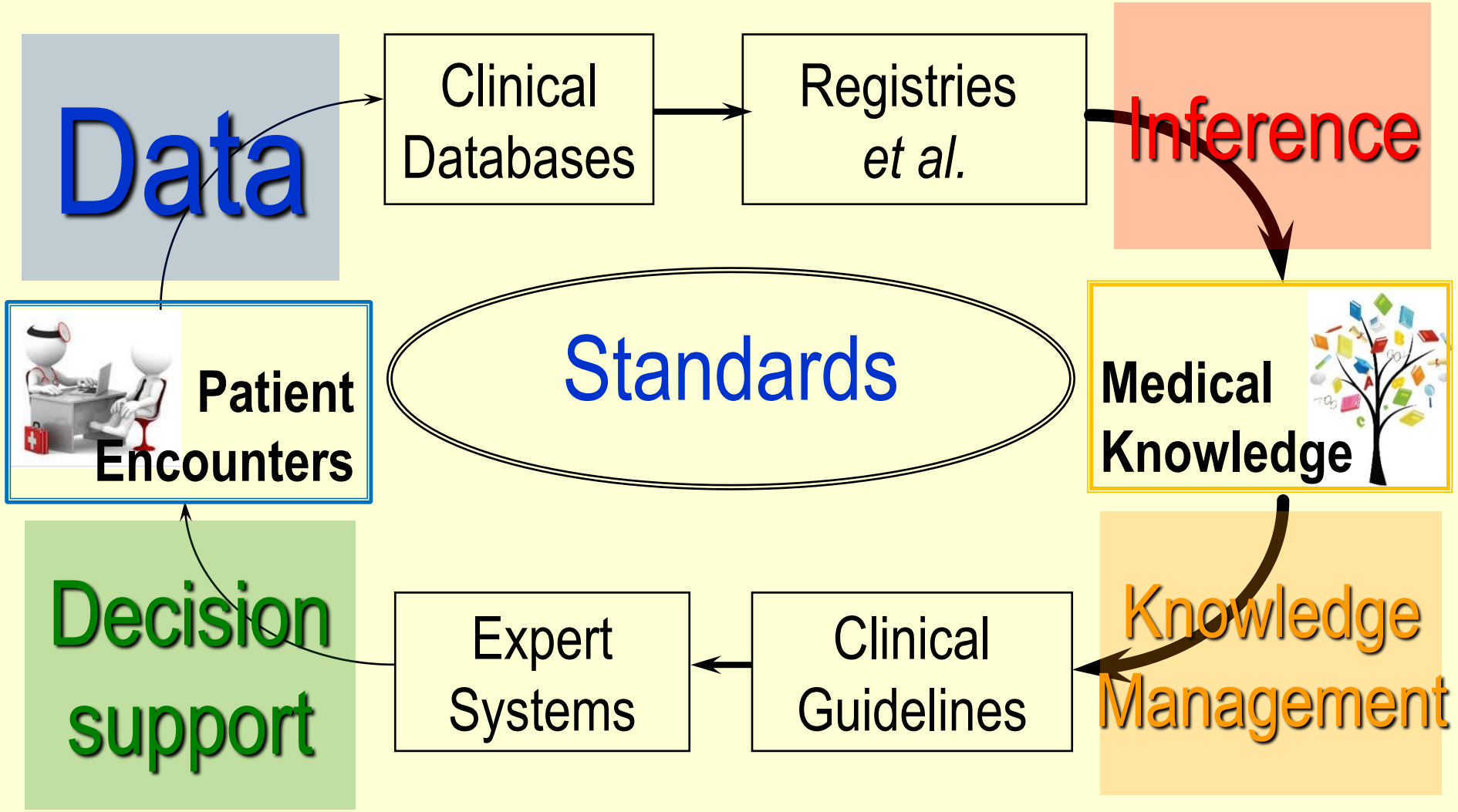


# Thanks!

@jessiet1023

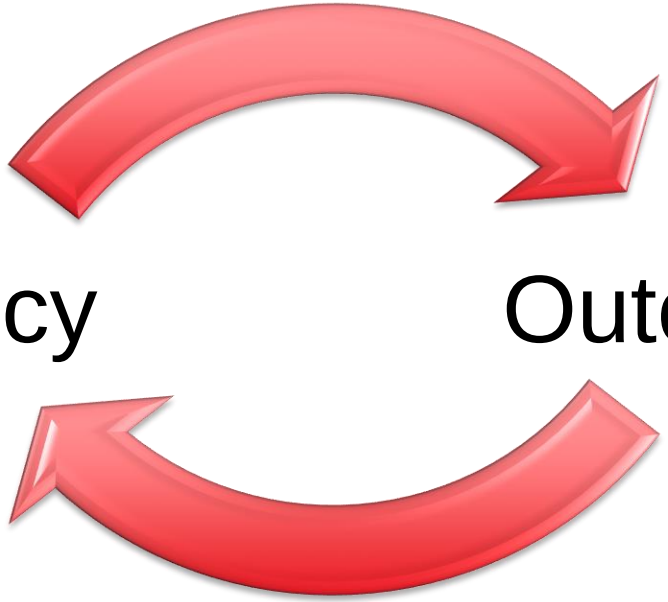
Adapted from Chris Chute

# Foundations for Learning Health System
# From "Evidence-based Medicine" to "Practice-based Evidence"



# The Learning DHHS

Policy impacts outcomes



Policy Outcomes

Outcomes inform policy