



CDAO Fall

Boston

AI in Real Life

A Case Study on how
UNIDOSUS is implementing AI
to make a real impact
15 October 2024

UnidosUS

Gen AI

Gen AI at UnidosUS

UNIDOSUS

Background

UNIDOSUS

UnidosUS

UnidosUS, previously known as NCLR (National Council of La Raza) is the nation's largest Hispanic civil rights and advocacy organization.

- Policy & Advocacy
- Programs
 - Partner with affiliated local organizations that work directly with community
 - Affiliate Network – 280+ affiliates across the US
- Digital Services
 - Technology Innovation not IT infrastructure
 - Current focus is on data

UNIDOSUS

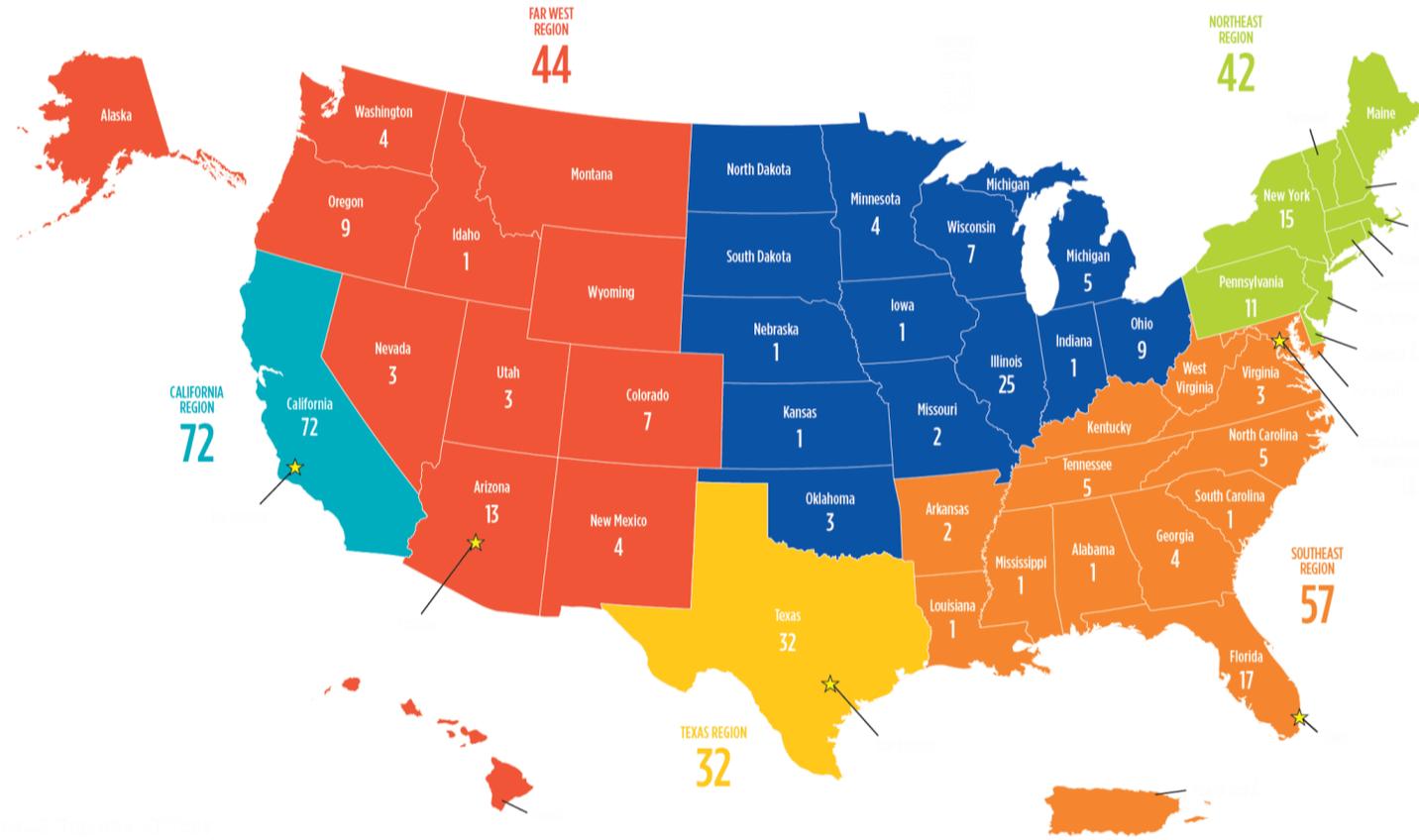


Strategic Framework

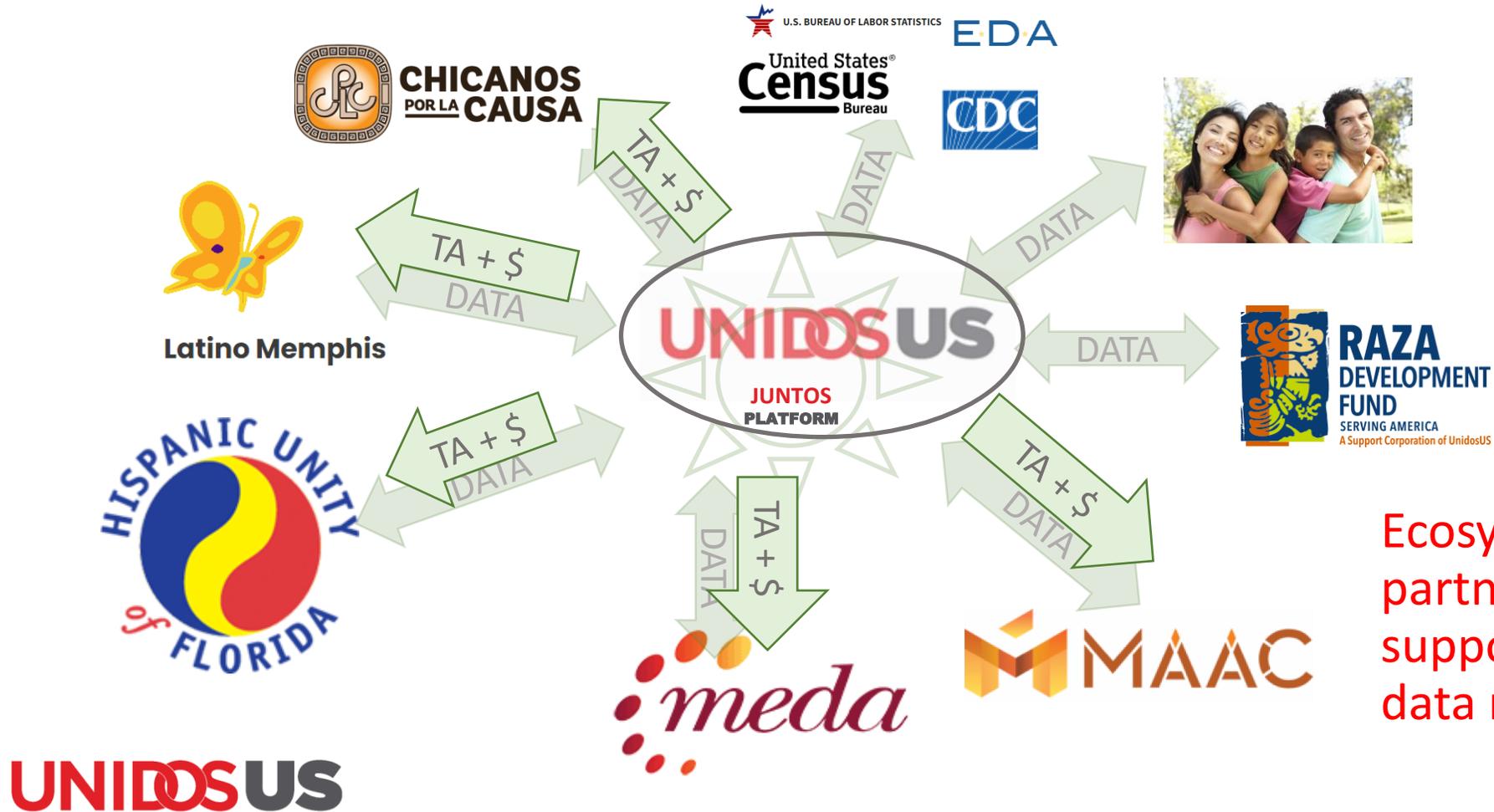


UnidosUS Affiliate Network

300+ Affiliates
Over \$4.5 billion
aggregate revenue
across network



Unidos US “platform” amplified by Juntos Platform



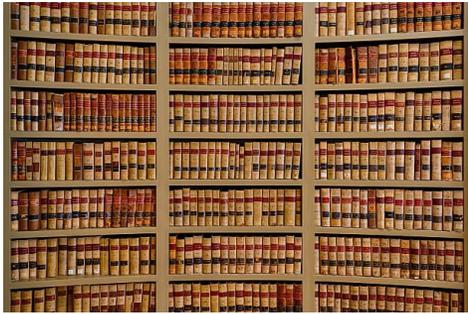
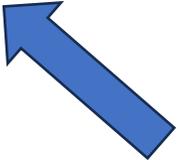
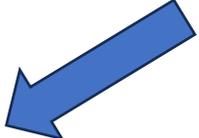
Ecosystem of community, partners and grantees supported by shared data resources

Gen AI

UNIDOSUS

Gen AI Working Analogy

Supreme Court Clerks



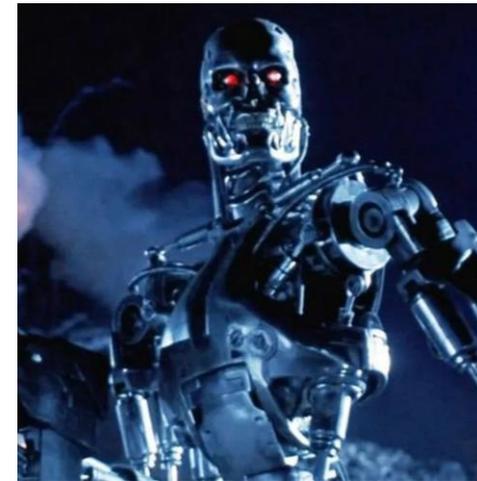
Challenges



Bias



Hallucination



“General AI”

Gen AI at UnidosUS

UNIDOSUS

Initiatives

Every staff member has their own AI “clerk”

- AI synthesizes information and summarizes documents
- New staff can benefit from organizational knowledge
- Existing staff can focus on value adding planning instead of research

Staff as Fellow Researchers

- No one is truly an expert on Gen AI Implementation
- Staff are not users but true partners
- Secondary goal of all tools is to learn about how best to use Gen AI

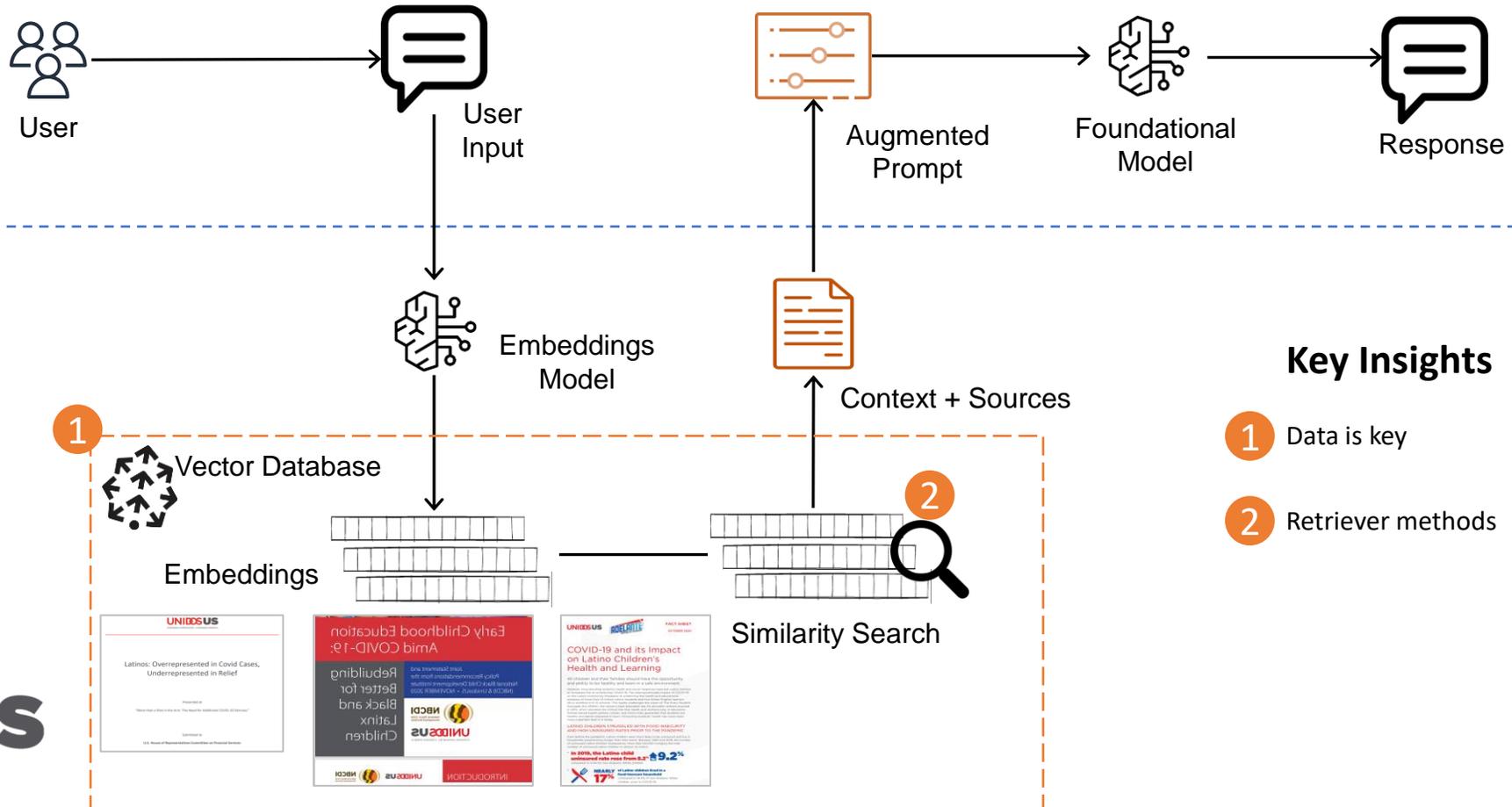
Tech stack

RAG Workflow

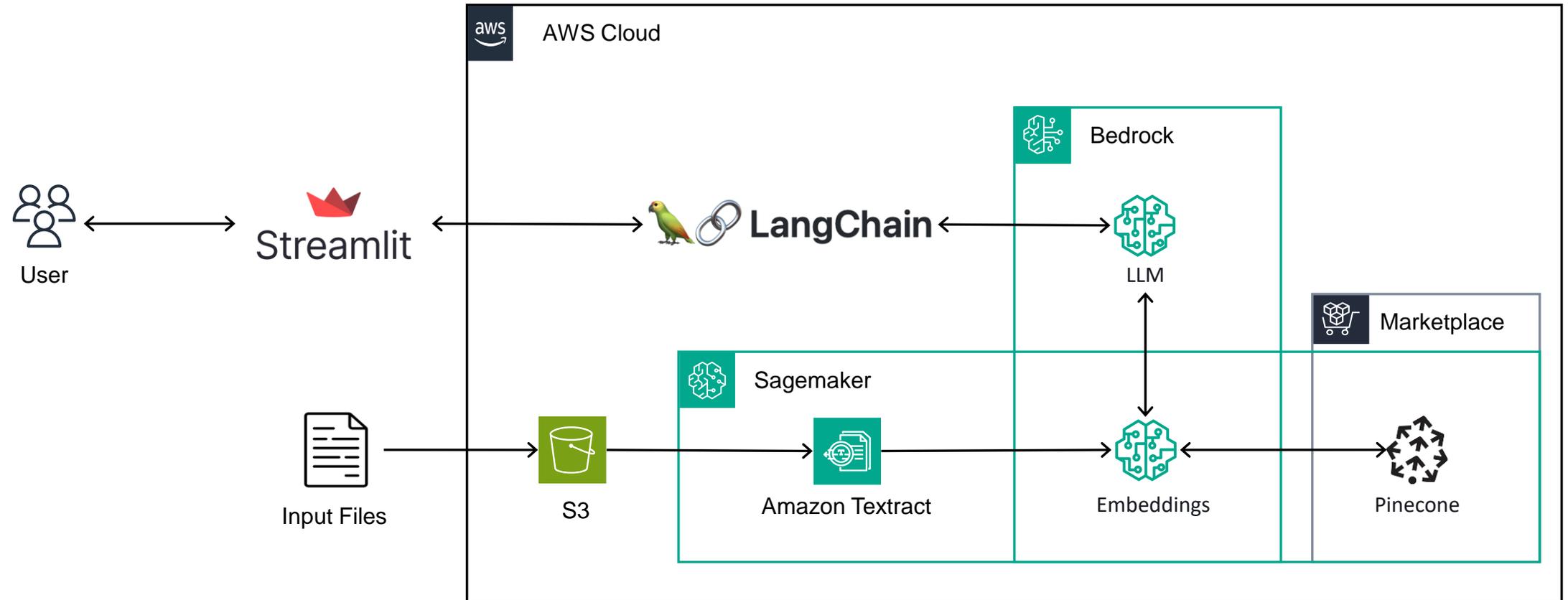
Orchestration
Workflow

RAG
Workflow

UNIDOSUS



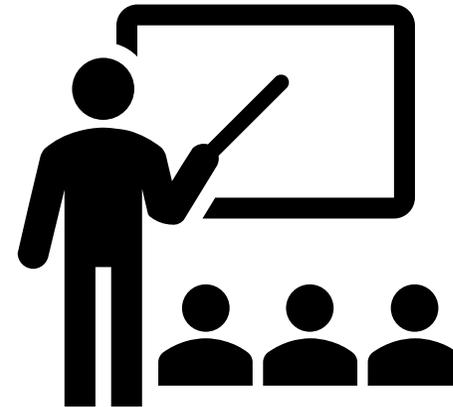
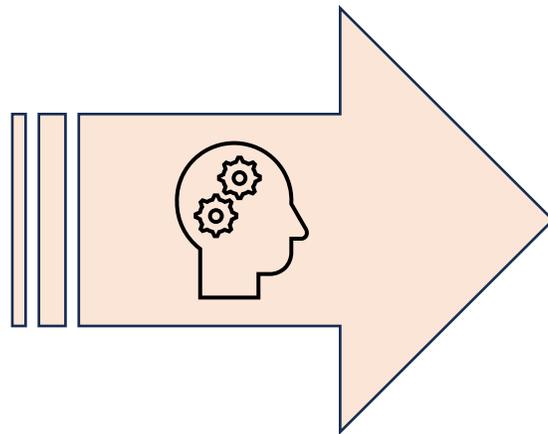
RAG Architecture



Curriculum dev with Gen AI

UUS program curriculum development

New programs require synthesize of curriculum from many source docs

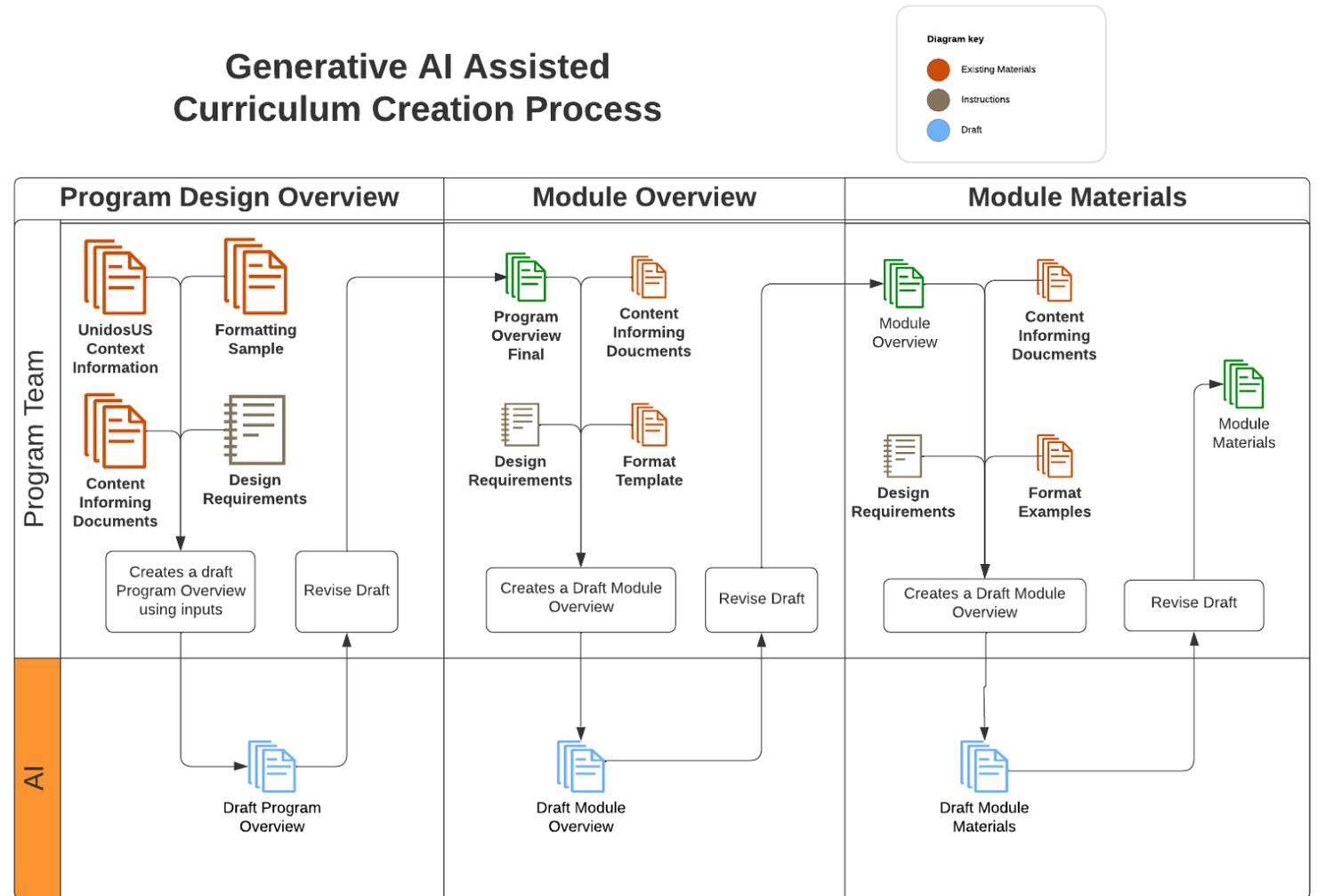


Curriculum Development Solution Flow

Using Generative AI and RAG (ie context documents) to create program curriculum.

Key is that Gen AI just creates drafts for program manager to use as raw material for actual document

Generative AI Assisted Curriculum Creation Process



Policy Library AI Tool

Innovation - Generative AI Policy Tool

Custom “UnidosUS ChatGPT”

- Addresses Hallucination/Bias, Privacy, Visibility
- Enables safe experimentation for UUS staff
- Increases UUS reputation for innovation

The screenshot displays the UnidosUS Policy Document Assistant interface. On the left is a sidebar with the UnidosUS logo and navigation options. The main content area is titled "Policy Document Assistant - v2" and includes a search bar, a knowledge base dropdown, and a list of generated text. A red crown icon is visible in the bottom right corner.

UNIDOSUS

Navigation
Select a page:
Policy App

Select Time Period
Year
2000 2012 2024

Select Document Type
Select Type:
Choose an option

Purpose of using AI tool
What is the purpose of using this tool?
generate draft remarks
Submit

Share ☆ 🌐 ⓘ

Policy Document Assistant - v2

How can I help you?

what was the impact of COVID on latino students education in the US

✓ Knowledge Base: impact of COVID on Latino students education in the US

✓ Complete!

The impact of COVID-19 on Latino students' education in the US has been significant and multifaceted. The pandemic has exacerbated existing inequities in the education system, disproportionately affecting Latino students in several ways. Key impacts include:

1. **Educational Disruption and Remote Learning Challenges:** Many Latino students faced hurdles with distance learning due to a lack of access to high-speed internet and digital devices. Approximately one in three Latino households lacked high-speed internet, making remote learning difficult.
2. **Increased Dropout Rates:** The pandemic has led to an increase in dropout rates among Latino students. Pre-pandemic strides in Latino enrollment and attainment are now at risk, with a notable decline in Latino postsecondary enrollment observed for the first time in a decade.
3. **Health and Economic Impacts:** Latino communities have been hit hard by the health and

What is UnidosUS? ➤

demo

Policy Library Project

A Digital Services Project

Welcome: apatricio@unidosus.org

What is your purpose for using this tool today?

Submit

UNIDOSUS

Visioning the future

Biography of Maria Perez

Early Life and Background

Maria Perez was born to a family of migrant workers from Mexico. Growing up in the San Joaquin Valley, California, Maria's early life was marked by the challenges and resilience typical of many immigrant families. Her parents toiled in the agricultural fields, instilling in her a strong work ethic and the value of education as a pathway to a better life.

Education

Maria attended a local high school in a school district with limited resources. Despite these challenges, she demonstrated a passion for learning and a particular interest in the sciences. Although she did okay academically, her determination and curiosity set her apart.

After high school, Maria enrolled in a community college, where she found her footing and excelled in her studies. Her academic performance and dedication earned her a place at California State University, where she pursued a degree in Biology. It was here that Maria's interest in marine biology began to take shape, thanks to an inspiring professor who introduced her to the wonders of the ocean and its ecosystems.

Advanced Studies and Career

Maria's burgeoning passion for marine biology led her to collaborate with her professor on various research projects. This hands-on experience was crucial in shaping her career path. After



Key takeaways

Gen AI use cases is about synthesizing info not making judgements

- ChatGPT is a toy, need to use RAG and related approaches
- Still many use cases for traditional AI

People and AI interaction needs to be designed in from the start

- AI cannot replace expertise
- People cannot synthesize information as efficiently or quickly

Gen AI is evolving technology not mature

- Clear Guardrails/Guidelines need to be put in place
- Experimentation and Iteration is the best approach

Questions?

UNIDOSUS

Andrew Patricio

Principal, Data & Analytics

apatricio@unidosus.org

<https://www.linkedin.com/in/andrewpatricio1/>

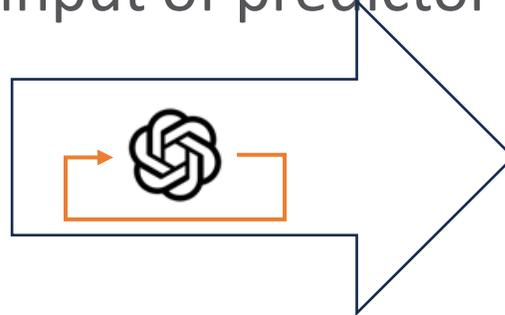
UNIDOSUS

Gen AI/RAG overview

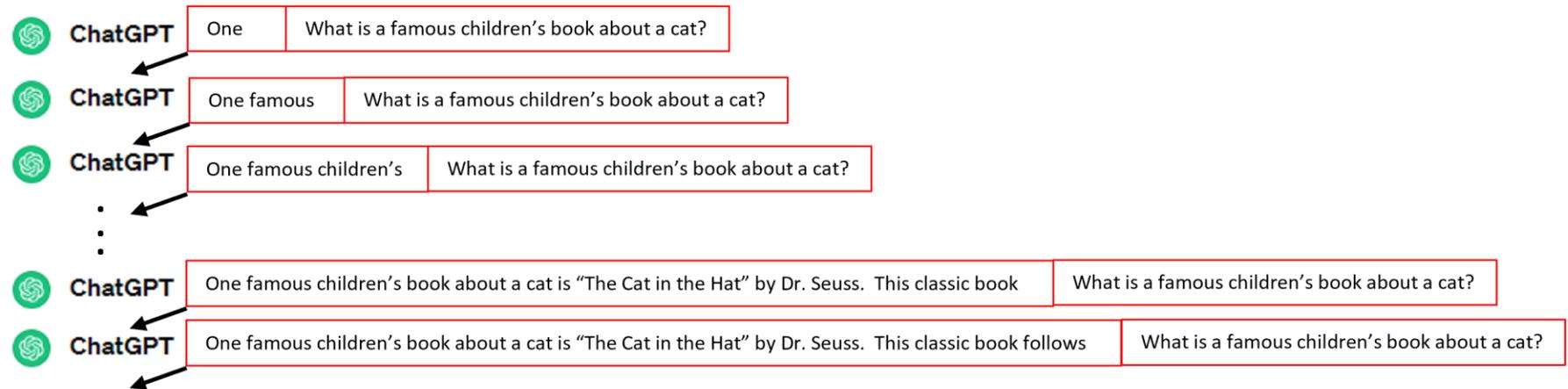
Generative AI – how it works

LLM is simply a predictor of what the next word is given previous words fed into it
Actual response from LLM based tools is created via multiple runs of output continuously fed back into input of predictor

You
What is a famous children's book about a cat?



ChatGPT
A famous children's book about a cat is "The Cat in the Hat" by Dr. Seuss. It was first published in 1957 and has since become a beloved classic in children's literature. The story follows the mischievous Cat in the Hat as he visits the home of two children on a rainy day, bringing chaos and fun into their lives with his antics. The book is known for its rhyming text, whimsical illustrations, and timeless appeal to young readers.



Generative AI - Hallucination

LLM is model of language NOT model of knowledge:
Designed to generate something that sounds correct not necessarily something that is correct.

If it seems like GenAI “knows” something it is because for humans talking about concepts and the actual concepts are one in the same

Gen AI can create plausible-sounding but wrong responses called “Hallucinations”



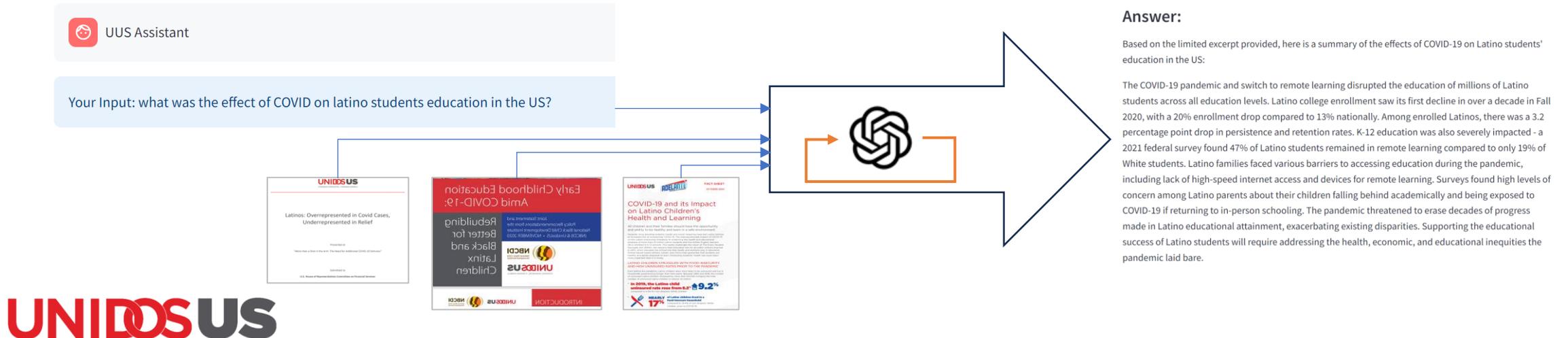
The screenshot shows a news article on the Insider website. The article title is "A law firm was fined \$5,000 after one of its lawyers used ChatGPT to write a court brief riddled with fake case references". The author is Grace Dean, and the article was published on June 23, 2023, at 6:55 AM CDT. The article includes a photo of a lawyer in a blue suit and red tie sitting at a desk with a gavel and scales of justice. Below the photo, there is a list of bullet points summarizing the article's content.

- A law firm was fined \$5,000 after one of its lawyers used ChatGPT to write a court brief.
- The document had included references to some cases and

Solution to Hallucination: Provide Context

Just as prompt and partial output are fed back into the LLM as it is creating a response, contextual documents can also be provided

- Document source can be preloaded database of contextual documents vetted by your organization or searches through the internet (latter not recommended)
- Though Gen AI doesn't "understand" documents it can still synthesize/summarize them
- Use Generative AI as an assistant to find and summarize relevant documents as a kind of personal "Smart Grad Student Intern", "Paralegal", "Clerk", etc



UNIDOS US

RAG Approach

what was the effect of COVID on latino students education in the US?

Question



UNIDOSUS

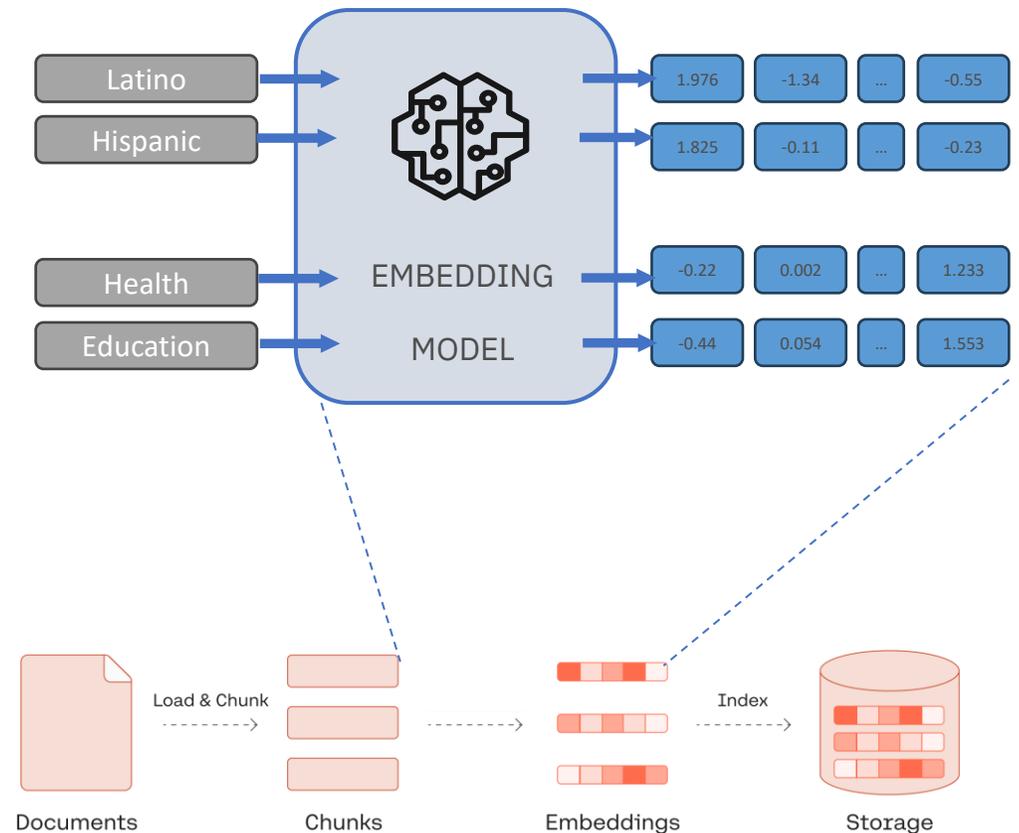
Context

Question



Embeddings and Vector Databases

- **Embeddings** are numerical representations of values or objects like text in a high-dimensional space. By representing data as vectors, embeddings capture the semantic relationships between data points, enabling tasks like similarity search, classification, and clustering more efficiently.
- A **vector database** is a data storage system that is specifically designed to both store and retrieve vectors quickly. This type of database is useful for storing the embeddings generated by a Model.



Retrieval Augmented Generation (RAG)

1. User enters prompt, which is first bumped up against a database storing documents to find those most relevant
2. Prompt and relevant documents all sent to LLM to generate response
3. Can be configured to consider documents as part of response or build response strictly from documents alone (so that “I don’t know” can be valid response)

