AI + DS +Healthcare Projects -

Exploring healthcare gaps through a data-driven patient-led lens

USHASREE JAKILINKI –ASPIRING HEALTH DATA SCIENTIST

Clinical Decision Support System (CDSS) –Medication Allergy Safety (EHR plugin/Webpage/Mobile App)

PROJECT -2

Who am I -TBD

My name is Ushasree Jakilinki (https://www.linkedin.com/in/ushasreejakilinki/). I bring a robust foundation of over two decades in software development and technical leadership across the Automotive and Mobile sectors, complemented by recent experience as a Software Program Manager. I hold a Bachelor's degree in Computer Science and Engineering, and I have been accepted into the UofM Master's of Applied Data Science program, starting in August 2025. My long-term aspiration is to pursue a Ph.D. in AI and Healthcare. My dedication to healthcare innovation stems from firsthand encounters with systemic challenges in patient safety, chronic disease diagnosis, and critical care delivery – experiences that have profoundly shaped my belief that AI and data science are uniquely positioned to address these gaps. Beyond healthcare, my passions include education (ignited by witnessing its transformative power firsthand) and leveraging my project management experience to drive complex data science initiatives. My core technical background also includes continued interest in the Automotive sector.

Who am I -TBD

In March 2025, I was suddenly struck by severe abdominal pain so debilitating that I could barely move. I had experienced similar reactions before when consuming dairy, but this time the pain was relentless and I could not figure out what I was doing in terms of food. I tried every diet imaginable— liquids diet, elimination diet, vegan, documenting every single thing I put in my mouth—yet nothing brought relief. I sought help from gastroenterologists, allergists, my PCP, and integrative care specialists. Tests – Endoscopy,MRI confirmed internal inflammation, but no one could explain why. They gave me more meds for inflammation and Pepcid for Gerd.The medications they prescribed only made me worse. The only option I was offered was symptom management.
The breakthrough came not from my doctors, but from a brief comment by a pharmacist: *"It could be your meds."* That night, I pored over every prescription I was taking, cross-checking both active and inactive ingredients. To my shock, I discovered that three of my new medications—and one I had been taking for years—contained lactose. While I had tolerated small amounts before, the cumulative exposure overwhelmed my system and

Who am I -TBD

What horrified me most wasn't just the pain, the trauma in a way but the systemic failure. My lactose anaphylaxis was clearly documented in all my medical records – MyChart, pharmacy databases but my doctors, nor the pharmacists had flagged the risk. When I asked why, one doctor admitted it wasn't her expertise; another said she couldn't research alternatives and that is outside her job scope, one local pharmacist said call the manufacturer in short each individual kept passing the ball, and I suffered. I must have spent at least 2-3 weeks just calling people, each call lasting more than 3 hours at a time.The real solution turned out to be simple but exhausting

- It was my job to debug myself in a way to figure out the problem with constant pain
- I had to identify lactose-free formulations myself—for example, replacing AbbVie's Levothyroxine (which contains lactose) with Pfizer's Levoxyl (which does not).
- I had to work with my doctor's office, get prescriptions, have them approved by insurance and coordinate with multiple providers to get the medication dispensed.

Who am I -TBD

The data was always there. The system simply failed to connect it. My experience revealed a larger truth:

- patients are often left to navigate dangerous gaps between EHRs, prescribers, pharmacists, and insurers.
- Until we build healthcare systems that integrate and act on the full scope of patient data—including inactive ingredients—others will continue to suffer needlessly.
- Most allergy people are good at reading labels but medication that too prescription labels are beyond the scope and knowledge of a normal person. To my knowledge the ingredients are not part of the medication bottles yet.
- I am an educated adult, but think if this was a child, a baby, a senior, a blind person – it would be classified as “Idiopathic” or colic or some such thing and symptom management is the only way out.

Why –

Why Allergy Safety Matters: August 12th 2025

- Food allergies affect around 33 million adults and children in the United States, with cow's milk, eggs, wheat, soy, peanuts, tree nuts, fish, shellfish, and sesame being the most common food allergens.[1]
- Each year, around 3.4 million adults and children seek emergency care for a severe allergic reaction related to food. This means that every 10 seconds a food allergy reaction sends a patient to the emergency room. [2]
- The Cost of Food Allergies: A $25B Epidemic.[3]
- Food allergies are rising at an alarming rate.[4,11]
  - **Overall rise:** The Centers for Disease Control and Prevention (CDC) reported that food allergies in children increased by 50% between 1997 and 2011.
  - **Continued growth:** A report from FoodAllergy.org states that childhood food allergy prevalence increased by another 50% between 2007 and 2021.
  - **Specific allergens:** The prevalence of self-reported peanut or tree nut allergies in U.S. children more than tripled between 1997 and 2008.

- Up to 60% of adult anaphylaxis cases are historically classified as idiopathic.[10]

Why –

Trends in Food Allergies and Allergic Syndromes

- **General Food Allergies**
  - Prevalence in children rose by **50% (1997–2011)** [CDC]
  - Increased another **50% between 2007–2021** [FoodAllergy.org]
  - Peanut & tree nut allergies in US children **more than tripled (1997–2008)**
- **Celiac Disease**
  - **Incidence rising ~7.5% annually** (Western countries, 2020 meta-analysis)
  - Partly due to better diagnostics/awareness
  - **6.4-fold increase** in Scottish children (1990–2009)
- **Alpha-gal Syndrome (AGS)**
  - First US cases: early 2000s; **rapid recent growth**
  - **110,000+ suspected cases (2010–2022); up to 450,000 possibly undiagnosed** (CDC)
  - Annual increase in positive lab tests since 2017

Why –Food allergens in medications

- According to [5] , **Sample 42,052 oral medications**

Ingredient	% occurrence in medications
Lactose/Milk/Casein [5]	44.86%
Corn starch/Wheat starch [5]	36.58%
With Gluten[6]	52.60%
With soy[6]	20.60%

- According to [6] sample size: 108

- Risks of dairy derived excipients in medications for lactose intolerant and cow milk protein allergic patients – PubMed TBD
- Egg TBD
- Gelatin TBD –Alpha girl

Current solutions

- **ADINA** act-The ADINA Act exists because of Adina Togal's frightening experience with an unlisted allergen in her antibiotic. Her story brought national attention to the risk for countless others, leading to a legislative push to require clear allergen labeling in medications—just as is already required for food.[9]

A middle schooler from Maple Grove, Minnesota, living with celiac disease and a severe dairy allergy.

While at summer camp, Adina suffered a **severe allergic reaction after taking an amoxicillin pill**. The reaction included vomiting, loss of consciousness, dangerously low heart rate, and severe abdominal pain.

Her mother, Jennifer Togal, tried to get an ingredient list from the drug manufacturer but encountered red tape: the company would only provide details with a physician's letter, and even then, the information was unclear or incomplete.

The process of getting allergen information took weeks, and, disturbingly, **medication ingredient lists are not required to disclose potential food allergens the way foods are.**

Medication labels are not legally required to list major food allergens (like gluten, milk, peanuts, egg, etc.). This makes it hard and sometimes dangerous for people with celiac disease or food allergies to know what's in their prescription or OTC medications.

Adina's reaction was one of her worst ever—prompting the realization that many other patients could face similar hidden dangers.

Current solutions –LLM solution

Step	What to Do	Details/Examples	Blockages/problems in following these steps
1. Know Risks	Be aware of common med allergens	Lactose, wheat/gluten, soy, peanut, egg, gelatin, fish, sesame	Great, most allergy sufferers are
2. Read Labels	Check “inactive ingredients”/excipients on each medication’s label or insert	Watch for names like starch, lecithin, arachis oil	As of 2025, labels on prescription medications do not exist
3. Ask Questions	Explicitly tell/ask your pharmacist and doctor about allergens	“Is the lactose dairy-based?” “Where does the starch come from?”	Who will answer these questions ? Doctors? Pharmascits ? Manufacturers ?Google ?Dailymed ?
4. Use Resources	Consult FDA Inactive Ingredient Database & advocacy orgs for up-to-date excipient info	FDA IIG, FARE, Beyond Celiac	Great but these databases have older data, they keep getting updated, and the med that one is taking might not be what daily med or openfda has
**5. Clarify Unclear	Contact manufacturer directly (via pharmacist if possible)	Get written confirmation if source is ambiguous	From who ?Manufacturer ? Pharmascits ?
6. Switch Safely	Be extra careful with generics/brand switches—compounds can change	Re-check with each new script or supplier	Where? How ?
7. Consider Compounding	Look into custom compounding for essential meds if standard forms can’t be made allergen-free	Ask your allergist and pharmacist	These can be very expensive, not covered by insurance and simply put put of reach to most common people
8. Emergency Ready	Carry needed emergency meds; wear a medical alert bracelet	Epinephrine auto-injector, antihistamines, action plan	Always
9. Report Problems	Report any suspected reaction to FDA’s MedWatch and your healthcare team	Increases safety for you and others	Great but these things take time

Current solutions

- But why should we do all the above manually when we are so technically advanced ? Can we create a technical solution to fix this problem ?—A system that is a Mobile or Web or part of a EHR system will notify when there is a allergen in the (hopefully) to be prescribed drug ??

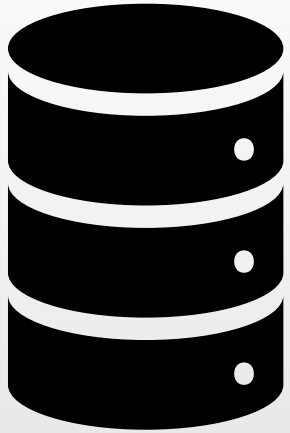
What is CDSS

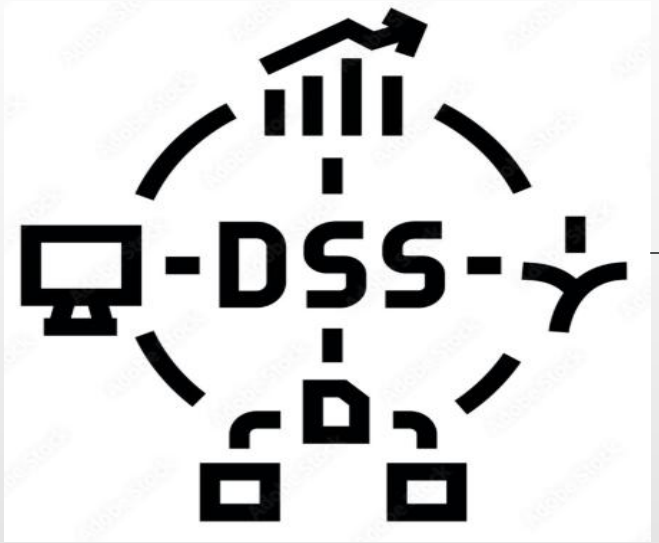
That's why I want to design a **Clinical Decision Support System (CDSS)** — a plugin for electronic health records (EHRs), a mobile app, or web-based tool that:

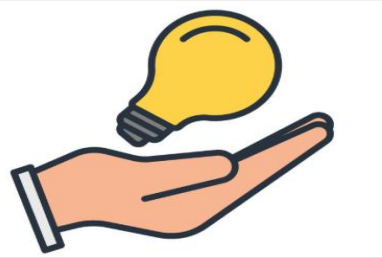
- **Flags known allergens** (including both active and inactive ingredients) based on a patient's documented allergy profile
- **Pulls accurate ingredient data** from trusted sources like DailyMed or RxNorm, OpenFDA (and hopefully one day official pharmacy data bases with version control)
- **Automatically alerts** prescribers and pharmacists — before harm occurs
- **Empowers patients** to verify medications independently, with a tool that is accessible and trustworthy

This solution is for anyone with food-based allergies. It's for caregivers, doctors, and pharmacists who want to do the right thing but don't have the tools. It's for every patient who deserves safe care without fear of allergic harm — especially when it's preventable.

User flow







Input –Allergen + Med name

Drug ingredients DB

Decision Support System(Rules Engine +Logic)

Safe/Unsafe/Suggestion

Is the safe suggestion supported by insurance ? Does the pharmacy have it in stock ?
Should we talk to the prescriber about this change or should it be automatic ?
Is this a Data science problem or data extraction and mapping problem ?? Where/how does AI/DS/ML come into picture ? Maybe while recommending new allergen free medication but is it ?? Should we preprocess for a user/patient based on his allergens and medication/diseases so output is faster ??
Should we calculate how much allergen per medication or multiple medications ?? This is definitely a later feature, but

Support or help

What do I need for this project ?	From whom
Pharmacy Knowledge –Official database access	Pharmacy school or Pharmacy database
Allergy knowledge	https://medresearch.umich.edu/labs-departments/centers/food-allergy
End user testing/System testing	Allergy patients, BeyondCeliac, Fare
Legal issues if independent Android or webapp. This should be used a initial startpoint rather than the authority	
Per	

Phases of project

Milestone	Title	Description	Key Skills/Tech
1	**Foundational Prototype-Epic**	Build a proof-of-concept with hardcoded data for 2 drugs (*Trazodone, Levothyroxine*), covering 5 doses, 5 manufacturers, active & inactive ingredients. Predefined allergen list (*Dairy*).	Python, Requests, JSON, REST APIs (mock), Data wrangling
2	**User-Facing MVP**	Expand to 10 drugs. Add a basic web interface for patient inputs (Streamlit). Allow users to select allergens and view flagged results. Give user a alternative option if drug is unsafe	Streamlit, Python, UI/UX principles
3	**Real-World Data Integration**	Replace hardcoded data with APIs & databases: DailyMed, FDA SPL. Scale drug dataset significantly. Build pipeline to manage this data.	API calls, Data pipeline, Database management
	Drug	Enhance functionality to also check for potential drug–drug interactions based	

Phases of project- **Foundational Prototype**

CPMAI Phase	What needs to be done(Foundational Prototype level)
Phase I: Business Understanding – Feature under Epic	MVP goal: Demonstrate feasibility of detecting allergens in 2 common medications for patients with known food allergies. Stakeholders : Patients & Caregivers Success criteria: *System correctly identifies allergens and gives SAFE –UNSAFE feedback*
Phase II: Data Understanding	Explore and document the hardcoded dataset: list of drugs, their doses, manufacturers, active/inactive ingredients, allergens of concern. Understand the structure & potential gaps.
Phase III: Data Preparation	Prepare a clean hardcoded dataset (e.g., JSON or CSV). Map allergens to ingredients for those 2 drugs. Validate data manually for correctness.
Phase IV: Modeling	Write the Python logic to: take user allergen input, take drug name input, check ingredients against allergens, and flag matches.
Phase V: Evaluation	Test your prototype against predefined test cases: e.g., patient allergic to lactose checks Trazodone, system flags it. Record results & confirm correct behavior.
Phase VI: Operationalization	Deliver a runnable prototype (e.g., Jupyter notebook or CLI Python script) + documentation. Show demo to a test user (me). Record feedback.

~~Phase V Evaluation Metrics for MVP~~

To ensure our Foundational Prototype is successful, we will measure:

- **Accuracy of Allergen Detection (%)**
  - *What it measures:* How often the system correctly identifies the presence or absence of a known allergen in a drug's ingredients for our hardcoded test cases.
  - *Goal:* Aim for 100% accuracy for the initial hardcoded dataset.
- **Precision/Recall/F1 (if applicable)**
  - *What they measure:*
    - **Precision:** Of all the times the system *flagged* an allergen, how many were actually correct? (Minimizing false positives).
    - **Recall:** Of all the *actual* allergens present, how many did the system correctly *flag*? (Minimizing false negatives – crucial for safety).
    - **F1-Score:** A balance between Precision and Recall.
  - *Goal:* High precision and recall, especially for safety-critical alerts.
- **Response Time**
  - *What it measures:* How quickly the system processes a drug and allergy input and returns a result.
  - *Goal:* Near-instantaneous response for a smooth user experience (e.g., < 1 second).

~~What – Features and Roadmap~~

- **In your CDSS Capstone "Product":**
  - **Features:** "Hidden lactose detection," "Cross-referencing with patient allergy list," "Risk level flagging," "Pharmacist override option."
  - **Roadmap:** "Phase 1: MVP for hidden lactose in oral medications. Phase 2: Expand to other hidden allergens (e.g., gluten, soy, protien). Phase 3: Integrate with specific pharmacy systems for real-time alerts. Phase 4: Incorporate advanced NLP for unstructured data beyond just inactive ingredients."

Thinking about your capstone with this "product mindset" will be incredibly valuable for your AI Product Manager aspirations!

CPMAI –AI Pattern

CDSS is primarily
a: Recognition System
- Its core AI task is to identify known allergens in drug ingredients.

Its primary value is: Decision Support
- It empowers users to make safer medication choices based on the recognition output.

Why this distinction matters for an AI PM:
- **Data Strategy:** Focus on acquiring and labeling comprehensive data for *known allergens* and *ingredient lists*.
- **Model Selection:** Rule-based systems (for MVP) or classification models are appropriate.
- **Evaluation Metrics:** Prioritize precision, recall, and F1-score for allergen detection.
- **User Experience:** Design clear, actionable alerts that effectively support user decisions.
- **Ethical Considerations:** Ensure robust accuracy and fairness in allergen identification across all patient demographics.

Data Challenges & Applied Data Science Solutions Mastering the Data Landscape

Key Challenges:

- Heterogeneous Sources: Pulling data from DailyMed, RxNorm, OpenFDA, FDA SPLs (varying formats: JSON, XML, unstructured text).
- Inconsistent Naming/IDs: Standardizing drug and ingredient names across databases.
- Identifying Hidden Excipients: Parsing active vs. inactive ingredients, especially for unlisted hidden components (e.g., lactose source from excipient lists).
- Data Volume: Managing and processing millions of drug records and their ingredients efficiently.

My Applied Data Science Solutions:

- Robust Data Pipelines: Developed modular Python scripts using requests, JSON parsing, XML parsing, regex for text extraction.
- Comprehensive Data Cleaning: Implemented techniques for standardization, de-duplication, and error handling.
- Schema Design: Designed efficient data models for storing drug, ingredient, and allergy information.
- Rule Engine Design: Built a precise rule-based system for allergy flagging, accounting for aliases and ingredient hierarchies.

Visual: A flowchart showing data flowing from various sources into your cleaning/processing pipeline, then into a structured database.

Subject Matter Experts (SMEs) testing :Pharmacists, allergist, patients knowledge testing in a way

~~When – Schedule -TBD~~

The **Schedule** answers the **"WHEN?"** questions.
It's all about the timing of delivery and execution.

•**What questions does it answer?**

- **"When will this product/feature be ready?"** (The estimated completion or launch date.)
- **"When will users be able to access it?"** (The release date.)
- **"What's the sequence of development?"** (What needs to be built first, second, etc., due to dependencies.)
- **"How long will each phase or feature take to develop?"** (Estimates for design, development, testing.)
- **"Are we on track to meet our deadlines?"** (Progress tracking against the plan.)
- **"What are the key milestones and delivery dates?"** (Important checkpoints for stakeholders.)
- **"When can we iterate or add more functionality?"** (Planning for future versions.)

•**In your CDSS Capstone "Product":** Your schedule would answer: "When will the prototype for hidden allergen detection be functional?", "When can I demonstrate the initial integration with DailyMed?", "When will the report generation feature be ready for review?"

Risks & Mitigation
Navigating Project Complexity TBD

Risk: Perceived Insufficiency of Core Data Science & Scope Management for Capstone.

Elaboration for presentation: Given the non-ML focus, there's a risk the project might be seen as primarily data engineering. Additionally, the vast scope of 'all drugs' could lead to a shallow implementation.

Mitigation Strategy:

Deep Dive on Data Engineering & Quality: Emphasizing the rigorous design of robust data pipelines, sophisticated data cleaning methodologies, and comprehensive data validation frameworks as core Applied Data Science contributions.

Strategic Scope Management: For the capstone, focusing on a deep, high-quality implementation for a representative subset of data/drugs, while designing the system to be scalable to the full dataset (demonstrating the architecture for large-scale data handling).

Sophisticated Rule-Based Logic: Highlighting the complexity of the allergy-matching rules, including handling ingredient aliases and complex chemical structures.

Future ML Roadmap: Laying a solid data foundation for future Machine Learning applications (e.g., personalized risk assessment, NLP for unstructured data) as opportunities beyond the capstone scope.

Visual: A simple risk/mitigation icon or a before/after diagram illustrating the refined focus.

Project Status & Key Accomplishments
Progress Towards a Safer Future TBD

Focus on the phases you've *completed* or are *currently focusing on* for the capstone.

Risks

ROAM Analysis - Capstone Scope Risk

R - Risks:

•**R1: Insufficient Core DS/ML Demonstration:** The project's initial phases (data integration, extraction, rule-based matching of ingredients to allergens) may be perceived as primarily **data engineering and rule-based logic**, rather than demonstrating advanced predictive modeling, statistical inference, or core machine learning algorithms typically expected of a Master of Applied Data Science (MADS) capstone.

- **Impact:** Project may not meet the MADS program's criteria for a capstone, potentially requiring significant re-scoping or additional work.

O - Opportunities:

•**O1: Future DS/ML Integration:** The foundational data work in early phases creates a robust platform for **future integration of advanced ML/DS components**, such as:

- **Drug Interaction Prediction:** Leveraging ML to identify complex, non-obvious drug-drug or drug-allergy interactions.
- **Personalized Risk Assessment:** Using patient data and ML to predict individual likelihood/severity of reactions.
- **Natural Language Processing (NLP):** Extracting nuanced allergen information from unstructured text in drug labels or patient notes.

•**O2: Practical Impact:** Even in its rule-based form, the system addresses a critical, real-world patient safety gap, providing a clear and valuable application of data.

A - Assumptions:

Future work & Broader impact

Headline: Beyond the Capstone: Expanding Impact

Future Enhancements (Opportunities from your RAM analysis):

- Full integration with all DailyMed/FDA data.
- Advanced drug-drug and drug-allergy interaction prediction (potential for ML here).
- Personalized risk assessment.
- NLP for unstructured drug label data.
- Integration with EHRs (MyChart, Walgreens, CVS) for real-time alerts.

Broader Impact:

- Enhanced Patient Safety: Preventing adverse drug reactions and anaphylaxis.
- Empowered Patients: Providing a tool for independent medication verification.
- Improved Clinical Workflow: Reducing manual cross-checking for healthcare providers.
- Addressing a National Health Gap: Contributing to transparency and safety in medication labeling.
- Visual: A vision-board style graphic or icons representing different aspects of future growth.

Visual: A vision-board style graphic or icons representing different aspects of future growth.

Q&A/Thank You

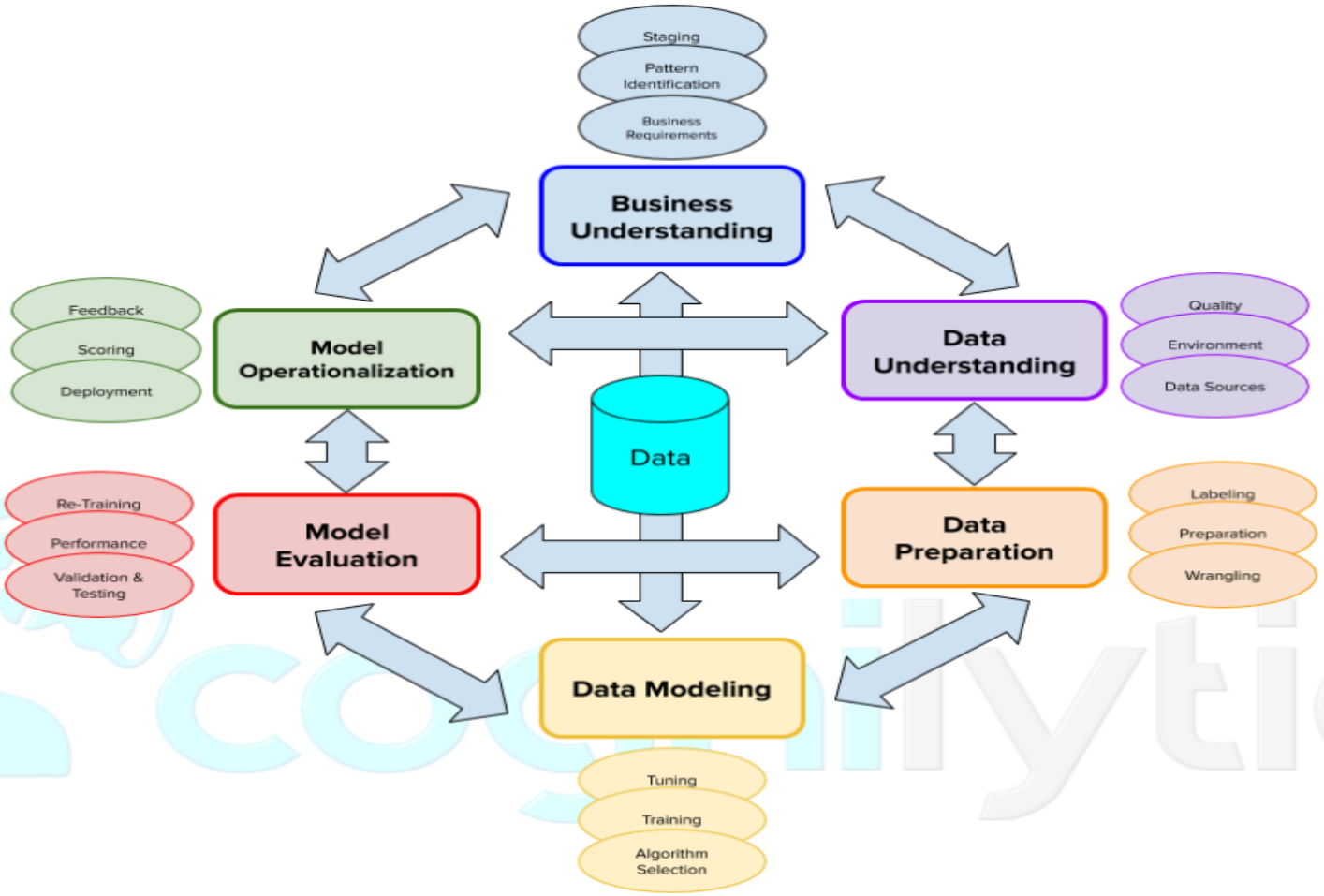
Questions ?
Thank you message
UofM logo

Potential Partners

- https://www.beyondceliac.org/living-with-celiac-disease/gluten-in-medication/
- https://www.foodallergy.org/

How – process CPMAI

Figure: CPMAI Phases Overview



Why –cont…

- Prevalence:
  - Approximately 6.2% of adults and 5.8% of children in the U.S. have food allergies. (2021)
  - Up to 50% of medication allergies go unreported or undiagnosed, significantly increasing adverse reaction risks.
- Emergency Room Visits & Hospitalizations:
  - Food allergy reactions lead to approximately 200,000 emergency room visits annually. [4]
  - Anaphylaxis, a severe allergic reaction, results in an estimated **45,000-50,000 emergency room visits annually** in the U.S.
  - Around 30,000 food allergy-related hospitalizations occur each year.
- Severity:
  - Over 40% of children with food allergies have experienced a severe allergic reaction, such as anaphylaxis.
- Economic Burden:
  - Caring for children with food allergies costs U.S. families nearly $25 billion annually (adjusted to $33 billion in 2024 dollars).
- The Hidden Danger:
  - Hidden allergens, particularly inactive ingredients (excipients), in medications are a significant and often overlooked cause of allergic reactions. For example, lactose is found in nearly 45% of medications. [9] [10]

Input from Seth –adina act TBD

https://celiac.org/gluten-free-living/gluten-in-medicine-vitamins-and-supplements/
https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2036.2008.03889.x
https://news.mit.edu/2019/inactive-ingredients-reactions-study-0313

https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2036.2008.03889.x

https://news.mit.edu/2019/inactive-ingredients-reactions-study-0313

https://www.jacionline.org/article/S0091-6749(14)00432-1/pdf
https://pubmed.ncbi.nlm.nih.gov/24878443/
https://news.mit.edu/2019/inactive-ingredients-reactions-study-0313
https://www.jacionline.org/article/S0091-6749(14)00432-1/pdf
https://pubmed.ncbi.nlm.nih.gov/24878443/
https://publications.aap.org/pediatricsinreview/article-abstract/27/3/118/76013/Egg-based-Vaccines?redirectedFrom=fulltext
https://www.cdc.gov/flu/vaccine-process/?CDC_AAref_Val=https://www.cdc.gov/flu/prevent/how-fluvaccine-made.htm

Research papers

1. Food Allergies Statistics and Facts You Should Know
2. FARE Food Allergy Facts and Statistics_April2024.pdf
3. The Cost of Food Allergies: A $25B Epidemic — Prevent Food Allergies
4. Epidemiology and the Growing Epidemic of Food Allergy in Children and Adults Across the Globe | Current Allergy and Asthma Reports
5. 'Inactive' ingredients in oral medications – PMC
6. Hidden Allergens Found in Common Medications – EMJ
7. Risks of dairy derived excipients in medications for lactose intolerant and cow milk protein allergic patients – PMC
8. Presence of gluten and soy derived excipients in medicinal products and their implications on allergen safety and labeling – PMC
9. ADINA Act — inspired by Minnesota girl's allergy scare — would label drug ingredients | FOX 9 Minneapolis-St. Paul
10. **When allergies have no name: is idiopathic anaphylaxis driven by co-factors?** https://pmc.ncbi.nlm.nih.gov/articles/PMC11527779/
11. **Trends in Allergic Conditions Among Children: United States, 1997–2011** https://www.cdc.gov/nchs/products/databriefs/db121.htm
12. Time trends in the epidemiology of food allergy in England: an observational analysis of Clinical Practice Research Datalink data **https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(24)00163-4/fulltext**