



# Leveraging AI to Advance Public Health Data Infrastructure

**Aug. 18, 2025 | 2025 ARC Preconference**

**SHAPING TOMORROW'S PUBLIC HEALTH TODAY.**



# Session Objectives

- Understand the foundational concepts of AI and how it can be applied to public health.
- Explore real-world applications of GenAI in public health through case studies.
- Recognize challenges and best practices in implementing GenAI solutions in public health settings.



# Agenda

1. Understanding AI in Public Health
2. Speaker Introductions
3. PHA Speaker presentations showcasing GenAI being used in Public Health
4. Q&A



# Understanding AI in Public Health

---



# Key Definitions

**Artificial Intelligence (AI)** is the development, implementation, and use of computer systems that can perform tasks that typically require human intelligence. AI describes technology that makes computers (seem to) act rationally.

**Machine learning (ML)** allows a computer to analyze data to do a task without being explicitly programmed. Common functions of machine learning are to (1) find patterns, like groupings of similar items and (2) to guess or predict an output based on a set of inputs.

**Narrow AI** also known as “weak” or “traditional” AI, focuses on performing specific tasks within a limited domain, such as image recognition, speech synthesis, or playing chess. Narrow AI has been in use for decades (decision support, google searches)

**General AI** refers to highly autonomous systems that feel like they possess human-level intelligence and can handle various cognitive tasks across different domains. Large language models (LLMs) are the major advancement in general AI.

- **Generative AI (GenAI)** refers to artificial intelligence systems that can create new, original content, such as text, images, or music, by learning patterns and structures from existing data.





# Key Definitions

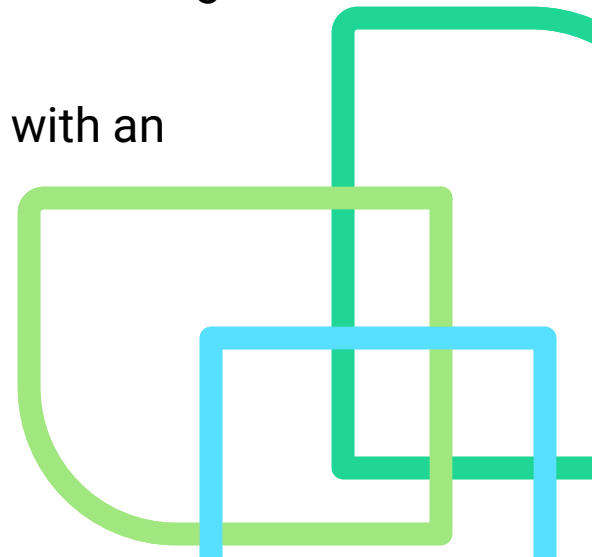
**Large language models:** These are advanced AI models that can understand and generate human-like text based on the patterns and structures they learn from vast amounts of data

**Text chunking and vectorization/vector store:** This involves breaking down text into smaller chunks and converting them into numerical vectors that can be stored and processed by AI models

**Retrieval augmented generation (or RAG) approach:** This is a method that combines the strengths of retrieval-based and generation-based models to produce more accurate and contextually relevant response

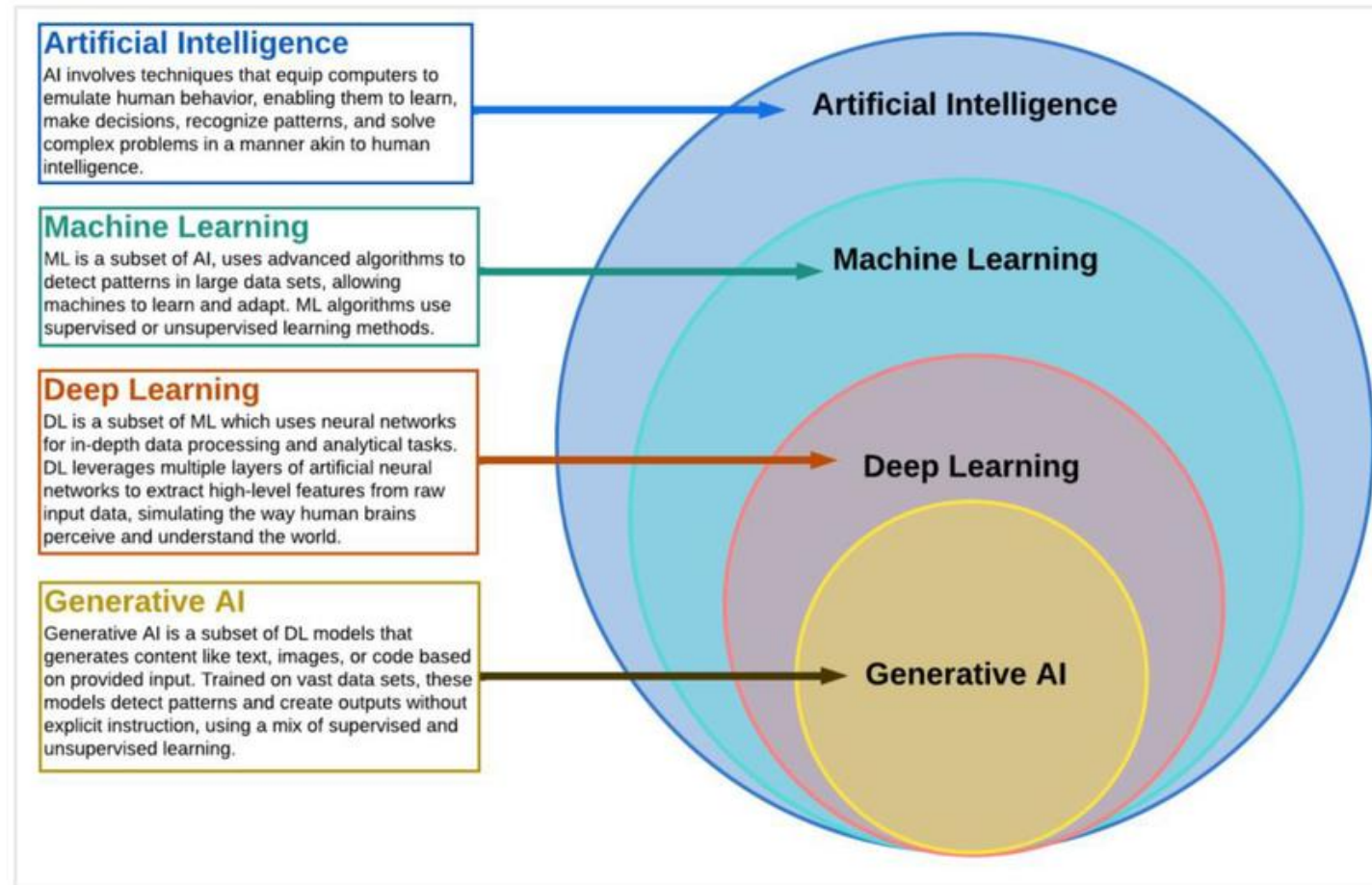
**Deterministic variables:** These are variables in a model that have a fixed value and do not change based on probability

**Probabilistic variables:** These are variables that have a range of possible values, each with an associated probability



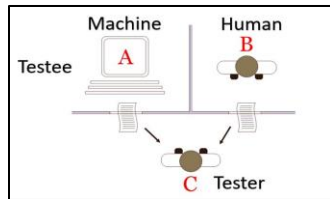


# A Comparative View of AI, ML, Deep Learning, & GenAI





# A Comparative View of AI, ML, Deep Learning, & GenAI



## Turing Test (1950)

A vision of what's to come



## IBM Watson (2011)

AI defeats humans in Jeopardy!



## OpenAI's GPT-3 Release (2020)

AI is capable of generating human-like text



## Proliferation of genAI (2023)

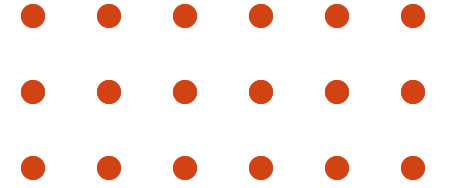
GenAI has expanded across industries and many people use AI daily



## America's AI Action Plan (2025)

White House unveiled comprehensive AI Strategy in July of 2025





# Meet the Speakers



**Dejan Jovanov**

Chief Data & Informatics Officer,  
Division of Patient Safety & Quality,  
Office of Policy Planning & Statistics  
Illinois Department of Public Health



**Stephanie Meyer, MPH**

Epidemiology & Data Unit Supervisor,  
Emerging Infectious Diseases  
Epidemiology & Response Section,  
Infectious Disease Epidemiology,  
Prevention, & Control Division  
Minnesota Department of Health



**Jim Daniel, MPH**

Leader of State and Local  
Public Health at Amazon  
Web Services (AWS)



# Measles Outbreak Simulator Dashboard

*Changing the response from reactive to proactive*





Dejan Jovanov (he/him)  
Chief Data and Informatics Officer  
Illinois Department of Public  
Health



# Setting the Stage:

## Purpose:

Projects and estimates the likelihood of measles outbreaks in Illinois schools (PreK–12) using:

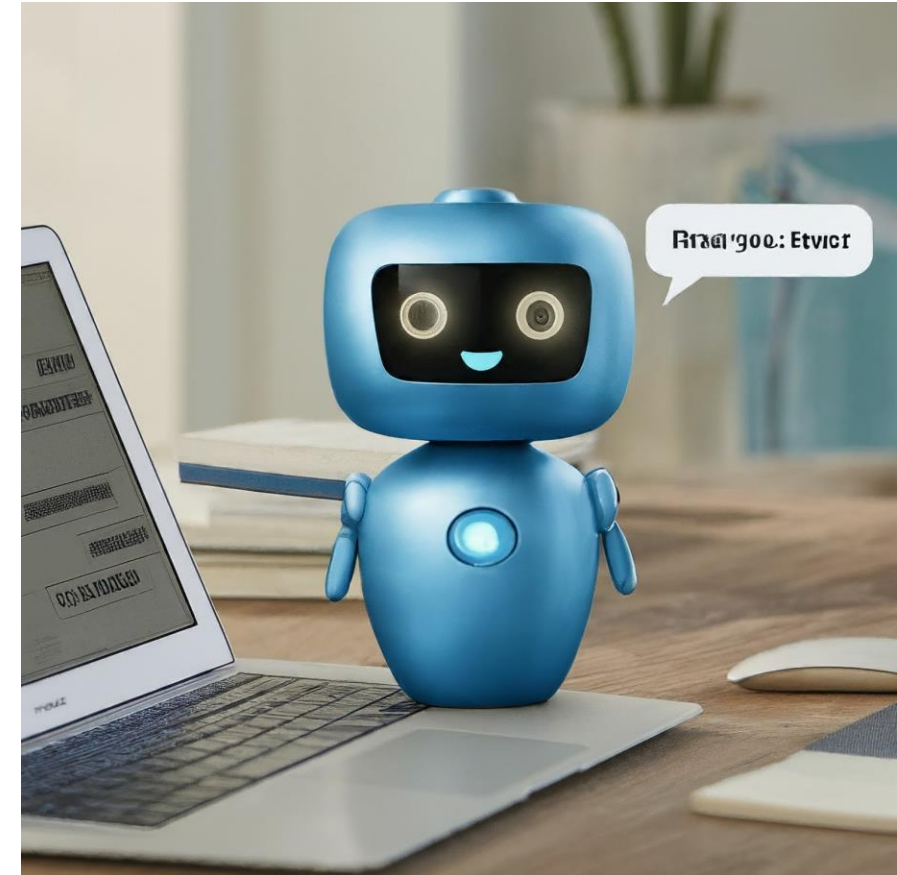
- School vaccination rates
- Enrollment data
- Additional model parameters

## Background:

- Released by the Illinois Department of Public Health (IDPH) in April 2025 due to a nationwide surge in measles cases
- Inspired by the University of Texas Austin *epiEngage Measles Outbreak Simulator*
- Part of our initiative to transform the IDPH data ecosystem from reactive to proactive

## Use:

- Supports school administrators and staff in outbreak prevention and control efforts
- Inform the public





# Measles Simulator Dashboard Data Components



Illinois State Board of Education (ISBE) School Immunization Data and Directory Information (PreK-12)\*

- Measles Vaccination Coverage %
- Enrollment Counts

1

Basic Reproduction Number ( $R_0$ )

- The average number of individuals in a susceptible population that become infected after exposure to an infectious individual with measles.

2

Average Latent Period (days)

- The number of days a person is infected with measles before they become infectious.

3

Average Infectious Period (days)

- The average number of days an individual with measles can continue to spread or infect others with measles.

4

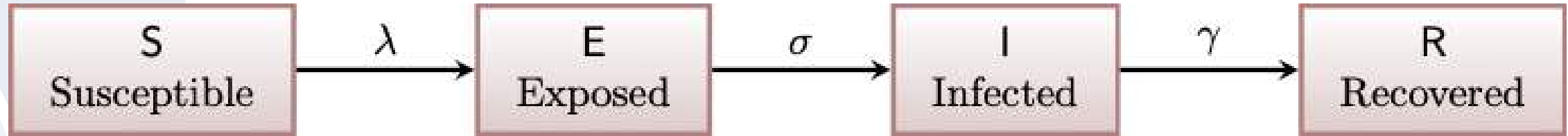
Minimum Outbreak Size (new infections)

- The minimum number of new infections needed to define an outbreak in the selected school.

Likelihood and  
Size of a  
Measles  
Outbreak in an  
Illinois School



# SEIR model used for Measles Prediction



- **Model Overview:**
- **SEIR** = *Susceptible* → *Exposed* → *Infected* → *Removed*
- Population moves between compartments based on defined rates.
- **Key Parameters:**
- **$\lambda$  (Exposure Rate):** How quickly susceptible individuals become exposed
- **$\sigma$  (Latent Rate):** How quickly exposed individuals become infected
- **$\gamma$  (Removed Rate):** How quickly infected individuals recover or are removed
- Also includes: vaccine rate, initial infection, contact rate, basic reproduction number ( **$R_0$** )
- Defaults follow CDC Measles Clinical Diagnosis Fact Sheet



# Measles Simulator Dashboard Calculation Process

Request  
from  
Dashboard

Infectious Parameters

- Precalculated schools with under 95% vaccination rate with default parameters
- Total of 7,374 cached results
- Peak of 2,488 new cached results on 4/9/25 when the dashboard was on the news
- Other options:
  - TabPy
  - Snowflake

AI : Machine Learning Services

Hash Infectious Parameters

Hash  
Exists

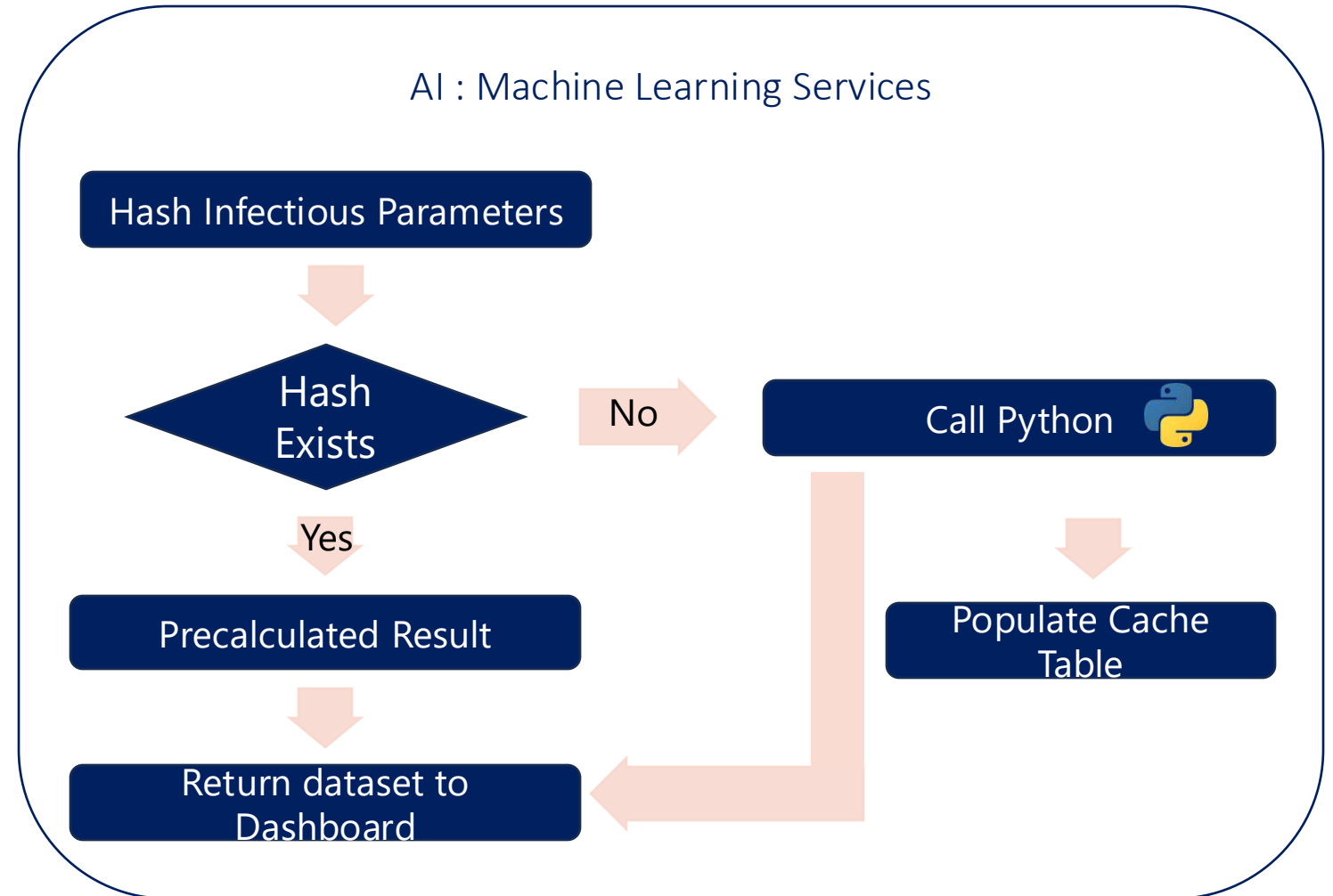
No

Call Python 

Populate Cache  
Table

Precalculated Result

Return dataset to  
Dashboard





# Measles Simulator Dashboard

## Dashboard Parameters



### 1. Choose Enrollment & Vaccination Rates

- Select manually or pick a specific school.

1

### Measles Outbreak Projections in Illinois Schools

Select the county, type of school funding and school district you want to see data for.

If no results are returned, confirm that the school district filter is either set to (All), or to a relevant school district in the selected county-school funding combination.

County	Funding Type	School District
City of Chicago	Public	(All)

Select a school from the list of relevant schools below to generate a simulation.

School District	Facility Name	Enrollment PreK-12	School Vaccination Rate
Burnham SD 154-5	Burnham Elem School	146	98.6%
Calumet Public SD 132	Burr Oak Academy	318	95.3%
	Burr Oak Elem School	288	97.2%
	Calumet Elem School	309	98.1%

2

### 2. Filter Schools

- Use County, Funding Type, and School District filters.
- If the county is changed, adjust the School District drop-down or set to (All).

3

### 3. Update & Select

- The table updates based on your selections.
- Click a school to use its enrollment and vaccination rate.

If you are a keyboard user, or if your school of interest is not present in the list above, you can type in the school model inputs below.

School Enrollment	School Vaccination Rate (%)	Students Initially Infected
500	85.0%	1

4

- Values can also be adjusted manually to model changes.

You can change the default epidemic parameters using the sliders below to generate a more customized projection.

Please note that the four parameters below are for use of local health departments and epidemiologists.

Interpretation of the information based on any adjustments to the epidemic parameters below is extremely nuanced and subject to misinterpretation.

Basic Reproduction Number (R0)	Average Latent Period (days)	Average Infectious Period (days)	Minimum Outbreak Size (new infections)
15	10.5	5	10

### 4. Adjust Epidemic Parameters

- Modify the minimum outbreak size to change the threshold for potential outbreaks.
- Other epidemic parameters can be updated for custom scenarios.

Chance of exceeding 10 new infections

58%

with selected model inputs and epidemic parameters

Likely outbreak size

45 - 72 cases

if there are over 10 new infections



# Measles Simulator Dashboard

## Dashboard Results

**1** Chance of exceeding 10 new infections  
**58%**  
with selected model inputs and epidemic parameters

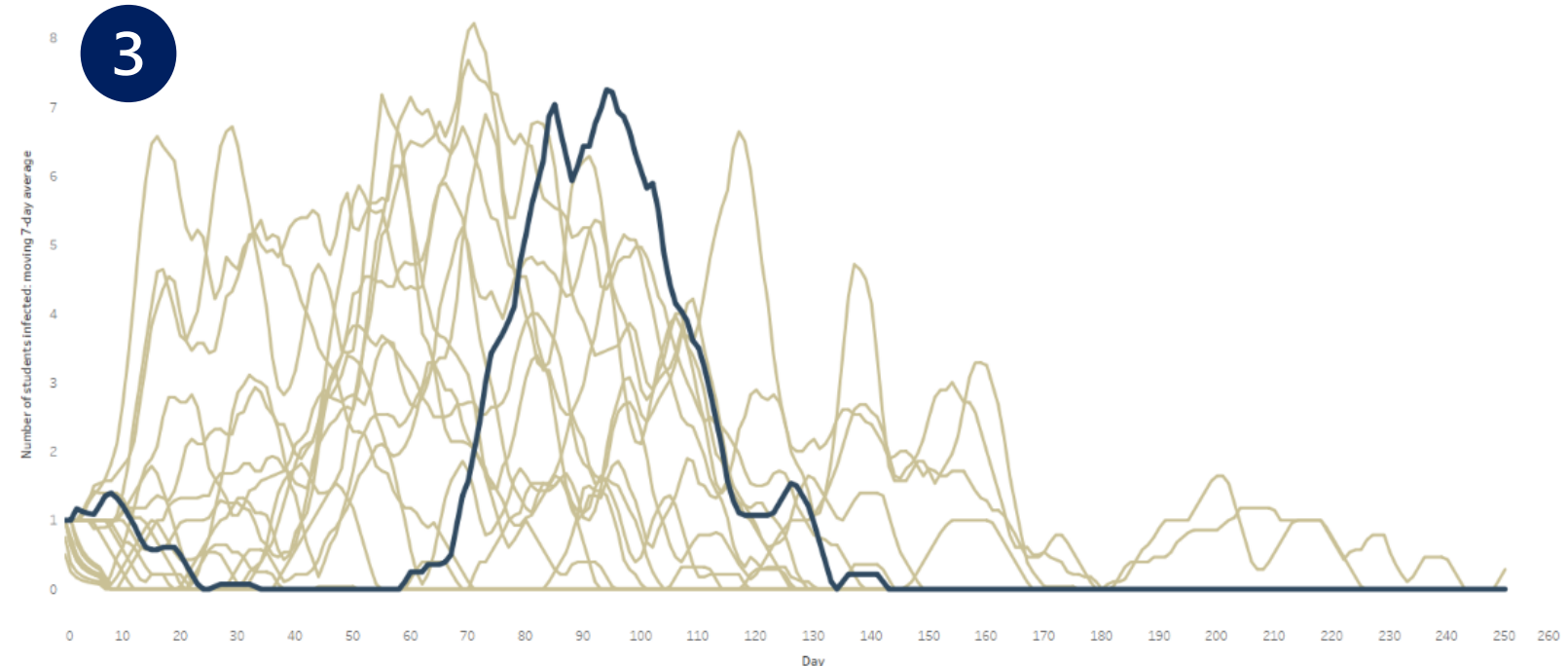
**2** Likely outbreak size  
**45 - 72 cases**  
if there are over 10 new infections

The results are calculated by running 200 stochastic simulations using the entered parameters.

1. The "Chance of exceeding X new infections" is the percentage of simulations that had more infections than the selected minimum (10 by default).
2. Likely outbreak size uses the middle 95% of the outbreaks that are larger than the selected minimum (10 by default).
3. The simulated projections show a random selection of 20 outbreak projections. The highlighted line is the projection with the outbreak size nearest to the median.

Projections and estimates for outbreaks provided below, are specific to the school population selected and incorporate the selected school vaccination coverage, enrollment, and additional model parameter values. The calculation of the proportion of the school population susceptible to measles does not account for community immunity outside schools, interventions, or breakthrough infections amongst vaccinated students.

### Simulated Projections



Simulation Color Legend

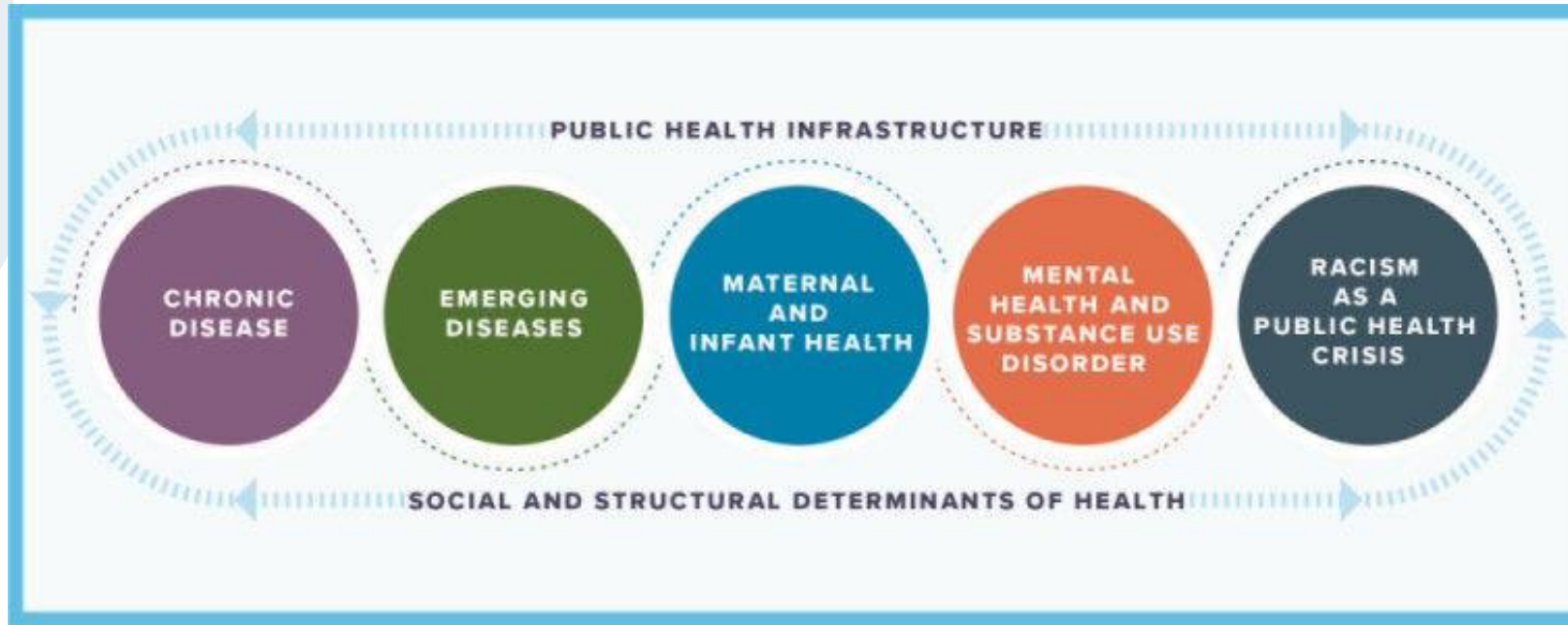
■ Projection closest to the median ■ All other projections

Use the highlighter dropdown below to highlight any one particular projection above.

Highlight Min. Simulation ID



# What's Next



- Adapt the SEIR model (adding compartments, age groups, vaccination status, seasonality) to fit the specific disease.
- Expand the use of the model on diseases with a clear incubation period (e.g., measles, COVID-19, influenza).
- Build additional dashboards





Questions?

Contact:

Dejan Jovanov: [Dejan.Jovanov@Illinois.gov](mailto:Dejan.Jovanov@Illinois.gov)





# **Project CLAIRE: Comprehensive Language Model Artificial Intelligence for Record Evaluation - Assessing the use of Artificial Intelligence for Infectious Disease Surveillance**

**Stephanie Meyer, MPH**

**Epidemiology Supervisor | Infectious Disease Epidemiology, Prevention, and Control Division  
Minnesota Department of Health**



# Background

- COVID-19 is one of many reportable diseases in Minnesota
- Minnesota is part of the CDC Emerging Infections Program
- Medical chart abstractions are often necessary
- COVID-NET monitors laboratory-confirmed, COVID-19-associated hospitalizations among children and adults

Reportable Diseases, MN Rule 4605.7000 to 4605.7900

## Diseases Reportable to the Minnesota Department of Health

651-201-5414 or 1-877-676-5414 24 hours a day, 7 days a week

### REPORT IMMEDIATELY BY TELEPHONE

Anthrax ( <i>Bacillus anthracis</i> )	Middle East Respiratory Syndrome (MERS)
Botulism ( <i>Clostridium botulinum</i> )	Ornithosis
Brucellosis ( <i>Brucella</i> spp.)	Plague ( <i>Yersinia pestis</i> )
Cholera ( <i>Vibrio cholerae</i> )	Polio
Diphtheria ( <i>Corynebacterium diphtheriae</i> )	Q fever ( <i>Coxiella burnetii</i> )
Free-living amoebic infection	Rabies (animal and human cases and suspected cases)
(including at least: <i>Acanthamoeba</i> spp., <i>Naegleria fowleri</i> , <i>Balamuthia</i> spp., <i>Sappinia</i> spp.)	Rubella and congenital rubella syndrome
Granuloma ( <i>Mycobacterium</i> spp.)	Severe Acute Respiratory Syndrome (SARS)
Hemolytic uremic syndrome	Smallpox ( <i>variola</i> )
Measles ( <i>rubeola</i> )	Tularemia ( <i>Francisella tularensis</i> )
Melioidosis ( <i>Burkholderia pseudomallei</i> )	Unusual or increased case incidence of any suspect infectious illness
Meningococcal disease ( <i>Neisseria meningitidis</i> ) (invasive)	Viral hemorrhagic fever (including but not limited to Ebola virus disease and Lassa fever)

### REPORT WITHIN ONE WORKING DAY

Amebiasis ( <i>Entamoeba histolytica/dispar</i> )	Listeriosis ( <i>Listeria monocytogenes</i> )
Anaplasmosis ( <i>Anaplasma phagocytophilum</i> )	Lyme disease ( <i>Borrelia burgdorferi</i> , and other <i>Borrelia</i> spp.)
Arboviral disease (including, but not limited to, La Crosse encephalitis, eastern equine encephalitis, western equine encephalitis, St. Louis encephalitis, West Nile virus disease, Powassan virus disease, and Jamestown Canyon virus disease)	Malaria ( <i>Plasmodium</i> spp.)
Babesiosis ( <i>Babesia</i> spp.)	Meningitis (caused by viral agents)
Brucellosis ( <i>Brucella</i> spp.)	Mumps
Campylobacteriosis ( <i>Campylobacter</i> spp.)	Neonatal sepsis
Candida auris	(bacteria isolated from a sterile site, excluding coagulase-negative <i>Staphylococcus</i> less than seven days after birth)
Carbapenem-resistant Enterobacteriaceae (CRE)	Peritonitis ( <i>Bordetella pertussis</i> )
Cat scratch disease (infection caused by <i>Bartonella</i> species)	Pittsboro ( <i>Chlamydia pneumoniae</i> )
Chancroid ( <i>Haemophilus ducreyi</i> )	Reproductive infections
Chikungunya virus disease	Salmonellosis, including typhoid ( <i>Salmonella</i> spp.)
Chlamydia trachomatis infections	Shigellosis ( <i>Shigella</i> spp.)
Coccidioidomycosis	Spotted fever rickettsiosis ( <i>Rickettsia</i> spp.)
Cryptosporidiosis (in infants under one year of age)	Staphylococcal disease (only <i>Staphylococcus aureus</i> (VISA), vancomycin-resistant <i>Staphylococcus aureus</i> (VRSA), and death or critical illness due to community-associated <i>Staphylococcus aureus</i> in a previously healthy individual)
Cytomegalovirus (CMV) infection	Streptococcal disease - invasive disease caused by Groups A and B streptococci and <i>S. pneumoniae</i>
Dengue virus infection	Streptococcal disease - non-invasive <i>S. pneumoniae</i> (urine antigen laboratory-confirmed pneumonia)
Diphtheria	Syphilis ( <i>Treponema pallidum</i> )
Enteric (EHEC) coli infection ( <i>E. coli</i> O157:H7, other Shiga toxin-producing <i>E. coli</i> , enterohemorrhagic <i>E. coli</i> , enterotoxigenic <i>E. coli</i> , or other pathogenic <i>E. coli</i> )	Tetanus ( <i>Clostridium tetani</i> )
Giardiasis ( <i>Giardia lamblia</i> )	Toxic shock syndrome
Gonorrhea ( <i>Neisseria gonorrhoeae</i> )	Toxoplasmosis ( <i>Toxoplasma gondii</i> )
Hepatitis A virus infection	Transmissible spongiform encephalopathy
Hepatitis B virus infection	Trichinosis ( <i>Trichinella spiralis</i> )
Hepatitis C virus infection	Tuberculosis ( <i>Mycobacterium tuberculosis</i> complex) (pulmonary or extrapulmonary sites of disease, including clinically diagnosed disease). Latent tuberculosis infection is not reportable.
Hepatitis E virus infection	Typhoid ( <i>Salmonella</i> spp.)
Histoplasmosis ( <i>Histoplasma capsulatum</i> )	Unexplained deaths and unexplained critical illness (positivity due to infectious cause)
Human immunodeficiency virus (HIV) infection, including Acquired Immunodeficiency Syndrome (AIDS)	Varicella ( <i>Varicella zoster</i> )
Influenza (unusual case incidence, critical illness, or laboratory-confirmed cases)	Vibrio spp.
Kawasaki disease	Yersinia
Legionnaires disease	Yersinia enterocolitica (regardless of specimen source)
Leptospirosis ( <i>Leptospira interrogans</i> )	Zika virus disease
Leptospirosis ( <i>Leptospira interrogans</i> )	Zoster (shingles)

### SENTINEL SURVEILLANCE

Diseases reportable through sentinel surveillance are reportable based on the incidence of the pathogen in the specific health care facility. Sentinel surveillance is not statewide reporting.

Staphylococcus aureus
Candidemia ( <i>Candida</i> spp.) (blood isolates only)
Carbapenem-resistant Acinetobacter spp. (CRAB)
and Pseudomonas aeruginosa (CR-PA)
Clostridium difficile
Respiratory syncytial virus (RSV)
Non-tuberculous Mycobacteria (NTM), pulmonary and extrapulmonary

### FOOTNOTES

- Submission of clinical materials required. Submit isolates or, if an isolate is not available, submit material containing the infectious agent in the following order of preference: a patient specimen; nucleic acid; or other laboratory material. Call the MDH Public Health Laboratory at 651-201-4953 for instructions.
- Invasive disease only: isolated from a normally sterile site, e.g., blood, CSF, joint fluid, etc.
- In the event of SARS or another severe respiratory outbreak, also report cases of health care workers hospitalized for pneumonia or acute respiratory distress syndrome.
- Also report a pregnancy in a person with Zika, or a person chronically infected with hepatitis B, HIV, or syphilis.

### TO REPORT

- For immediate reporting call: 651-201-5414 or 1-877-676-5414.
- Report forms can be downloaded at [www.health.state.mn.us/diseases](http://www.health.state.mn.us/diseases)

**MDH DEPARTMENT OF HEALTH**  
Infectious Disease Epidemiology, Prevention and Control  
Phone: 651-201-5414 or 1-877-676-5414 | Fax: 651-201-5443  
www.health.state.mn.us/diseases



# What is a Medical Chart Review?

## ED Vitals:

Patient Vitals for the past 24 hrs:

	BP	Temp	Pulse	Resp	SpO2	Height	Weight
09/02/23 0202	127/85	99 °F (37.2 °C)	97	22	94 %	—	—
09/02/23 0000	(!) 115/91	—	(!) 110	(!) 34	95 %	—	—
09/01/23 2345	128/78	—	(!) 103	(!) 32	95 %	—	—
09/01/23 2330	—	—	(!) 103	(!) 26	95 %	—	—
09/01/23 2315	(!) 128/116	—	(!) 108	(!) 32	94 %	—	—
09/01/23 2306	136/88	99.5 °F (37.5 °C)	(!) 107	20	97 %	5' 4.02" (1.626 m)	84.8 kg

## MDM:

██████████ is a 76 y.o. female who presents for the above. The differential includes pneumonia, sepsis, viral syndrome, antibiotic reaction. The patient has generalized weakness here and recent diagnosis of pneumonia. Patient is given albuterol treatment. Sepsis protocol was started and she is given some lactated Ringer's. Patient is given levofloxacin after extensive conversation with her about her allergies and given the pneumonia. Chest x-ray does show significant left-sided infiltrate. She has significant hyponatremia as well. I did discuss the case with renal on-call who recommended fluid resuscitation and recheck of sodium. Patient is transferred to hospitalist service for further work-up and care.

## DIAGNOSIS:

	ICD-10-CM
1. Pneumonia of left lung due to infectious organism, unspecified part of lung	J18.9
2. Hyponatremia	E87.1

Medical charts are read and reviewed to determine:

- Course of illness
- Course of hospitalization
- Treatments
- Underlying Conditions



## I Notes (group 1 of 2) (continued)

g: Take 25 mg by mouth once daily. molecalciferol, vitamin D3, 25 mcg, 1000 nit, 25 mcg (1,000 unit) oral tablet g: Take 25 mcg by mouth once daily.	9/29/2022	Patient	Yes	Yes
g: Take 10 mg by mouth every evening. clobenzaprine (FLEXERIL) 10 mg oral blet g: Take 10 mg by mouth every evening.	9/29/2022	Patient	Yes	Yes
g: Inject 1.5 mg under the skin every 7 (seven) days. abapentin (NEURONTIN) 300 mg oral apsule g: Take 300 mg by mouth three times a day.	9/29/2022	Patient	Yes	Yes
g: Take 5 mg by mouth twice a day. ipizide (GLUCOTROL XL) 5 mg oral extended release tablet 24 HR g: Take 5 mg by mouth twice a day.	9/29/2022	Patient	Yes	Yes
g: Take 1 capsule by mouth once daily. multivit-mins/iron/folic/lycop (CENTRUM EN ORAL) g: Take 1 capsule by mouth once daily.	9/29/2022	Patient	Yes	Yes
g: Instill 4 mg into one nostril as needed. Seek emergency care immediately after use. aloxone (NARCAN) 4 mg/actuation Nasal pray g: Instill 4 mg into one nostril as needed. Seek emergency care immediately after use.	PRN	Patient	Yes	Yes
g: Take 20 mg by mouth twice a day. xyCODONE (OXYCONTIN) 20 mg oral extended release tablet 12 HR g: Take 20 mg by mouth twice a day.	9/29/2022	Patient	Yes	Yes
g: Take 15 mg by mouth every 4 (four) hours as needed. xyCODONE, immediate release, TOXICODONE) 15 mg oral tablet g: Take 15 mg by mouth every 4 (four) hours as needed.	9/29/2022	Patient	Yes	Yes
g: Take 40 mg by mouth once daily. antoprazole (PROTONIX) 40 mg oral elayed release tablet g: Take 40 mg by mouth once daily.	9/29/2022	Patient	Yes	Yes
g: Take 20 mg by mouth once daily. varoxaban (XARELTO) 20 mg oral tablet @0200 g: Take 20 mg by mouth once daily.	9/30/2022	Patient	Yes	Yes
g: Take 0.4 mg by mouth at bedtime. msulosin (FLOMAX) 0.4 mg oral capsule g: Take 0.4 mg by mouth at bedtime.	9/29/2022	Patient	Yes	Yes
g: Take 100 mg by mouth at bedtime. aZODone (DESYREL) 100 mg oral tablet g: Take 100 mg by mouth at bedtime.	9/29/2022	Patient	Yes	Yes

# Medical Chart Layout

- Medications or treatments may be listed over multiple pages or multiple days
- Each treatment may have been administered already or something that the patient will take home
- Details and nuance about dosage and timing can be very important in relation to an infection



# Chart Abstractions

<b>4. Height:</b> <input type="text" value="65"/> <input checked="" type="radio"/> in <input type="radio"/> cm <a href="#">reset</a>	<b>5. Weight:</b> <input type="text" value="116.3"/> <input type="radio"/> lbs <input checked="" type="radio"/> kgs <a href="#">reset</a>	<b>6. BMI (non-pregnant cases and cases 2+ years old):</b> <input type="text" value="42.7"/> <input type="text" value="42.7"/> <a href="#">View equation</a>
<b>7. Smoker (tobacco):</b> <input checked="" type="radio"/> Current <input type="radio"/> Former <input type="radio"/> No/ Unknown <a href="#">reset</a>		<b>7a. Smoker (marijuana/vaping) [MN question]</b> <input type="radio"/> Current <input type="radio"/> Former <input checked="" type="radio"/> No/ Unknown <a href="#">reset</a> <b>Type of non-tobacco smoking (select all that apply):</b>
<b>8. Environmental tobacco smoke exposure (for pediatric patients ≤12years):</b>		<b>9. Alcohol misuse:</b> <input type="radio"/> Current <input checked="" type="radio"/> Former <input type="radio"/> No/ Unknown <a href="#">reset</a>
<b>10. Substance misuse:</b> <input checked="" type="radio"/> Current <input type="radio"/> Former <input type="radio"/> No/Unknown <a href="#">reset</a>		<b>11. Substance MisuseType (current use only) (Select all that apply):</b> <input type="checkbox"/> IVDU <input checked="" type="checkbox"/> Opioids <input type="checkbox"/> Cocaine <input type="checkbox"/> Methamphetamines <input type="checkbox"/> Marijuana <input type="checkbox"/> Polysubstance abuse - not otherwise specified <input type="checkbox"/> Unknown <input type="checkbox"/> Other, specify
<b>12. Code status on admission:</b> <input checked="" type="radio"/> Full code <input type="radio"/> DNR (Do Not Resuscitate) / DNI (Do not Intubate) / CMO (Comfort measures only) <input type="radio"/> Unknown <a href="#">reset</a>		

ICU and Other interventions	
<b>Invasive mechanical ventilation?</b> <input checked="" type="radio"/> Yes <input type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>	<b>BiPAP or CPAP?</b> <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>
<b>High flow nasal cannula (e.g. Vapotherm)?</b> <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>	<b>Supplemental Oxygen?</b>
<b>4. ECMO?</b> <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>	
<b>6. Renal Replacement Therapy (RRT) or Dialysis?</b> Includes Peritoneal Dialysis (PD), Hemodialysis (HD), Continuous Venovenous Hemofiltration (CVVH), Continuous Venovenous Hemodialysis (CVVHD), and Slow Continuous Ultrafiltration (SCUF) <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>	
<b>7. Was the patient admitted to an intensive care unit (ICU)?</b> <input checked="" type="radio"/> Yes <input type="radio"/> No <input type="radio"/> Unknown <a href="#">reset</a>	
<b>7a. Date of ICU admission:</b> <input type="text" value="12-24-2024"/> <input type="text" value="31"/> M-D-Y <a href="#">reset</a>	<b>7b. Date of ICU discharge:</b> <input type="text" value="12-25-2024"/> <input type="text" value="31"/> M-D-Y <a href="#">reset</a>
F. Outcome	
<b>1. What was the outcome of the patient upon discharge?</b> <input type="radio"/> Alive <input checked="" type="radio"/> Died during hospitalization <input type="radio"/> Unknown <a href="#">reset</a>	
<b>2. Date of death:</b> <input type="text" value="12-25-2024"/> <input type="text" value="31"/> M-D-Y <a href="#">reset</a>	



# Medical Chart Review

- Manual chart review involves multiple staff
- COVID-NET includes a sample of hospitalized cases for chart review

Season*	2020-2021	2021-2022	2022-2023	2023-2024
Total COVID-NET Cases	39,318	34,154	14,908	12,971
Chart Reviews Completed	20,019 (50%)	7,006 (20%)	2,368 (16%)	2,669 (21%)
Review Staff	17	12	11	10

\* Season starts October 1 through Sept 30 with exception of 2020-2021, which began in March 2020



# Medical Chart Review

- Manual chart review involves multiple staff
- COVID-NET includes a sample of hospitalized cases for chart review

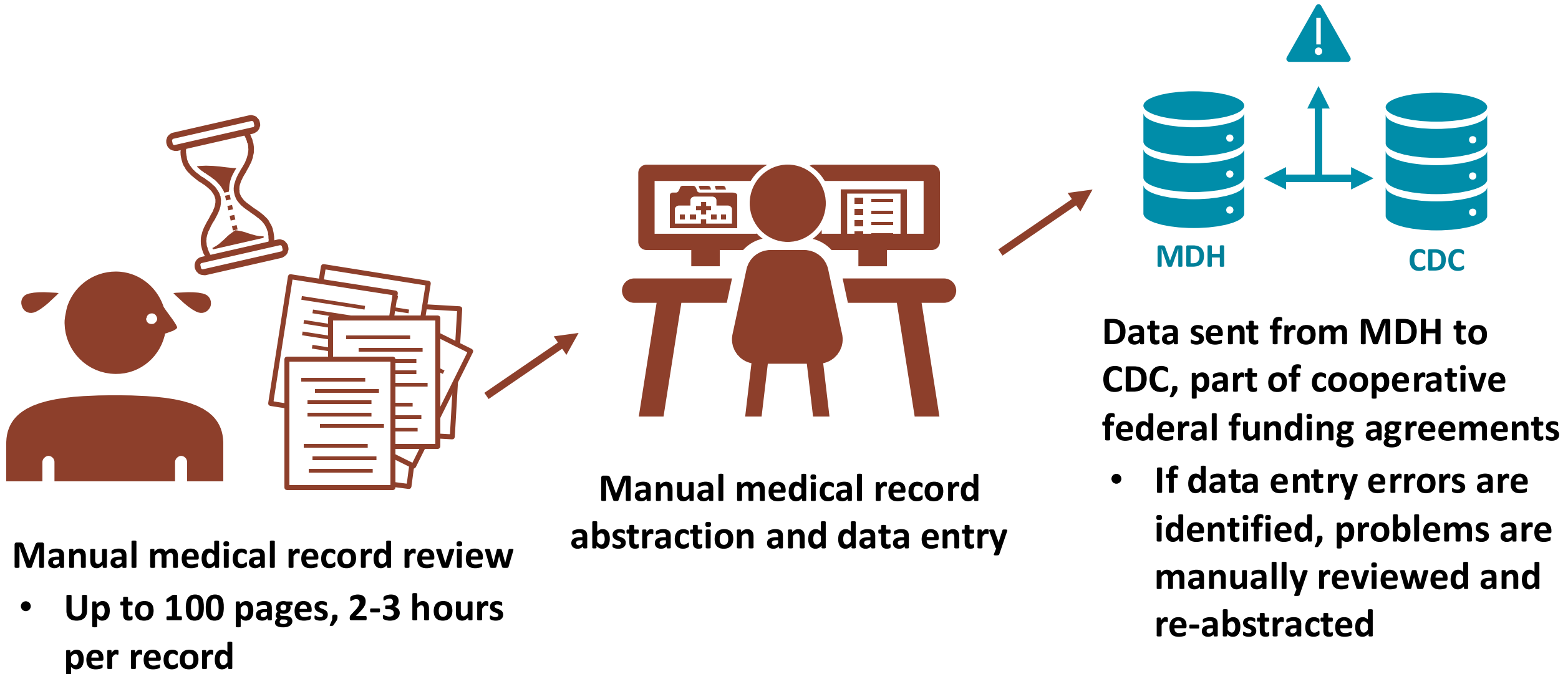
Season*	2020-2021	2021-2022	2022-2023	2023-2024
Total COVID-NET Cases	39,318	34,154	14,908	12,971
Chart Reviews Completed	20,019 (50%)	7,006 (20%)	2,368 (16%)	2,669 (21%)
Review Staff	17	12	11	10

- The cost of full-time staff to review thousands of charts is not tenable

\* Season starts October 1 through Sept 30 with exception of 2020-2021, which began in March 2020

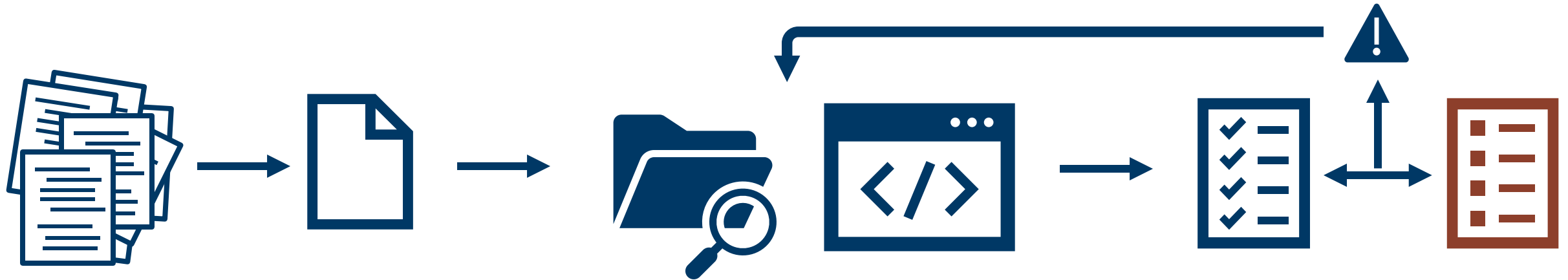


# Current State – Manual Process





# Future State



**A document is generated  
with all the text from the  
medical record**

**Use large language models  
to “read” the chart and  
extract key data elements  
Includes code, prompt  
engineering, and large  
language model selection**

**Data is validated by  
comparison to manually  
collected data  
Results are used to improve  
the abstraction pipeline**

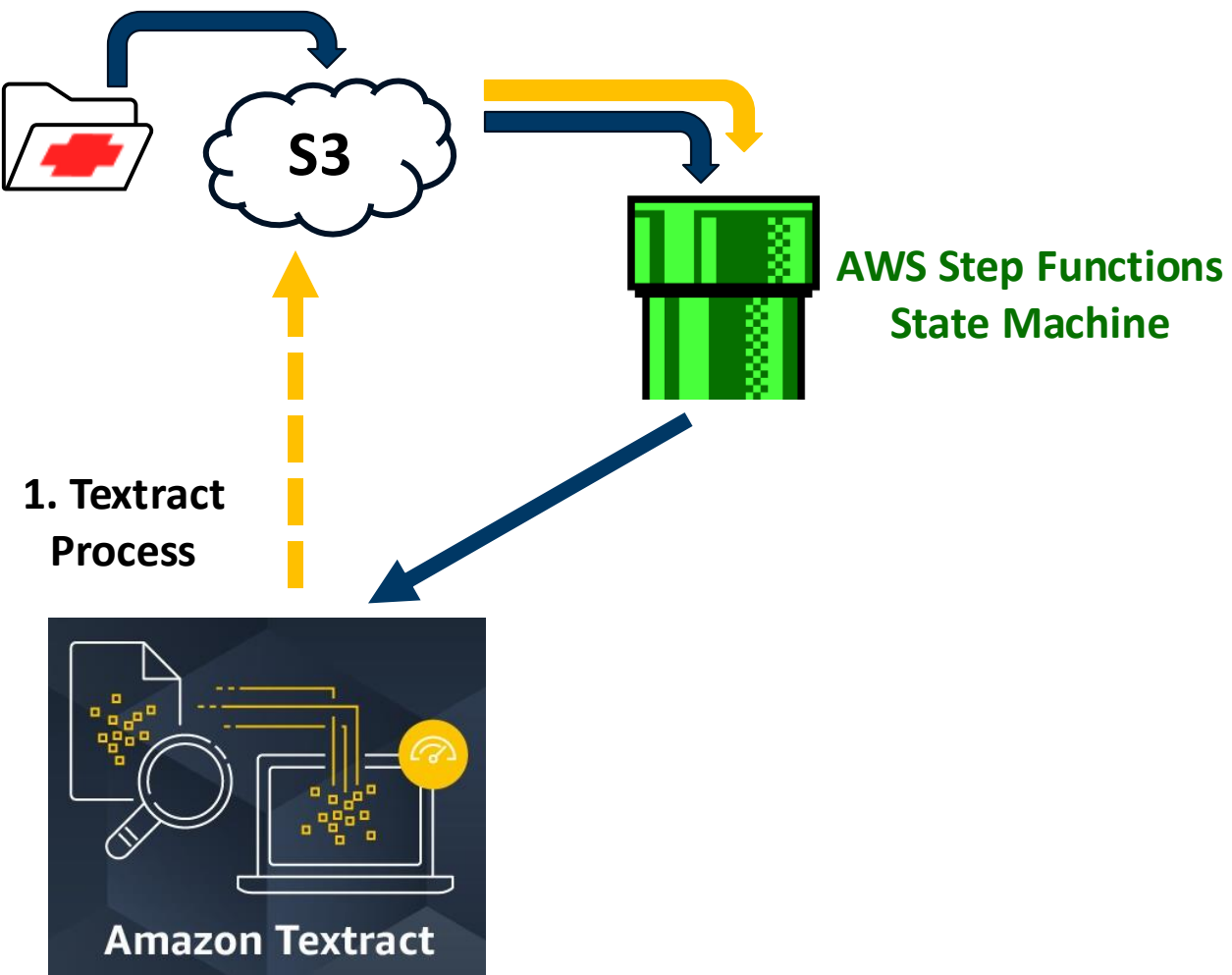


# CLAIRE Pilot Project

- **We identified 150 COVID-19 medical charts from one health system from the 2022-2023 respiratory season**
- **All charts were redacted PDF files**
- **All charts had been manually reviewed by Minnesota Department of Health (MDH) staff with answers entered into a REDCap database**
- **MDH partnered with MNIT MDH and Amazon Web Services (AWS) for evaluation of options**
- **Analysis of the results from text extracts of abstracted charts from the AWS pipeline were compared with manual abstractions**

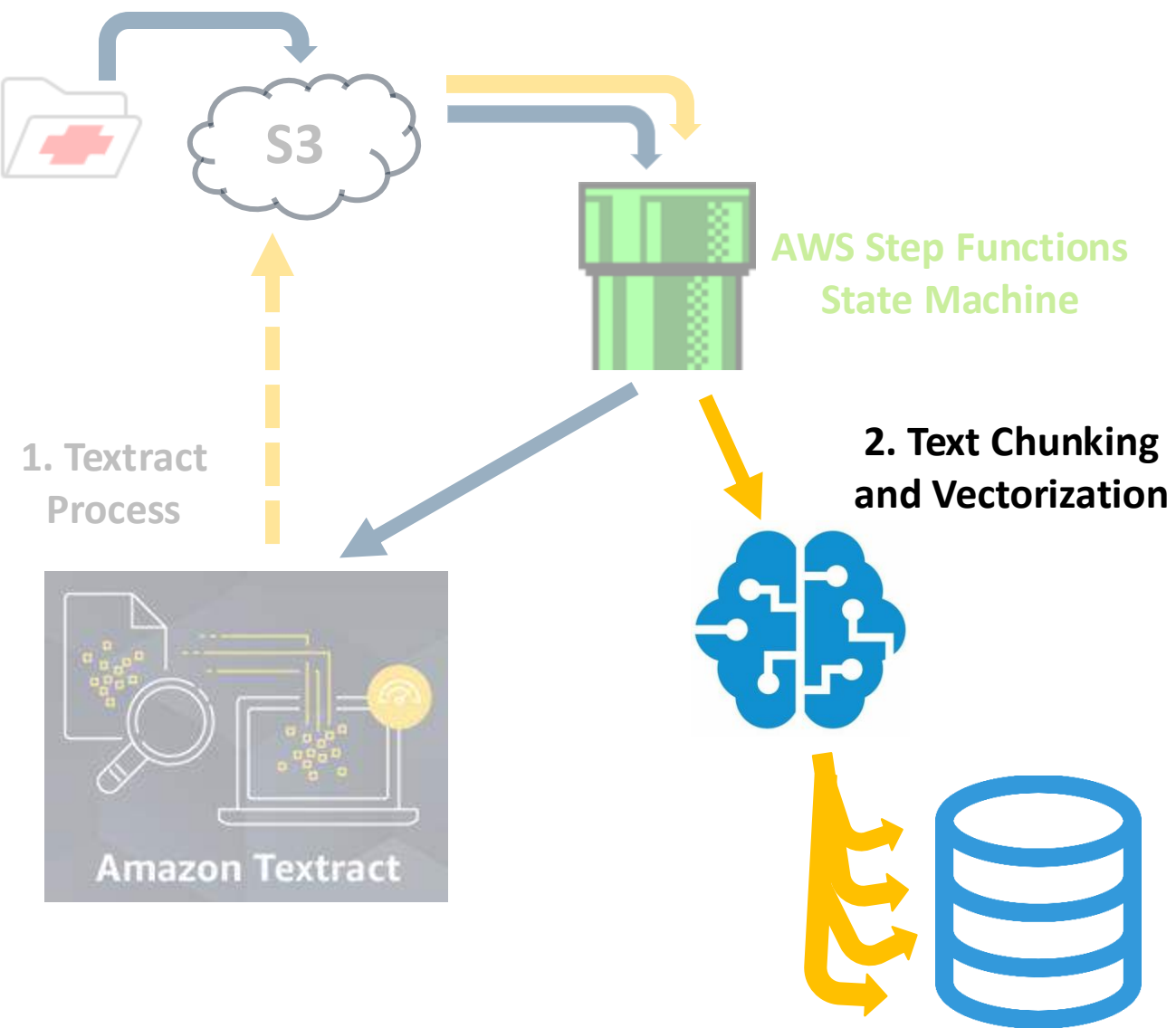


# CLAIRE Pilot Project



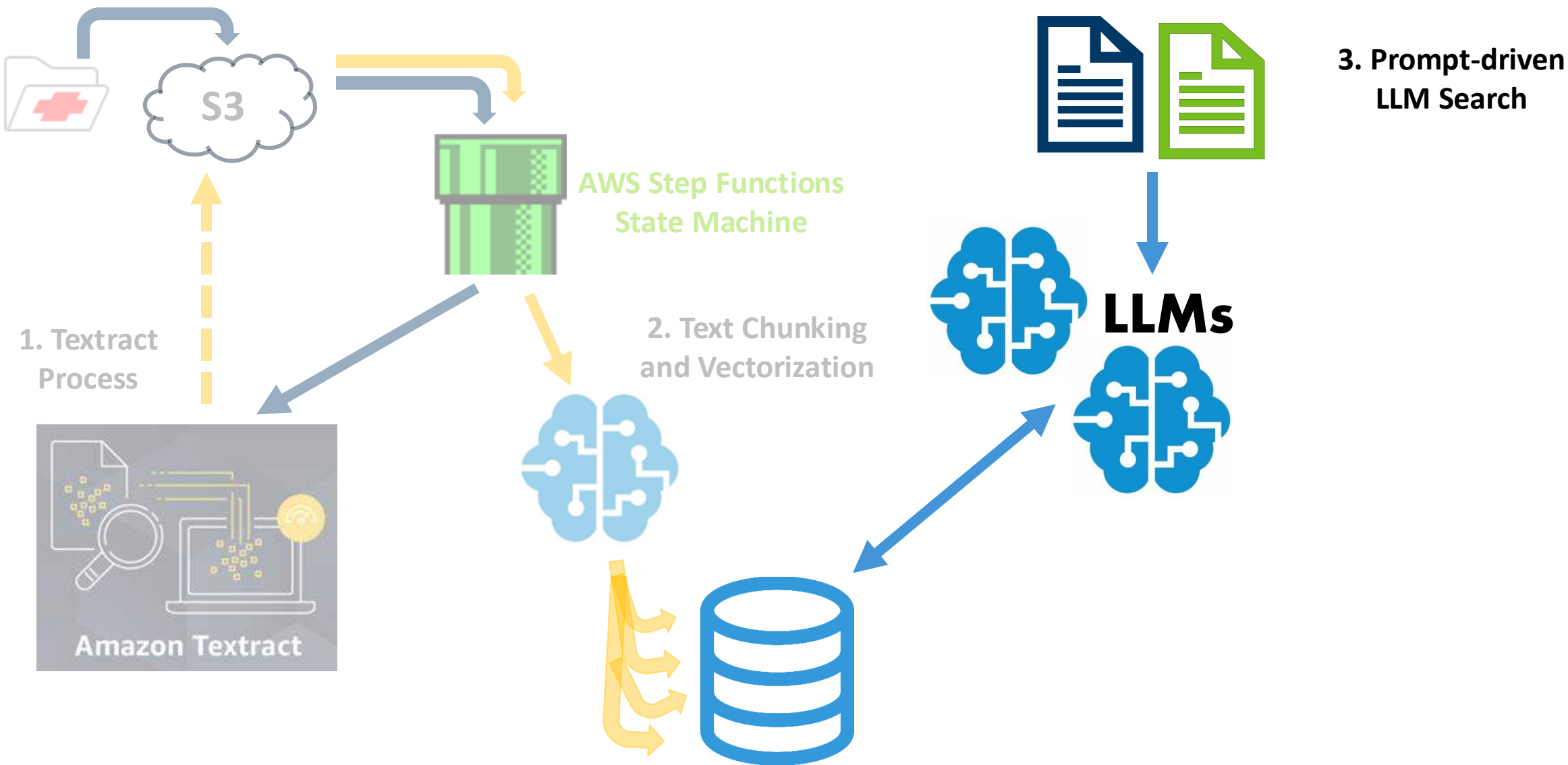


# CLAIRE Pilot Project



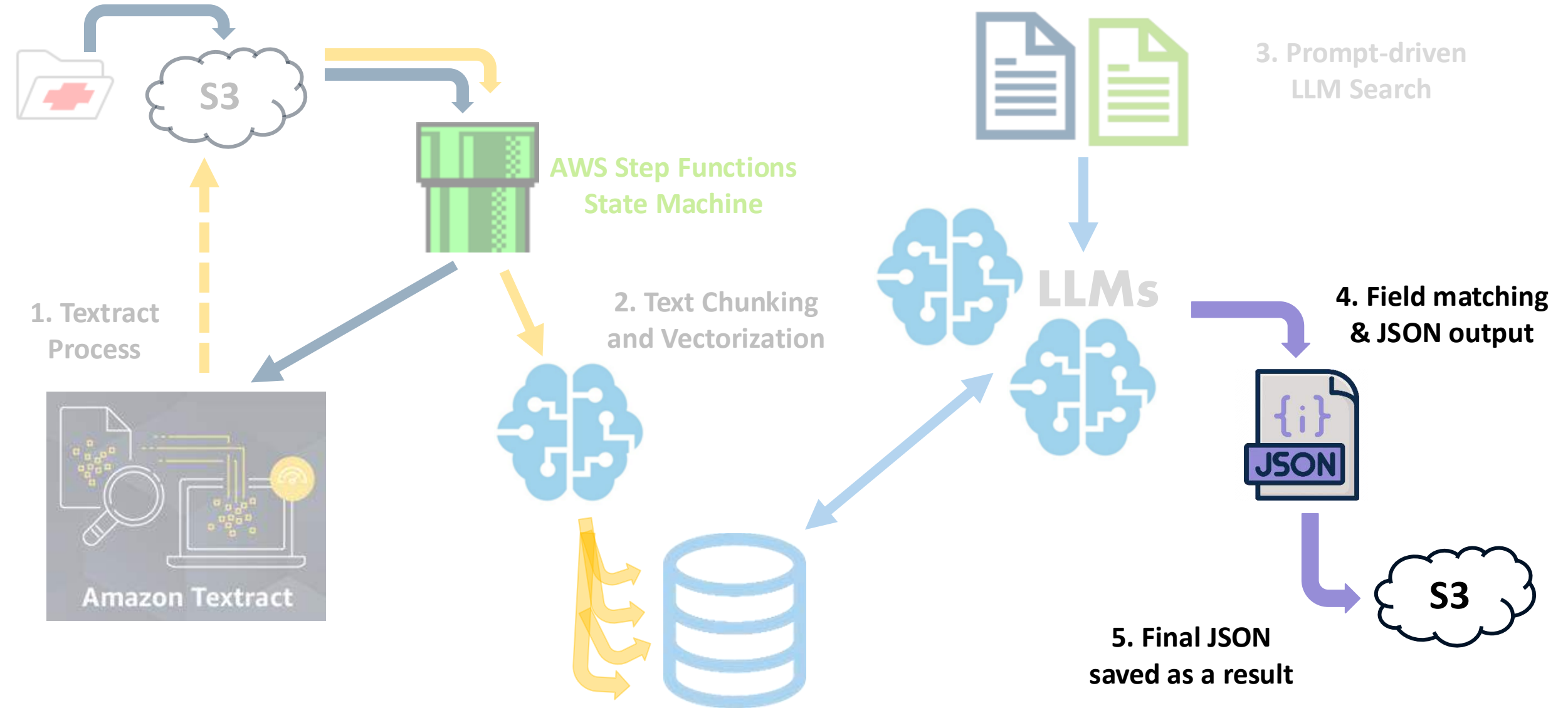


# CLAIRE Pilot Project



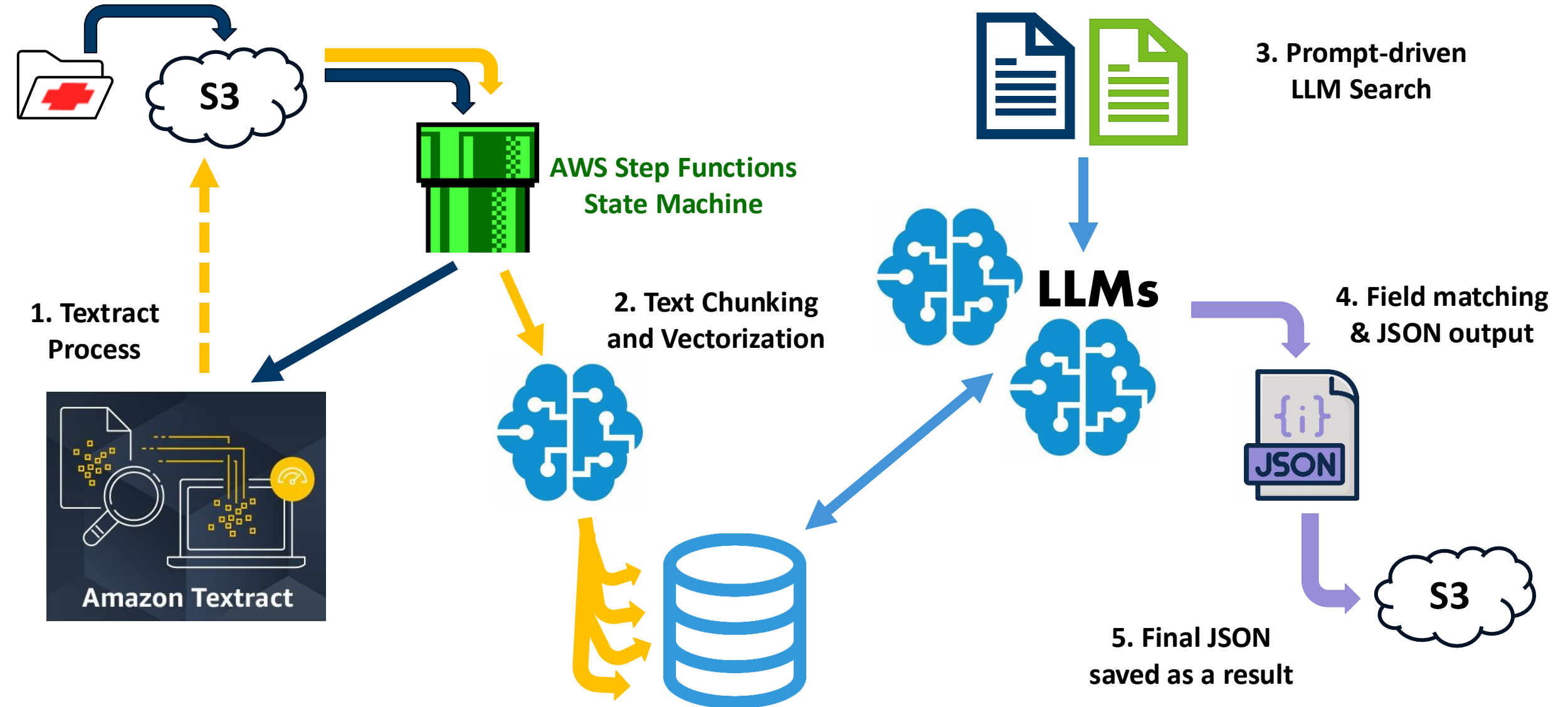


# CLAIRE Pilot Project





# CLAIRE Pilot Project





# Restructuring the Data

- Throughout the pilot project we looked at various data elements and corresponding accuracy
- We divided data into four categories:
- **STRUCTURED DATA** – data collection better suited to standard code
- **PROMPT DRIVEN SEARCH** – data collection that might require decision-making based upon additional information
- **PROBABILISTIC MATCHING** – data collection that we thought might be best-suited for LLMs
- **SUMMARIZATION** – data collection that might require manual review by staff, but a summary with citations could be helpful and streamline the work



# Beneficial Early Results

- We looked at the accuracy by category

Category	Number of variables	Average accuracy across the category
STRUCTURED DATA	76	49.7%
PROMPT DRIVEN SEARCH	50	95.5%
PROBABILISTIC MATCHING	220	92.7%
SUMMARIZATION	33	88.1%

- STRUCTURED DATA variables had the lowest accuracy via large language model assessment
- SUMMARIZATION variables also had a lower accuracy score



# Next Steps

- **With adequate funding:**
  - **Continue CLAIRE pilot – more cases, more file types, more providers**
  - **Other medical chart reviews for other disease areas**
  - **eCR (Electronic Case Report) use**
- **Large language models (LLMs) combined with deterministic code may be a smart, hybrid approach to medical chart review, and could result in a more efficient use of staff time**
- **LLM use could expand our ability to use medical data directly from providers to better understand population health**



# Thank You!

This project was supported by Cooperative Agreements NU50CK000648-01-00 and NU50CK000508-05-02 from the Centers for Disease Control and Prevention.

**Stephanie Meyer**

*stephanie.meyer@state.mn.us*



# Public Health Generative AI Use Cases



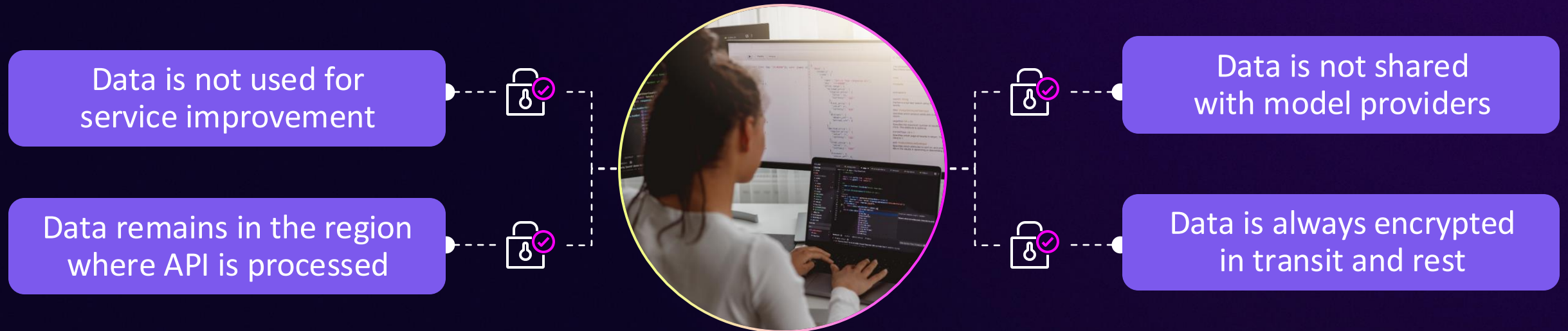
# Agenda – Public Health Use Cases for GenAI

- Electronic Case Reports
- Intelligent Document Processing
- Chatbots
- Generative Business Intelligence and Data Analytics



# Data Privacy & Security

## Securely build generative AI applications with your data



Build with comprehensive data protection & privacy;  
leverage AWS security services and best practices



# Electronic Case Reports

- Problem: Public Health receives thousands of large clinical documents (CDA and HL7) with actionable public health data buried in non-standard data elements, non-standard vocabularies and critical information in free text fields
- Solution: Use Generative AI to ask the questions that public health needs from electronic case reports and extract actionable public health data in a standard format
  - Pregnancy Status (Syphilis in a woman of child bearing age)
  - Occupation (Is the person with Hepatitis A food handler?)



# Electronic Case Reports

```
<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:cda="urn:hl7-org:v3" xmlns:sdctc="urn:hl7-org:sdctc"
xmlns:voc="http://www.lantanagroup.com/voc" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:hl7-org:v3 ../schema/infrastructure/cda/CDA_SDTC.xsd">
  <realmCode code="US"/>
  <typeId extension="POCD_HD000040" root="2.16.840.1.113883.1.3">
    <!-- [C-CDA R1.1] US Realm Header -->
  </typeId>
  <templateId root="2.16.840.1.113883.10.20.22.1.1">
    <!-- [C-CDA R2.1] US Realm Header (V3) -->
  </templateId>
  <templateId extension="2015-08-01" root="2.16.840.1.113883.10.20.22.1.1">
    <!-- [eICR R2 STU1.1] Initial Public Health Case Report Document (eICR) (V2) -->
  </templateId>
  <templateId extension="2016-12-01" root="2.16.840.1.113883.10.20.15.2"/>
  <id root="38e6a983-38ad-484f-a7d2-b294cde5435">
    <!-- Globally unique document ID (extension) is scoped by vendor/software -->
  </id>
  <code code="55751-2" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC" displayName="Public
Report">
    <!-- Document Code -->
  </code>
  <title>Initial Public Health Case Report</title>
  <effectiveTime value="20200505110515-0400"/>
  <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25" displayName="Normal"/>
  <languageCode code="en-US"/>
  <setId extension="8d86218e-0fea-11eb-8216-a80388425cfb" root="1.2.840.114350.1.13.380.3.7.1.1"/>
  <versionNumber value="1"/>
  <recordTarget>
    <!-- ***** recordTarget: The patient ***** -->
  </recordTarget>
  <patientRole>
    <id extension="PT-470127" root="2.16.840.1.113883.19.5">
      <!-- Patient ID -->
    </id>
    <!-- SSN -->
    <id extension="222-22-2222" root="2.16.840.1.113883.4.1"/>
    <addr use="H">
      <streetAddressLine>2222 Home Street</streetAddressLine>
      <city>Sacramento</city>
      <state>CA</state>
      <postalCode>94203</postalCode>
      <county>Sacramento County</county>
      <country>US</country>
    </addr>
    <telecom use="HP" value="tel:555-555-2003"/>
    <telecom use="HP" value="mailto:jose@email.com"/>
    <patient>
      <name use="L">
        <given>Joseph</given>
        <family>Patient</family>
      </name>
      <administrativeGenderCode code="M" codeSystem="2.16.840.1.113883.5.1" displayName="Male"/>
      <birthTime value="19890730"/>
      <sdctc:deceasedInd value="false"/>
      <religiousAffiliationCode code="1013" displayName="Christian (non-Catholic, non-specific)"
codeSystem="2.16.840.1.113883.5.1076" codeSystemName="HL7 Religious Affiliation"/>
      <raceCode code="1002-5" codeSystem="2.16.840.1.113883.6.238" codeSystemName="Race & Ethnicity"
displayName="American Indian or Alaska Native"/>
      <ethnicGroupCode code="2135-2" codeSystem="2.16.840.1.113883.6.238" codeSystemName="Race & Ethnicity - CDC"
displayName="Hispanic or Latino"/>
      <languageCommunication>
        <languageCode code="es"/>
        <preferenceInd value="true"/>
      </languageCommunication>
    </patient>
  </patientRole>
```



```
{
  "patient_identifier": "PT-470127",
  "patient_gender": "Male",
  "travel_details": "Recent travel for vacation to Wuhan China from April 22, 2020 to May 5, 2020",
  "medications_received": "Remdesivir 100 mg Injection, 100 mg every day for 4 days",
  "pregnancy_status": "Not specified",
  "symptoms": "Difficulty Breathing (finding)",
  "symptoms_onset_date": "05/13/2020",
  "occupation": "Sales",
  "does_job_involve_food_respond_yes_no_or_unknown": "Unknown",
  "has_chronic_conditions": "Unknown",
  "has_asthma": "Unknown",
  "diagnoses": "Coronavirus as the cause of diseases classified elsewhere (B97.2)",
  "was_hospitalized": "Yes",
  "date_hospitalized": "05/13/2020",
  "date_discharged": "05/13/2020",
  "is_patient_expired": "No",
  "all_medications": "Remdesivir 100 mg Injection",
  "medications_per_condition": "Remdesivir for Coronavirus",
  "eicr_trigger_data": "Coronavirus as the cause of diseases classified elsewhere (B97.2), SARS coronavirus 2 N gene detected",
  "is_this_person_homeless_or_prisoned": "Unknown"
}
```



# Intelligent Document Processing

- Problem: Public Health receives paper forms across multiple domains including Vital Records, Electronic Laboratory Reports, Communicable Disease Reports, and WIC applications
- Solution: Use Generative AI to automatically extract the data from paper forms and transform it into formats easily consumable by the systems of record.



Mail to: California Department of Public Health  
Immunization Branch  
880 Marina Bay Parkway  
Building 2, 8th Floor, MS 7915  
San Francisco, CA 94134-9412  
Or Fax to: (415) 620-3948

## MEASLES (RUBEOLA) CASE REPORT

### PATIENT DEMOGRAPHICS

Patient name—last first middle initial Date of birth Age (enter age and check one) Months ☐ Years ☒ Gender ☒ Male ☐ Female

Address—number, street City State ZIP code County

Home (916) 555 1212 Work ( ) Email: spongebob@squarepants.com

RACE (check all that apply)

☐ Hispanic/Latino ☐ Asian: Please specify: ☐ Thai ☐ Pacific Islander: Please specify:

☒ Non-Hispanic/Non-Latino ☐ Native American/Alaskan Native ☐ Chinese ☐ Vietnamese ☐ Guamanian

☐ Unknown ☐ White ☐ Korean ☐ Samoan

☐ Other: ☐ Filipino ☐ Laotian ☐ Other Asian: ☐ Other Pacific Islander:

Country of birth USA Country of residence USA

### COMMON LAB TRACKING DATA

CMRID Number U2B Case ID Number WebCMR ID Number

Date reported to county Date investigation started Person/clinician reporting case Reporter telephone

Case investigator completing form Investigator telephone Investigator's jurisdiction

### SIGNS AND SYMPTOMS

Rash ☒ Yes ☐ No ☐ Unknown Rash onset date 6/1/2025 Rash duration 5 days Generalized rash ☒ Yes ☐ No ☐ Unknown Origin on body Chest Direction of spread out

Fever ☒ Yes ☐ No ☐ Unknown Fever onset date 6/1/2025 Was temperature taken ☒ Yes ☐ No ☐ Unknown Was temperature >101°F (38.3°C) ☒ Yes ☐ No ☐ Unknown If temperature not taken, skin was Hot Warm Normal Unknown

Cough ☒ Yes ☐ No ☐ Unknown Runny nose (coryza) ☒ Yes ☐ No ☐ Unknown Conjunctivitis ☒ Yes ☐ No ☐ Unknown Koplik's spots ☒ Yes ☐ No ☐ Unknown

Other symptoms Describe other symptoms ☐ Yes ☒ No ☐ Unknown

Does case meet clinical criteria for further investigation? ☐ Yes ☒ No ☐ Unknown

CASE MEETS CDC/CSTE CLINICAL CRITERIA? (FOR STATE USE ONLY) ☐ Yes ☒ No ☐ Unknown

### COMPLICATIONS AND OTHER SYMPTOMS

Hospitalized Days hospitalized Pneumonia Encephalitis Death If yes, date of death

Other complications If yes, describe other complications

### LABORATORY TESTS

Lab tests done for measles CASE LAB CONFIRMED (FOR LHD USE) CASE LAB CONFIRMED (FOR STATE USE ONLY) LAB RESULT CODES

☒ Yes ☐ No ☐ Unknown ☒ Yes ☐ No ☐ Unknown ☐ Yes ☐ No ☐ Unknown

Serology performed Specimen date Result interpretation

☒ Yes ☐ No ☐ Unknown ☐ P ☐ N ☐ I ☐ E ☐ X ☐ U

IgM ☐ P ☐ N ☐ I ☐ E ☐ X ☐ U

IgG (acute) ☐ P ☐ N ☐ I ☐ E ☐ X ☐ U

IgG (convalescent) ☐ P ☐ N ☐ I ☐ E ☐ X ☐ U

Specimen obtained for virus isolation Specimen source Specimen date Virus isolated

☒ Yes ☐ No ☐ Unknown ☒ Nasopharyngeal ☐ Urine ☐ Other ☐ Unknown 6/2/2025 ☐ Yes ☐ No ☒ Unknown

Specimen sent to CDC for genotyping Date sent Virus genotype

☐ Yes ☐ No ☒ Unknown ☐ Yes ☐ No ☐ Unknown

Other lab tests performed Other lab test specimen date Specify other lab tests Other lab test results

☐ Yes ☐ No ☒ Unknown ☐ Yes ☐ No ☐ Unknown

CDPH 8345 (8/08)

Page 1 of 2

Measles (Rubella) Case Report—CDPH 8256

**VACCINATION/MEDICAL HISTORY**

Received one or more doses of measles containing vaccine (MCV)

Number of doses

☐ Yes ☒ No ☐ Unknown

Dates of vaccination—Dose 1

Dose 2

Dose 3

\_\_\_\_/\_\_\_\_/\_\_\_\_

Reason not vaccinated (check all that apply)

1 ☐ Personal Beliefs Exemption (PBE)

4 ☐ Lab confirmation of previous disease

7 ☐ Delay in starting series or between doses

2 ☐ Permanent Medical Exemption (PME)

5 ☐ MD diagnosis of previous disease

8 ☐ Other

3 ☐ Temporary Medical Exemption

6 ☐ Under age for vaccination

9 ☐ Unknown

Prior MD diagnosed measles (see reason 5)

Pregnant

Immunocompromised

☐ Yes ☐ No ☐ Unknown

☐ Yes ☒ No ☐ Unknown

☐ Yes ☒ No ☐ Unknown

**EPIDEMIOLOGICAL EXPOSURE HISTORY**

Spread Setting (check all that apply)

1 ☐ Day care

4 ☐ Hospital Ward

7 ☐ Home

10 ☐ College

13 ☐ Church

2 ☐ School

5 ☐ Hospital ER

8 ☐ Work

11 ☐ Military

14 ☐ International travel

3 ☐ Doctor's office

6 ☐ Outpatient hospital clinic

9 ☐ Unknown

12 ☐ Correctional facility

15 ☐ Other

Recent travel or arrival from other country or state within 18 days of rash onset? ☐ Yes ☐ No ☐ Unknown

Countries or states visited

Dates in countries or states visited

Date of arrival in California

\_\_\_\_/\_\_\_\_/\_\_\_\_

Close contact with person(s) with rash 8-17 days before rash onset? ☒ Yes ☐ No ☐ Unknown

	Name	Rash onset date	Relationship	Age (Years)	Same household
1	Sandy Squirel	6/3/2015	friend	13	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Unknown
2					<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown
3					<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown

Please list other contacts on a separate sheet or use the contact tracing work sheet.

Epi-linked to a lab-confirmed case?

☒ Yes ☐ No ☐ Unknown

Import status

☒ Indigenous ☐ Out-of-state import

☐ International import

If case is indigenous, is case

☐ Import-linked (linked to imported case) ☐ Endemic ☒ Unknown Source

☐ Imported virus (viral genetic evidence indicates an imported genotype)

Outbreak location

Bikin Bottom

If case is imported, describe source

**CONTACT INVESTIGATION**

Spread Setting (check all that apply)

1 ☐ Day care

4 ☐ Hospital Ward

7 ☐ Home

10 ☐ College

13 ☐ Church

2 ☐ School

5 ☐ Hospital ER

8 ☐ Work

11 ☐ Military

14 ☐ International travel

3 ☐ Doctor's office

6 ☐ Outpatient hospital clinic

9 ☐ Unknown

12 ☐ Correctional facility

15 ☐ Other

Number of susceptible contacts

Close contacts who have rash 8-17 days after exposure to case (list below)

☐ Yes ☐ No ☐ Unknown

	Name	Rash onset date	Relationship	Age (Years)	Same household
1	Eugene Krebs	6/4/2015	co-worker	25	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Unknown
2					<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown
3					<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown

Please list other contacts on a separate sheet or use the contact tracing work sheet.

**CASE CLASSIFICATION (FOR LHD USE)**

☒ Confirmed ☐ Probable ☐ Suspect ☐ Not a case ☐ Unknown

**CASE CLASSIFICATION (FOR STATE USE ONLY)**

☐ Confirmed ☐ Probable ☐ Suspect ☐ Not a case ☐ Unknown

**MEASLES CASE DEFINITION**

Clinical case definition: An illness characterized by all the



# Generative AI Chatbots

- Problem: Public Health departments receive multiple inquiries. Providers using Vital Records and Immunization Information Systems call asking basic questions about how to use these systems. The general public may also call Public Health with simple questions about how to obtain an immunization record or how to enroll in a program like WIC. Although this information is on the Public Health
- Solution: Using Generative AI, resources from public health like user manuals and public facing websites can be the knowledge bases to enable chatbots to automatically answer questions from end users and the general public.



# AI Chatbots

## Communicable Disease Epi on Call - Guide to Surveillance, Reporting and Control

Interactive Generative AI Assistant

### Welcome to the GenAI Epi Assistant

This is a test chatbot to test GenAI to answer questions from content that can be found in the following

#### Reference Pages

- [Guide to Surveillance, Reporting and Control](#)
- [Clinical and Laboratory Testing Guidance for Monkeypox](#)

#### How to Use the Assistant

Our AI documentation assistant is available in the chat window. You can:

- Ask specific questions about the content in the reference pages
- Request information about specific topics
- Get help understanding complex information
- Find relevant sections in the documentation





### Epi on Call

Customer has joined the chat

BOT

11:25 AM

Welcome to the Self Service Chatbot. How may I assist you with inquiry?

**B** *I*    

What do I do if I find a bat in my room?



End chat

 Start a Call





# Generative Business Intelligence

- Problem: Public Health staff do not always have the technical and analytic skills to interact with their data to generate the graphs and stories from their siloed systems to help improve population health.
- Solution: Generative BI enables business users to interact with their data using natural language prompts to build graphs, dashboards and stories that describe their data



# AI-assisted storytelling

## Interpret data for others

Help others derive meaning from data and reach conclusions to drive decisions

## Generate stories using AI

Produce cohesive, powerful, and insightful narratives

## Create refined content

Control AI verbosity, customize narrative text, and apply stunning visual themes to bring content to life

## Augment with unstructured data

Upload files with related content to enrich the story with additional insights

## Share up-to-date governed data

Quickly update and disseminate data at any time

The screenshot displays the Amazon QuickSight 'Build story' interface. The main dashboard, titled 'Interests Analysis', features a radar chart labeled 'Revenue by primary interest and source'. The chart has 12 axes representing different interest categories: Unspecified, Travel, Technology, Sports, Outdoor, Home and, Health, Food and Dining, Finance, Fashion, Entertainment, and Wellness. The chart shows varying levels of revenue across these categories, with 'Unspecified' and 'Travel' showing higher values. To the right of the chart, a text block explains that health, beauty/wellness, and travel generated significant lead revenues (\$46,113, \$31,041, and an unspecified amount respectively), suggesting these interests have a significant impact on performance and that marketing campaigns should target these high-value segments. It also notes that arts/culture, automotive, and unspecified interests attracted lower revenues, indicating they may not align as well with campaign objectives or the target demographic. The text concludes by stating that leveraging customer interest data can help refine marketing strategies and drive stronger returns.

The 'Build story' sidebar on the right contains the following sections:

- Describe your data story**: A text box with the prompt: 'Build a story about marketing campaign performance over time. Describe top campaigns and account managers. How can we improve overall campaign success?'
- Select visuals**: A section with an '+ ADD' button and three selected visual types: 'Total impressions', 'Monthly impressions', and 'Most effective impressions'. A fourth option, 'Comparing revenue', is also visible.
- Select documents (optional)**: A section with a dashed box containing the text: 'Drag a file here or [upload a file](#)'. Below this, it states: 'Up to 5 files can be uploaded in .pdf, .txt, .rtf, .docx, .ppt, or similar formats, with a limit of ~10MB each.' Two document thumbnails are shown: 'Marketing Team Meeting Minutes.docx' and 'Campaign Effectiveness Analysis.docx'. A checkbox labeled 'Use insights from Amazon Q Business' is checked.
- BUILD**: A blue button to generate the story, with a 'Learn more' link next to it.



# Generative Business Intelligence

## Build visuals

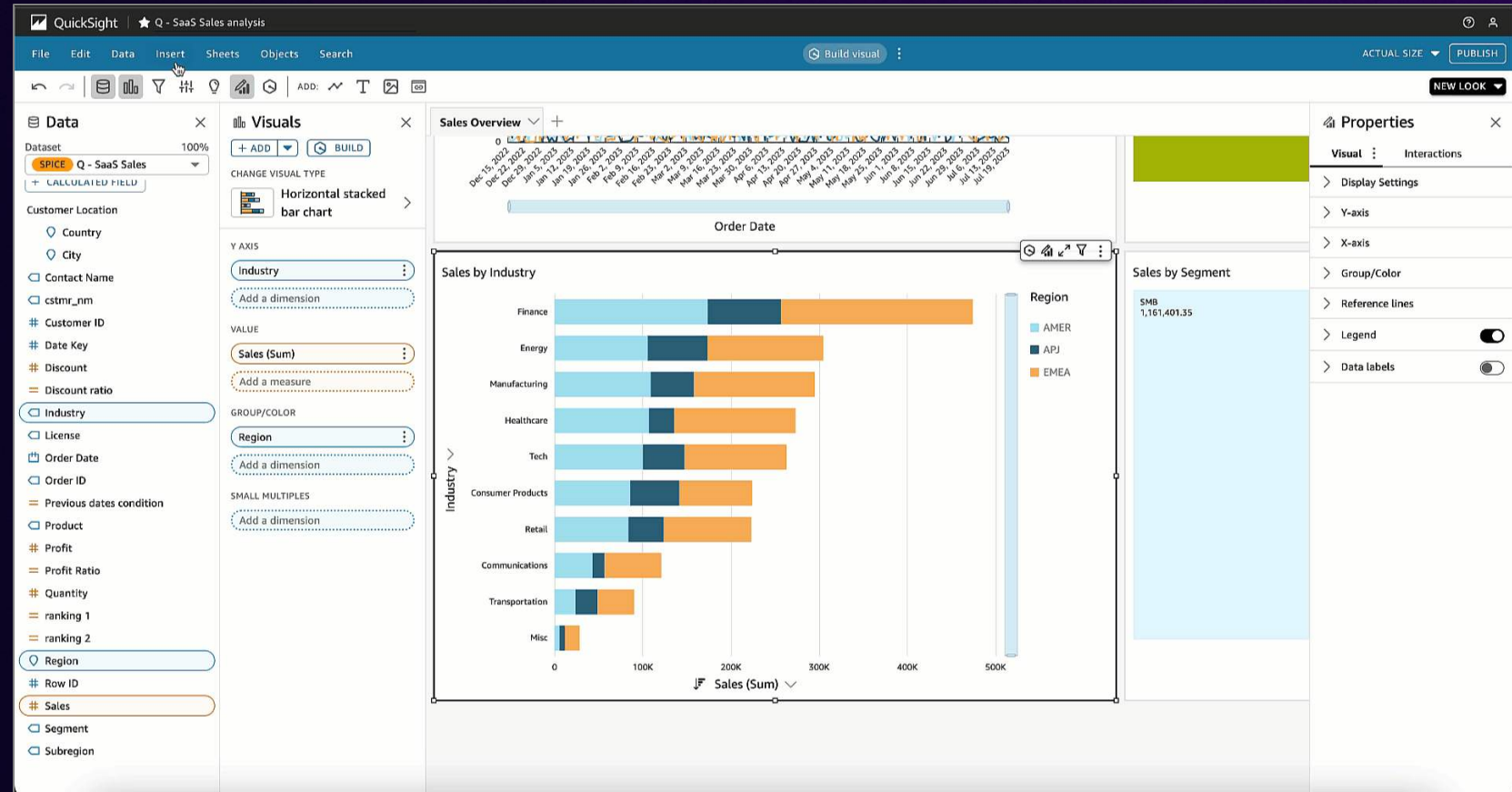
Use natural language to quickly build visuals for dashboards and reports

## Build calculations

Easily create calculations using natural language without looking up or learning specific syntax

## Refine visuals

Quickly update visuals by describing desired formats using natural language







**Blog: Transforming electronic case reporting with generative AI: Unlocking faster public health responses**



**Blog: Scalable intelligent document processing using Amazon Bedrock**





# Thank you!





# Audience Q & A





# Thank You!

## UP NEXT...

**5-6pm Central Time** | Join us for our *Data Modernization in Motion Networking Reception* –  
Park View, 4<sup>th</sup> Floor

## TOMORROW...

**8:30-9:30am Central Time** | Hashbrowns & High Tech  
Peer Networking Breakfast

**9:45am Central Time** | Concurrent sessions continue,  
followed by preconference closing plenary