Unlocking Opportunities

Leveraging Generative AI in University-Sponsored Programs Operations | March 2024

Lisa A. Wilson, Interim AVP Research Administration Optimization, ORAO, ORA
Tubal Yisrael, Project Support Specialist, ORAO, ORA
Alex Wagner, Director of Research Analytics, RDA, ORA

Today's Discussion Topics…

- **The State, Challenges, and Solutions of RA Post-Pandemic**
- **AI Use Cases Across Industries**
- **AI in Research Administration Use Cases**
- **Emory's ORA Private PoC Chatbot**
  - Phase 1 Use Cases
  - **Internal build:** architecture, development platform, project team
  - **External build:** commercial prototype
  - Associated costs, lessons learned, build or buy decision
- **Phase Two: The Case for an Enterprise-level Solution**
- **Data Analytics**
  - **Metrics**
  - **Technology Adoption Model (TAM)**
  - **Change Management**
- **Q&A**

Office of Research Administration & Optimization

The Current State of Research Administration Operations at Research Universities

Operational Challenges: Data Analytics and AI-Powered Solutions

General Challenges Managing Research Administration Operations

Capacity

Recruiting and retention

Engagement

Procedures, operations

Onboarding

Training

Data Analytics

Emory's Proactive Approach to Addressing the Challenges

Emory's Challenges

- **Building Capacity:**
  - FY23 Goal: $1B in new awards, cont., supps
  - FY24 Goal: $1B sponsored research expenditures
  - Supporting growth with existing resources
- **Recruitment and Retention**
- **Staff Engagement**
- **Operations:** outdated workflows and SOPs
- **Data Analytics:** Utilization
- **Onboarding and Training**: 180 + days

Emory's Solutions

- **Recruiting & Retaining:** offered fully remote, salary adjustments, promotions, and competitive salaries.
- **Engagement:** two in-person retreats/yr.
- **Operations**: Dragon Team (NASA Tiger Team) - continuous process improvements
- **Analytics Hub:** tracking and analyzing RA data for informed decision-making
- **Revitalized Training:** strategies for shortening the time to full portfolio management and filling the skills gap
- **AI-powered solutions:** 24/7 instantaneous answers to policy and procedure queries. SOP generation, prompt engineering training. ***An ORA Chatbot.***

Chatbots Use Cases Across Industries

Retail

website bots offer personalized product recommendations

Healthcare

Babylon Health's chatbot offers medical consultations based on personal medical history.

Banking

BoA's virtual assistant offers financial advice using NLP, and Vanguard now has a digital advisor

Higher Ed

GT professor built a platform that manages the high volume of questions from students in an online course, in 2016.

Office of Research Administration & Optimization

Virtual Teaching Assistant

Jill Watson

https://youtu.be/WbCguICyfTA

Office of Research Administration & Optimization

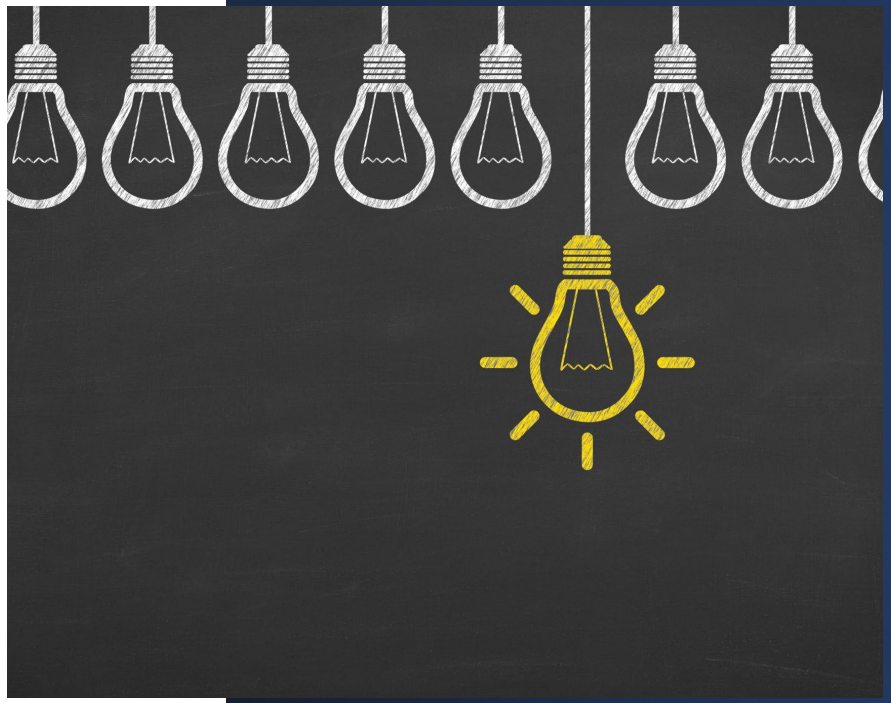
AI in Research Administration

Emory's ORA ChatGPT PoC Project

Phase One: Use Cases, Architecture, and Build or Buy Decision

AI in Research Administration @ Emory University

Imagine every research administrator having access to a digital assistant that comprehensively understands how we conduct research administration at Emory. This assistant could promptly answer pre- and post-award task-related policies, procedures, and systems inquiries and generate new content in seconds.



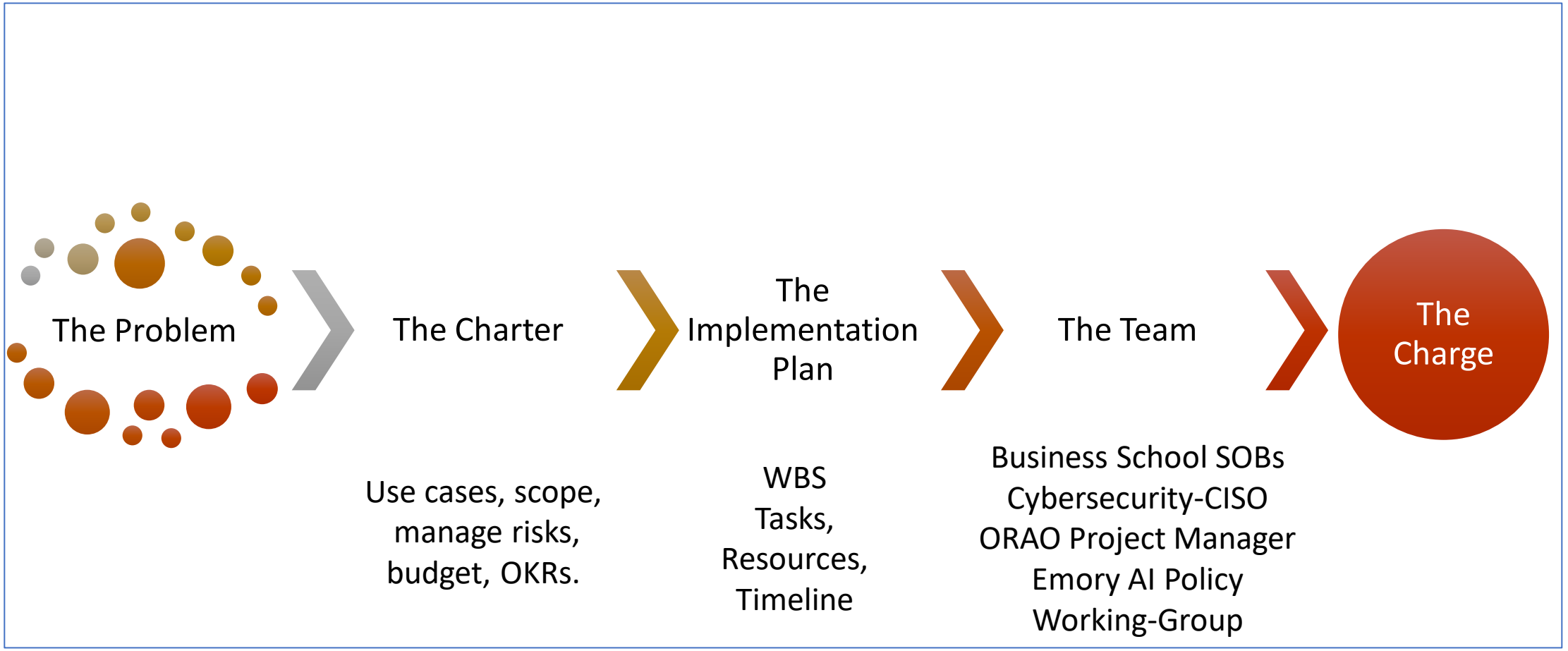
Now, imagine giving all university principal investigators access to an even more robust, responsible, trustworthy research administration digital assistant about a year later.

That’s our big, lofty, bodacious goal in ORA at Emory. Here’s how we began.

Pathway to Build or Buy

The Internal Build

The Business Case



Office of Research Administration
Chatbot Proof-of-Concept
Project Charter

Project Name: ORA Large Language Model (LLM) and Chatbot Proof-of-Concept

Project Co-Leads: Davion Johnson, Strategic Operations Associate, SOT (Strategic Operations and Training); Tubal Yisrael, Project Support Specialist, SOT

Project Sponsor: Lisa Wilson, Sr. Director, Strategic Operations & Training, ORA

Project Collaborators: Benn Konsynski, Professor, Goizueta Business School, and Students of Benn (SOBs)

Charter Prepared By: Lisa Wilson, Sr. Director, SOT

Project Manager/Lead Responsibilities

The project lead and manager are responsible for the following tasks:

1. Ensuring alignment with the Charter's scope of work
2. Leading the development of the project action plan
3. Facilitating the rollout of the project plan

Purpose

This project will develop an ORA Chatbot using a large language model, drawing from the ORA knowledge repository. The aim is to validate the feasibility and benefits of integrating an LLM-backed chatbot into ORA operations. The Chatbot will provide virtual assistant services that will quickly answer research administration queries, aid in updating and drafting policies, guidelines, SOPs, and training resources, reduce time on tasks, enhance staff productivity, and streamline efficiency.

The Need

1. New hires say it takes too long to get answers to unit, role, or task-specific questions.
2. The learning curve for systems and tools should be shorter.
3. The learning curve is steep and long, more than 180 days (more than six months)
4. It takes too long to update standard operating procedure manuals.
5. It takes too long to update or author new policies.
6. The greatest need is RAS, specifically RAS PEDS.

Use Cases

Real-world cases where the ORA LLM-backed Chatbot can be employed include the following:

1. Virtual Assistant – to respond instantly to the unit, role, task, or system-specific queries.
2. Document generation or content creation – create original text to develop new SOPs (Standard Operation Procedures), policies, training materials, and job aids.
3. Productivity – significantly reduces the time it takes to perform tasks and increases efficiency.
4. Proof-of-concept for build or buy decision.

Deliverables

1. An ORA Knowledge Repository – all SOPs, process maps, policies, job aids, and training materials.
2. An ORA ChatGPT-like virtual assistant and AI text generation tool.
3. An AI Panel concurrent session at RW23.
4. A prompting engineering Master Class for research administrators.
5. Implementation Plan in PPM (Project Portfolio Management) Pro.
6. New RA operations business model.

Benefits

1. Continuous Learning: 27/7 access to instantaneous, consistent, and accurate responses. Optimizes training.
2. Builds Capacity: addresses staff and skill shortages by reducing time to onboarding and training new hires.
3. Optimizes SOP Updates: generates new content for SOPs that SMEs can fine-tune.
4. Positive Impact on Productivity: reduces time on tasks; allows more time to focus on complex, value-added tasks.

Scope

In Scope: Training the LLM on ORA's guidelines, job aids, training materials, policies, SOPs (Standard Operation Procedures), and the new grants management systems training documents (next phase). Fed agency data - Uniform Guidance, National Institutes of Health (NIH), and the National Science Foundation (NSF)-PAPPG. Anonymized data.

Out of Scope: Other Fed agency and industry applications and data sources (BioRaft or OnBase). Confidential, sensitive, proprietary, proposal, award, and intellectual property content, and personally identifiable information. This is not a Robotic Process Automation (RPA) project.

Risks and Mitigation

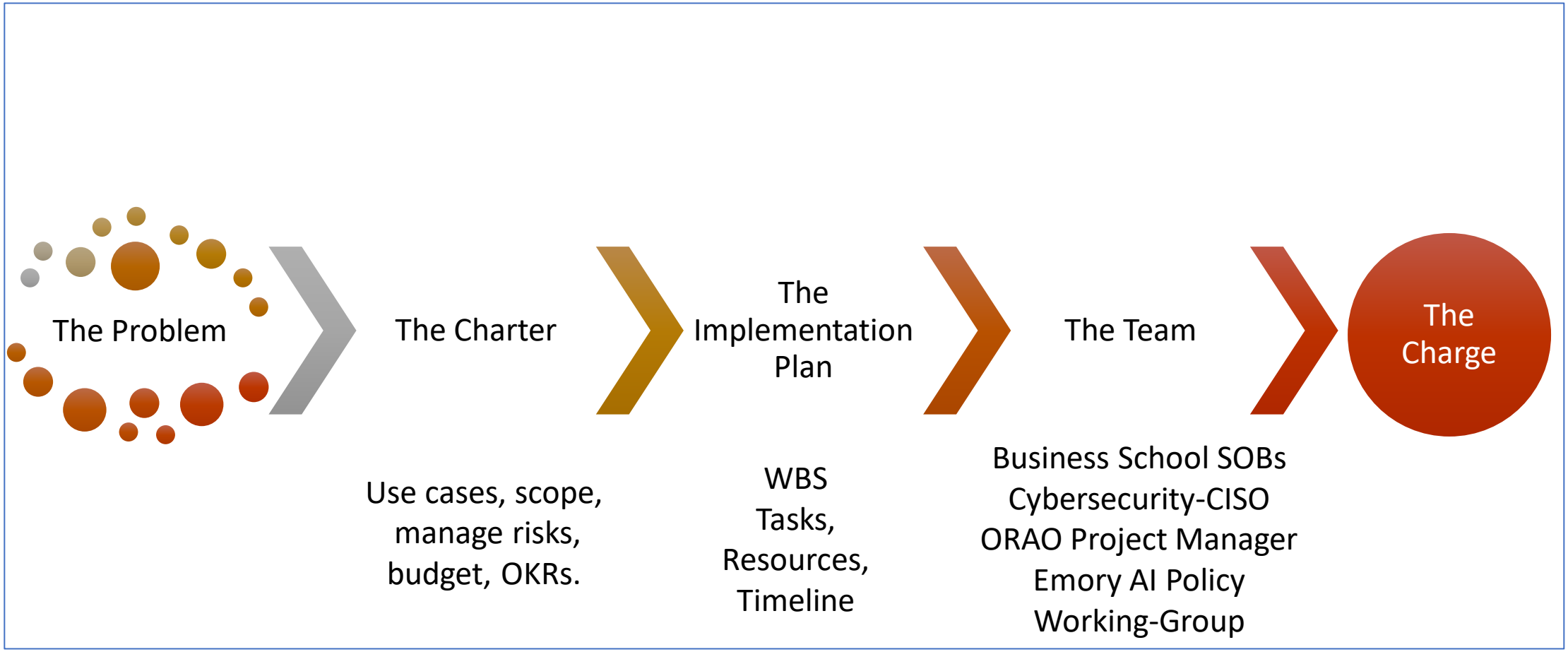
The potential risks of building a proof-of-concept LLM are minimal.

Potential Challenges	Preventative Measures
Network Security	ORA LLM will reside on Emory's network—only ORA staff users. (Confirm with Benn)
Responsible Use & Compliance	Adhere to government and Emory guidelines that are under development and ethical use.
Data Privacy & Security	Data protection/security, inaccuracies, and copyright infringement. The pre-training materials will include only 30+ vetted ORA documents and four fed agency URLs.
Late Adoption	Emory could be at a serious competitive disadvantage if we do not quickly adopt GenAL in research and research administration.

Implementation Plan

Title	% Complete	Start Date	Target Date
FY23 Project Initiation	100.00%	7/31/2023 9:00 AM	8/2/2023 5:00 PM
Create Project Charter	100.00%	7/31/2023 9:00 AM	8/2/2023 5:00 PM
Define the project's objectives and scope	100.00%	7/31/2023 9:00 AM	7/31/2023 5:00 PM
Define End User and high-level use case	100.00%	7/31/2023 9:00 AM	7/31/2023 5:00 PM
Define the high-level requirements for the Chatbox	100.00%	7/31/2023 9:00 AM	7/31/2023 5:00 PM
Set up the project team, including roles and responsibilities	100.00%	7/31/2023 9:00 AM	8/2/2023 5:00 PM
Phase 1 (Student Build)	100.00%	7/31/2023 9:00 AM	10/20/2023 5:00 PM
Requirements & Design Phase (exploration)	100.00%	9/14/2023 9:00 AM	9/27/2023 5:00 PM
Development Phase	100.00%	7/31/2023 9:00 AM	10/20/2023 5:00 PM
Model Training Phase	100.00%	7/31/2023 9:00 AM	10/20/2023 5:00 PM
Testing and Quality Assurance Phase	100.00%	10/13/2023 9:00 AM	10/20/2023 5:00 PM
Deployment Phase	100.00%	9/27/2023 9:00 AM	10/20/2023 5:00 PM
Create short film of the demonstration and query I/O	100.00%	7/31/2023 9:00 AM	10/20/2023 5:00 PM
User Training, Demo, and Documentation	0.00%	8/23/2023 9:00 AM	12/31/2023 5:00 PM

The TEAM


The Problem
The Charter
The Implementation Plan
The Team
The Charge
Use cases, scope, manage risks, budget, OKRs.
WBS
Tasks,
Resources,
Timeline
Business School SOBs
Cybersecurity-CISO
ORAO Project Manager
Emory AI Policy
Working-Group

Objectives

Design

AI model with generative capabilities (ORAGPT)

Development

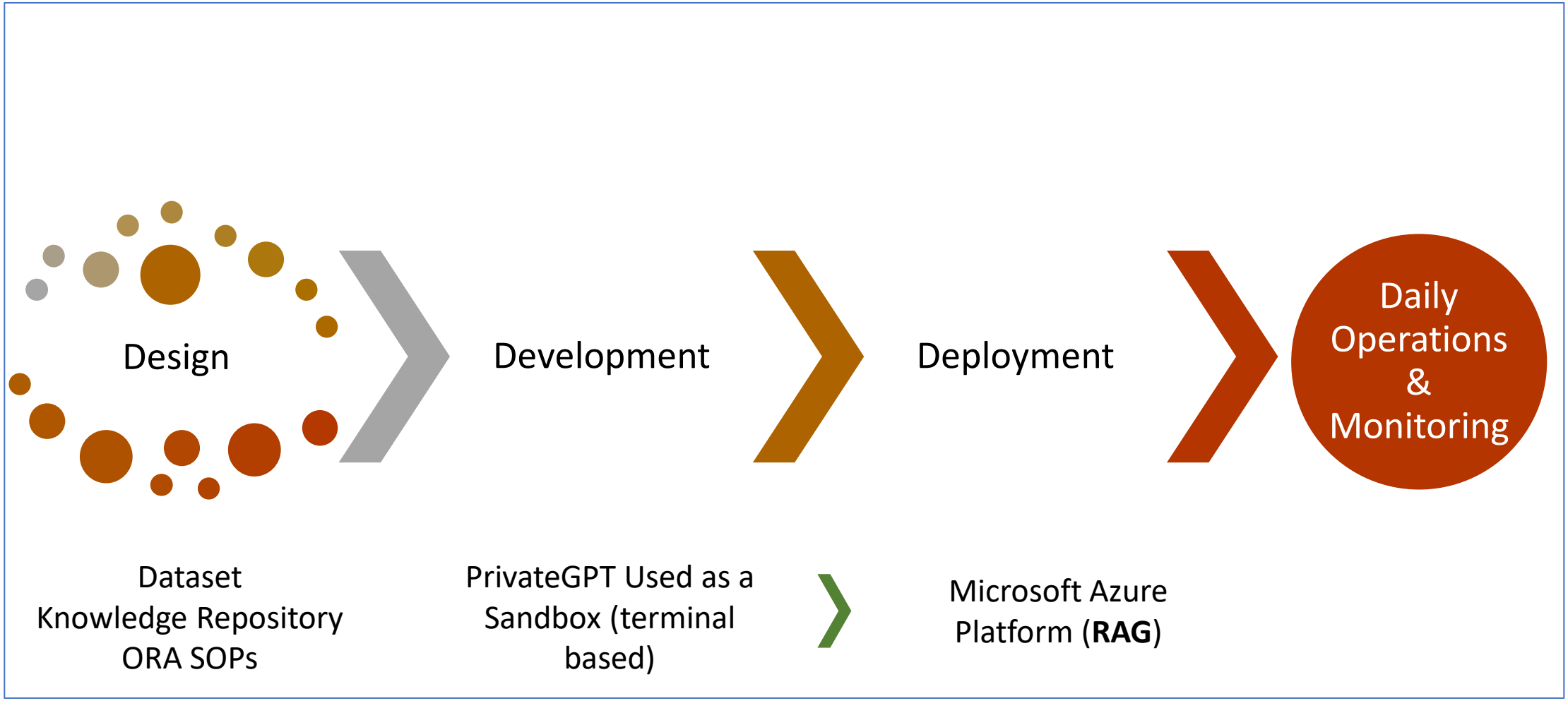
Scalable, compatible, and meets Emory University's data protection policies

Deployment

Instantaneous, 24/7 answers to task-related questions.
Write new policies, procedures, and training materials.

Daily Operations

Phase One: Selecting Our Approach



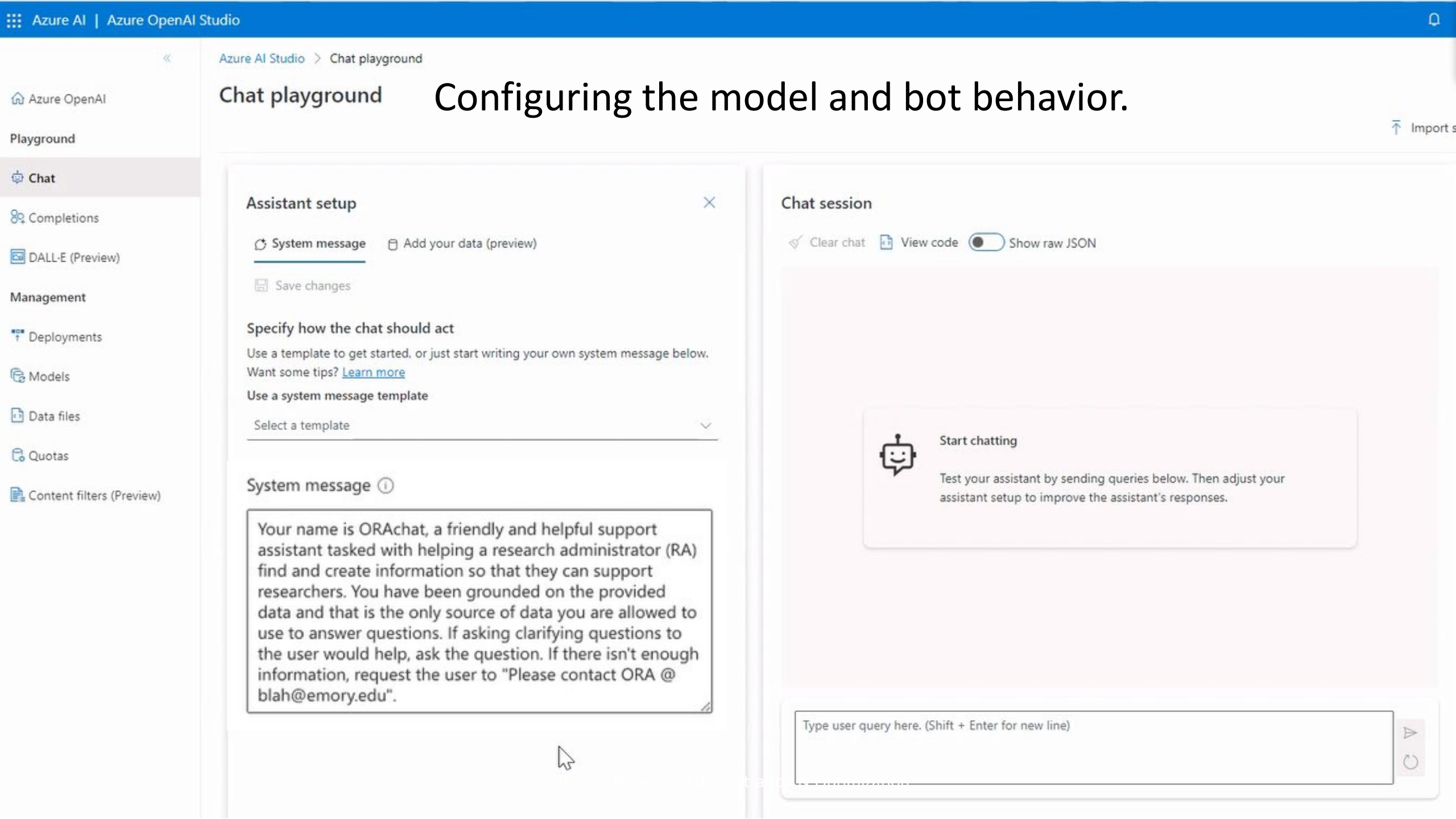
Design: Corpus and **RAG**

In Scope

- A subset of the ORA Knowledge Repository
  - **SOPs, policies, guidelines, training manuals, and task checklists.**
- 2 CFR 200 – Uniform Guidance
- NSF PAPPG
- NIH guidelines

Out of Scope

- Confidential, Sensitive, Proprietary, Private
- Personally Identifiable Information (PII)
- Proposals, Awards, IP
- Integration of third-party applications
- Access to the Internet knowledge


Azure AI | Azure OpenAI Studio
Azure AI Studio > Chat playground
Chat playground
Configuring the model and bot behavior.
Azure OpenAI
Playground
Chat
Completions
DALL-E (Preview)
Management
Deployments
Models
Data files
Quotas
Content filters (Preview)
Import
Assistant setup
System message
Add your data (preview)
Save changes
Specify how the chat should act
Use a template to get started, or just start writing your own system message below.
Want some tips? Learn more
Use a system message template
Select a template
System message
Your name is ORAchat, a friendly and helpful support assistant tasked with helping a research administrator (RA) find and create information so that they can support researchers. You have been grounded on the provided data and that is the only source of data you are allowed to use to answer questions. If asking clarifying questions to the user would help, ask the question. If there isn't enough information, request the user to "Please contact ORA @ blah@emory.edu".
Chat session
Clear chat
View code
Show raw JSON
Start chatting
Test your assistant by sending queries below. Then adjust your assistant setup to improve the assistant's responses.
Type user query here. (Shift + Enter for new line)

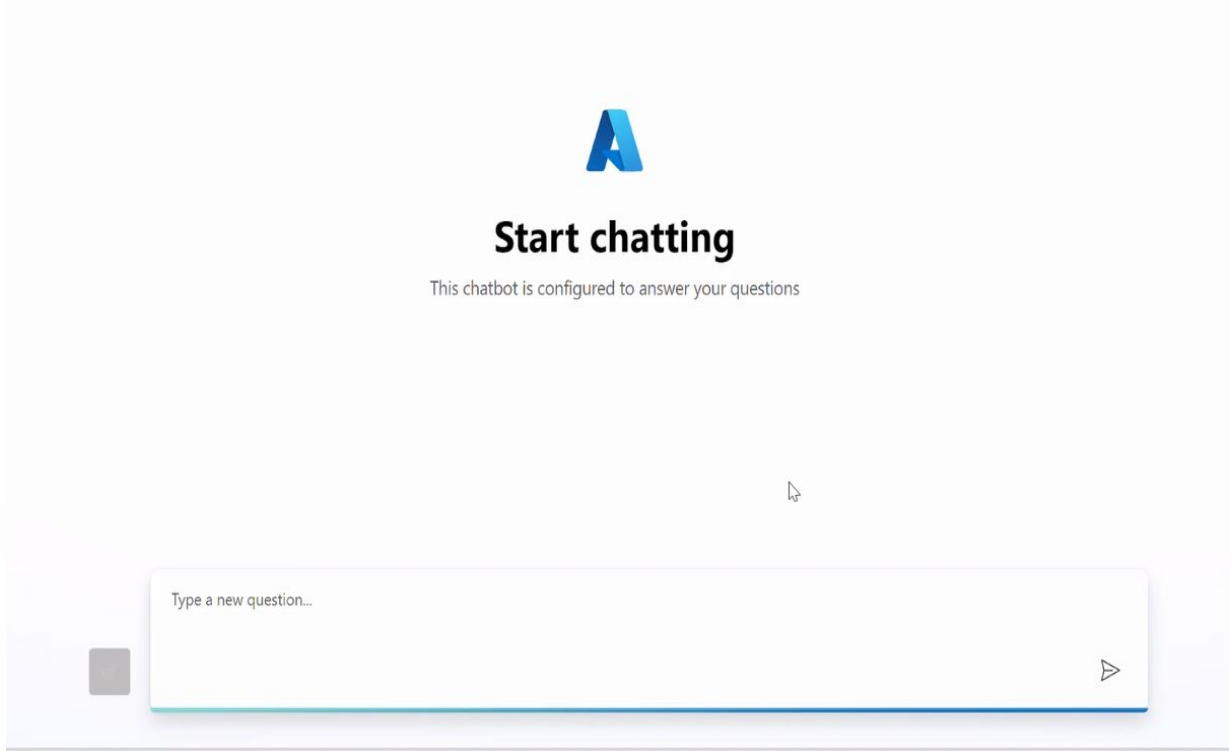
TEVV System Responses

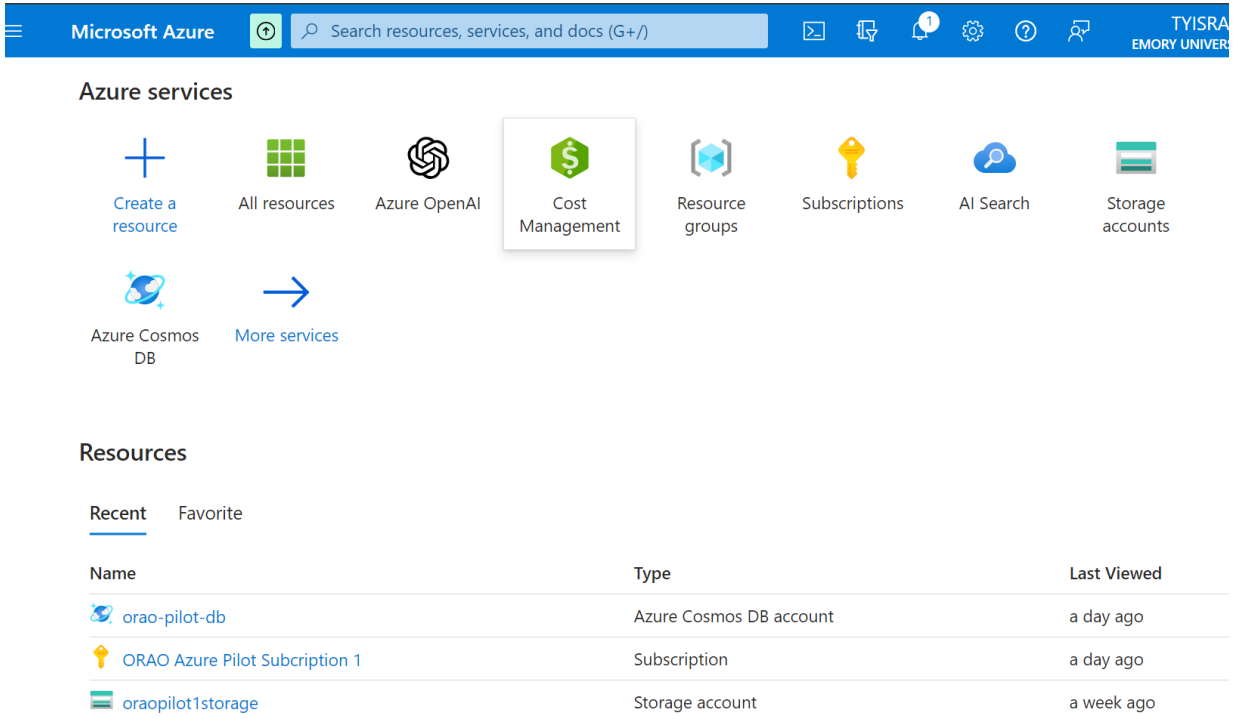
Testing | Evaluating | Validating | Verifying

How did our bot perform?

POC Q&A bot response document

PoC End User & Admin Interface

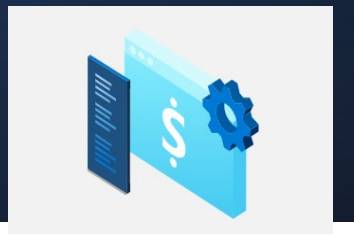


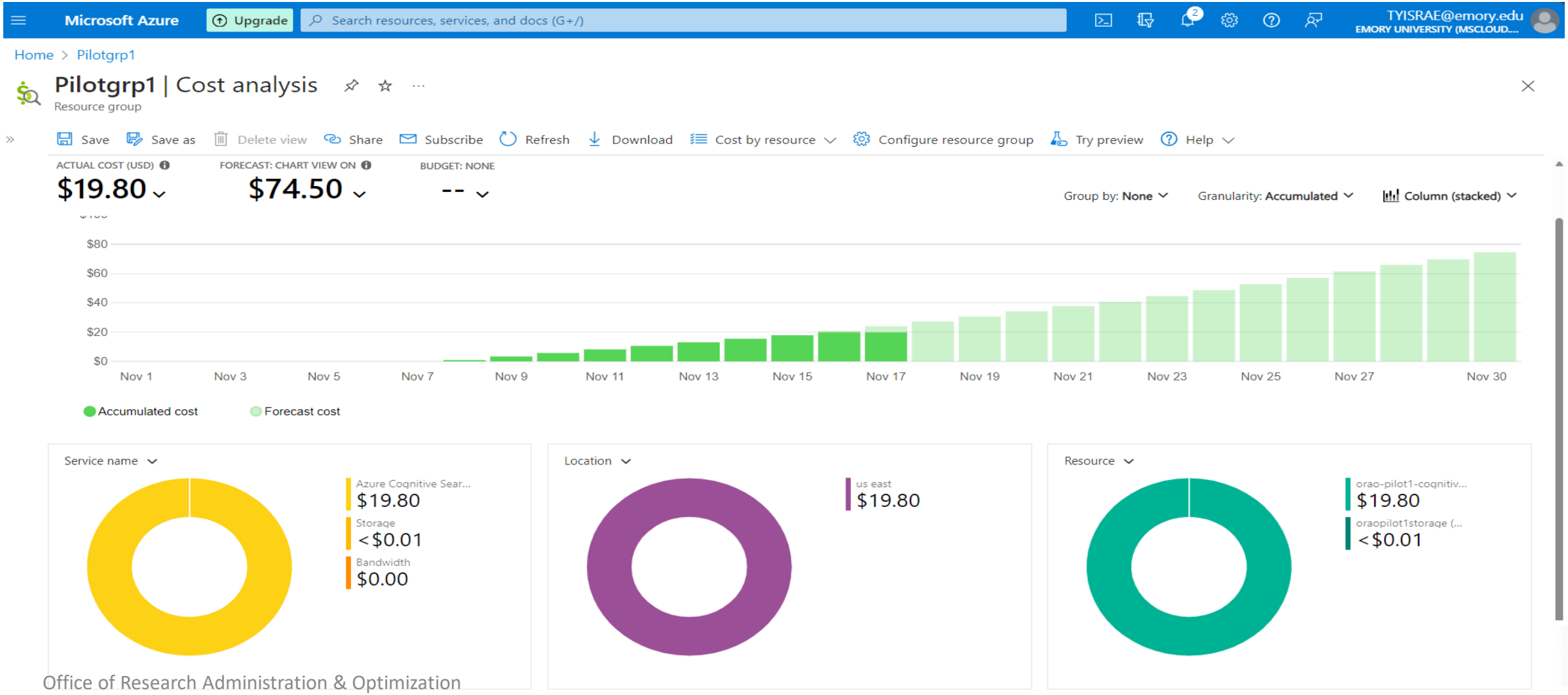


Office of Research Administration & Optimization

ORA ChatGPT Demo

Associated Costs





Success of Phase One

Accomplishments

- Created an ORA Knowledge Repository
- Developed an ORA ChatGPT-like Virtual Assistant
- Hosted AI in Research Administration Panel Sessions
- Created an implementation Plan in PPM Pro
- Internally built a PoC prototype in six weeks
- Learned Azure while building the bot

Model Performance

Accuracy and reliability, as indicated by the Q&A bot responses document captured in video and Google doc link below:

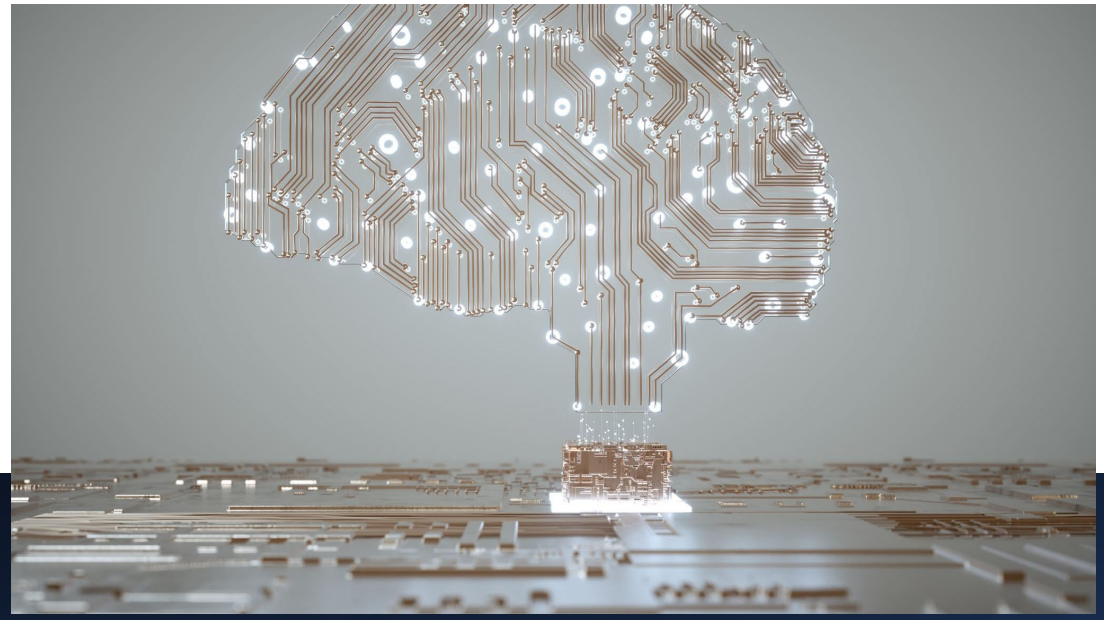
https://docs.google.com/spreadsheets/d/1zdTv6nPIW_vSyHiy4gA43g-Xer4Moh02W9th7Mdr0hA/edit

The next phase will implement complete formal evaluation (TEVV), metrics, and analysis.

Pathway to Build or Buy

The Commercial Build

Leveraging ChatGPT and Generative Artificial Intelligence in Sponsored Programs Operations



Lessons Learned

Office of Research Administration & Optimization

The Journey to MS Azure

https://www.youtube.com/watch?v=wCo7MuBPbRI&list=PLbsVkwJx50CNvKxq15juCx7M6Ccoa6vuL&index=2

Lessons Learned, Challenges, and Concerns

- Adapt to the rapidly changing AI market
- Gauge organizational readiness
- Formalize robust testing, evaluating, verifying, and validating
- Develop guiding principles: responsible, ethical, trustworthy, transparent, explainable
- Implement robust change management and technology adoption
- Develop robust performance metrics and analytics
- Keep humans in the loop during all lifecycle stages
- Vet the source materials better for accuracy, reduced hallucinations, and bias

Lessons Learned

- **Augments to improve GPT responses:**
  - Vector Search, Semantic Search, Text-Embedding
- **Document intelligence:**
  - Allows the model to respond to conversational questions about specific file types: spreadsheets, slide decks, invoices, contracts, CAD, PDF, Word, etc.
  - Include various file types in the training dataset.
- **Transparent cost management:**
  - Costs for running the bot – clear and understandable, no hidden fees, predictable and trackable expenses to manage the budget effectively.
  - The most significant cost comes from using Azure Search services for data search and indexing.
  - Highest cost comes from data Search (Cognitive – Azure AI search)

ORA ChatGPT PoC Farewell

https://www.youtube.com/watch?v=xcmUk2Hv8BM&list=PLbsVkwJx50CNvKxq15juCx7M6Ccoa6vuL&index=3

Office of Research Administration & Optimization

The Case for Phase Two

ORA Chatbot: Building a More Robust, Enhanced Enterprise Model and the Growing List of Business Use Cases

The Road to a More Robust Enterprise Model.

Proven foundation and potential for scaling up

Expand and enhance the project for broader institutional benefit

Gain institutional support for an enterprise solution

Leverage partnerships for a myriad of use cases

Integrate the Data Analytics Team, TAM, and Change Management

Metrics: Engaging the ORA Research Data Analytics Team

- Current applications of RDA in RA
- Explain the role of RDA in phase two
- Use AI to predict behavior
- Benefits of engaging the data analytics team

Thanks for your time and attention!

Questions