Leveraging Custom AI Solutions for California Department of Finance

09/26/2024

About My Custom AI

- Presenter: Dr. Amine Allouah, *My Custom AI co-founder*
  - **PhD in Applied Mathematics and Computer Science** from Columbia University
  - **Former Meta Engineer** Led development of custom algorithms enhancing team efficiencies

About My Custom AI

- Presenter: Dr. Amine Allouah, *My Custom AI co-founder*
  - **PhD in Applied Mathematics and Computer Science** from Columbia University
  - **Former Meta Engineer** Led development of custom algorithms enhancing team efficiencies

- My Custom AI
  - **Our Focus**: Crafting end-to-end custom AI solutions to tackle specific organization challenges
  - **Our Team**: Comprised of experts from Tech companies and leading Silicon Valley innovators
  - **Our Goal**: Leveraging AI to unlock new opportunities and drive efficiency in organizations

Outline of the presentation

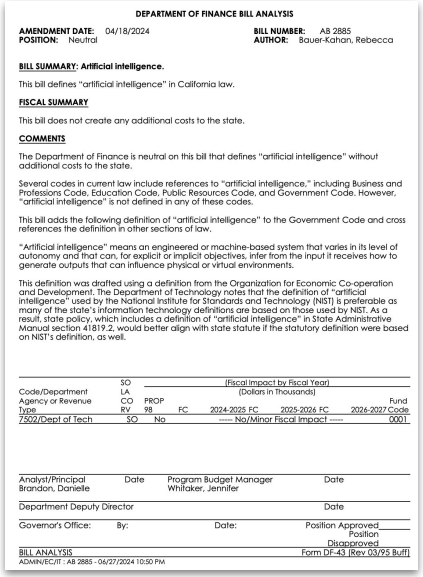
1. **Overview of challenges**
2. Relevant past/current use cases
3. Proposed solutions
4. Benefits and Risks

Overview of Challenges

Challenge 1: Streamline Legislative Research and Analysis

- **Problem**: High volume of legislative bills requires thorough fiscal and policy analysis, competing with urgent budget tasks

- **Need**: Support staff by automating the extraction and validation of relevant background and fiscal data to support analysts in creating accurate, timely analyses



DEPARTMENT OF FINANCE BILL ANALYSIS

AMENDMENT DATE: 04/18/2024 **BILL NUMBER:** AB 2885
POSITION: Neutral **AUTHOR:** Bauer-Kahan, Rebecca

BILL SUMMARY: Artificial intelligence.

This bill defines "artificial intelligence" in California law.

FISCAL SUMMARY

This bill does not create any additional costs to the state.

COMMENTS

The Department of Finance is neutral on this bill that defines "artificial intelligence" without additional costs to the state.

Several codes in current law include references to "artificial intelligence," including Business and Professions Code, Education Code, Public Resources Code, and Government Code. However, "artificial intelligence" is not defined in any of these codes.

This bill adds the following definition of "artificial intelligence" to the Government Code and cross references the definition in other sections of law.

"Artificial intelligence" means an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.

This definition was drafted using a definition from the Organization for Economic Co-operation and Development. The Department of Technology notes that the definition of "artificial intelligence" used by the National Institute for Standards and Technology (NIST) is preferable as many of the state's information technology definitions are based on those used by NIST. As a result, state policy, which includes a definition of "artificial intelligence" in State Administrative Manual section 41819.2, would better align with state statute if the statutory definition were based on NIST's definition, as well.

	SO		(Fiscal Impact by Fiscal Year)		
Code/Department	LA		(Dollars in Thousands)		
Agency or Revenue	CO	PROP			Fund
Type	RV	98	FC	2024-2025 FC	2025-2026 FC	2026-2027 Code
7502/Dept of Tech	SO	No	------ No/Minor Fiscal Impact ------		0001

Analyst/Principal Date Program Budget Manager Date
Brandon, Danielle Whitaker, Jennifer

Department Deputy Director Date

Governor's Office: By: Date: Position Approved
Position Disapproved

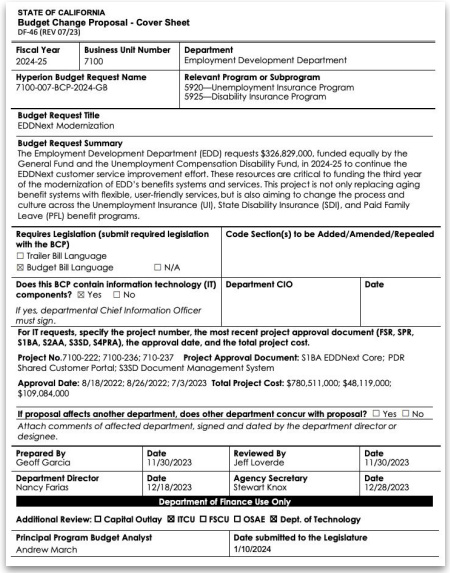
BILL ANALYSIS Form DF-43 (Rev 03/95 Buff)
ADMIN/EC/IT - AB 2885 - 06/27/2024 10:50 PM

Overview of Challenges

Challenge 2: Identify Statewide Efficiency Opportunities through Budget Change Proposal (BCP) Analysis

- **Problem**: BCPs must be analyzed for trends and shared resource opportunities across departments and fiscal year(s)

- **Need**: AI-driven insights to support analysts in uncovering patterns, forecasting needs, and maximizing efficiency across state resources.



STATE OF CALIFORNIA
Budget Change Proposal - Cover Sheet
DF-46 (REV 07/23)

Fiscal Year	Business Unit Number	Department
2024-25	7100	Employment Development Department

Hyperion Budget Request Name	Relevant Program or Subprogram
7100-007-BCP-2024-GB	5920—Unemployment Insurance Program 5925—Disability Insurance Program

Budget Request Title
EDDNext Modernization

Budget Request Summary
The Employment Development Department (EDD) requests $326,829,000, funded equally by the General Fund and the Unemployment Compensation Disability Fund, in 2024-25 to continue the EDDNext customer service improvement effort. These resources are critical to funding the third year of the modernization of EDD's benefits systems and services. This project is not only replacing aging benefit systems with flexible, user-friendly services, but is also aiming to change the process and culture across the Unemployment Insurance (UI), State Disability Insurance (SDI), and Paid Family Leave (PFL) benefit programs.

Requires Legislation (submit required legislation with the BCP)	Code Section(s) to be Added/Amended/Repealed
☐ Trailer Bill Language ☒ Budget Bill Language ☐ N/A	

Does this BCP contain information technology (IT) components? ☒ Yes ☐ No *If yes, departmental Chief Information Officer must sign.*	Department CIO	Date

For IT requests, specify the project number, the most recent project approval document (FSR, SPR, S1BA, S2AA, S3SD, S4PRA), the approval date, and the total project cost.

Project No. 7100-222; 7100-236; 710-237 **Project Approval Document:** S1BA EDDNext Core; PDR Shared Customer Portal; S3SD Document Management System

Approval Date: 8/18/2022; 8/26/2022; 7/3/2023 **Total Project Cost:** $780,511,000; $48,119,000; $109,084,000

If proposal affects another department, does other department concur with proposal? ☐ Yes ☐ No
Attach comments of affected department, signed and dated by the department director or designee.

Prepared By	Date	Reviewed By	Date
Geoff Garcia	11/30/2023	Jeff Loverde	11/30/2023
Department Director	**Date**	**Agency Secretary**	**Date**
Nancy Farias	12/18/2023	Stewart Knox	12/28/2023

Department of Finance Use Only

Additional Review: ☐ Capital Outlay ☒ ITCU ☐ FSCU ☐ OSAE ☒ Dept. of Technology

Principal Program Budget Analyst	Date submitted to the Legislature
Andrew March	1/10/2024

Outline of the presentation

1. Overview of challenges
2. **Relevant past/current use cases**
3. Proposed solutions
4. Benefits and Risks

Relevant past/current use cases

1. Large Scale AI-Powered Document Processing and Quality Control

2. Cost-Effective Data Consolidation for Business Insights

3. Generative AI for Marketing Analytics and Insights

AI-Powered Document Processing and Quality Control

Goal

- Collaborating with a book publishing company that has millions of scanned historical books

- Enable accurate and cost-effective restoration of these books, ensuring high-quality reprints and digital versions

AI-Powered Document Processing and Quality Control

Goal

- Collaborating with a book publishing company that has millions of scanned historical books

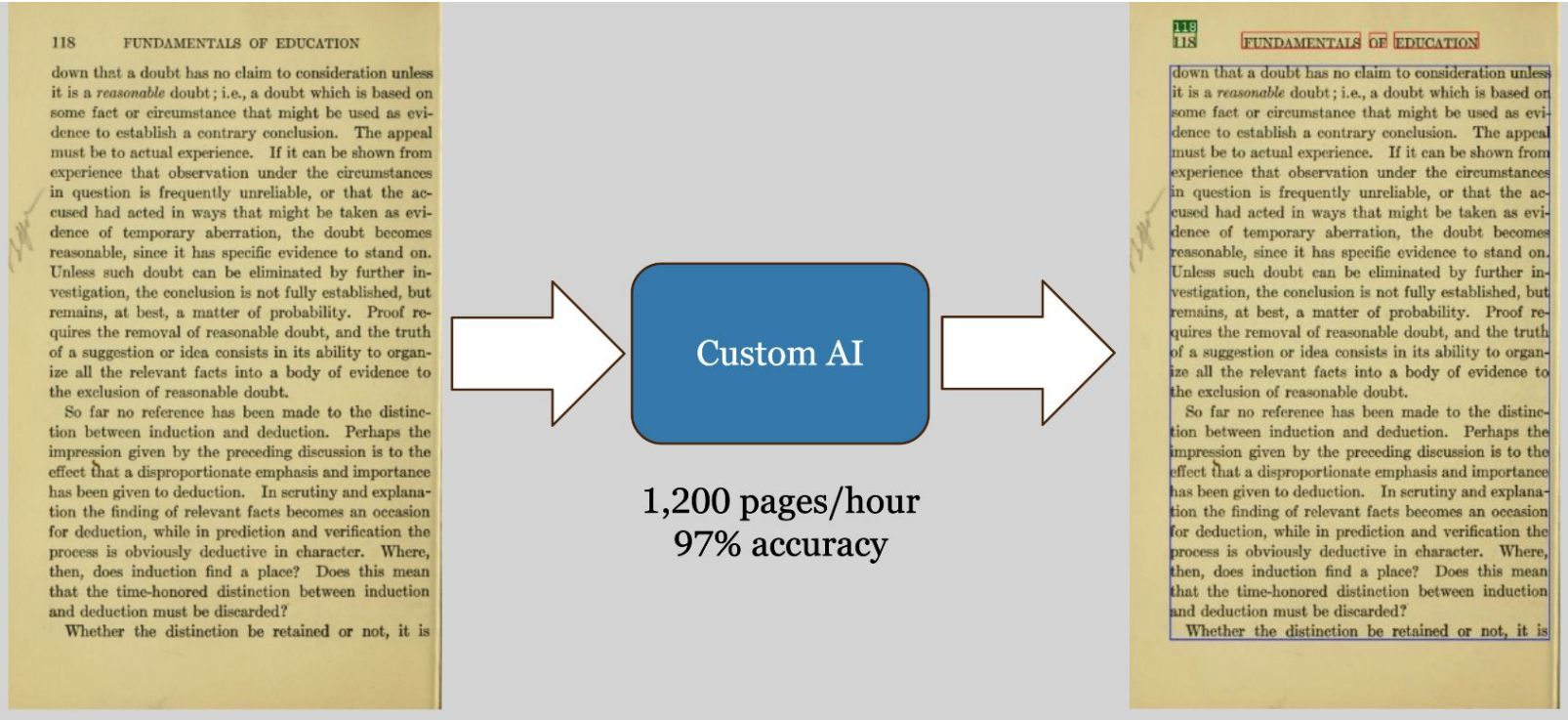
- Enable accurate and cost-effective restoration of these books, ensuring high-quality reprints and digital versions

Problem

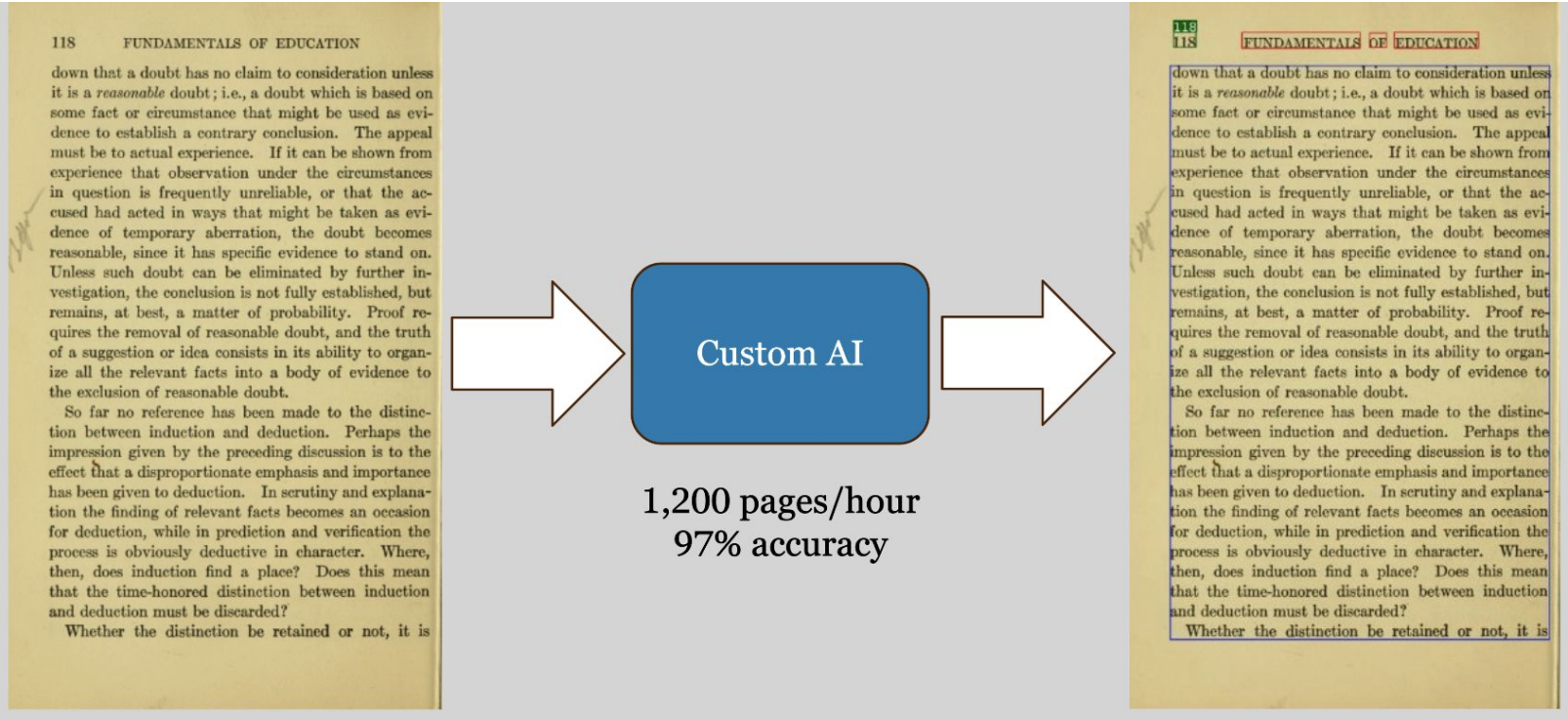
- Public technologies (Google, Microsoft) are cost-prohibitive for large-scale historical book restoration (3M)

- Custom extraction needed and not supported by existing solutions

AI-Powered Document Processing and Quality Control



AI-Powered Document Processing and Quality Control



Our current solution costs **90%** less than public solutions with **similar accuracy**

Relevant past/current use cases

1. Large Scale AI-Powered Document Processing and Quality Control

2. Cost-Effective Data Consolidation for Business Insights

3. Generative AI for Marketing Analytics and Insights

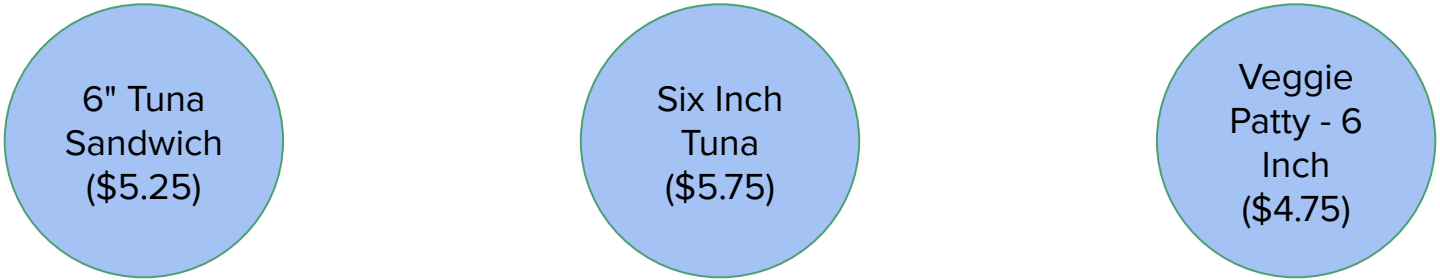
Cost-Effective Data Consolidation for Business Insights

Problem

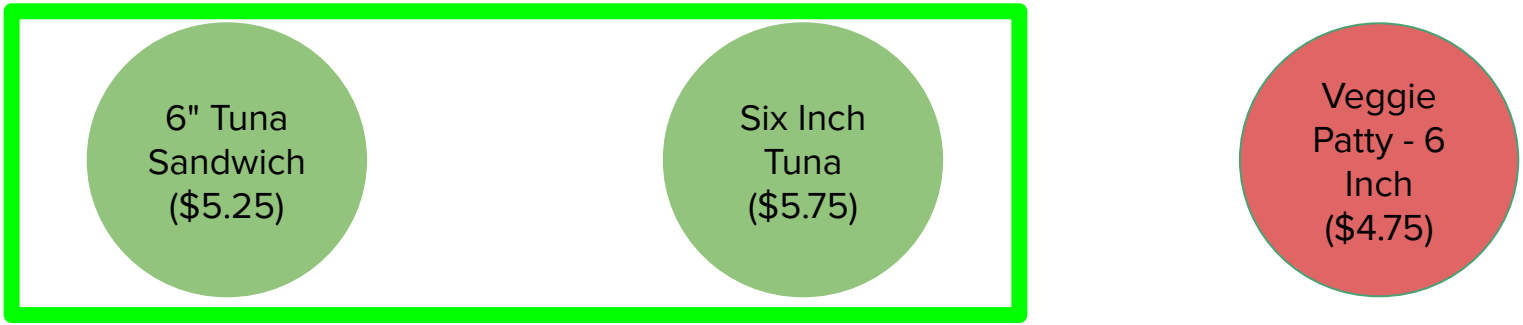
- A large-scale data aggregator collects vast amounts of pricing data with many variations of the same items across regions, making consolidation difficult

- Small errors in item grouping can lead to inaccurate business insights, such as flawed inflation or regional price analysis

Cost-Effective Data Consolidation for Business Insights



Cost-Effective Data Consolidation for Business Insights



Cost-Effective Data Consolidation for Business Insights

Problem

- A large-scale data aggregator collects vast amounts of pricing data with many variations of the same items across regions, making consolidation difficult

- Small errors in item grouping can lead to inaccurate business insights, such as flawed inflation or regional price analysis

Solution

- Provide a scalable, cost-effective AI solution to ensure accurate price comparisons and reliable business insights.

Cost-Effective Data Consolidation for Business Insights

Problem

- A large-scale data aggregator collects vast amounts of pricing data with many variations of the same items across regions, making consolidation difficult

- Small errors in item grouping can lead to inaccurate business insights, such as flawed inflation or regional price analysis

Solution

- Provide a scalable, cost-effective AI solution to ensure accurate price comparisons and reliable business insights.

Our current solution costs **100X** less than most accurate solutions with **1% less accuracy**

Relevant past/current use cases

1. Large Scale AI-Powered Document Processing and Quality Control

2. Cost-Effective Data Consolidation for Business Insights

3. Generative AI for Marketing Analytics and Insights

Outline of the presentation

1. Overview of challenges
2. Relevant past/current use cases
3. **Proposed solutions**
4. Benefits and Risks

Proposed solution: Challenge 1

Proposed solution: Challenge 1

Automatic Parsing of different datasets → Validation of Information Accuracy → Summarizes and Prioritizes Content

- Other previous analyses/studies

- Other programs (public or private) or laws (state or federal).

- Historical Finance recommendations

=> Similar AI-Powered Document Processing and Quality Control

Proposed solution: Challenge 1

Automatic Parsing of different datasets	Validation of Information Accuracy	Summarizes and Prioritizes Content
- Other previous analyses/studies - Other programs (public or private) or laws (state or federal). - Historical Finance recommendations *=> Similar AI-Powered Document Processing and Quality Control*	- Cross reference datasets - Ranking system *=> Similar to Data consolidation use case*	

Proposed solution: Challenge 1

Automatic Parsing of different datasets	Validation of Information Accuracy	Summarizes and Prioritizes Content
- Other previous analyses/studies	- Cross reference datasets	- Given a specific analyst issue, summarize and prioritize the content
- Other programs (public or private) or laws (state or federal).	- Ranking system	
- Historical Finance recommendations		-This step relies on the previous blocks
=> Similar AI-Powered Document Processing and Quality Control	*=> Similar to Data consolidation use case*	

Proposed solution: Challenge 2

3 **Strategic Recommendations**

Pattern recognition 2

1 **Data Aggregation**

Outline of the presentation

1. Overview of challenges
2. Relevant past/current use cases
3. Proposed solutions
4. **Benefits and Risks**

Benefits and risks of applying AI

Challenge	Feasibility	Risks	Benefits of Applying AI
1	Yes	Medium	High
2	Yes	High	Medium

Benefits and risks of applying AI

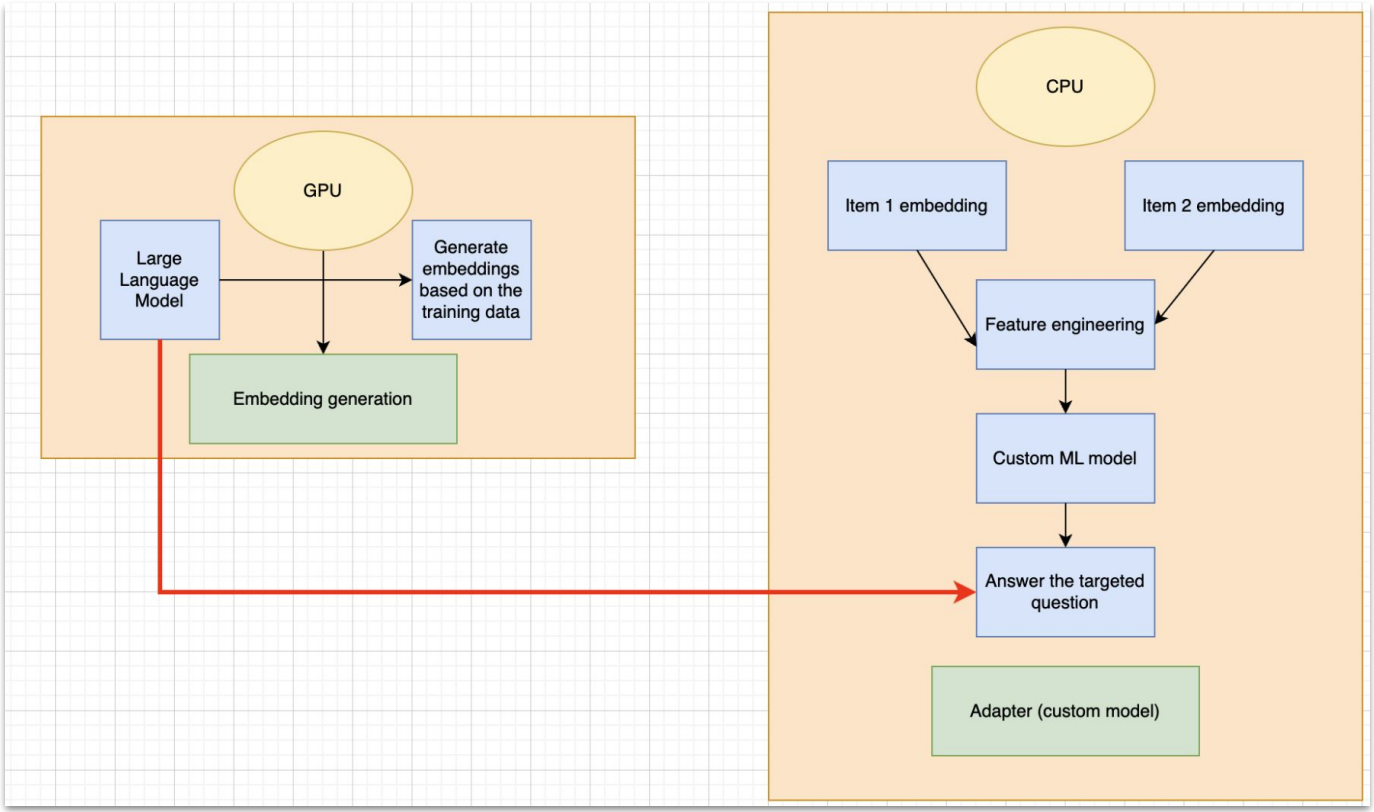
Challenge	Feasibility	Risks	Benefits of Applying AI
1	Yes	Medium	High
2	Yes	High	Medium

- **Challenge 1 risks:** Quantification tasks (like impact estimation) current GenAI solution are not good, need very custom approaches.
- **Challenge 2 risks**: The amount of data, e.g. hundreds of BCPs per year

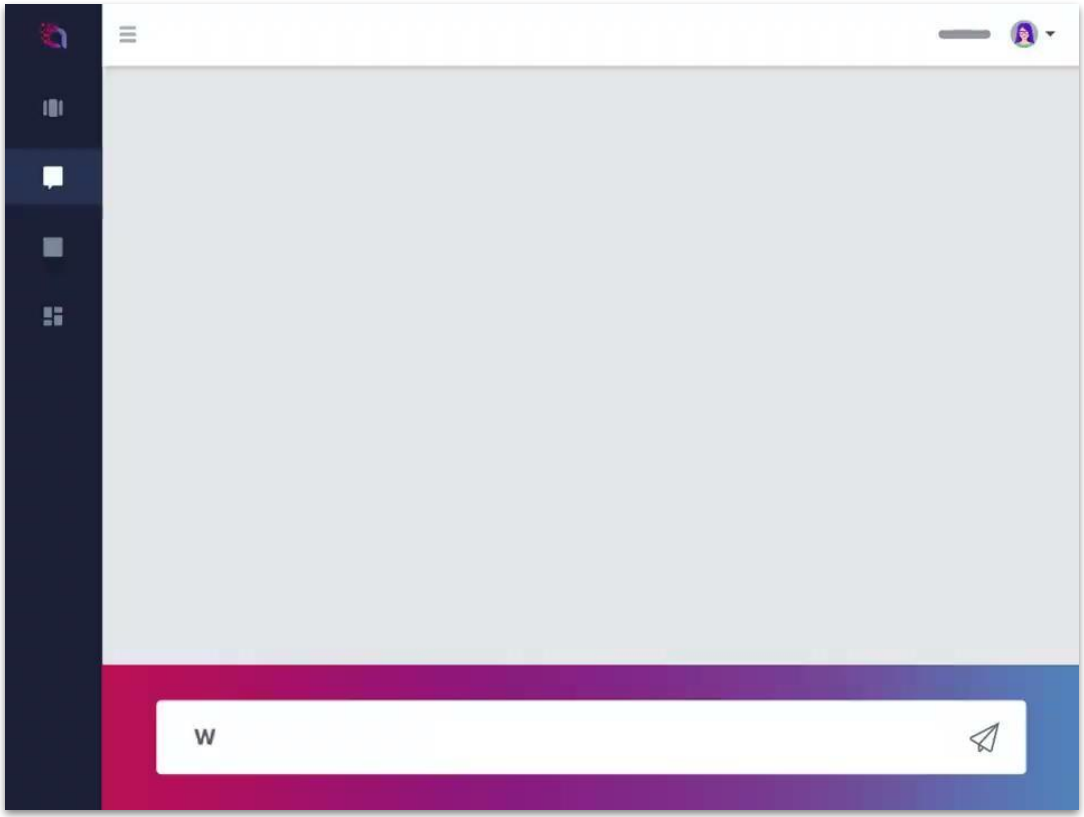
Q&A Session

Backup slides

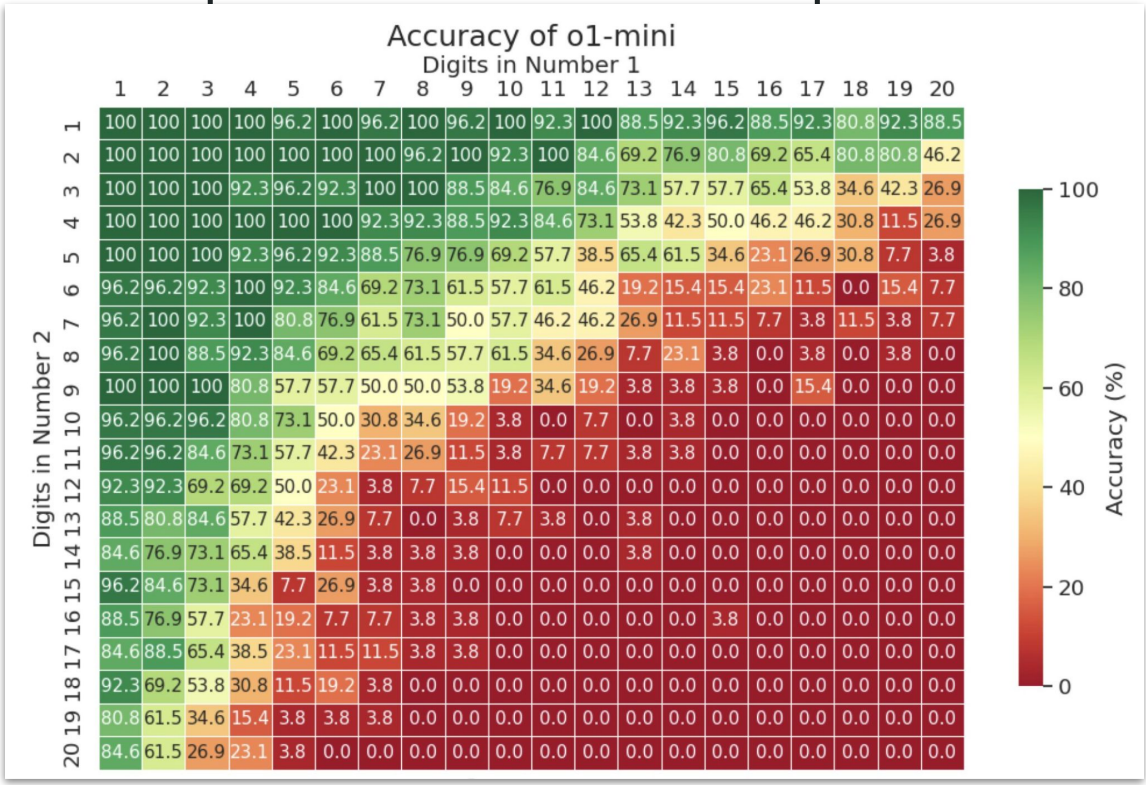
Cost-Effective Data Consolidation for Business Insights



Generative AI for Marketing Analytics and Insights

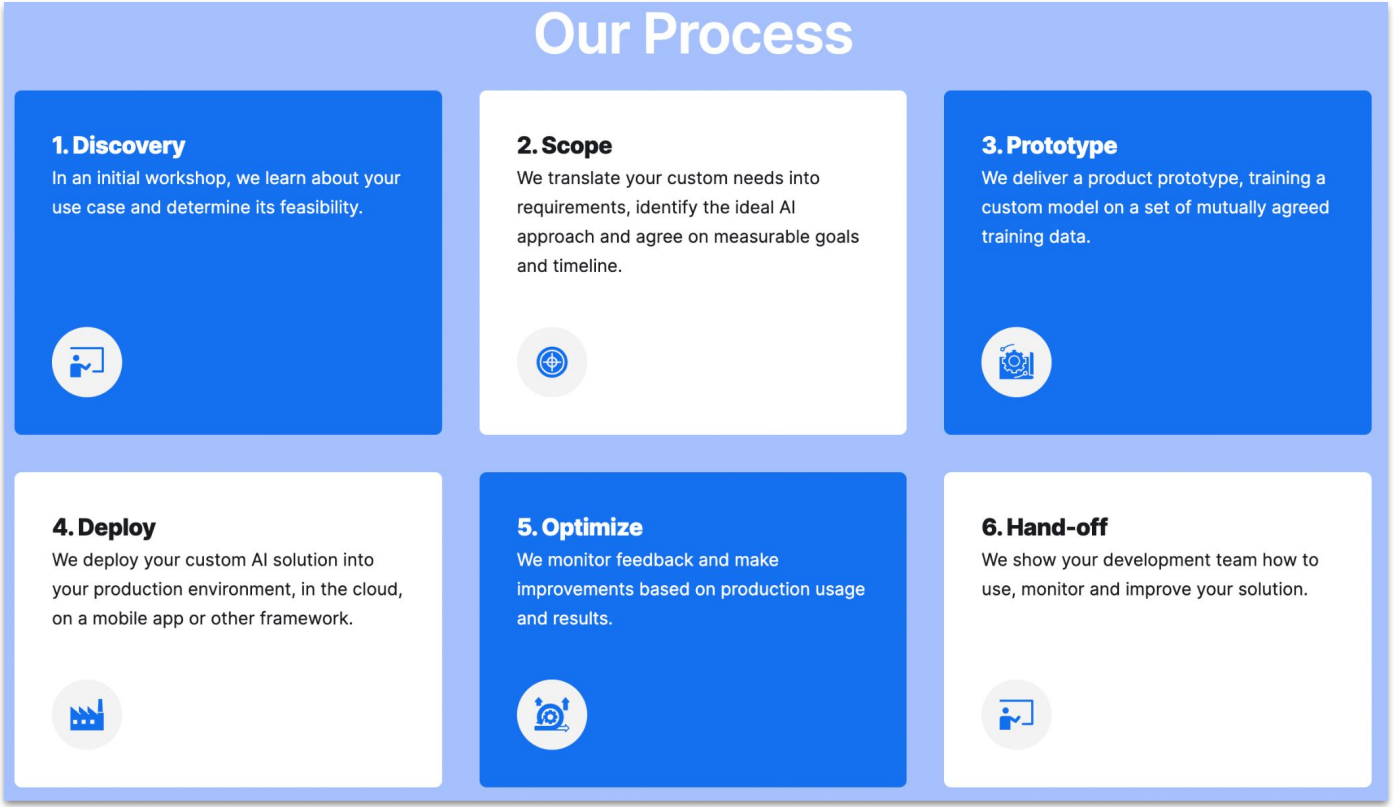


LLM performance at multiplication

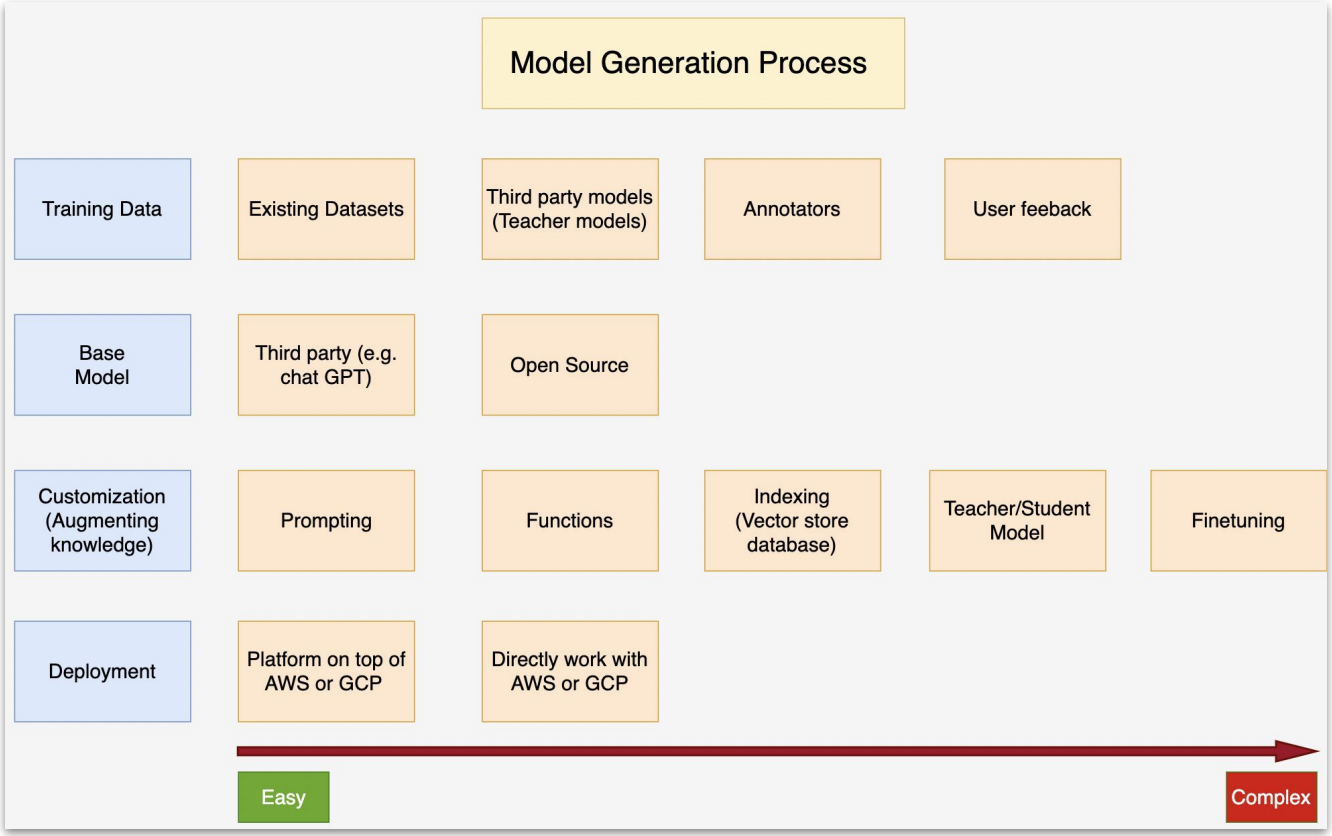


Credits to Prof. Yuntian Deng Link

My Custom AI: Process



My Custom AI: End to End model customization



My Custom AI Services

- **AI Feasibility Study**:
  - Identify if a use case is feasible using AI and the potential risks.
  - If it is feasible, identify different approaches and their tradeoffs.

- **AI Development & Deployment**:
  - Training Data generation and annotation
  - Model: Customization, measurement and optimization
  - Deployment on the partner infrastructure.

- **AI Professional Training**:
  - Workshops regarding a specific area in AI.
  - Brainstorming session on how AI can be applied internally.