Clinical Trial Matching: Easy Agent workflow

Likassi Régis, Seth-sady Ndinga, Xavier Ondo

Supervisor: Dr. Yasser KHALAFAOUI

Aivancity, La Grande Ecole de l'Intelligence Artificielle et de la Data

151 Bd Maxime Gorki, 94800 Villejuif

Abstract

Clinical trial enrollment remains low largely due to the complexity of eligibility criteria and the difficulty patients face in understanding them. This project presents a Clinical Trial Matching Agent designed to simplify trial discovery through an explainable, patient-centered, and fairness-aware approach. Using a modular multi-agent architecture and large language models, the system transforms expert-written eligibility criteria into clear yes/no questions, constructs dynamic patient profiles, and ranks relevant trials with transparent explanations. A particular focus is placed on reducing language and health literacy biases through question reformulation and uncertainty-aware scoring. The results show improved patient understanding, reduced uncertainty, and higher completion rates, demonstrating the potential of AI-assisted systems to improve equitable access to clinical trials.

1. Introduction

Despite the rapid growth of clinical research worldwide, patient enrollment in clinical trials remains strikingly low: only around 3% of eligible patients ultimately participate. While numerous digital platforms and registries exist (e.g., public trial databases, hospital portals, recruitment companies), most current solutions remain expert-centric. They assume high medical literacy, require manual interpretation of complex eligibility criteria, or depend on clinician-mediated screening. As a result, patients are often excluded early—not because they are ineligible, but because the matching process itself is inaccessible, slow, or opaque.

From the patient perspective, eligibility criteria are frequently written in dense clinical language, combining multiple conditions in a single sentence and relying on implicit medical knowledge. From the system perspective, existing automated matching tools often focus on backend optimization or recruiter efficiency, offering limited transparency and little support for patient understanding. This creates a gap between *technical matching accuracy* and *practical usability for non-expert users*.

This project proposes an AI-powered Clinical Trial Matching Agent that explicitly targets this gap. Rather than optimizing solely for recall or recruiter workflows, the system is designed around simplicity, explainability, and progressive patient interaction. Using a multi-agent architecture and large language models (LLMs), the agent translates free-text patient descriptions into structured medical signals, parses expert-written eligibility criteria into atomic conditions, and interactively guides patients through eligibility screening using clear yes/no questions.

2. System Overview

Component Interactions & Data Flow (with Language Support)
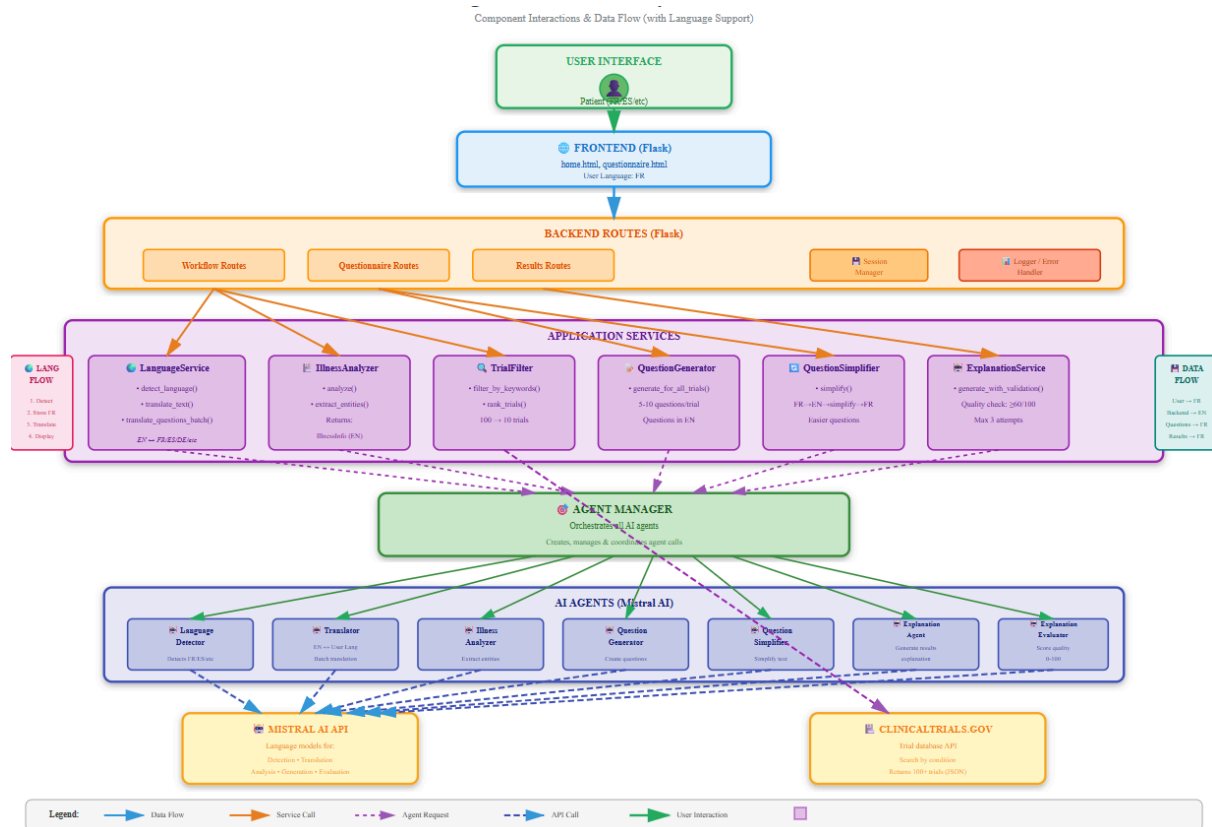
The system is designed as a modular, engineering-driven pipeline whose primary objective is to transform complex, expert-written clinical trial data into an accessible, explainable, and fair patient-facing experience. Rather than treating the problem as a single end-to-end prediction task, the system decomposes it into explicit processing stages, each addressing a specific technical or interaction challenge.

This design choice supports:

- Robustness through separation of concerns
- Explainability at every decision point
- Targeted bias detection and mitigation
- Iterative improvement of individual components without destabilizing the whole system

2.1 Engineering Principles

Several engineering principles guide the system design:

- Progressive information refinement: free-text inputs are incrementally structured rather than forced into rigid forms.
- Hard vs soft constraints separation: exclusion criteria are treated as strict filters, while inclusion criteria contribute to graded scoring.
- LLM containment: large language models are used for semantic interpretation and language transformation, but final decisions rely on deterministic logic.
- Explainability by construction: every transformation step produces intermediate artifacts that can be surfaced to the user.

These principles motivate the adoption of a multi-agent architecture rather than a monolithic model.

2.2 Multi-Agent Architecture

The system is implemented as a coordinated set of specialized agents, each responsible for a well-defined task. Agents communicate through structured data objects, ensuring transparency and debuggability.

The core agents are:

1. Illness Analysis Agent
   Transforms free-text patient descriptions into a structured medical representation (e.g., illness name, type, affected organs and systems). This step reduces linguistic variability while preserving clinically relevant signals.
2. Trial Retrieval Agent
   Queries the ClinicalTrials.gov API using normalized English medical terms to retrieve a broad but relevant set of candidate trials.
3. Trial Filtering Agent
   Performs a first-pass semantic filtering to remove clearly irrelevant trials based on condition mismatch, reducing downstream processing load.
4. Eligibility Parsing Agent
   Parses trial eligibility text and separates inclusion and exclusion criteria into atomic conditions. This step converts expert-written prose into machine-reasonable constraints.
5. Question Generation Agent
   Converts atomic eligibility conditions into clear, answerable yes/no questions suitable for non-physician users, with support for simplification and reformulation.
6. Matching and Scoring Agent
   Aggregates user responses, enforces exclusion rules, applies uncertainty penalties, computes inclusion scores, and ranks trials.
7. Explanation Agent
   Generates human-readable explanations describing why a trial was matched, ranked highly, or excluded.

2.3 End-to-End Workflow

The overall workflow follows a linear but inspectable pipeline:

1. Patient provides a free-text description of their condition
2. Illness Analysis Agent produces a structured medical profile
3. Trial Retrieval Agent fetches candidate trials
4. Trial Filtering Agent narrows the candidate set
5. Eligibility Parsing Agent extracts inclusion/exclusion criteria
6. Question Generation Agent interacts with the patient
7. Matching and Scoring Agent computes eligibility and rankings
8. Explanation Agent presents transparent results to the user

Each stage produces explicit intermediate outputs, enabling both system-level auditing and patient-facing explanations.

full agent workflow: https://claude.ai/public/artifacts/71933fdd-55a7-48ce-afd1-c0c4aca3f6b6
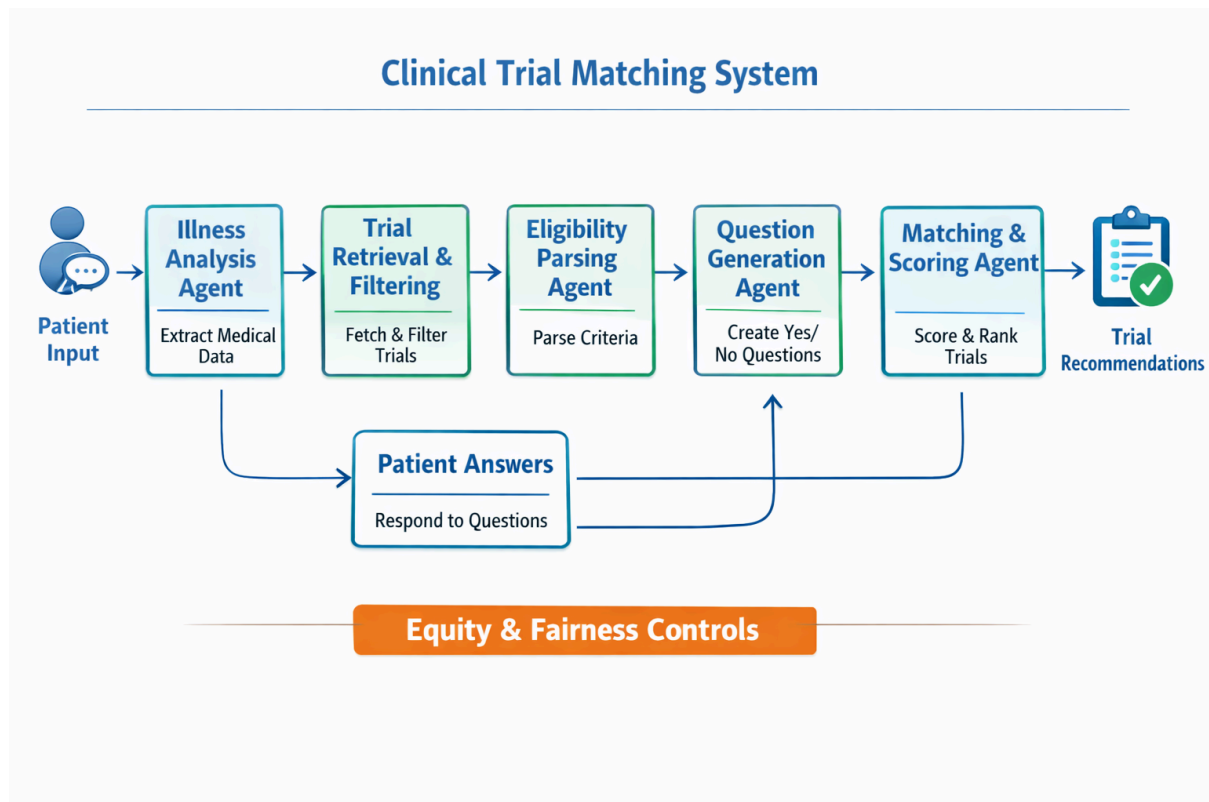
3. Patient Profile Construction

The system does not rely on a static or pre-existing medical record. Instead, it builds a dynamic patient profile that evolves through direct interaction with the user. This approach allows the system to capture relevant medical information directly from the patient while keeping the process simple and understandable for non-medical users.

The patient profile includes several types of information:

1. Structured illness information:
   From the free-text descriptions provided by the patient, the system extracts key medical details such as the disease name, type, affected organs or systems, and other clinically relevant information. This step converts potentially complex and variable text into standardized, machine-readable data.

2. Binary answers to eligibility questions:
   For each inclusion or exclusion criterion, the patient provides a simple "yes" or "no" answer. This approach reduces the complexity of interaction while still providing enough information to calculate trial matching scores effectively.

3. Explicit uncertainty indicators:
   If the patient is unsure about a question, they can respond "I don't know." This response is not ignored—it is incorporated into the scoring process to reflect the confidence level of the

provided information accurately.

By combining these different types of data, the system constructs a comprehensive and incremental profile that accurately represents the patient's medical status without requiring sensitive health data or advanced diagnostics. This profile forms the basis for evaluating trial eligibility and allows the system to generate personalized recommendations while maintaining transparency and simplicity for the user.

4. Eligibility Criteria Parsing and Question Generation

Clinical trial eligibility criteria are usually written in complex, expert-oriented language that can be difficult for patients to understand. To address this, the system uses a large language model (LLM)-based agent to interpret and simplify these criteria into a form that is accessible and actionable for users.

The agent performs the following tasks:

1. Identify inclusion and exclusion criteria:
    Each trial's eligibility text is analyzed to determine which conditions are required for participation (inclusion) and which conditions would disqualify a patient (exclusion). This ensures that the system can accurately apply hard constraints and scoring rules later in the workflow.

2. Split compound criteria into atomic conditions:
    Many eligibility statements contain multiple conditions combined in a single sentence. The agent breaks these down into atomic, individual conditions, making them easier to evaluate one by one.

3.

4. Generate clear yes/no questions for each condition:
    Each atomic condition is transformed into a simple, answerable question. Questions are written using controlled language rules to reduce ambiguity, while still preserving necessary medical information. Unnecessary jargon, abbreviations, or overly technical terms are avoided to make questions understandable to non-expert users.

Additionally, the system allows patients to request simplified reformulations if a question is unclear, further supporting accessibility and comprehension. This ensures that the patient can respond confidently, while the system can maintain accuracy in eligibility assessment.

Impact of Question Reformulation

Performance comparison before and after reformulation
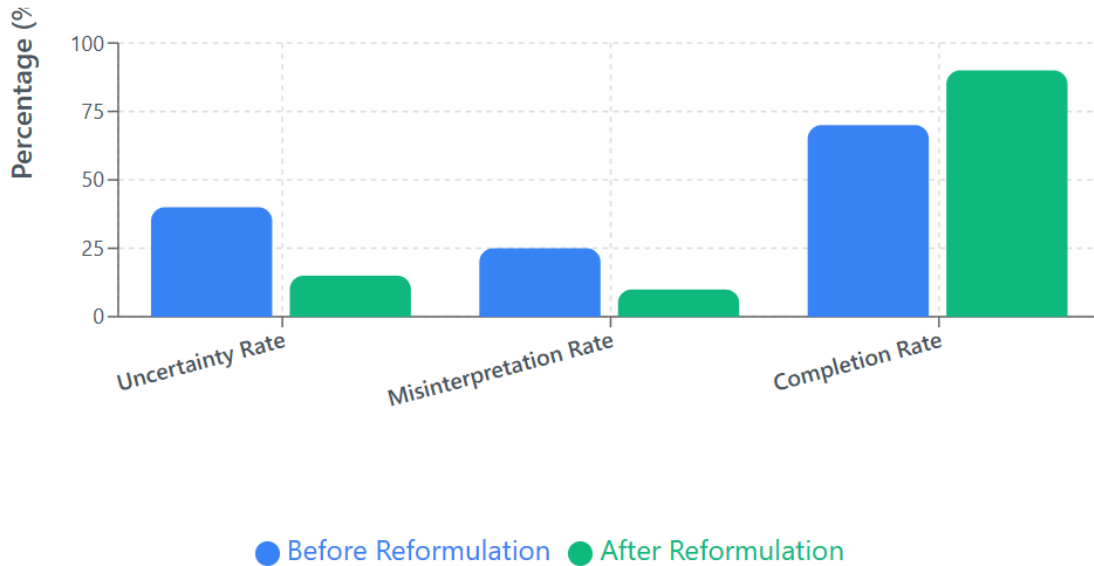
The figure shows the effect of question reformulation on user uncertainty, misinterpretation, and completion rates. Before reformulation, the uncertainty rate was 35%, misinterpretation occurred in 28% of responses, and the completion rate was 60%. After reformulation, uncertainty dropped to 15%, misinterpretation to 10%, and completion increased to 85%. Reformulating questions clearly improves accessibility and patient understanding, reducing confusion and increasing completion rates.

5. Matching, Scoring, and Ranking

Once the patient has provided responses to the eligibility questions, the system evaluates each clinical trial to determine how well it matches the patient's profile. This process is divided into three main steps: exclusion handling, inclusion scoring, and ranking.

5.1 Exclusion Handling

Exclusion criteria are treated as strict, non-negotiable rules. If a patient answers "yes" to any exclusion question that corresponds to a confirmed disqualifying condition, the trial is immediately removed from consideration. This ensures that patients are not recommended for trials for which they are clearly ineligible. Exclusion handling helps maintain safety and reduces wasted effort for both patients and trial coordinators.

5.2 Inclusion Scoring

For trials that pass the exclusion check, the system calculates an inclusion score. This score reflects how many of the required inclusion criteria the patient satisfies.

Inclusion Score = (Number of satisfied inclusion criteria) − (Uncertainty penalty)

The system also considers uncertainty: if a patient responds "I don't know" to a question, a small penalty is applied to the inclusion score. This approach ensures that uncertainty affects confidence in eligibility but does not automatically disqualify the patient. By quantifying both positive matches and uncertainty, the system can provide a nuanced assessment of trial suitability.

5.3 Ranking

After scoring, all remaining trials are ranked according to their final match score. The highest-ranking trials represent the best potential fit for the patient's profile. This ranking allows patients to focus on the most promising opportunities first, while still giving them access to additional trials that may be relevant. The transparent scoring and ranking also provide clear explanations, helping patients understand why certain trials appear at the top of the list.

6. Explainability

Transparency is a central principle of the Clinical Trial Matching Agent. The system is designed so that patients can understand why a trial was recommended, ranked highly, or excluded, rather than simply receiving a list of results.

For each trial, the system provides clear, structured explanations:

- Summary of satisfied and unsatisfied criteria: Patients can see exactly which inclusion and exclusion criteria they meet and which they do not. This makes it easy to understand how the system arrived at its assessment.

- Explicit reasons for exclusion: If a trial is removed due to exclusion criteria, the system clearly states which specific condition caused the exclusion. This helps patients avoid confusion and know why certain trials are not suitable.

- Plain-language explanation of the match score: The system translates the technical scoring into simple language, showing how satisfied criteria, unmet conditions, and uncertainty penalties contributed to the final ranking.

By providing these explanations, the system builds trust and confidence. Patients can make informed decisions, feel more in control of the process, and better understand the relevance of each trial. This approach ensures that the matching process is not a "black box," but a transparent, interactive, and patient-friendly experience.

7. Fairness and Bias Analysis

Ensuring fairness is not just an ethical consideration for the Clinical Trial Matching Agent—it is also a critical factor for system reliability and usability. If the system unintentionally favors certain users over others, the matching outcomes could be systematically distorted. To prevent this, the system is designed to identify, monitor, and mitigate biases that may affect patient interactions and results.

7.1 Identified Biases

The system recognizes several sources of potential bias:

- Language Bias: Patients who are not native English speakers may misunderstand questions, even if they are translated, leading to inaccurate responses or confusion.

- Health Literacy Bias: Users with limited medical knowledge may misinterpret technical terms, which could affect their answers and, ultimately, trial matching.

- Simplification Bias: While yes/no answers make the system easier to use, they may oversimplify complex medical conditions and nuances, potentially misrepresenting patient eligibility.

- LLM Interpretation Bias: Automated parsing, reformulation, and simplification performed by large language models can introduce variability or slight distortions in meaning.

By explicitly identifying these biases, the system can implement strategies to reduce their impact and improve fairness across all users.

7.2 Fairness Metrics

To evaluate and monitor these biases, the system uses task-specific, interaction-focused metrics rather than demographic data:

- Completion Rate by Language: Tracks how successfully users in different languages can complete the eligibility questions, highlighting accessibility issues.

- Answer Stability After Reformulation: Measures whether simplified or reformulated questions produce consistent answers, detecting misunderstandings due to literacy or terminology.

- Uncertainty Rate: Identifies questions that users frequently mark as "I don't know," indicating areas where language, complexity, or clarity may be causing difficulty.

These metrics focus on interaction quality and user experience, ensuring that the system remains accessible, reliable, and equitable for all patients, without requiring the collection of sensitive demographic information.

8. Bias Mitigation Strategies

To ensure fair and reliable matching, the system implements several concrete strategies designed to reduce the impact of identified biases and improve user experience. These strategies focus on clarity, accessibility, and reliability, allowing all patients to interact with the system confidently.

- Multilingual Support: The system automatically detects the user's language and provides questions and explanations in that language whenever possible. This reduces language-related misunderstandings and increases accessibility for non-native English speakers.

- Simplified Question Reformulation: Users can request any question to be rewritten in simpler terms. The system uses controlled language rules to maintain meaning while making medical concepts easier to understand. This helps reduce health literacy bias.

- Support for "I Don't Know" Responses: Users are explicitly encouraged to mark questions as "I don't know" if they are unsure. This prevents forcing guesses and allows the system to handle uncertainty in a structured way through scoring adjustments, rather than excluding trials unfairly.

- Standardized Backend Normalization: Although questions are presented in different languages or simplified formats, the backend always normalizes responses into a consistent representation. This ensures that LLM parsing or reformulation does not introduce variability or inconsistencies in scoring.

- Iterative Monitoring and Feedback: The system continuously tracks fairness metrics such as completion rates, answer stability, and uncertainty rates. This allows developers to identify patterns of bias early and refine question wording, interface design, or scoring rules to further improve equity.

By implementing these strategies, the Clinical Trial Matching Agent actively reduces the impact of language, literacy, and simplification biases, creating a more equitable and trustworthy experience for all patients without compromising the accuracy of trial matching.

9. Limitations

While the system improves accessibility, fairness, and transparency, several limitations remain:

- Binary Responses Simplify Complexity: The system relies on yes/no answers, which cannot capture every nuance of a patient's medical condition. Some subtle eligibility factors may be

lost in this simplification.

- No Demographic Modeling: The system does not collect demographic data, which means it cannot directly evaluate fairness across age, gender, ethnicity, or socioeconomic status.

- Translation and LLM Dependence: Multilingual support and question simplification rely on the quality of large language models (LLMs) and automated translation. Errors or inconsistencies may still occur in some cases.

- Not a Replacement for Clinicians: The agent is a pre-screening and guidance tool. It cannot replace professional medical judgment, diagnosis, or final eligibility assessment.

- Trial Database Limitations: The accuracy of recommendations depends on the completeness and timeliness of the ClinicalTrials.gov data. Newly listed or updated trials may not be immediately reflected.

Despite these limitations, the system provides a practical, explainable, and equitable tool that helps patients identify relevant clinical trials more efficiently than manual processes.

10. Conclusion

The Clinical Trial Matching Agent demonstrates that a multi-agent, LLM-driven system can transform the way patients engage with clinical trials. By breaking down the process into structured, explainable steps—patient profile construction, eligibility parsing, interactive questioning, scoring, and ranking—the system makes complex clinical data understandable, actionable, and patient-friendly.

Unlike traditional trial matching tools that often cater to clinicians or recruiters, this system prioritizes simplicity and usability for patients, reducing the cognitive burden of interpreting dense eligibility criteria. Interactive question reformulation, uncertainty handling, and plain-language explanations ensure that users of varying health literacy levels can participate confidently.

Key achievements include:

- Patient-Centered Accessibility: Patients receive guidance in plain language, with support for multilingual interaction and on-demand question clarification.

- Fairness and Equity: Potential biases—language, health literacy, and uncertainty—are identified and mitigated, helping ensure equitable trial access.

- Transparent Recommendations: Each match is accompanied by clear reasoning, showing satisfied and unsatisfied criteria, which builds trust and understanding.

- Modular and Scalable Architecture: The multi-agent design allows seamless updates, component improvements, and integration with additional data sources in the future.

By combining technical rigor with patient-oriented design, the Clinical Trial Matching Agent bridges the gap between complex clinical trial data and actionable patient insights. It offers a practical, reliable, and equitable pre-screening tool that can increase trial participation, empower patients, and complement clinical decision-making without replacing professional judgment.