

## README: Question 6 ClinicalTrialDataAgent

```
## Q6 - AI Clinical Trial Data Agent

**Purpose:** Column-agnostic AI assistant that turns natural-language questions about the ADAE dataset into filters and returns matching subjects. Supports multi-criteria questions (e.g. "Who had severe events involving Pruritus?"). Uses OpenAI (LangChain) and optional pharmaverse AE data fallback.

**Requirements:** OPENAI_API_KEY must be set in the environment (e.g. export OPENAI_API_KEY='your-key'). Without it, the agent exits with instructions to set the key.

### Contents

| Path | Description |
|-----|-----|
| `ADAE_ClinicalTrialDataAgent.py` | Main agent: loads ADAE (local CSV or pharmaverse ae.rda), parses questions to JSON filters, applies filters (AND), optional spelling/value fallback |
| `test_agent_queries.py` | Runs three example queries: "Who died?", "Who had fractures?", "Who had severe events involving cancer?" |
| `requirements.txt` | pandas, langchain-openai, langchain-core, pyreadr (for pharmaverse .rda fallback) |
| `data/adae.csv` | Optional. If missing, downloads and uses pharmaverse ae.rda and notifies the user. |

### How to run

Set OPENAI_API_KEY (required), then run the agent or test script:

```bash
cd Q6
export OPENAI_API_KEY='your-key'    # required
pip install -r requirements.txt
python ADAE_ClinicalTrialDataAgent.py    # Interactive CLI
# or
python test_agent_queries.py          # Run the three example queries
```

- If data/adae.csv is not present, the script uses the pharmaverse default AE file and prints a note to stderr.
```

Download the entire Q6 folder from git.

To use this agent, **set your OpenAI API key.**

Option 1 - Export in your shell (bash/zsh):

```
export OPENAI_API_KEY='your-api-key-here'
```

Option 2 - In the same terminal before running:

```
OPENAI_API_KEY='your-api-key-here'
```

```
python ADAE_ClinicalTrialDataAgent.py
```

Get an API key at: <https://platform.openai.com/api-keys>

cd <"to the Q6 folder with code ADAE\_ClinicalTrialDataAgent.py">

**Install requirements:**

pip install -r requirements.txt

**RUNNING the test script, test\_agent\_queries.py.**

At a terminal prompt execute the following:

**python test\_agent\_queries.py**

```
● (base) ambika@ambikas-mbp Q6 % python test_agent_queries.py
Running 3 example queries with ClinicalTrialDataAgent
=====
--- Query 1: Who died? ---
Parsed filters: 1 criterion/criteria
  1. AESDTH equals 'Y'
Unique subjects: 3
Subjects: ['01-701-1211', '01-704-1445', '01-710-1083']

--- Query 2: Who had fractures? ---
Parsed filters: 1 criterion/criteria
  1. AETERM contains 'fracture'
Unique subjects: 4
Subjects: ['01-709-1238', '01-710-1006', '01-710-1070', '01-710-1368']

--- Query 3: Who had severe events involving cancer? ---
Parsed filters: 2 criterion/criteria
  1. AESEV equals 'SEVERE'
  2. AETERM contains 'cancer'
Unique subjects: 2
Subjects: ['01-705-1393', '01-716-1189']

=====
Done.
○ (base) ambika@ambikas-mbp Q6 % █
```

**RUNNING the agent:**

1. At a terminal prompt execute the following:
  - 1) **python ADAE\_ClinicalTrialDataAgent.py**
2. Ask questions.

Suggested questions:

- 1) Patients with fractures
- 2) Who died? (sudden death and also implying death as in suicide and myocardial infarction)
- 3) Who had diarrhea? (spelt DIARRHOEA in AETERM)
- 4) Who had severe events?
- 5) Who had severe events involving cancer? (AND clause)
- 6) Who had severe events within 15 days of study?

Example outputs:

```
(base) ambika@ambikas-mbp Q6 % python ADAE_ClinicalTrialDataAgent.py
ClinicalTrialDataAgent Ready

Ask a question (or type exit): █
```

```
(base) ambika@ambikas-mbp Q6 % python ADAE_ClinicalTrialDataAgent.py
ClinicalTrialDataAgent Ready

Ask a question (or type exit): patients with fractures
=====
patients with fractures
=====
USUBJID      AETERM      AELLT      AESEV      AESTDY     AESEQ      AEDECOD
01-709-1238   HIP FRACTURE  HIP FRACTURE MODERATE  17.0       5          HIP FRACTURE
01-710-1006   FACIAL BONES FRACTURE  FRACTURED NOSE MODERATE 180.0      11         FACIAL BONES FRACTURE
01-710-1070   HIP FRACTURE  HIP FRACTURE  SEVERE    137.0      5          HIP FRACTURE
01-710-1368   HIP FRACTURE  FRACTURED HIP  SEVERE    64.0       1          HIP FRACTURE
01-710-1368   HIP FRACTURE  FRACTURED HIP  SEVERE    64.0       4          HIP FRACTURE
=====

Unique subjects (sorted): 4
-----
01-709-1238
01-710-1006
01-710-1070
01-710-1368
-----

Ask a question (or type exit): █
```

Ask a question (or type exit): who had headache?

```
=====  
who had headache?  
=====
```

| USUBJID     | AETERM   | AELLT                 | AESEV    | AESTDY | AESEQ | AEDECOD  |
|-------------|----------|-----------------------|----------|--------|-------|----------|
| 01-701-1180 | HEADACHE | HEADACHE              | MODERATE | Nan    | 4     | HEADACHE |
| 01-701-1302 | HEADACHE | HEADACHE              | MILD     | 18.0   | 13    | HEADACHE |
| 01-701-1363 | HEADACHE | HEADACHE              | MODERATE | Nan    | 2     | HEADACHE |
| 01-701-1363 | HEADACHE | HEADACHE              | MILD     | Nan    | 4     | HEADACHE |
| 01-701-1383 | HEADACHE | FULLNESS HEAD         | MILD     | 141.0  | 11    | HEADACHE |
| 01-704-1388 | HEADACHE | HEADACHE              | MODERATE | -106.0 | 1     | HEADACHE |
| 01-708-1019 | HEADACHE | FRONTAL HEADACHE      | SEVERE   | 12.0   | 1     | HEADACHE |
| 01-708-1087 | HEADACHE | HEADACHE              | MILD     | 118.0  | 4     | HEADACHE |
| 01-708-1272 | HEADACHE | HEADACHE              | MILD     | 30.0   | 7     | HEADACHE |
| 01-708-1272 | HEADACHE | HEADACHE              | MILD     | 30.0   | 10    | HEADACHE |
| 01-708-1316 | HEADACHE | HEADACHE              | MILD     | 160.0  | 1     | HEADACHE |
| 01-709-1217 | HEADACHE | HEADACHE              | MILD     | 75.0   | 9     | HEADACHE |
| 01-709-1339 | HEADACHE | HEADACHE              | MODERATE | Nan    | 1     | HEADACHE |
| 01-710-1137 | HEADACHE | HEADACHE              | MILD     | 34.0   | 4     | HEADACHE |
| 01-710-1142 | HEADACHE | HEADACHE              | MILD     | 3.0    | 2     | HEADACHE |
| 01-710-1315 | HEADACHE | HEADACHE              | MILD     | -1.0   | 1     | HEADACHE |
| 01-711-1036 | HEADACHE | INTERMITTENT HEADACHE | MILD     | 1.0    | 1     | HEADACHE |
| 01-716-1418 | HEADACHE | HEADACHE              | MODERATE | Nan    | 5     | HEADACHE |
| 01-716-1418 | HEADACHE | HEADACHE              | MODERATE | Nan    | 7     | HEADACHE |
| 01-716-1418 | HEADACHE | HEADACHE              | MILD     | 145.0  | 9     | HEADACHE |
| 01-716-1418 | HEADACHE | HEADACHE              | MILD     | 145.0  | 10    | HEADACHE |

Unique subjects (sorted): 16

01-701-1180  
01-701-1302  
01-701-1363  
01-701-1383  
01-704-1388  
01-708-1019  
01-708-1087  
01-708-1272  
01-708-1316  
01-709-1217  
01-709-1339  
01-710-1137  
01-710-1142  
01-710-1315  
01-711-1036  
01-716-1418

Ask a question (or type exit): ■

Ask a question (or type exit): Which subjects had diarrhea?

=====

Which subjects had diarrhea?  
(No exact AETERM match; used closest value from data: "DIARRHOEA")

```
=====
```

| USUBJID     | AETERM    | AELLT        | AESEV    | AESTDY | AESEQ | AEDECOD   |
|-------------|-----------|--------------|----------|--------|-------|-----------|
| 01-701-1015 | DIARRHOEA | DIARRHEA     | MILD     | 8.0    | 3     | DIARRHOEA |
| 01-701-1387 | DIARRHOEA | DIARRHEA     | MILD     | 7.0    | 1     | DIARRHOEA |
| 01-701-1415 | DIARRHOEA | DIARRHEA     | MILD     | 168.0  | 6     | DIARRHOEA |
| 01-703-1042 | DIARRHOEA | DIARRHEA     | MILD     | 3.0    | 1     | DIARRHOEA |
| 01-703-1403 | DIARRHOEA | DIARRHEA     | MODERATE | 2.0    | 2     | DIARRHOEA |
| 01-704-1010 | DIARRHOEA | DIARRHEA     | MILD     | 135.0  | 3     | DIARRHOEA |
| 01-704-1065 | DIARRHOEA | DIARRHEA     | MILD     | 53.0   | 9     | DIARRHOEA |
| 01-704-1041 | DIARRHOEA | DIARRHEA     | MILD     | 8.0    | 2     | DIARRHOEA |
| 01-708-1272 | DIARRHOEA | DIARRHEA     | MILD     | 6.0    | 1     | DIARRHOEA |
| 01-709-1007 | DIARRHOEA | DIARRHEA     | MILD     | 31.0   | 4     | DIARRHOEA |
| 01-709-1020 | DIARRHOEA | DIARRHEA     | MILD     | 34.0   | 2     | DIARRHOEA |
| 01-709-1238 | DIARRHOEA | DIARRHEA     | MODERATE | 14.0   | 1     | DIARRHOEA |
| 01-709-1264 | DIARRHOEA | DIARRHEA     | MILD     | 21.0   | 1     | DIARRHOEA |
| 01-710-1195 | DIARRHOEA | LOOSE STOOLS | MILD     | 17.0   | 1     | DIARRHOEA |
| 01-710-1195 | DIARRHOEA | LOOSE STOOLS | MILD     | 17.0   | 6     | DIARRHOEA |
| 01-714-1195 | DIARRHOEA | DIARRHEA     | MILD     | 11.0   | 2     | DIARRHOEA |
| 01-716-1044 | DIARRHOEA | DIARRHEA     | MILD     | 25.0   | 1     | DIARRHOEA |
| 01-716-1044 | DIARRHOEA | DIARRHEA     | MILD     | 25.0   | 2     | DIARRHOEA |
| 01-717-1004 | DIARRHOEA | SOFT STOOLS  | MILD     | -1.0   | 2     | DIARRHOEA |
| 01-718-1139 | DIARRHOEA | DIARRHEA     | MILD     | 182.0  | 3     | DIARRHOEA |

Unique subjects (sorted): 18

01-701-1015  
01-701-1387  
01-701-1415  
01-703-1042  
01-703-1403  
01-704-1010  
01-704-1065  
01-706-1041  
01-708-1272  
01-709-1007  
01-709-1020  
01-709-1238  
01-710-1264  
01-714-1195  
01-714-1288  
01-716-1044  
01-717-1004  
01-718-1139

Ask a question (or type exit): -■

Ask a question (or type exit): who had severe events

=====

who had severe events

=====

| USUBJID     | ATERM  | AELLT | AESEV | AESTDY  | AESEQ | AEDECOD                             |
|-------------|--|-------|-------|---|-------|-------------------------------------|
| 01-701-1211 | SUDDEN DEATH                                   |       |       | 61.0  | 9     | SUDDEN DEATH                        |
| 01-703-1086 | APPLICATION SITE IRRITATION                    |       |       | 12.0  | 3     | APPLICATION SITE IRRITATION         |
| 01-703-1119 | APPLICATION SITE IRRITATION                    |       |       | 25.0  | 2     | APPLICATION SITE IRRITATION         |
| 01-703-1119 | PRURITUS                                       |       |       | 74.0  | 4     | PRURITUS                            |
| 01-703-1175 | ATROVENTRICULAR BLOCK SECOND DEGREE            |       |       | -2.0  | 1     | ATROVENTRICULAR BLOCK SECOND DEGREE |
| 01-704-1008 | TRANSIENT ISCHAEMIC ATTACK                     |       |       | 40.0  | 2     | TRANSIENT ISCHAEMIC ATTACK          |
| 01-704-1135 | RASH   |       |       | 39.0  | 2     | RASH                                |
| 01-704-1445 | COMPLETED SUICIDE                              |       |       | 174.0   | 1     | COMPLETED SUICIDE                   |
| 01-705-1393 | COLON CANCER                                   |       |       | 137.0   | 3     | COLON CANCER                        |
| 01-706-1049 | SYNCOPE  |       |       | 36.0  | 2     | SYNCOPE                             |
| 01-708-1019 | HEADACHE                                       |       |       | FRONTAL HEADACHE SEVERE                               | 12.0  | 1                                   |
| 01-708-1019 | STUPOR   |       |       | STUPOR SEVERE   | 13.0  | 2                                   |
| 01-708-1019 | TRANSIENT ISCHAEMIC ATTACK                     |       |       | TRANSIENT ISCHAEMIC ATTACKS SEVERE                    | 13.0  | 3                                   |
| 01-708-1178 | ATRIAL FIBRILLATION                            |       |       | ATRIAL FIBRILLATION SEVERE                            | 86.0  | 7                                   |
| 01-708-1178 | DIZZINESS                                      |       |       | LIGHT-HEADED SEVERE                                   | 86.0  | 6                                   |
| 01-708-1272 | MUSCLE SPASMS                                  |       |       | LEG CRAMPS SEVERE                                     | 13.0  | 2                                   |
| 01-708-1428 | APPLICATION SITE ERYTHMA                       |       |       | APPLICATION SITE REDNESS SEVERE                       | 32.0  | 6                                   |
| 01-708-1428 | APPLICATION SITE PRURITUS                      |       |       | APPLICATION SITE ITCHING SEVERE                       | 32.0  | 5                                   |
| 01-709-1007 | APPLICATION SITE WARMTH                        |       |       | APPLICATION SITE WARMTH SEVERE                        | 32.0  | 7                                   |
| 01-709-1259 | BENIGN PROSTATIC HYPERPLASTA                   |       |       | APPLICATION SITE RASH SEVERE                          | 22.0  | 3                                   |
| 01-710-1070 | GASTROINTESTINAL HEMORRHAGE                    |       |       | ENLARGED PROSTATE SEVERE                              | 51.0  | 6                                   |
| 01-710-1070 | HIP FRACTURE                                   |       |       | GI BLEED SEVERE                                       | 151.0 | 6                                   |
| 01-710-1077 | ARTHRITIS                                      |       |       | HIP FRACTURE SEVERE                                   | 137.0 | 5                                   |
| 01-710-1083 | MYOCARDIAL INFARCTION                          |       |       | ARTHRITIS SEVERE                                      | Nan   | ARTHRITIS                           |
| 01-710-1154 | SKIN IRRITATION                                |       |       | MYOCARDIAL INFARCT SEVERE                             | 12.0  | 1                                   |
| 01-710-1271 | MYOCARDIAL INFARCTION                          |       |       | SKIN IRRITATION SEVERE                                | 21.0  | 2                                   |
| 01-710-1368 | HIP FRACTURE                                   |       |       | MYOCARDIAL INFARCT SEVERE                             | 56.0  | 2                                   |
| 01-710-1368 | HIP FRACTURE                                   |       |       | FRACTURED HIP SEVERE                                  | 64.0  | 1                                   |
| 01-711-1143 | NASOPHARYNGITIS                                |       |       | HIP FRACTURE SEVERE                                   | 64.0  | 4                                   |
| 01-714-1195 | APPLICATION SITE ERYTHMA                       |       |       | COLD SEVERE   | 26.0  | 3                                   |
| 01-714-1195 | APPLICATION SITE ERYTHMA                       |       |       | APPLICATION SITE REDNESS SEVERE                       | 18.0  | 3                                   |
| 01-716-1103 | BLISTER  |       |       | APPLICATION SITE REDNESS SEVERE                       | 18.0  | 8                                   |
| 01-716-1103 | BLISTER  |       |       | BLISTERS SEVERE                                       | 151.0 | 4                                   |
| 01-716-1189 | PROSTATE CANCER                                |       |       | BLISTERS SEVERE                                       | 151.0 | 5                                   |
| 01-717-1174 | RASH   |       |       | PROSTATE CANCER SEVERE                                | 122.0 | 4                                   |
| 01-718-1066 | SYNCOPE  |       |       | RASH SEVERE   | 94.0  | 4                                   |
| 01-718-1066 | SYNCOPE  |       |       | SYNCOPE SEVERE  | 10.0  | 1                                   |
| 01-718-1079 | APPLICATION SITE IRRITATION                    |       |       | SYNCOPE SEVERE  | 24.0  | 3                                   |
| 01-718-1170 | AGITATION                                      |       |       | APPLICATION SITE IRRITATION SEVERE                    | 17.0  | 3                                   |
| 01-718-1170 | SYNCOPE  |       |       | INCREASED AGITATION SEVERE                            | 31.0  | 6                                   |
| 01-718-1371 | PARTIAL SEIZURES WITH SECONDARY GENERALISATION |       |       | SYNCOPE SEVERE  | 27.0  | 5                                   |
| 01-718-1427 | NAUSEA   |       |       | PARTIAL SEIZURES WITH SECONDARY GENERALISATION SEVERE | 38.0  | 4                                   |
|             |  |       |       | NAUSEA SEVERE   | 36.0  | 8                                   |
|             |  |       |       |   |       | NAUSEA                              |

Unique subjects (sorted): 31

=====

01-701-1211

01-703-1086

01-703-1119

01-703-1175

01-704-1008

01-704-1135

01-704-1445

01-705-1393

01-706-1049

01-708-1019

01-708-1178

01-708-1272

01-708-1428

01-709-1007

01-709-1259

01-710-1070

01-710-1077

01-710-1083

01-710-1154

01-710-1271

01-710-1368

01-711-1143

01-714-1195

01-716-1103

01-716-1189

01-717-1174

01-718-1066

01-718-1079

01-718-1170

01-718-1371

01-718-1427

Ask a question (or type exit): who had severe events involving HYPERPLASIA

=====

who had severe events involving HYPERPLASIA

=====

| USUBJID     | ATERM                        | AELLT | AESEV | AESTDY                   | AESEQ | AEDECOD                        |
|-------------|------------------------------|-------|-------|--------------------------|-------|--------------------------------|
| 01-709-1259 | BENIGN PROSTATIC HYPERPLASIA |       |       | ENLARGED PROSTATE SEVERE | 51.0  | 6 BENIGN PROSTATIC HYPERPLASIA |

Unique subjects (sorted): 1

=====

01-709-1259

Ask a question (or type exit): who had severe events involving cancer

=====

who had severe events involving cancer

=====

| USUBJID     | ATERM           | AELLT | AESEV                  | AESTDY | AESEQ | AEDECOD         |
|-------------|-----------------|-------|------------------------|--------|-------|-----------------|
| 01-705-1393 | COLON CANCER    |       | COLON CANCER SEVERE    | 137.0  | 3     | COLON CANCER    |
| 01-716-1189 | PROSTATE CANCER |       | PROSTATE CANCER SEVERE | 122.0  | 4     | PROSTATE CANCER |

Unique subjects (sorted): 2

=====

01-705-1393

01-716-1189

```
Ask a question (or type exit): Who died?  
=====  
Who died?  
=====
```

| USUBJID     | AETERM                | AELLT        | AESEV   | AESTDY | AESEQ | AEDECOD                 |
|-------------|-----------------------|--------------|---------|--------|-------|-------------------------|
| 01-701-1211 | SUDDEN DEATH          | SUDDEN DEATH | SEVERE  | 61.0   | 9     | SUDDEN DEATH            |
| 01-704-1445 | COMPLETED SUICIDE     | SUICIDE      | SEVERE  | 174.0  | 1     | COMPLETED SUICIDE       |
| 01-710-1083 | MYOCARDIAL INFARCTION | MYOCARDIAL   | INFARCT | SEVERE | 12.0  | 1 MYOCARDIAL INFARCTION |

```
=====
```

```
Unique subjects (sorted): 3  
=====
```

```
01-701-1211  
01-704-1445  
01-710-1083  
=====
```

```
Ask a question (or type exit): ■
```

```
Ask a question (or type exit): who had stomach problems  
=====
```

```
who had stomach problems  
=====
```

| USUBJID     | AETERM             | AELLT   | AESEV | AESTDY | AESEQ | AEDECOD              |
|-------------|--------------------|---------|-------|--------|-------|----------------------|
| 01-701-1302 | STOMACH DISCOMFORT | STOMACH | UPSET | MILD   | 2.0   | 3 STOMACH DISCOMFORT |

```
=====
```

```
Unique subjects (sorted): 1  
=====
```

```
01-701-1302  
=====
```

```
Ask a question (or type exit): who had toothache  
=====
```

```
who had toothache  
=====
```

```
(No matching records)  
=====
```

```
Ask a question (or type exit): who had eye problems  
=====
```

```
who had eye problems  
=====
```

```
(No exact AETERM match; used closest value from data: "EYE ALLERGY")  
=====
```

| USUBJID     | AETERM      | AELLT       | AESEV | AESTDY | AESEQ | AEDECOD     |
|-------------|-------------|-------------|-------|--------|-------|-------------|
| 01-701-1130 | EYE ALLERGY | EYE ALLERGY | MILD  | 84.0   | 5     | EYE ALLERGY |

```
=====
```

```
Unique subjects (sorted): 1  
=====
```

```
01-701-1130  
=====
```

Ask a question (or type exit): who had severe events within 15 days of study?

=====

who had severe events within 15 days of study?

=====

| USUBJID<br>AEDECOD  | AETERM                      | AELLT                       | AESEV             | AESTDY | AESEQ         |
|---|-----------------------------|-----------------------------|-------------------|--------|---------------|
| 01-703-1086 APPLICATION SITE IRRITATION<br>ON SITE IRRITATION | APPLICATION SITE IRRITATION | APPLICATION SITE IRRITATION | SEVERE            | 12.0   | 3 APPLICATION |
| 01-708-1019<br>HEADACHE                                       | HEADACHE                    | FRONTAL                     | HEADACHE SEVERE   | 12.0   | 1             |
| 01-708-1019<br>STUPOR   | STUPOR                      |                             | STUPOR SEVERE     | 13.0   | 2             |
| 01-708-1019 TRANSIENT ISCHAEMIC ATTACK<br>T ISCHAEMIC ATTACK  | TRANSIENT ISCHAEMIC ATTACK  | TRANSIENT ISCHEMIC ATTACKS  | SEVERE            | 13.0   | 3 TRANSIE     |
| 01-708-1272<br>MUSCLE SPASMS                                  | MUSCLE SPASMS               |                             | LEG CRAMPS SEVERE | 13.0   | 2             |
| 01-710-1083 CARDIAL INFARCTION                                | MYOCARDIAL INFARCTION       | MYOCARDIAL INFARCT          | SEVERE            | 12.0   | 1 MY          |
| 01-718-1066<br>SYNCOPE  | SYNCOPE                     |                             | SYNCOPE SEVERE    | 10.0   | 1             |

=====

Unique subjects (sorted): 5

-----

01-703-1086  
01-708-1019  
01-708-1272  
01-710-1083  
01-718-1066

-----