

Power + Potential of Multimodal AI: Practical Examples

AI Hub

Christopher Martin

September 10 2024

New Orleans, LA USA



What is Multimodal AI?

Multimodal AI is capable of simultaneously processing information from multiple modalities, including images, videos, and text.¹

- 1) Image:** Knowledge-informed Q+A on an image
- 2) Audio:** Summarize an audio recording
- 3) Video:** Extract significant highlights from a video



Ground Rules

All examples demonstrated with publicly accessible platforms and models using personal computers and accounts.

Platform: Google Vertex AI Studio

Model: Gemini-1.5-pro-preview-0409

To maximize simplicity and reproducibility, all examples shown start with the preconfigured examples provided by Google.



Vertex AI

The Gemini logo, featuring the word 'Gemini' in a blue-to-purple gradient font with a purple starburst above the 'i'.

Gemini




Image

Knowledge-informed Q+A

01

Prompt CLEAR PROMPT INSERT MEDIA ?



You are an expert on answering questions briefly and accurately about the world.

Where is this mountain located? How long would it take the average person to climb it?

279 tokens **SUBMIT**

Response Markdown

The model will generate a response after you click Submit

Model
gemini-1.5-pro-preview-0409 ?

Region *
us-central1 (Iowa) ?

Temperature ?
0 2

Output token limit ?
0 8192

Add stop sequence ?
Press Enter after each sequence

SAFETY SETTINGS

> **Advanced**

RESET PARAMETERS

i We want your [feedback](#).

VIEW SERVICE DISCLAIMERS

REPORT INAPPROPRIATE RESPONSES



Audio

Summarize a recording

02

Prompt

CLEAR PROMPT INSERT MEDIA ?

audio_summary_clean_...

Please provide a summary for the audio, and tell me how many speakers are present

3,008 tokens SUBMIT

Response

Markdown

The model will generate a response after you click Submit

Model gemini-1.5-pro-preview-0409

Region * us-central1 (Iowa)

Temperature 0 2 1

Output token limit 0 8192 8192

Add stop sequence Press Enter after each sequence

SAFETY SETTINGS

Advanced

RESET PARAMETERS

We want your feedback

VIEW SERVICE DISCLAIMERS REPORT INAPPROPRIATE RESPONSES



Video

Extract significant highlights

03

Prompt CLEAR PROMPT INSERT MEDIA ?

google_sustainability.m... Provide a description of the video. The description should also contain anything important which people say in the video. Include timestamps.

70,231 tokens SUBMIT

Response Markdown

The model will generate a response after you click Submit

Model gemini-1.5-pro-preview-0409 ?

Region * us-central1 (lowa) ?

Temperature ? 0 2 1

Output token limit ? 0 8192 8192

Add stop sequence ?
Press Enter after each sequence

SAFETY SETTINGS

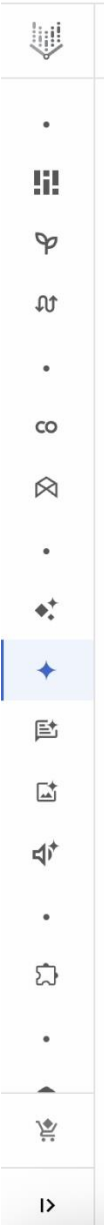
> Advanced

RESET PARAMETERS

i We want your [feedback](#).

VIEW SERVICE DISCLAIMERS

REPORT INAPPROPRIATE RESPONSES



A Practical Example: Reconciliation

Definition

Bank reconciliation is the process that companies use to make sure that cash balances they show on their books match the actual cash they have in the bank.

NOTE: The following examples are used for concept demonstration purposes only and are not intended to represent BNYM procedures.

Example 1: General Theory

Results provided by a ChatGPT-like text query

Example 2: Tailored Instructions

Detailed insights observed from video of complex tasks





General Theory

...

01

- TOOLS
 - Dashboard
 - Model Garden
 - Pipelines
- NOTEBOOKS
 - Colab Enterprise
 - Workbench
- VERTEX AI STUDIO
 - Overview
 - Multimodal
 - Language**
 - Vision
 - Speech
- BUILD WITH GEN AI
 - Extensions
- DATA
 - Marketplace

Prompt CLEAR PROMPT

You are a business analyst documenting processes for bank reconciliation. Bank reconciliation is the process of comparing your company's bank statements to your own records, ensuring all transactions are accounted for.

Your task is to create a detailed set of instructions for bank reconciliation:

- * List reconciliation actions
- * Give specific instructions on how to perform each action
- * Explain why the user should perform these actions

78 tokens **SUBMIT**

Response Markdown

The model will generate a response after you click Submit

Model
gemini-1.5-pro-preview-0409

Region *
us-central1 (Iowa)

Temperature ?
0 ————— 2

Output token limit ?
0 ————— 8192

Add stop sequence ?
Press Enter after each sequence

SAFETY SETTINGS

> Advanced

RESET PARAMETERS

i We want your [feedback](#).

VIEW SERVICE DISCLAIMERS

REPORT INAPPROPRIATE RESPONSES



Tailored Instructions

...

02

BANK RECONCILIATION

[Company Name]

Date: Sep-30-2019

Bank Statement Adjustments

Balance as at Sep-30-2019 \$1,036,141.66 \$1,036,141.66 \$0.00

Add Deposits in Transit

| Batch-Entry | Deposit Date | Description | Reference | Amount |
|-------------|--------------|---------------------|--------------------------|-------------|
| 4092-08 | Jul-30-2019 | CREDIT CARD PAYMENT | AR Receipt Batch #: 2690 | \$162.97 |
| 4092-08 | Jul-30-2019 | CREDIT CARD PAYMENT | AR Receipt Batch #: 2691 | \$106.75 |
| 4137-12 | Jul-30-2019 | DEPOSIT | AR Receipt Batch #: 2713 | \$2,454.14 |
| 4138-12 | Jul-30-2019 | DEPOSIT | AR Receipt Batch #: 2712 | \$2,285.06 |
| 4141-13 | Jul-30-2019 | DEPOSIT | AR Receipt Batch #: 2726 | \$10,680.41 |
| 4142-13 | Jul-30-2019 | DEPOSIT | AR Receipt Batch #: 2727 | \$3,972.48 |
| 4147-13 | Jul-30-2019 | CREDIT CARD PAYMENT | AR Receipt Batch #: 2724 | \$162.97 |
| 4148-13 | Jul-30-2019 | CREDIT CARD PAYMENT | AR Receipt Batch #: 2725 | \$482.40 |

Total Deposits in Transit \$20,307.18

Less Outstanding Checks/Wires

| Batch-Entry | Check/Wire Date | Description | Reference | Check Number | Amount |
|-------------|-----------------|-------------|----------------|--------------|-------------|
| 2808-74 | May/16/2019 | VENDOR 851 | Pr000000465429 | 44654 | -\$703.23 |
| 3491-74 | May/05/2019 | VENDOR 887 | PY000000465429 | 44717 | -\$715.33 |
| 3419-41 | May/09/2019 | VENDOR 273 | PY000000465426 | 44943 | -\$1,414.98 |
| 2419-41 | May/08/2019 | VENDOR 409 | Pr000000465425 | 44934 | -\$299.98 |
| 2463-45 | Jun/22/2019 | VENDOR 103 | PY000000465426 | 44987 | -\$164.13 |
| 2469-46 | Jun/20/2019 | VENDOR 543 | PY000000465434 | 44574 | -\$155.83 |
| 2631-62 | Jun/06/2019 | VENDOR 353 | Pr000000465435 | 45034 | -\$109.98 |
| 2631-62 | Jun/06/2019 | VENDOR 509 | PY000000465433 | 54564 | -\$355.61 |
| 3126-11 | Jun/04/2019 | VENDOR 692 | PY000000465435 | 45025 | -\$433.58 |
| 3250-24 | Jun/08/2019 | VENDOR 486 | Pr000000465434 | 45272 | -\$219.98 |
| 3683-67 | Jun/26/2019 | VENDOR 376 | PY000000565440 | 45487 | -\$1,879.98 |
| 4112-10 | Jun/21/2019 | VENDOR 642 | PY000000565431 | 45608 | -\$245.25 |
| 4112-10 | Jun/21/2019 | VENDOR 202 | Pr000000565438 | 45605 | -\$137.41 |
| 4112-10 | Jun/21/2019 | VENDOR 011 | PY000000565440 | 45604 | -\$562.45 |
| 4112-10 | Jun/21/2019 | VENDOR 257 | PY000000565441 | 45599 | -\$2,892.25 |
| 4112-10 | Jun/21/2019 | VENDOR 087 | Pr000000565437 | 45600 | -\$46.25 |
| 4112-10 | Jul/21/2019 | VENDOR 479 | PY000000565436 | 45614 | -\$453.25 |
| 4112-10 | Jul/25/2019 | VENDOR 232 | PY000000565441 | 45616 | -\$405.25 |
| 4112-10 | Jul/25/2019 | VENDOR 787 | Pr000000565440 | 45619 | -\$805.25 |
| 4112-10 | Jul/14/2019 | VENDOR 473 | PY000000565446 | 45585 | -\$791.09 |
| 4112-10 | Jul/21/2019 | VENDOR 107 | PY000000565441 | 45611 | -\$453.25 |
| 4112-10 | Jul/24/2019 | VENDOR 727 | Pr000000565442 | 45621 | -\$1,750.21 |

Vertex AI

Recon Exp 10 Mins v1 HISTORY NOTES API REFERENCE

SAVE DELETE GET CODE

- TOOLS
 - Dashboard
 - Model Garden
 - Pipelines
- NOTEBOOKS
 - Colab Enterprise
 - Workbench
- VERTEX AI STUDIO
 - Overview
 - Multimodal
 - Language
 - Vision
 - Speech
- BUILD WITH GEN AI
 - Extensions
- DATA
 - Marketplace

Prompt CLEAR PROMPT INSERT MEDIA ?

recon-10-mins-720p.mov

You are a business analyst documenting processes for bank reconciliation. Bank reconciliation is the process of comparing your company's bank statements to your own records, ensuring all transactions are accounted for. The video is of an expert using Microsoft Excel to reconcile transactions across a spreadsheet. Your task is to use this video to create a detailed set of instructions for bank reconciliation using only examples included in this video:

- * List the reconciliation actions shown in the video
- * Give specific instructions on how to perform each action
- * Explain why the user should perform these actions
- * List start and end times in the video where each of these actions occurred
- * Write the instructions in the form of an executable script

249,119 tokens **SUBMIT**

Response Markdown

The model will generate a response after you click Submit

Model gemini-1.5-pro-preview-0409 ?

Region * us-central1 (Iowa) ?

Temperature ? 0

Output token limit ? 8192

Add stop sequence ?
Press Enter after each sequence

SAFETY SETTINGS

> Advanced

RESET PARAMETERS

i We want your [feedback](#).

VIEW SERVICE DISCLAIMERS

REPORT INAPPROPRIATE RESPONSES

Preparation:

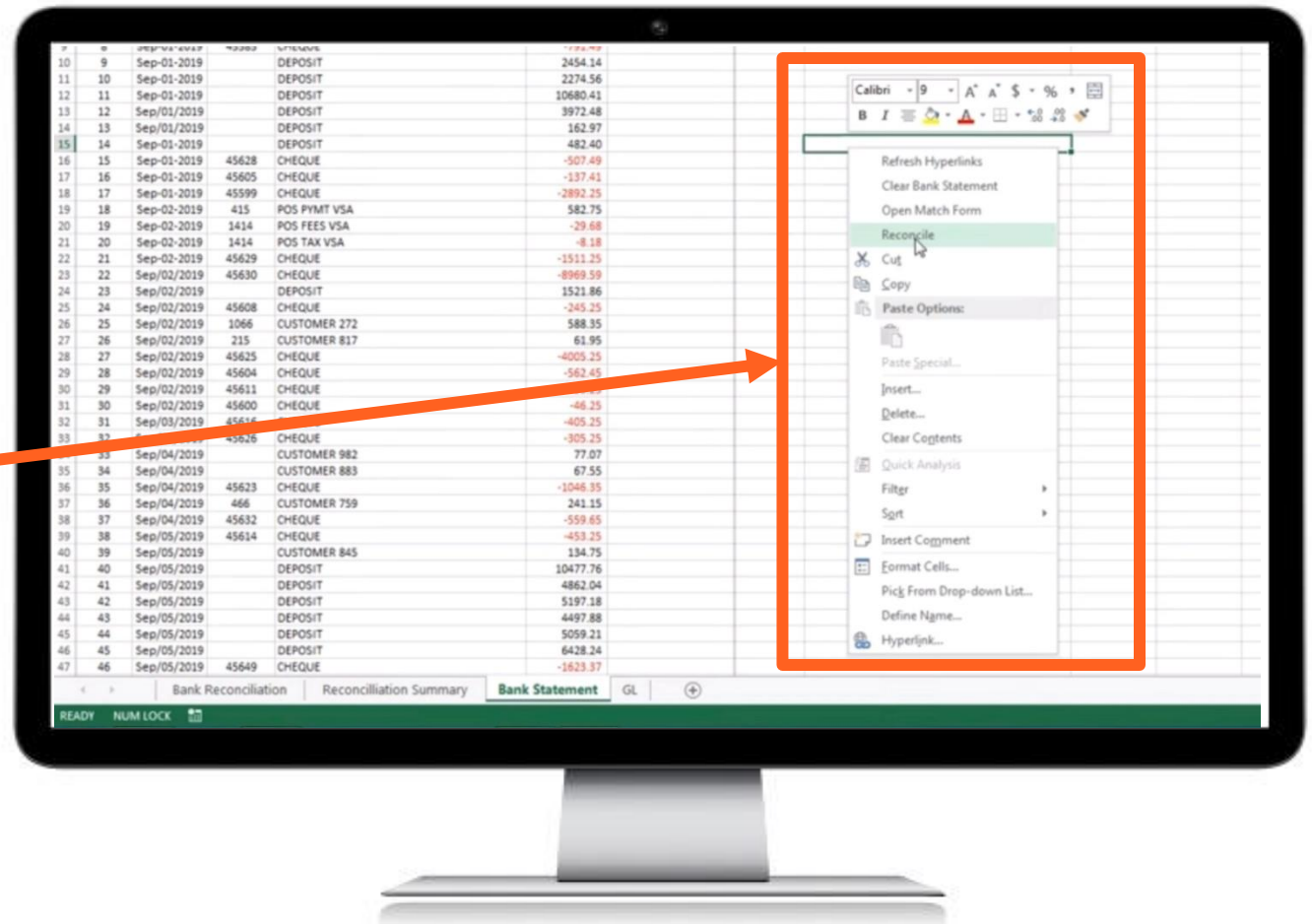
1. Open your bank statement and general ledger data in Excel: Ensure both datasets are for the same period.
2. (00:26 - 00:31) Prepare your reconciliation worksheet: This sheet should have columns for date, transaction description, amount, and a status column (e.g., "Reconciled," "Outstanding Check," "Deposit in Transit," "Difference"). You may also include additional columns for notes, references, etc.

Automatic Reconciliation (01:41 - 02:04):

1. Select all transactions in both the bank statement and general ledger tabs.
2. Right-click and choose "Reconcile." This triggers Excel's built-in logic to match transactions based on specific criteria (likely amount and date).
3. Review the "Reconciliation Status" column: Matched transactions will be marked as "Reconciled."

Manual Reconciliation (02:22 - 04:50):

1. Filter the "Reconciliation Status" column to show only blank entries. These represent transactions that Excel couldn't automatically match.
2. (02:56 - 03:39) Analyze unmatched transactions:
 - Identify potential errors: Look for duplicate entries or discrepancies in amounts.
 - Investigate missing transactions: Verify if transactions are missing from either the bank statement or the general ledger.

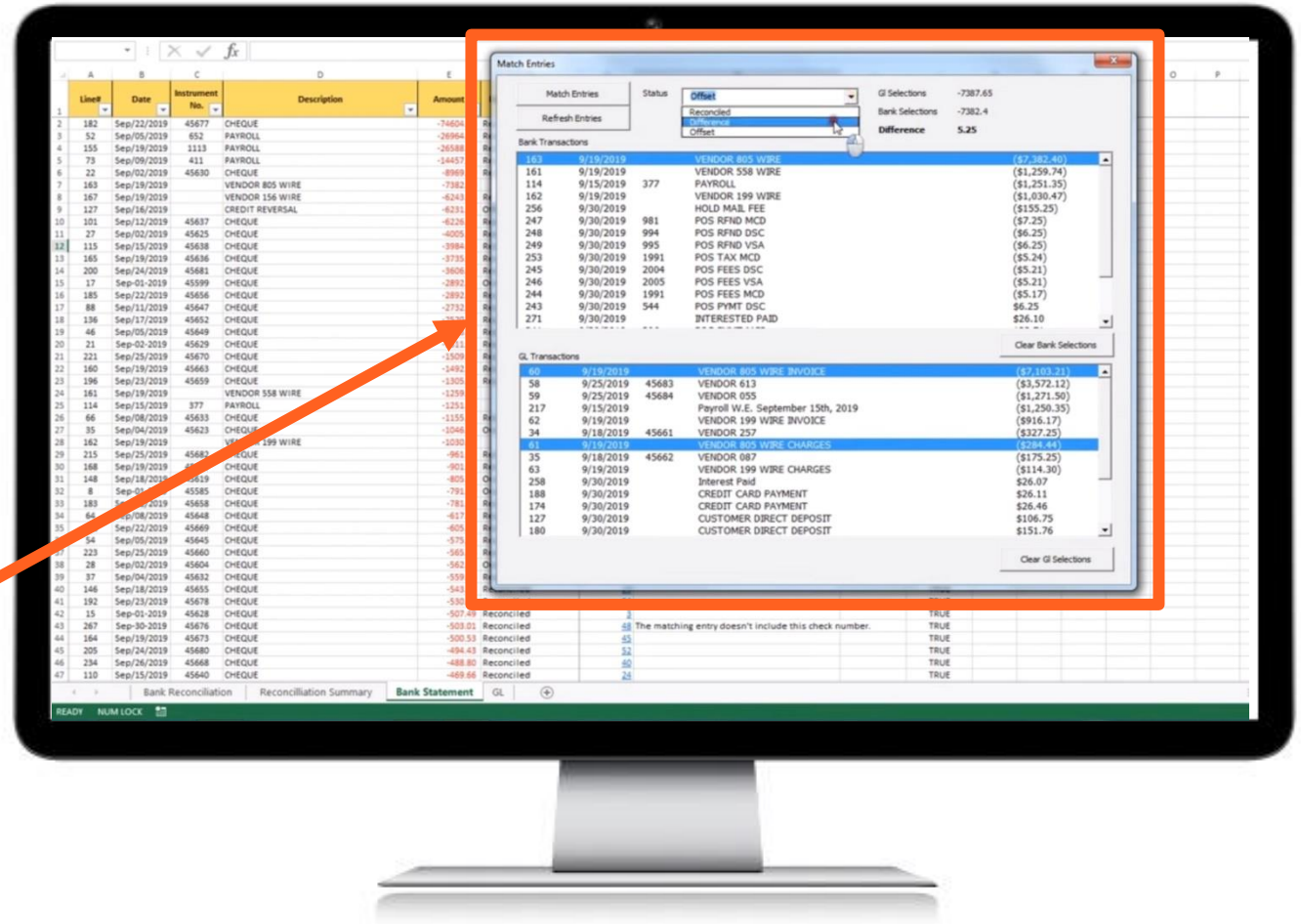


4. (05:40 - 07:51) Within the "Match Form":

- Review unmatched transactions from both the bank statement and the general ledger.
- Identify matching pairs based on additional information (e.g., check number, description).
- Select the matching pairs and choose the appropriate status:
 - a. "Reconciled": Marks the pair as matched and removes them from the "Match Form."
 - b. "Offset": Marks the pair as offsetting each other and removes them from the "Match Form."
 - c. "Difference": Flags the pair as having a discrepancy requiring further investigation.

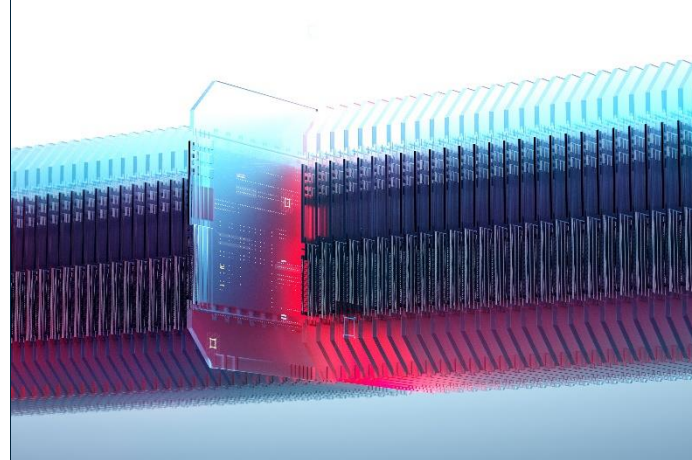
Review and Finalization (07:51 - 13:55):

1. (07:51 - 08:46) Filter the "Reconciliation Status" column to show only entries marked as "Difference." Investigate these discrepancies to identify the cause and make necessary corrections in your records.



The future... but when?

Multimodal models already demonstrate enormous power and potential, but are not without their own limitations.



Advantages

- Powerful
- Expressive
- Evolving

Drawbacks

- Expensive
- Slow
- Not Mature



ChatGPT is mildly embarrassing
at best ...

GPT-4 is the dumbest model any
of you will ever have to use
again, by a lot.

Sam Altman

CEO, OpenAI



