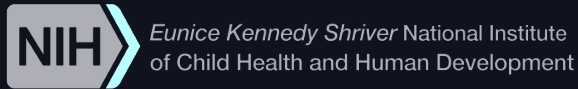


# Layered Intelligence: Generative AI Meets Classical Decision Sciences



**Danielle Heymann<sup>1</sup>**  
**June 12, 2024**

---

Contributors: Sheran Law<sup>1</sup>, Nancy Zhou<sup>1</sup>  
<sup>1</sup>National Institutes of Health



# CONTENTS

---

Introduction

---

Beyond Conversational AI

---

Generative AI in Topic Modeling

---

Predictive Analytics in Action

---

Enhanced Classification Algorithms with AI Agents

---

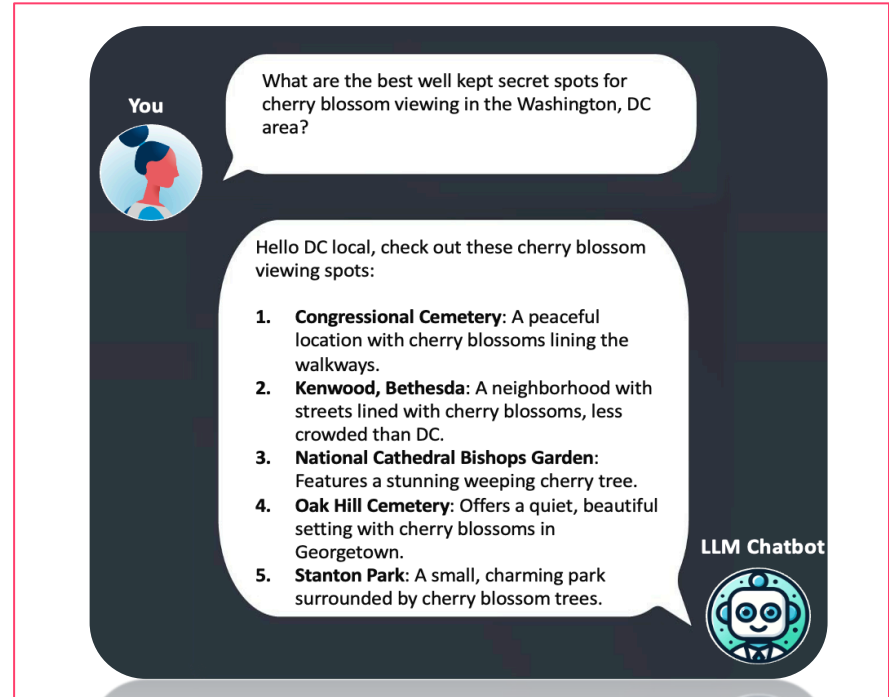
Conclusion and Discussion

# INTRODUCTION

## The Evolution Beyond Chat

While LLM-based chatbots have been revolutionary in changing how we interact with information daily, their true potential is far beyond.

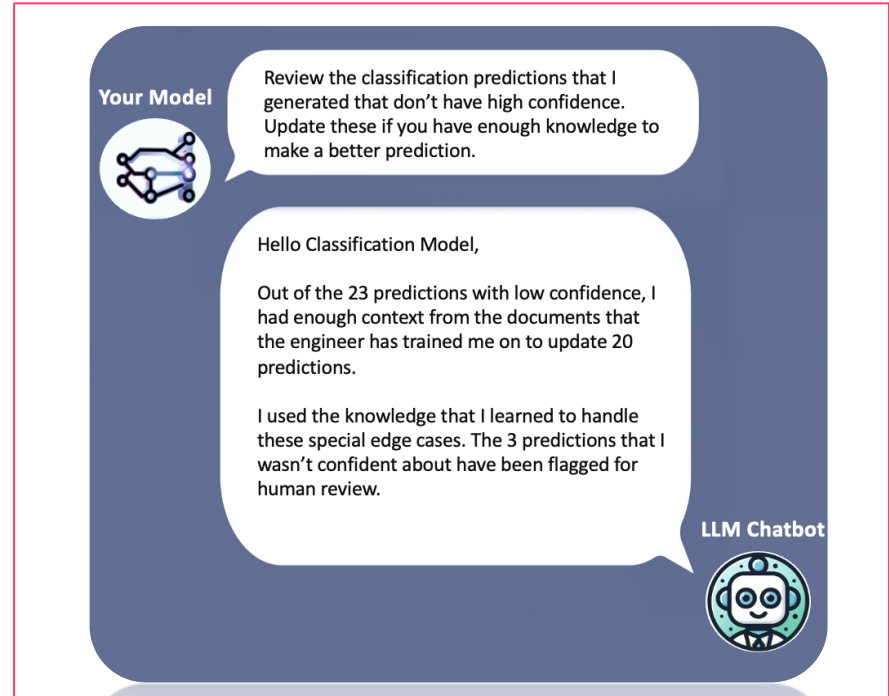
In decision sciences and analytics, generative AI agents can go beyond answering questions; they can process data, predict trends, and develop into sophisticated analytical engines



# INTRODUCTION

## Why AI Agents are Essential in Modern Analytics

- Dissect complex data
- Reveal underlying patterns and forecast future scenarios
- Transform language to logic
- Scale ability to interpret vast datasets
- Handle challenging edge cases
- Provide review and quality assurance (QA)



The image shows a chat interface with a dark blue background. On the left, there is a circular icon labeled "Your Model" containing a neural network diagram. On the right, there is a circular icon labeled "LLM Chatbot" containing a stylized robot head. Two white speech bubbles are shown. The top bubble, from "Your Model", says: "Review the classification predictions that I generated that don't have high confidence. Update these if you have enough knowledge to make a better prediction." The bottom bubble, from the "LLM Chatbot", says: "Hello Classification Model, Out of the 23 predictions with low confidence, I had enough context from the documents that the engineer has trained me on to update 20 predictions. I used the knowledge that I learned to handle these special edge cases. The 3 predictions that I wasn't confident about have been flagged for human review."



# GEN AI: OVERCOMING ANALYTICS BOTTLENECKS

We will explore how generative AI can address these common bottlenecks



## Volume

Overwhelmed by volume of data and required preprocessing and cleaning



## Surface Insights

Conventional analytics usually provide insights that lack depth in data interpretation and context



## Rigidity

Inflexibility in adapting to new or changing data and system logic



## Time

From data processing to evaluation and review, the traditional pipeline is timely

## The role of Generative AI

---

Streamline data handling, enhancing speed and efficiency

Delve deeper to uncover more challenging patterns and discoveries

Dynamically help with adjustments and evolving

Accelerate processing time and conduct QA



# GENERATIVE AI IN LDA TOPIC MODELING

Demo with NIH RePORTER Data

*How can we uncover hidden themes within NIH RePORTER's data and generate coherent, descriptive labels for each theme?*

# TOPIC MODELING

## Enhanced by Large Language Models

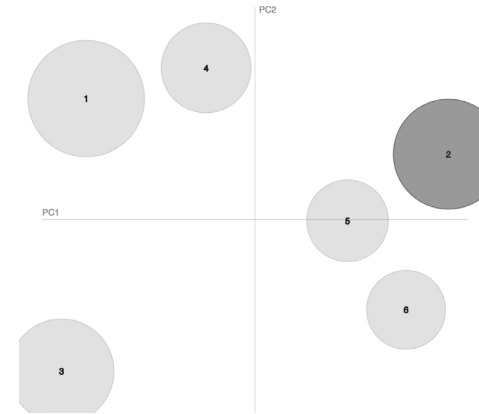
Topic modeling uses natural language processing methods to uncover hidden patterns and insights in data

Techniques like Latent Dirichlet Allocation (LDA), a probabilistic method, can be used to identify clusters within datasets, **but often leave these clusters unlabeled**

By integrating Generative AI, particularly Large Language Models (LLMs), we can enhance this process. A Gen AI agent can generate meaningful labels for the identified clusters, adding the missing context and clarity

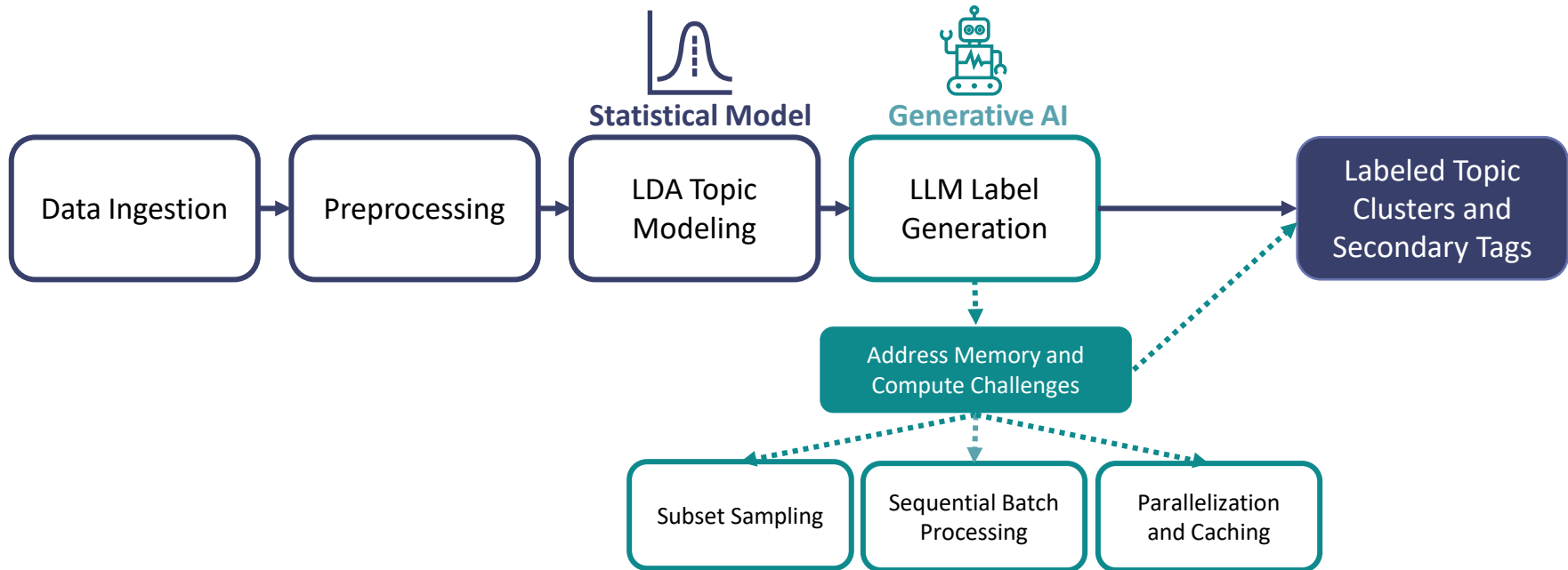
**A Gen AI agent transforms the traditional LDA results, turning clusters with no topic labels into actionable insights.** This is just one way to enhance descriptive modeling techniques with a Gen AI integration

Intertopic Distance Map (via multidimensional scaling)



# TOPIC MODELING WITH LDA AND LLMS

## A Framework





# DEMO

## Topic Modeling with LDA and LLMs

Dataset: NIH Reporter abstracts from National Institutes of Health (NIH) funded projects.

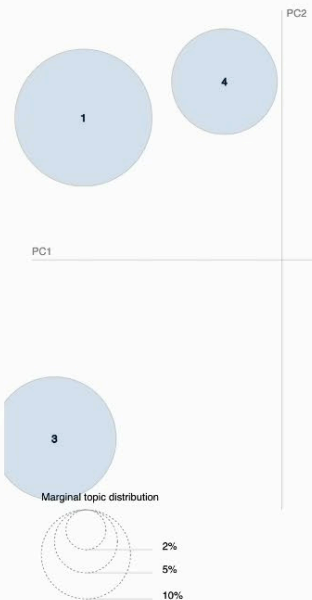
Source: <https://reporter.nih.gov/search/NU7NRcgedUO4EMr8itHS1A/projects>

Search Constraints: Fiscal Year: Active Projects; Admin: Yes; Agency/Institute/Center: NICHD; Activity Code: R01  
Equivalents; Project Start Date: On or After: 4/1/2023 (until 4/1/2024)"

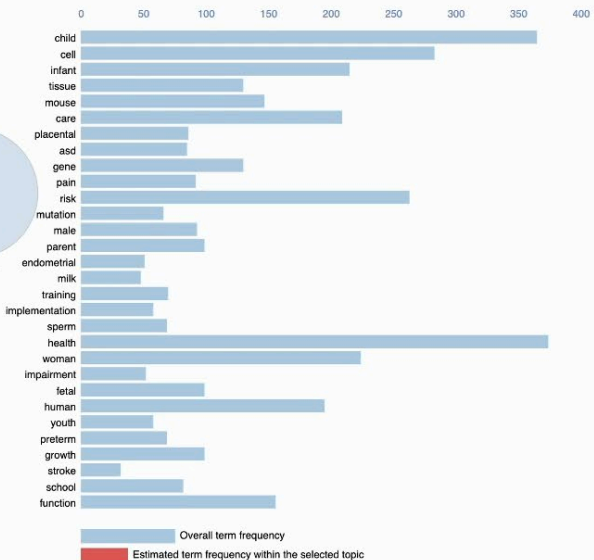
Selected Topic:

Slide to adjust relevance metric:<sup>(2)</sup>   $\lambda = 1$

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Salient Terms<sup>1</sup>



1. saliency(term w) = frequency(w) \* [sum\_t p(t|w) \* log(p(t|w)/p(t))] for topics t; see Chuang et. al (2012)  
 2. relevance(term w | topic t) =  $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$ ; see Sievert & Shirley (2014)

```
topic_labels = []
i = 0
for topic, topic_abstracts in zip(topic_words, topic_topics_from_LDA):
    topic_abstracts = ' '.join(topic_abstracts)
    response = client.chat.completions.create(
        model="gpt-3.5-turbo",
        messages=[{"role": "system",
                    "content": "You are a label gene",
                    {"role": "user", "content": "Generate a label for this topic and you should also reference the topic abstracts."}
                ]
    )
    label = response.choices[0].message.content
    print(f'Topic {i}: {label}')
    topic_labels.append(label)
    i += 1

# visualize the topics with the labels
prepared_data = pyLDAvis.lda_model.prepare(lda, doc_term_df, topic_labels)
pyLDAvis.display(prepared_data)
```

[66] 0.7s Python



# RANDOM FOREST CLASSIFICATION MODEL PROTOTYPE

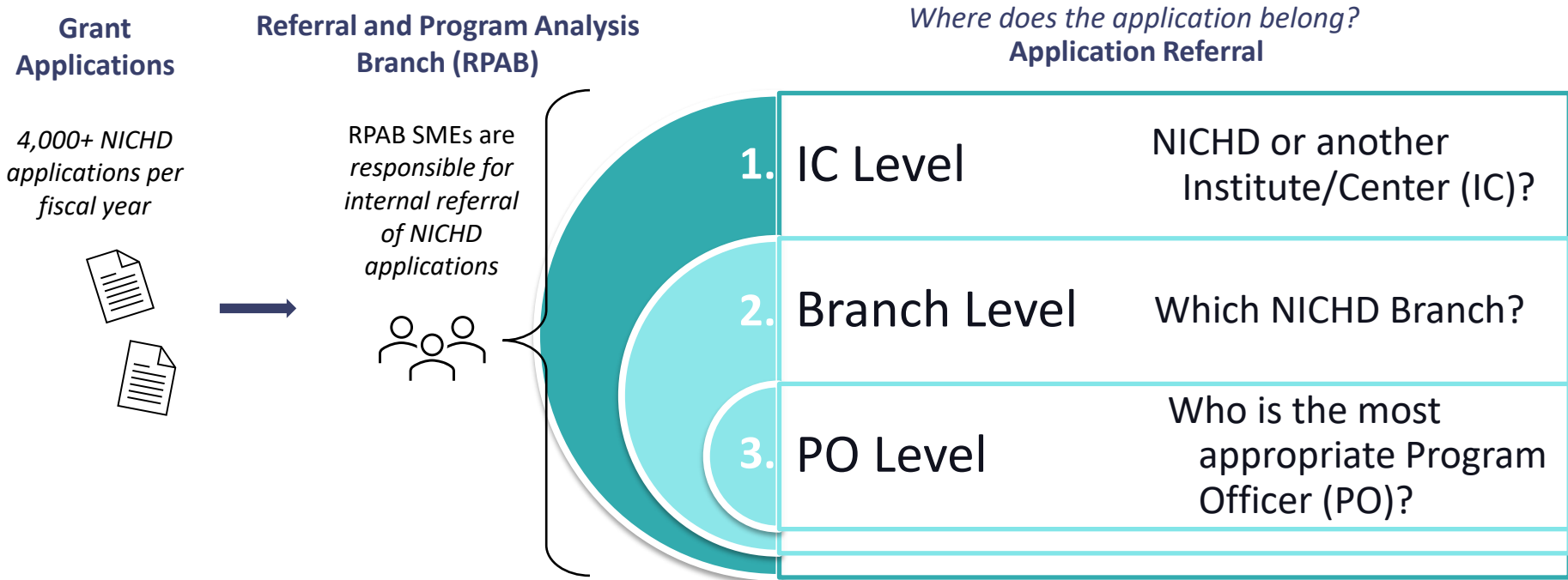
CLASSIFYING SCIENTIFIC BRANCHES OF GRANT  
APPLICATIONS AT NIH NICHD

NICHD RPAB Use Case

*How can an NIH Institute be more efficient in referring >3000 applications annually while maintaining accuracy?*

# GRANT APPLICATION REFERRAL PROCESS OVERVIEW

## NIH National Institute of Child Health and Human Development (NICHD)



# NICHD APPLICATION REFERRAL CHALLENGES

At NIH NICHD, our Referral and Program Analysis Branch (RPAB) AI/ML application referral system project is modernizing the grant application referral process.

The primary aim is to develop a semi-automated referral system to reduce the burden associated with the existing manual referral process. For example, this involves empowering our system to generate initial recommendations regarding the best match between an application and a branch along with more specific branch and program codes.

## The Challenge



### Time

Application referral is a time-consuming process for a high volume of applications.

In FY 2023, RPAB manually referred **4415** new applications



### Expertise

Requires considerable scientific expertise to accurately match applications to **one of 13 appropriate scientific branches and center**



### Complexity

Complex overlapping of scientific domains and policy-related considerations within the extramural branches.

# AI/ML APPROACH FOR NICHD REFERRALS

---

## **Solution Overview**

Developing an RPAB AI/ML Referral System prototype to streamline NICHD referral processes. This approach involves developing and refining advanced algorithms and NLP techniques to semi-automate decision making and enhance data-driven insights.

---

## **Key Components**

Data preprocessing, feature engineering, development of high performing classification models, analysis of evaluation metrics, user acceptance testing and review

---

## **Benefits**

Our AI/ML solution can deliver tangible benefits, including faster referral, enhanced accuracy, reduction in manual errors and administrative burden, real-time insights for improved decision-making

---

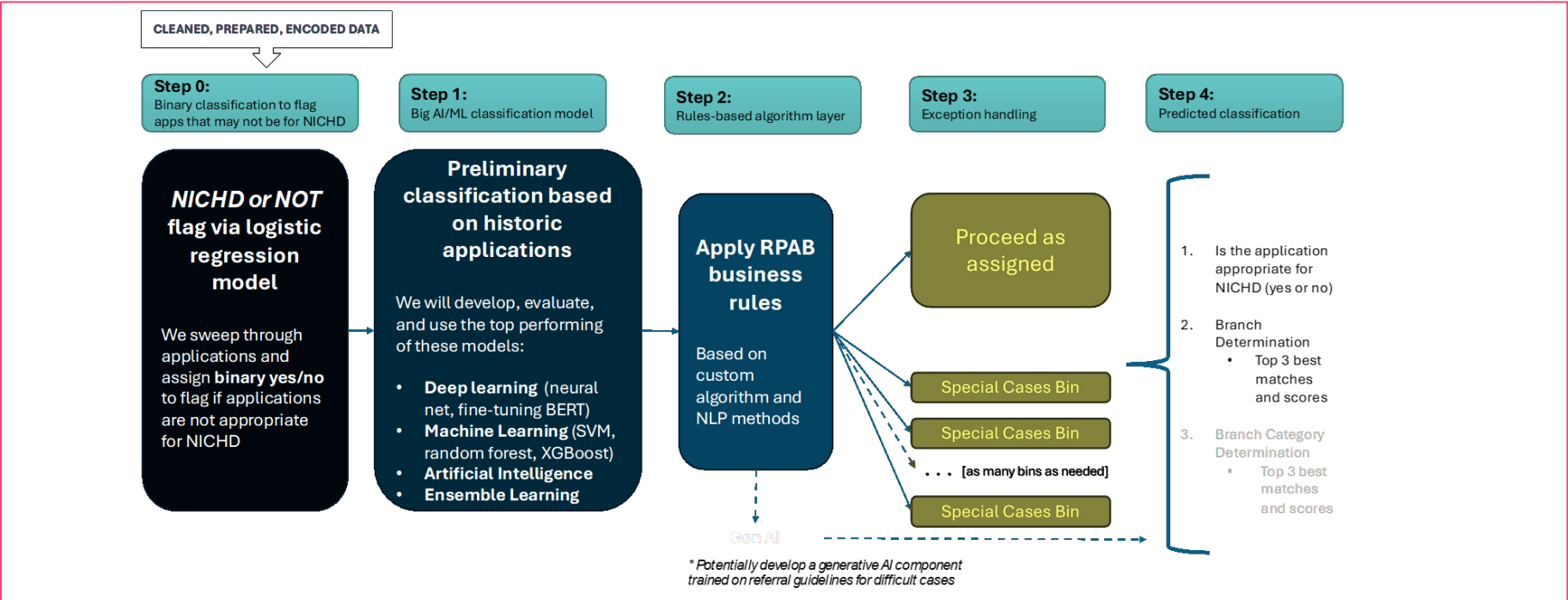
## **Scalability and Adaptability**

Designed to adapt to evolving scientific landscapes and changing NICHD research priorities, our solution ensures flexibility and scalability to meet future requirements and expand as needed.

---

# MODEL DEVELOPMENT AND INTEGRATION

## RPAB AI/ML Referral System: a multilayered system design



# PREDICTING GRANT APPLICATION CLASSIFICATION

## At NICHD RPAB

### Prototype Model Overview

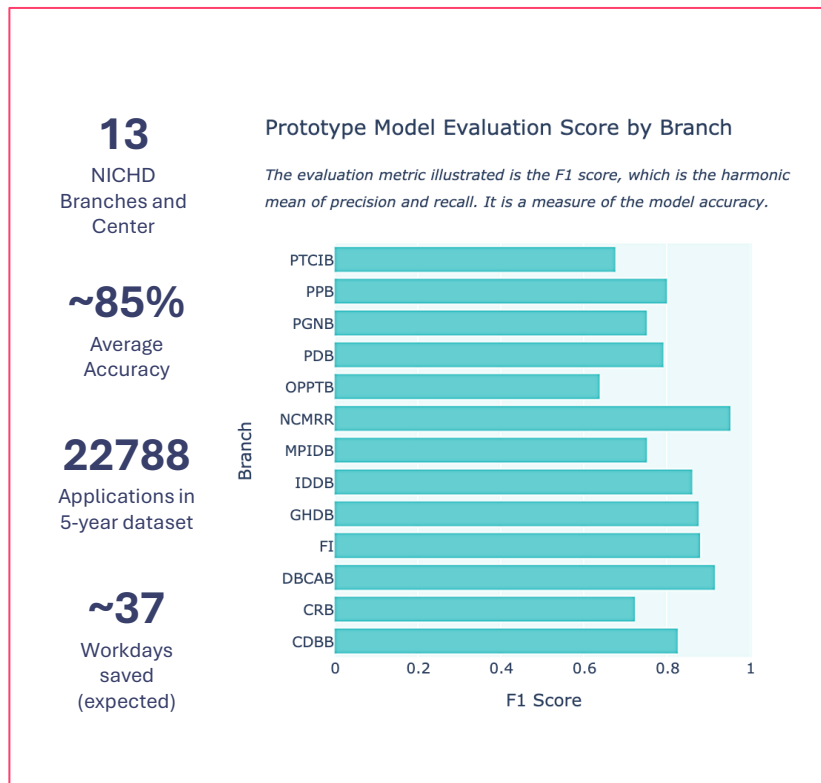
- Random forest classification algorithm with notable success to predict a grant application's scientific branch within NICHD
- Aimed at enhancing RPAB decision-making processes

### Model Confidence and Future Development

- High initial success rates
  - Achieved accuracies ranging from 64-95% for the various branches
  - Average accuracy of approximately 85%
- Future Development: System refinement by incorporating RPAB SME business rule logic to enhance performance; exploration of neural network classification model with Generative AI QA Bot

### Expected ROI

- If the AI/ML system saved just 20% of the time, this is estimated to be 4/20 minutes of time saved for manual application referral. With ~4400 applications per year, this adds up to **~37 workdays per fiscal year.**





# NEURAL NETWORK CLASSIFICATION WITH GENERATIVE AI QA BOT

A FRAMEWORK AND PROTOTYPE DEMO

Demo with NIH RePORTER Data

*How can we boost production-ready accuracy and confidence in classification of complex NIH RePORTER data?*

# REFINING UNCERTAINTY IN MACHINE LEARNING

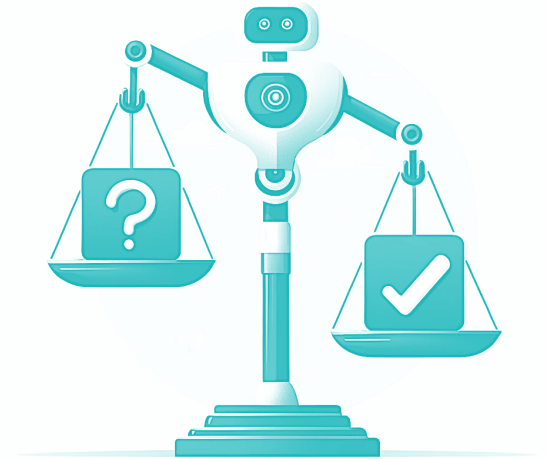
## Building on Machine Learning Foundations

Whether using neural networks or random forest, every classifier encounters predictions with low confidence.

### The Generative AI QA Bot

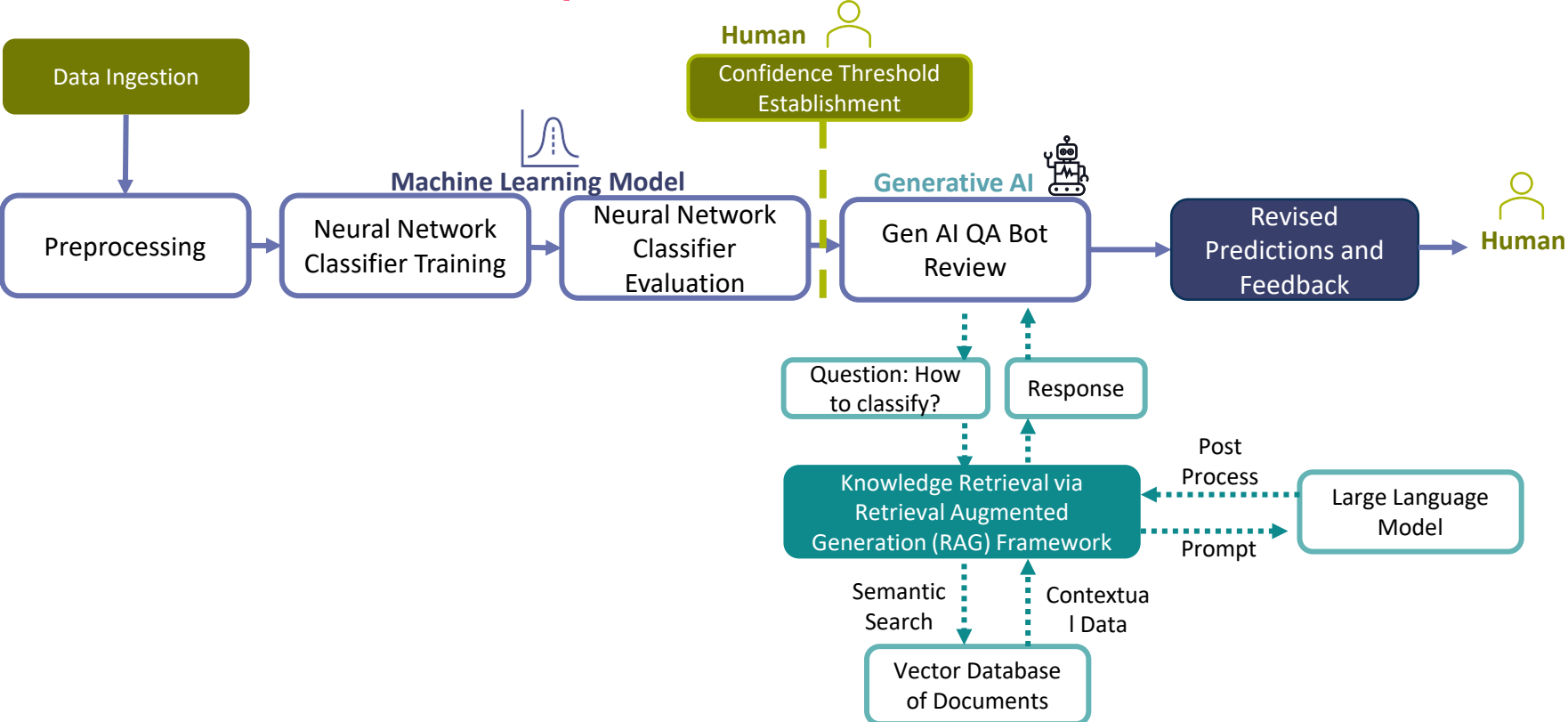
By embedding a Generative AI QA Bot in a model system, it is possible to:

- Review and refine uncertain outcomes at scale
- Improve the model's reliability and trust
- Provide the system with access to external context for larger knowledge base
- Provide additional notes for feedback to human-in-the-loop



# NEURAL NETWORK CLASSIFICATION + GEN AI

## A Framework with Gen AI QA Bot



# DEMO

## Neural Network Classification with Gen AI QA Bot

Dataset: NIH Reporter abstracts from National Institutes of Health (NIH) funded research projects during 2024. The most recent 15000 funded projects are used.

Source: <https://reporter.nih.gov/search/p8bbCVlgUEqsYZwl9yRufg/projects>

Search Constraints: Search Criteria Fiscal Year: 2024Activity Code: Research Projects

# Interactive Gen AI QA Bot Demonstration

## Predicting the most relevant NIH Institute or Center (IC)

This application predicts which NIH Institute or Center is most relevant for a grant application based on the project details provided.

### Neural Network + Gen AI QA Bot

Scenario Overview ▼

*NIH RePORTER public data is used for this demonstration.*

## ● Reviewing Low Confidence Prediction #1

### Project Details and Predictions

#### Project Information

The Imprinted Gene Network in the programming of Non-Alcoholic Fatty Liver Disease by early life cadmium exposure

#### Neural Network Prediction

Confidence Score

0.45

Predicted IC: National Institute of General Medical Sciences (NIGMS)



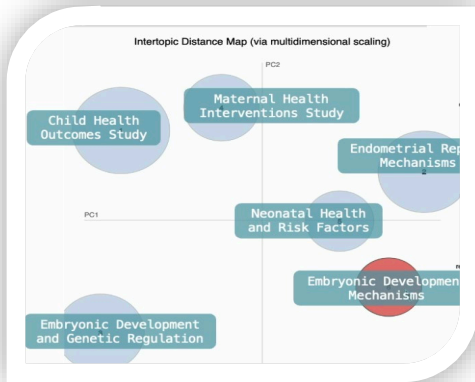
# CONCLUSION



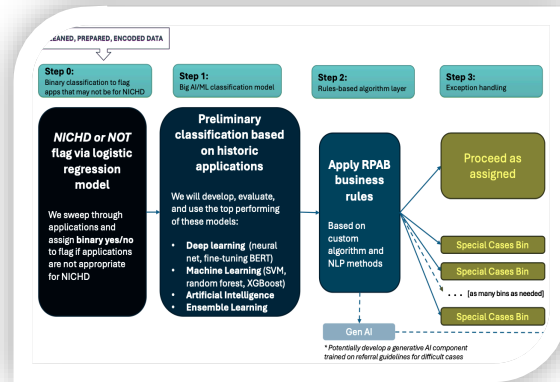
# CONCLUSION

We've explored the application of generative AI in topic modeling, the NIH RPAB AI/ML referral prototype, and neural network classification with a QA bot

## Generative AI in Topic Modeling



## RPAB AI/ML Application Referral Prototype



## Neural Network Model with Gen AI QA Bot

Reviewing Low Confidence Prediction #2

**Project Details and Predictions**

**Project Information**

**Body Composition and Energy Expenditure in Youth with Spina Bifida**

Project Number: 3R01HD096885-05S1

Abstract: The project abstract focuses on the prevalence of obesity in youth with spina bifida (SB) and aims to accurately measure body composition and energy expenditure in this population. The study's outcome will provide a method for accurately measuring these parameters in youth with SB.

Public Health Relevance: The public health relevance of this project lies in addressing the issue of obesity in youth with SB. By accurately measuring body composition and energy expenditure, the study can contribute to developing effective interventions and strategies to prevent and manage obesity in this population.

**Gen AI QA Bot Insight**

Generate Insight

Gen AI QA Bot Insight: This research project aims to investigate the prevalence of obesity in youth with spina bifida (SB) and develop accurate methods for measuring body composition and energy expenditure in this population. The study's focus on addressing obesity in youth with SB is of significant public health relevance.

**Neural Network Prediction**

Confidence Score: 0.40

Predicted IC: National Heart, Lung, and Blood Institute (NHLBI)

**Ground Truth Administering IC:**

National Institute of Child Health and Human Development (NICHD)

Success: The Gen AI QA Bot's suggestion aligns with the ground truth.



# REFERENCES AND RESOURCES

## Software, Images, Data

- Software and packages used in prototypes and demos
  - Data preprocessing: pandas, numpy, re, nltk
  - Topic modeling: gensim, pyLDAvis
  - Generative AI: Hugging Face Transformers, OpenAI GPT-3.5, LlamaIndex
  - Model training: TensorFlow, Keras, scikit-learn
  - Web app development: Streamlit
- Images generated by DALL-E 3
- Data from NIH RePORTER





# THANK YOU!

## QUESTIONS AND DISCUSSION



# DATA+AI SUMMIT

