Project Proposal: AI-Augmented Clinical Trial Explorer (ACTE) Implementation

Internal Development Proposal for Osteosarcoma Now

December 14, 2025

Project Overview

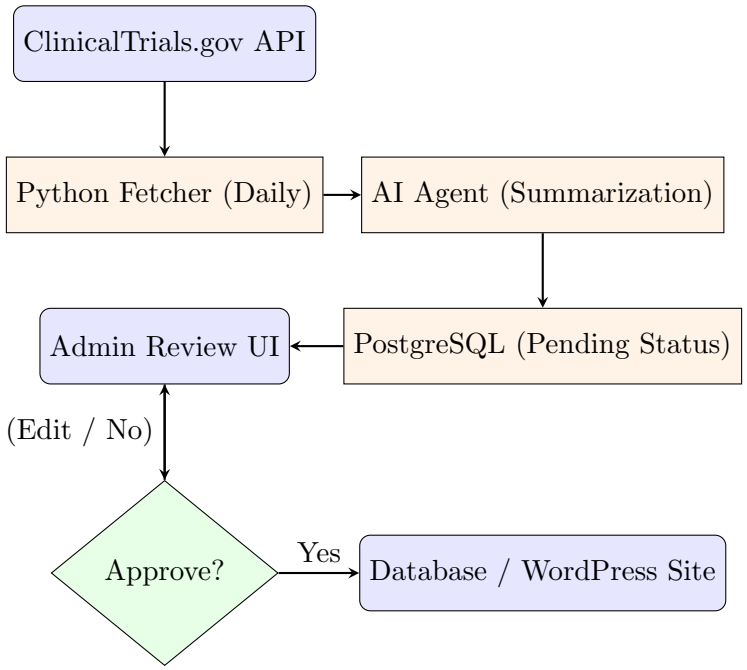
The current Osteosarcoma Clinical Trial Database relies on a legacy API (`jayr57.sg-host.com`). To gain full autonomy and ensure the highest quality of patient-friendly information, we propose building an AI-integrated system. This system will automatically fetch global trial data, use Generative AI to summarize technical details, and provide a human-in-the-loop interface for validation before publishing to the database and hence the website.

Proposed System Architecture

The system will be comprised of four primary components:

1. **Data Ingestion:** A Python engine that syncs with ClinicalTrials.gov daily.
2. **AI Processing Layer:** An agentic workflow utilizing OpenAI or Google Gemini to generate "Custom Summaries" and "Key Information."
3. **Curation Dashboard:** A secure web interface for staff to review, edit, and approve trial additions/changes/removals.
4. **The Data Delivery API:** A REST API built with FastAPI that replaces the current legacy endpoint.

Workflow Diagram



Technical Specification

- **Database:** PostgreSQL (Managed Hosting) to store official and custom trial fields.
- **Backend:** Python, FastAPI for the delivery API, and SQLAlchemy for database management.
- **AI Engine:** Integration with OpenAI API (GPT-4) using structured output (JSON mode) to ensure compatibility with existing frontend PHP templates.
- **Review UI:** A lightweight dashboard with password for approval / edit of AI-generated content. The dashboard should be buildt thinking of the one who is approving, showing how all fileds were made etc..

  This dashboard should also track approval / change statitics so that the AI Engine can be improved over time.

Implementation Phases

Phase 1: Database Setup Designing the schema to mirror current WordPress JSON requirements (NCTId, CustomBriefSummary, key_information, etc.).

Phase 2: AI Workflow Developing the Python scripts to automate database populatin with trials (trial summarization, and general info using agentic workflow).

Phase 3: The Human-in-the-Loop UI Building the dashboard for internal staff review.

Phase 4: Site Migration Updating `page-result.php` and `single-study.php` to point to the new API.

Conclusion

Transitioning to this modern stack will eliminate the risk of third-party failure, significantly reduce the manual workload for the Osteosarcoma Now team, and provide patients with the most current and accessible clinical trial information available globally.

Appendix

API Request for Osteosarcoma Trials

To programmatically fetch all clinical trials from the ClinicalTrials.gov database where "Osteosarcoma" is listed as a primary condition. This forms the foundational step of the daily data synchronization process.

B.1 API Endpoint and Query Construction

We will perform a `GET` request to the primary studies endpoint. The query will be constructed using the `query.expr` parameter, which allows for precise searching within specific data fields.

Base URL: `https://clinicaltrials.gov/api/v2/studies`

Search Expression: To find all relevant trials, we will search within the `Condition` area for the term `Osteosarcoma`. The expression is:

```
SEARCH[Study](AREA[Condition]Osteosarcoma)
```

B.2 Field Selection and Formatting

To ensure efficiency, we will specify only the necessary data fields for our application using the `fields` parameter. The response format will be explicitly set to JSON.

- **Fields Requested:** `NCTId, BriefTitle, BriefSummary, OverallStatus, Phase, EligibilityCri`
  `LocationCountry, LocationCity, LastUpdatePostDate`
- **Format:** `json`
- **Page Size:** `1000` (The maximum allowed, to minimize API calls)

B.3 Example Full Request URL

The following URL combines the endpoint, search expression, and control parameters into a single, complete API call. The Python ingestion agent will execute this request and handle pagination by using the `nextPageToken` provided in each response until all trial data is retrieved.

```
https://clinicaltrials.gov/api/v2/studies?
format=json
&pageSize=1000
&query.expr=SEARCH[Study](AREA[Condition]Osteosarcoma)
&fields=NCTId,BriefTitle,BriefSummary,OverallStatus,Phase,
EligibilityCriteria,LocationCountry,LocationCity,LastUpdatePostDate
```

(Note: URL is split across lines for readability.)

API Reference: Clinical Trials Endpoint

C.1 Overview

This section specifies the RESTful API that will serve curated clinical trial data to the Osteosarcoma Now WordPress frontend. The API is designed to be a direct, backward-compatible replacement for the legacy `jayr57.sg-host.com` endpoint. It will provide two primary endpoints for fetching lists of trials and single trial details.

Base URL: `https://your-new-api-domain.com/api/v1`

Authentication: No authentication is required for these public, read-only endpoints.

C.2 Endpoint: Search Trials

Retrieves a paginated list of approved clinical trials based on a set of search filters. This endpoint is used by the main search/results page (`page-result.php`).

Request

Method: `GET`

URL: `/trials`

Query Parameters (What is Sent)

Parameter	Type	Description
`term`	string	A keyword search (e.g., drug name, condition). Corresponds to the "Keyword" field.
`country`	string	The full name of the country to filter by (e.g., "United States").
`status`	string	Comma-separated list of statuses (e.g., "Recruiting,Available").
`phase`	string	Comma-separated list of phases (e.g., "Phase 1,Phase 2").
`type`	string	Comma-separated list of trial types (e.g., "Interventional,Observational").
`age`	string	Comma-separated list of age groups (e.g., "Child,Adult").
`page_no`	integer	The page number for pagination. Defaults to 1. Each page contains 10 results.

Success Response (What is Returned)

The API returns a JSON object containing the total count of found trials and a paginated list of trial objects in the `result` array.


```
{
  "count": 127, // Total number of trials found for the query
  "result": [
    {
      "NCTId": "NCT01234567",
      "BriefTitle": "Official Title of the Study",
      "CustomBriefTitle": "Patient-Friendly Title of the Study",
      "BriefSummary": "The official, technical summary...",
      "CustomBriefSummary": "The easy-to-read summary of the trial's aim.",
      "OverallStatus": "Recruiting",
      "CustomOverallStatus": "Recruiting now",
      "Phase": "Phase 2",
      "CustomPhase": "Phase 2",
      "LocationCountry": "United States",
      "CustomLocationCountry": "United States",
      "LocationCity": "Houston",
      "CustomLocationCity": "Houston"
    },
    {
```

```
      "NCTId": "NCT76543210",
      "BriefTitle": "Another Trial Title",
      "CustomBriefTitle": "Another Patient-Friendly Title",
      "BriefSummary": "Another technical summary...",
      "CustomBriefSummary": "Another easy-to-read summary.",
      "OverallStatus": "Completed",
      "CustomOverallStatus": "Finished trials",
      "Phase": "Phase 3",
      "CustomPhase": "Phase 3",
      "LocationCountry": "Canada, United Kingdom",
      "CustomLocationCountry": "Canada, United Kingdom",
      "LocationCity": "Toronto, London",
      "CustomLocationCity": "Toronto, London"
    }
    // ... up to 10 trial objects per page
  ]
}
```

C.3 Endpoint: Get Single Trial Details

Retrieves the full, detailed information for one specific clinical trial, identified by its NCT ID. This is used by the trial detail page (`single-study.php`).

Request

Method: `GET`

URL: `/trials/{nct_id}`

Path Parameters (What is Sent)

Parameter	Type	Description
{nct_id}	string	**Required.** The unique identifier for the trial (e.g., "NCT04132895").

Success Response (What is Returned)

The API returns a JSON object containing a single `result` array with one object, which includes all fields for the requested trial.

```
{
  "count": 1,
  "result": [
    {
      "NCTId": "NCT04132895",
      "BriefTitle": "Official Title of the Study",
      "CustomBriefTitle": "Patient-Friendly Title of the Study",
      "BriefSummary": "The official, technical summary...",
      "CustomBriefSummary": "The easy-to-read summary of the trial's aim.",
      "OverallStatus": "Recruiting",
      "CustomOverallStatus": "Recruiting now",
      "Phase": "Phase 2",
      "CustomPhase": "Phase 2",
```

```
      "LocationCountry": "United States, Canada",
      "CustomLocationCountry": "United States, Canada",
      "LocationCity": "Houston, Toronto",
      "CustomLocationCity": "Houston, Toronto",
      "StudyType": "Interventional",
      "CustomStudyType": "New treatments",
      "MinimumAge": "18 Years",
      "CustomMinimumAge": "18 Years",
      "MaximumAge": "65 Years",
      "CustomMaximumAge": "65 Years",
      "CentralContactName": "Dr. Smith",
      "CustomCentralContactName": "Dr. Smith",
      "CentralContactPhone": "555-123-4567",
      "CustomCentralContactPhone": "555-123-4567",
      "CentralContactEMail": "contact@trial.com",
      "CustomCentralContactEMail": "contact@trial.com",
      "LastUpdatePostDate": "January 1, 2024",
      "CustomLastUpdatePostDate": "January 1, 2024",
      "InterventionDescription": "Official description of the drug...",
      "CustomInterventionDescription": "Patient-friendly description of how the treatment wo
      "EligibilityCriteria": "Official inclusion and exclusion criteria text...",
      "CustomEligibilityCriteria": "Summarized inclusion/exclusion criteria.",
      "key_information": "This is the bulleted list of key info points."
    }
  ]
}
```