

# Housing Prices Study

SDCTA

Bindu Priyanka Achalla

Vedant Singh

Kevin Liang

Tuo Jiao





# About the Team



**Bindu Priyanka  
Achalla**



**Vedant Singh**



**Tuo Jiao**



**Kevin Liang**

# About SDCTA

- Nonpartisan association promoting effective and efficient government for San Diego taxpayers.
- Engages civic leadership and the public to develop well-researched policy solutions, while aiming for better ROI on public assets and services for San Diego citizens.
- Advocates for policies reducing tax revenue needs and improving public service efficiencies.



## Mission / Vision

- Supports public, civic leaders, and policymakers with independent research.
- Focuses on economic and quality-of-life issues impacting taxpayers and ratepayers countywide.

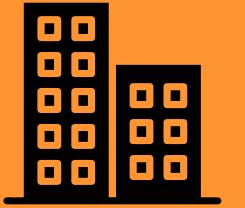
# Major Focus



**HOMELESSNESS**



**BEHAVIORAL HEALTH  
INITIATIVES**



**HOUSING**

# Project Overview

## Title: **Housing Prices Study**

- Taxpayer-funded school construction projects optimize student achievement.
- Effective and efficient project execution benefits students and taxpayers.
- Funded by municipal bonds repaid through property taxes.



## **Key Objectives**

- Develop a system to evaluate school construction projects against our rubric.
- Build a user-friendly interface to analyze the performance of bond programs.
- Investigate government policies affecting local housing prices and rental rates.

# Data Received from SDCTA

- RFQs/RFPs, agreements, and procurement documents for contractors and consultants.
- Change orders, budget documents, and financial information for construction projects.
- Board resolutions, meeting agendas, and minutes related to construction decisions.
- Project management records and governance documents for Fallbrook High School and Miracosta Community College.



# Key Performance Indicators (KPIs)

# KPI #1: Program / Project Budgets

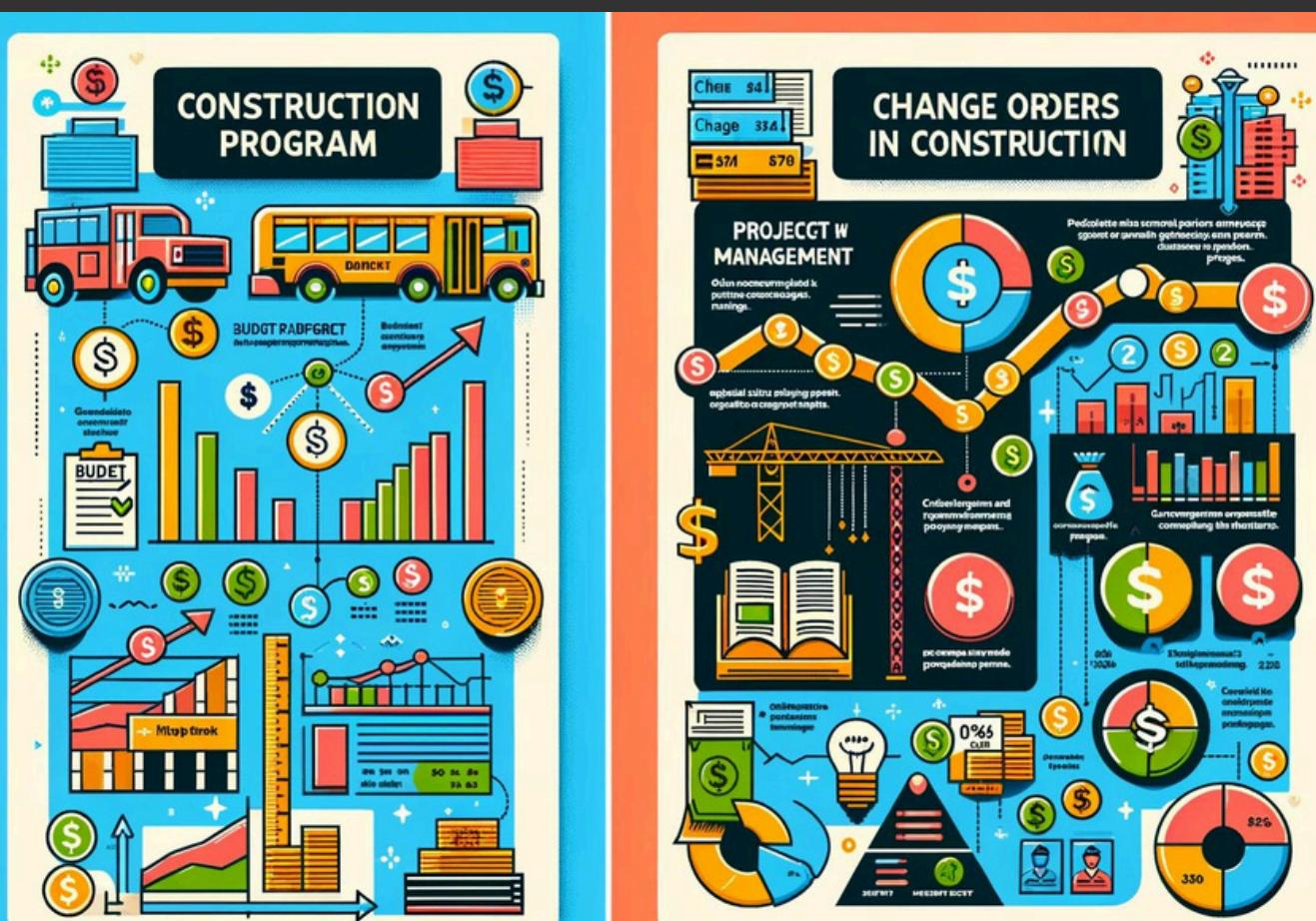
- Exceeds Standards: Construction Costs  $\geq$  80% of Budget, Comprehensive Cost Report
  - Needs Improvement: Construction Costs  $<$  70% of Budget, Inadequate Reporting

## KPI #2: Change Orders

- Exceeds Standards: Change Order Rate < 5%, Robust Processes
  - Needs Improvement: Change Order Rate > 8%, Poor Tracking

## KPI #3: Procurement of Contractors & Professional Services

- Exceeds Standards: Full Compliance, High Consultant Selection Criteria
  - Needs Improvement: Minimal Compliance, Inefficient Processes



# Key Performance Indicators (KPIs)

## KPI #4: Construction Management

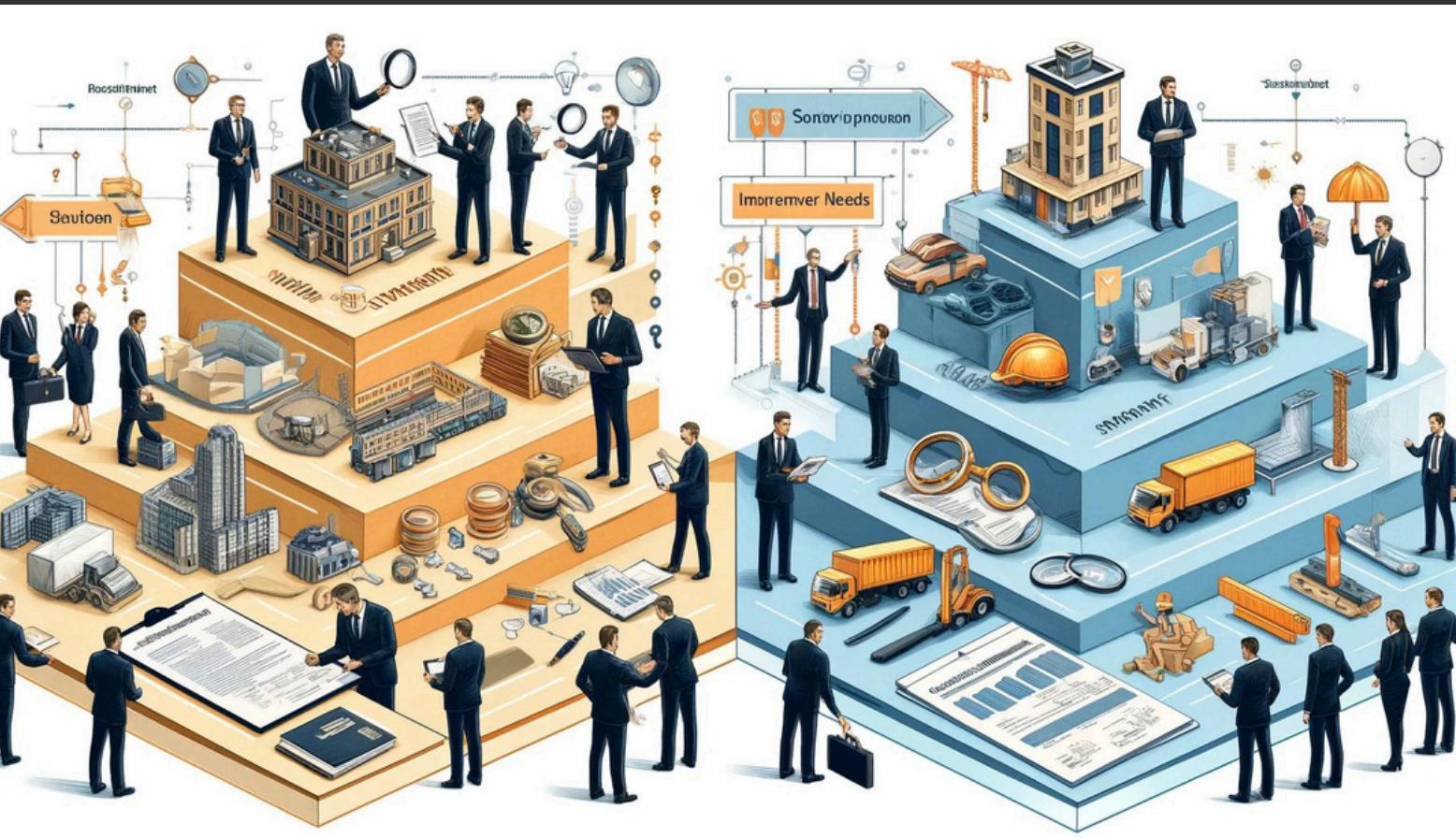
- Exceeds Standards: Manageable Workloads, Dedicated PMs
- Needs Improvement: Overburdened PMs, Lack of Oversight

## KPI #5: Overall Transparency

- Exceeds Standards: Comprehensive Reporting, Well-Documented CBOC Involvement
- Needs Improvement: Limited Documentation of Governance and Oversight

## KPI #6: Performance vs. Bond Criteria

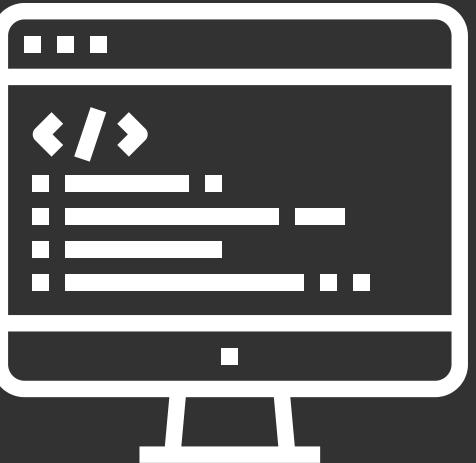
- Exceeds Standards: Full Alignment with Master Plan, Accurate Spending
- Needs Improvement: Poor Alignment, Significant Deviations in Execution



# Expected Deliverables

**The business desires a method/process/algorithm to:**

- Extract and evaluate quantitative data from construction project documents.
- Assess compliance with standards and measure overall project performance based on KPIs.
- Identify patterns and relationships in construction data to optimize project management.

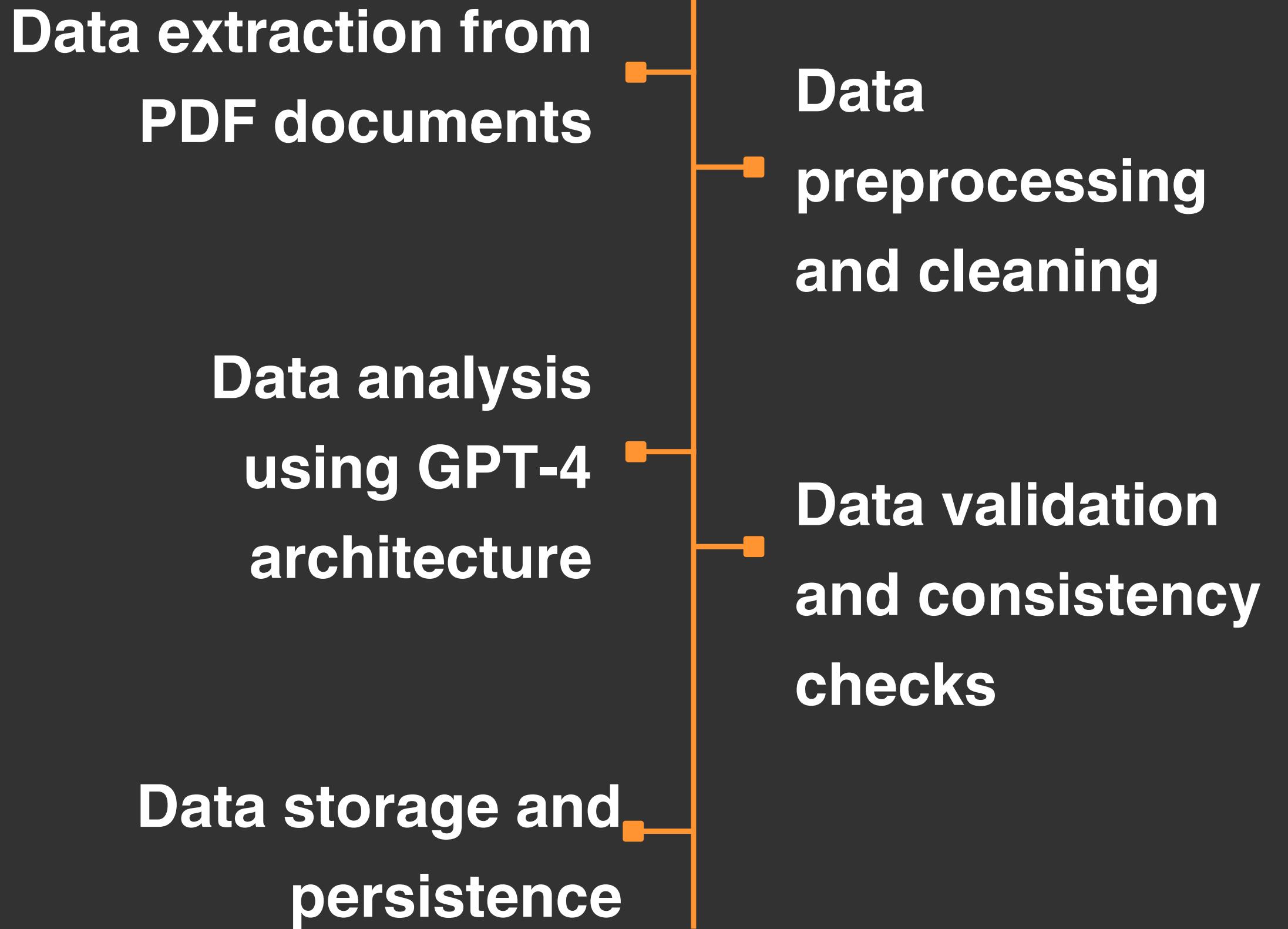


**The method the business will use to judge the usefulness of your solution:**

- The accuracy and comprehensiveness of KPI evaluations.
- The efficiency of data extraction and analysis processes.
- The ability to generate actionable insights for improving construction project outcomes.

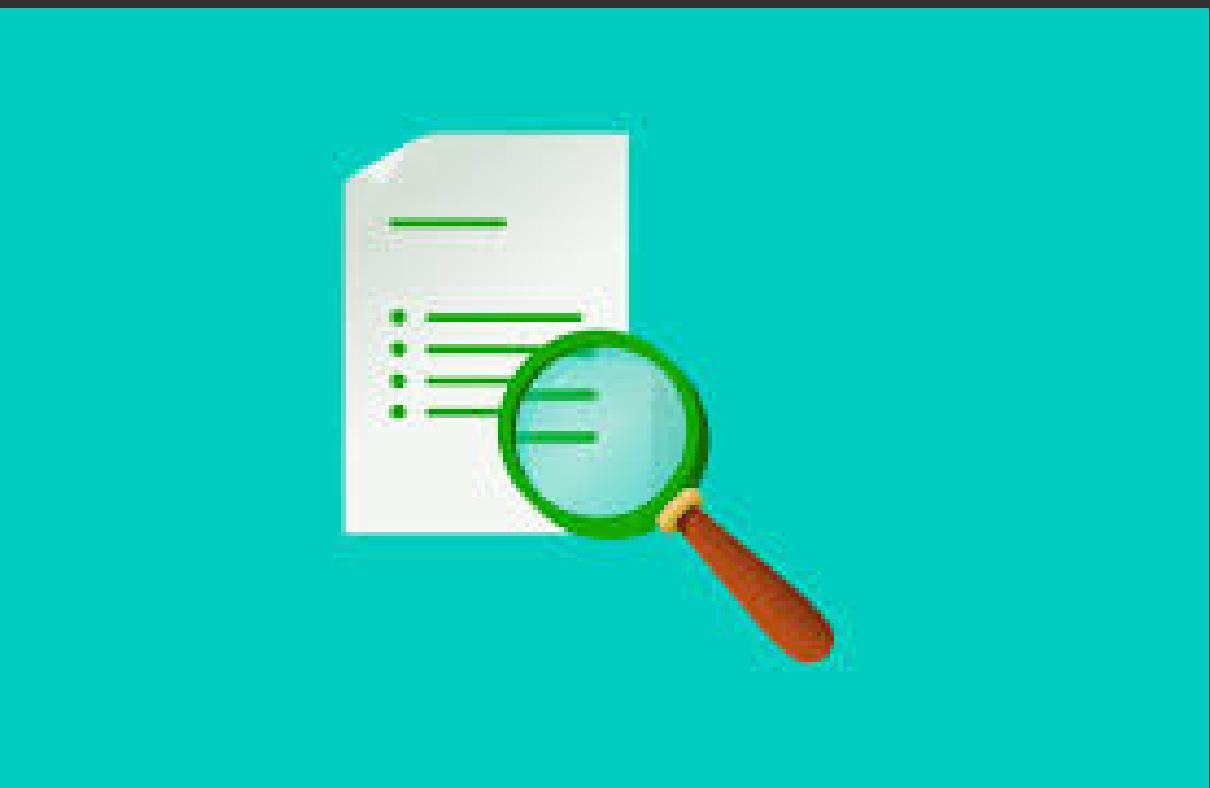


## Solution Path Overview



# Data Extraction, Preprocessing, and Analysis

- pdfplumber **library and extract\_text\_and\_images\_from\_pdf() function**
- **GPT-4 architecture and get\_info() function**
- **Natural Language Processing and valuable insights**



# Data Validation, Storage, and Deliverables

- **evaluate\_extraction\_consistency() function and consistency score**
- **GPT-4 model for independent verification**
- **JSON file storage and data persistence**
- **Key deliverables list**

# KPI Calculation, Evaluation, and Overall Rating

- **calculate\_kpi\_1() and calculate\_kpi\_2() functions**
- **evaluate\_kpi\_3(), evaluate\_kpi\_4(), and other evaluation functions**
- **calculate\_overall\_rating() function and overall project rating**
- **Quantitative and qualitative assessment of project performance**

# Project Insights and Future Outlook

- **integrate\_kpi\_results() function for generating final report**
- **Data-driven decision support**
- **Stakeholder engagement and feedback**
- **Continuous improvement and expansion to other projects**

# Open AI API:

```
import openai
import pdfplumber
import pytesseract
import pandas as pd
from PIL import Image
import warnings
from glob import gl "pytesseract": Unknown word. cSpell
from tqdm import tq (module) pytesseract
import os
import json
import re
import shutil
warnings.filterwarnings("ignore")

def extract_text_and_images_from_pdf(pdf_path):
    pages = []
    with pdfplumber.open(pdf_path) as pdf:
        for page in pdf.pages:
            image = page.to_image(resolution=300).original
            text = pytesseract.image_to_string(image)
            pages.append((image, text))
    return pages

openai.api_key = "sk-a3b51790b3534859af1631d60338b2f2"
openai.api_base = "https://api.deepseek.com/v1"
```

```
def get_info(page:str, prompt:str=''):
    response = openai.ChatCompletion.create(
        model="deepseek-chat",
        messages=[
            {"role": "system", "content": prompt},
            {"role": "user", "content": f"{page}"}
        ]
    )
    return response.choices[0].message['content']

prompt = """
    Imagine you are an intelligent assistant tasked with extracting all quantitative data from a construction p
    ...
"""

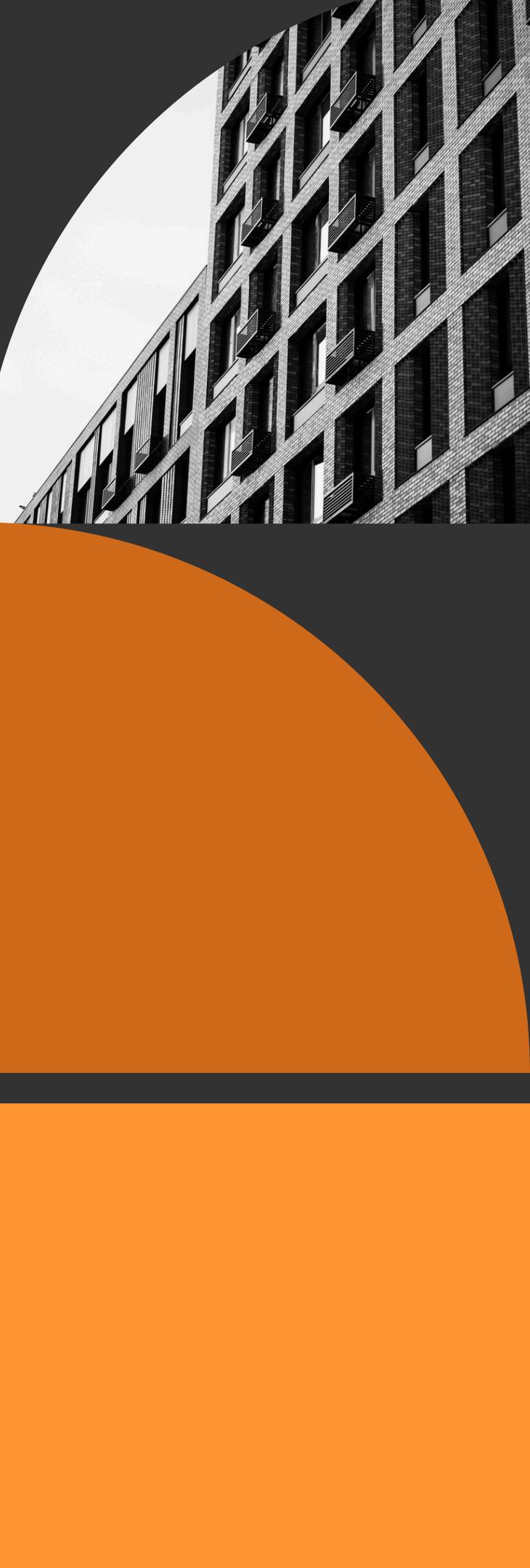
pdf_paths = glob("/Users/moonlet/Downloads/drive-download-20240507T231719Z-001/*.pdf")
results = []
for pdf_path in tqdm(pdf_paths):
    pages = extract_text_and_images_from_pdf(pdf_path)
    page_str = "\n".join([page[1] for page in pages])
    res = get_info(page_str, prompt=prompt)
    results.append(res)

clean_results = []
for i in range(len(results)):
    try:
        result = eval(results[i])
        if isinstance(result, dict):
            clean_results.append(result)
    except:
        if "```" in results[i]:
            extracted_dict = extract_dict_from_code_block(results[i])
            if isinstance(extracted_dict, dict):
                clean_results.append(extracted_dict)
```

Then you can talk about the previous codes, the problems of not being able to extract files, not finding keywords, etc. Then we use chatgpt to filter and extract keywords, and finally get our model. Through this AI big model, we can better extract keywords. This is our main problem. Later, it will be easy to calculate KPI according to the company's requirements.

# How Our Scoring System Helps Fallbrook High School

- Better Planning and Actions
- Saves Time
- Sets Standards for Future Projects



# Tips for Using the Scoring System

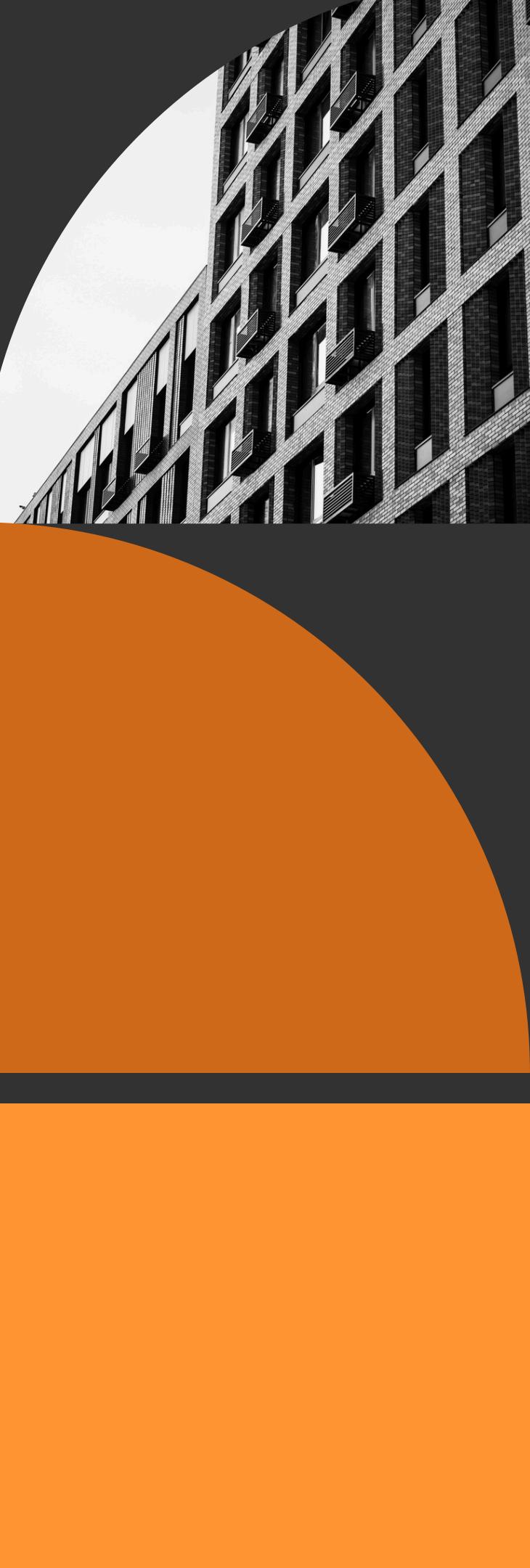
- In-depth Analysis for Low Scores: "Needs Improvement"
- Expand Use Cautiously



"Project Name": "Fallbrook High School",  
"KPI #1": "Needs Improvement",  
"KPI #2": "Needs Improvement",  
"KPI #3": "Exceeds Standards",  
"KPI #4": "Exceeds Standards",  
"KPI #5": "Exceeds Standards",  
"KPI #6": "Exceeds Standards",  
"Overall": "Outstanding"

## Business Recommendations Based on the Model

- Enhance Training and Support
- Policy and Procedure Updates



# Project Recap

- **Project Overview:** Automate evaluation to optimize taxpayer funds and improve outcomes
- **Achievements:** met "Exceeds Standards" in 4 of 6 KPIs, deemed "Outstanding"



**THANK  
YOU!**