Health Facility Location in Ethiopia: Leveraging LLMs to Integrate Expert Knowledge into Algorithmic Planning

Yohai Trabelsi*
Department of Data Science and AI, School of Computer Science, Ariel University
Ariel, Israel
yohai.trabelsi@gmail.com

Guojun Xiong†
John A. Paulson School of Engineering and Applied Sciences, Harvard University
Cambridge, MA, United States
xionggj1@gmail.com

Fentabil Getnet
National Data Management and Analytics Center for Health, Ethiopian Public Health Institute
Addis Ababa, Ethiopia
b.infen4ever@gmail.com

Stéphane Verguet
Department of Global Health and Population, Harvard T.H. Chan School of Public Health
Boston, MA, United States
verguet@hsph.harvard.edu

Milind Tambe
John A. Paulson School of Engineering and Applied Sciences, Harvard University
Cambridge, MA, United States
milind_tambe@harvard.edu

ABSTRACT

Ethiopia's Ministry of Health is upgrading health posts to improve access to essential services, particularly in rural areas. Limited resources, however, require careful prioritization of which facilities to upgrade to maximize population coverage while accounting for diverse expert and stakeholder preferences. In collaboration with the Ethiopian Public Health Institute and Ministry of Health, we propose a hybrid framework that systematically integrates expert knowledge with optimization techniques. Classical optimization methods provide theoretical guarantees but require explicit, quantitative objectives, whereas stakeholder criteria are often articulated in natural language and difficult to formalize. To bridge these domains, we develop the Large language model and Extended Greedy (LEG) framework. Our framework combines a provable approximation algorithm for population coverage optimization with LLM-driven iterative refinement that incorporates human-AI alignment to ensure solutions reflect expert qualitative guidance while preserving coverage guarantees. Experiments on real-world data from three Ethiopian regions demonstrate the framework's effectiveness and its potential to inform equitable, data-driven health system planning.

KEYWORDS

Health Facility Location; Optimization; Human Expert Knowledge; Human-AI Alignment; LLM

ACM Reference Format:
Yohai Trabelsi, Guojun Xiong, Fentabil Getnet, Stéphane Verguet, and Milind Tambe. 2026. Health Facility Location in Ethiopia: Leveraging LLMs to Integrate Expert Knowledge into Algorithmic Planning . In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. https://doi.org/10.65109/AIFQ1928

1 INTRODUCTION

Ensuring equitable access to essential health services remains a central challenge for Ethiopia's Ministry of Health (MOH). Since the launch of the Health Extension Program (HEP) in 2003/04 [31], the country has made significant progress in expanding basic care delivery to rural populations. The HEP Optimization Roadmap (2020–2035)[22] further envisions upgrading selected health posts into comprehensive ones that can provide advanced services such as childbirth and postnatal care (see Figure 1). However, upgrading facilities is expensive, and the available public budget is severely constrained [5]. Determining which facilities to upgrade thus becomes a complex optimization problem involving limited resources, heterogeneous population needs, and diverse stakeholder opinions.

Recent studies have applied optimization and geospatial methods to identify locations where comprehensive facilities are needed [3, 5, 9]. These efforts primarily focus on maximizing population coverage under distance and capacity constraints (See Figure 2). Yet, in practice, the final allocation decisions in the stepwise construction of comprehensive health posts remain dominated by expert judgment and stakeholder negotiation, rather than by algorithmic outputs. While human expertise captures important contextual knowledge—such as terrain accessibility or local socio-political considerations—it is often expressed in natural language and difficult to encode into a mathematical objective function [6, 35]. It is possible that even the advisors themselves may struggle to provide their preferences as a fully coherent set of recommendations.

Recent advances in Large Language Models (LLMs) [4, 35] offer a promising avenue for bridging the gap between qualitative human judgment and quantitative optimization. LLMs can interpret and structure unformalized expert advice, enabling algorithms to incorporate contextual and domain-specific reasoning that traditional models often overlook. Nevertheless, these methods typically lack formal theoretical guarantees on performance or stability—an issue

*Work conducted while the author was affiliated with Harvard University, USA.
†Correspondence to: Guojun Xiong (xionggj1@gmail.com).

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/AIFQ1928

Figure 1: Basic health post (left); Design of a comprehensive health post that provides more essential services (right). Source: Ministry of Health, Ethiopia.

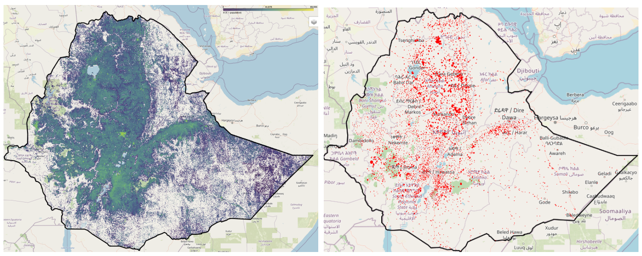

Figure 2: ([5]) Map of Ethiopia overlaid with projected population estimates for 2026 (log scale). Brighter yellow areas on the left indicate higher population densities. The map on the right shows the locations of health facilities capable of providing essential health services (red points). Many populations currently lack 2-hour access to such facilities [11].

that is especially critical in high-stakes domains such as public health or infrastructure planning. In the absence of such guarantees, language-based systems risk producing allocations that appear reasonable linguistically but fail to satisfy fairness, transparency, or policy-alignment criteria required for real-world adoption.

To address this gap, we develop a hybrid framework that couples algorithmic optimization with language-based expert reasoning. Our approach builds upon classical submodular maximization, ensuring provable guarantees on coverage, while leveraging LLMs to interpret and iteratively integrate human advice expressed in text. The LLM acts as a bridge between formal optimization and informal domain knowledge—translating verbal recommendations into structured allocation adjustments that preserve theoretical performance bounds. This integration allows the system to remain both rigorous and human-aligned.

We demonstrate the framework in collaboration with the Ethiopian Public Health Institute and Ministry of Health, focusing on three regions. Empirical results show that LLM-guided iterations significantly improve the alignment between algorithmic allocations and expert advice, while maintaining high coverage efficiency. Beyond this specific application, the proposed method provides a general blueprint for embedding qualitative human preferences into quantitative decision-making processes.

1.1 Process Description

In Figure 3 we present the proposed method in a high level. The process begins with the problem inputs, an initial greedy allocation, and a list of advice sentences. These inputs are provided to a large language model (e.g., Gemini, ChatGPT), which iteratively refine

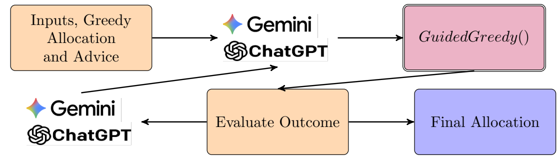

Figure 3: Overview of the proposed method.

the allocation strategy. The model proposes an allocation that is optimized using a guided greedy procedure parameterized by (α, β). The resulting outcome is then evaluated, and the evaluation feedback is transformed into a new prompt for further refinement. This iterative loop continues until the maximum number of iterations is reached, producing the final optimized allocation.

1.2 Contributions

In collaboration with the Ethiopian Public Health Institute and Ministry of Health, this work introduces the unified LEG framework for upgrading health facilities in Ethiopia that jointly optimizes population coverage and alignment with expert guidance. Our main contributions are summarized as follows:

▹ *LLM-integrated multi-objective optimization.* We propose a novel framework that embeds LLMs into a submodular optimization pipeline, enabling the joint optimization of quantitative coverage objectives and qualitative human advice.

▹ *Theory–practice trade-off with controllable guarantees.* We formalize a tunable mechanism to balance provable approximation guarantees and adherence to expert-driven recommendations.

▹ *Online budget adaptability.* We extend the framework to handle sequential or dynamically arriving budgets, a critical requirement for multi-year public-sector planning.

▹ *Real-world validation in Ethiopia.* We conduct extensive experiments on three regional datasets (Afar, Somali, and Benishangul-Gumuz), demonstrating that our method maintains strong coverage performance while significantly improving interpretability and stakeholder alignment.

2 RELATED WORKS

2.1 Plan-and-Execute

Many works are discussing the plan-and-execute approach where a high-level plan is computed before the execution of the tasks (e.g., [27, 36, 37]). Some of these works [7, 13] explore LLM-based hierarchical planning. However, their hierarchical design primarily serves to simplify complex problems, whereas our setting is closer to a multi-objective framework- requiring both high-level (abstract) performance and low-level (concrete) performance to be valuable independently.

2.2 Multi-Objective Optimization

The field of multi-objective optimization (for overviews, see [23, 24]) provides a principled foundation for handling competing objectives and evaluating trade-offs among them. A widely adopted strategy is weighted scalarization [10, 15, 19], which collapses multiple objectives into a single composite objective using pre-specified weights. Despite its broad applicability, scalarization presents a long-standing difficulty: effectively traversing the Pareto frontier typically demands manual or iterative tuning of these weights to reveal different trade-off solutions. To overcome this limitation, recent work such as [35] proposes a language-driven framework for exploring Pareto-optimal solutions. Instead of explicitly setting scalarization coefficients, an LLM interprets verbal feedback and adjusts the reward structure accordingly, enabling a more natural and user-centric interaction with the optimization process. In contrast, our approach operates within a hierarchically coordinated framework. Here, the LLM and the solver interact symbiotically—the LLM contributes high-level reasoning and structural insights, while the solver provides robust low-level optimization capabilities. This interplay allows the system to retain theoretical rigor while integrating expert intuition in a seamless and interpretable manner.

2.3 Using LLMs for Reward Shaping

Recent research has increasingly explored the use of LLMs for reward shaping and specification [18, 35]. Early systems such as Eureka [18] demonstrated that LLMs can automatically synthesize reward functions directly from natural-language descriptions of desired behaviors. Similarly, [17] shows that language interfaces can simplify the design of proxy rewards through interactive refinement. The Decision-Language Modeling framework [4, 29] extends this paradigm to societal domains, using LLMs to propose reward formulations for public-health resource allocation. [16] proposed a method to fine-tune LLM to efficiently translate natural-language human preferences into reward functions. Although these approaches highlight the expressive power of language-driven reward design, they often risk distorting the underlying task utility, since optimization is performed purely with respect to human-specified or textual preferences.

To mitigate this issue, [35] integrates language-derived shaping rewards into an existing solver's intrinsic objective, maintaining the base utility while progressively improving alignment with user intent. Yet, the requirement to explicitly construct a numerical reward function for every form of expert advice remains a practical bottleneck—translating nuanced guidance into precise reward formulations can itself demand significant effort and domain expertise.

2.4 Health Facility Location

Many studies have addressed health facility location problems (see, e.g., the survey by [2]). [30] examined the allocation of emergency health facilities while incorporating expert preferences into their model. [8] examined the placement of health facilities in the Philippines, explicitly considering trade-offs between equity and efficiency in their optimization approach. [25] optimized clinic locations in Kuala Langat, Malaysia, to maximize population coverage within 3–5 km. Other studies, such as [3, 5], focused on optimizing health facility locations in Ethiopia. However, most of these studies do not account for alignment with human expert preferences.

2.5 Human-AI Alignment

Recent work has focused on aligning AI decisions with human expertise (see [14, 26] for surveys) and aligning LLMs to better suit human-oriented tasks and expectations (see [32] for a survey). Our work contributes to the literature on aligning algorithmic decision-making with human preferences (e.g., [1, 21, 35]). However, whereas most existing approaches rely on the explicit construction of a numerical reward function, which is not always practical, our method enables LLMs to directly guide the allocation process, bypassing the need for an intermediate reward-function construction step.

3 SETTINGS AND PROBLEM FORMULATION

3.1 Settings

Let $r \in \mathbb{N}_{>0}$ denote the number of districts. Let the ground set be $V = T_1 \uplus \ldots \uplus T_r$, where each T_i represents the set of candidate grid cells in district i, and $\uplus$ denotes a disjoint union (See the example in Section 1.1). For simplicity, we assume that all grid cells in these districts are candidates for locating a facility.[1] A **grid-cell allocation** is any subset $S \subseteq V$.

For any grid-cell allocation $S \subseteq V$, we define the *district allocation* $h(S) \in \mathbb{N}^r$ by

$$h_i(S) = |S \cap T_i|, \quad \text{for } i = 1, \ldots, r,$$

so that $h_i(S)$ represents the number of facilities placed in district i. We denote by b the total budget, i.e., the number of health posts to be upgraded in a given year.

Coverage function. Let covered(S) be the set of grid cells served by the facilities in S. The total coverage associated with S is defined as

$$f(S) = \sum_{c \in \text{covered}(S)} w_c, \tag{1}$$

where w_c denotes the population of grid cell c. This monotone submodular function quantifies the cumulative population covered by the selected facilities.

Advice alignment function. Because the notion of "alignment with expert advice" is inherently qualitative, defining a corresponding numerical objective is nontrivial. We represent this alignment through a function $g : 2^V \to \mathbb{R}$ that assigns a score to each grid-cell allocation, reflecting how well the allocation adheres to the provided human or LLM-generated advice. (Here, 2^V denotes the set of all possible subsets of V.)

For evaluation, we approximate this alignment function using a set of advice sentences A. For each advice sentence $a \in A$, we define an auxiliary scoring function $g_a : 2^V \to \mathbb{R}$ that measures the extent to which the advice a is satisfied. An LLM is used to define these component functions and aggregate them into an overall alignment metric:

$$g_{\text{eval}}(S) = \sum_{a \in A} g_a\big(h(S)\big). \tag{2}$$

This construction serves as a proxy for human evaluation, enabling language-based advice to be incorporated into a quantitative optimization framework.

[1] This assumption simplifies exposition; our framework remains valid even if only a subset of cells within each district are eligible, as the optimization operates on any predefined candidate set.

3.2 Problem Formulation

The multi-objective problem. Given the functions f and g, and a total budget b, we define the following multi-objective formulation that jointly optimizes *population coverage* and *alignment with expert advice*:

$$\max_{S \subseteq V,\, |S|=b} \left\{ f(S),\, g(S) \right\}. \tag{3}$$

This formulation captures the fundamental trade-off between quantitative performance (coverage) and qualitative consistency (alignment).

The formulation in (3) defines an idealized bi-objective optimization problem that simultaneously maximizes population coverage $f(S)$ and alignment with expert advice $g(S)$. However, since the problem is intractable (there is no natural way to scalarize the two objectives) it is often necessary to provide guarantees with respect to the more measurable coverage objective. To make the problem operational, we introduce a constrained bi-objective relaxation that preserves theoretical guarantees on coverage while allowing flexible adaptation to both coverage and advice alignment.

The α–β guarantee problem. To provide theoretical guarantees while maintaining flexibility in alignment, we introduce two control parameters $\alpha, \beta \in [0, 1]$. Let OPT_b denote the optimal allocation of size b that maximizes the coverage function f. We then define the constrained optimization problem as:

$$\max_{S \subseteq V,\, |S|=b} \left\{ f(S),\, g(S) \right\},$$

$$\text{s.t.} \quad f(S) \geq (1 - e^{-\alpha\beta})\, f(OPT_b). \tag{4}$$

The constraint ensures that the final allocation retains at least a fraction $(1 - e^{-\alpha\beta})$ of the optimal coverage value. By adjusting α and β, one can explicitly control the balance between maintaining strong theoretical coverage guarantees and achieving closer alignment with expert guidance. The parameter α defines the proportion of selections made by the greedy approach relative to the LLM-based approach, while β controls the greedy approach's flexibility to deviate from the LLM allocation. The detailed operational meaning of these parameters is discussed further in Section 4.

Remark 1 (Challenges of multi-objective optimization). *Classical scalarization methods—such as weighted sums or Pareto front exploration [10, 19]—require predefined numerical weights for each objective, which are often unavailable when stakeholder preferences are expressed in natural language. Moreover, these approaches treat objectives as static, whereas expert opinions may evolve iteratively during the planning process. Our formulation differs by explicitly integrating language-based feedback within the optimization loop, thereby allowing the trade-off between coverage and alignment to be dynamically adjusted while preserving provable performance guarantees.*

4 PROPOSED METHOD

4.1 Algorithm Overview

Our framework integrates optimization algorithms with LLM-driven reasoning to balance two competing objectives: maximizing population coverage and aligning with qualitative expert advice as in Eq. (3). The overall process proceeds iteratively across five stages (Algorithm 1), alternating between algorithmic updates that ensure theoretical guarantees and language-based refinements that incorporate human guidance. At a high level, it contains 5 steps, i.e., Step 1: A classical greedy algorithm produces an initial coverage-maximizing (low level) grid-cell allocation. Step 2: Given the grid-cell allocation, an LLM produces a (high-level) district allocation using expert advice expressed in natural language. Step 3: A constrained greedy procedure refines the district level allocation into a grid-cell allocation while maintaining an α–β coverage guarantee. Step 4: The LLM receives structured feedback to improve alignment. Step 5: The prompts themselves are iteratively optimized through a form of "prompt gradient descent." We detail each step below.

Step 1: Initial grid-cell Allocation (Greedy in line 3 of Algorithm 1). We first allocate facilities to grid-cells using the standard greedy algorithm of Nemhauser et al. [20], which provides a $(1 - 1/e)$-approximation for maximizing monotone submodular functions. The resulting allocation of facilities to grid-cells, S_0, serves as the initial baseline. From this grid-cell allocation, we extract the district allocation $h(S_0)$, which encodes how many facilities are assigned to each district (without the information about the exact cells), and will be used to guide the subsequent district-level updates.

Step 2: LLM-Powered district allocation (line 6 of Algorithm 1). Next, we refine the district allocation $d = h(S)$ using an LLM that processes both quantitative and qualitative contextual information. The model considers the current and previous district allocations, budget, population statistics, and a set of expert advice sentences $\mathbf{A}$. To maintain stability, we restrict the LLM to modify the districts of at most two facilities per iteration. The full prompt used in this step is in the extended version of the paper [28].

This step produces a revised district allocation that respects both expert intent and practical feasibility. A full description of the prompting strategy is provided in the extended version [28].

Step 3: GuidedGreedy (line 7 of Algorithm 1;Algorithm 2). We consider the district allocation d, as a vector of per-district budget and perform a guided greedy selection (Algorithm 2) to determine specific facility locations. In line 3 of Algorithm 2, we initialize the set $\hat{S}$ as an empty set.

In the loop on lines 4–12 of Algorithm 2, the algorithm greedily adds grid-cells to the set $\hat{S}$. On lines 5–6, we compute the cells with the maximum marginal gains, either ignoring or considering the district allocation d, respectively. The condition in line 8 ensures that at least $\lceil \alpha b \rceil$ cells attain a marginal gain of at least β times the maximum marginal gains computed with no restrictions (i.e., ignoring d). A running example demonstrating Algorithm 2 is given in Table 1.

After obtaining the set S_i, we compute the values $f(S_i)$ and per-district contributions $f(S_i \cap T_j)$ for all $1 \leq j \leq r$. These values constitute the quantitative feedback signals, Δf, and Δh which are then used to guide the next LLM iteration.

Step 4: Verbal Reinforcement via scalar, and vectors' Comparison. (lines 8-9 of Algorithm 1). We quantify the change in both coverage and allocation between iterations by computing:

$$\Delta f = f(S_{i+1}) - f(S_i), \quad \Delta h = h(S_{i+1}) - h(S_i). \tag{5}$$

Table 1: Running example of Algorithm 2, inputs: $\alpha = 0.25, \beta = 0.5, d = \{1:3, 2:0\}$ **; columns: *Iter* = algorithm iteration,** $|S|$**= size of grid-cell allocation at iteration start,** d**= per district budgets at iteration start,** $c(L5)$ **and** $c_d(L6)$ **= values in lines 5 and 6, *L8 Check* = condition outcome (line 8), *Action* = selected action (line 9 or 12).**

Iter	$\|S\|$	d	c (L5)	c_d (L6)	L8 Check	Action
1	0	$\{1:3, 2:0\}$	10	4	$0 > 0$? **No** $4 \geq 0.5 \cdot 10$? **No**	Pick c (10; T_2)
2	1	$\{1:3, 2:0\}$	8	4	$1 > 0$? **Yes**	Pick c_d (4; T_1)
3	2	$\{1:2, 2:0\}$	8	3	$2 > 0$? **Yes**	Pick c_d (3; T_1)

These differences form the basis of a new prompt to the LLM, requesting verbal feedback to improve future alignment. Specifically, the $Feedback$ in the prompt instructs the LLM to (i) describe the observed differences and (ii) propose adjustments that enhance both alignment with expert advice and overall coverage.

Step 5: Prompt optimization. (line 10 of Algorithm 1). Following [35], we replace conventional parameter-space gradient descent with an analogous update in the prompt space. We decompose each prompt into a static and editable part:

$$\text{Prompt}_i = P_{\text{Fix}} \parallel P_{\text{Editable},i}, \tag{6}$$

where P_{Fix} encodes the task template and $P_{\text{Editable},i}$ accumulates iteration-specific refinements. After receiving new feedback, we update

$$P_{\text{Editable},i+1} = P_{\text{Editable},i} + Feedback_{i+1}, \tag{7}$$

yielding the new prompt $\text{Prompt}_{i+1} = P_{\text{Fix}} \parallel P_{\text{Editable},i+1}$. This procedure gradually improves prompt quality and alignment consistency over time. The high level of the suggested approach is given in Algorithm 1.

Algorithm 1 LLM-Enhanced District Resource Allocation

1: **Input:** Budget b, advice set **A**, balance parameters $\alpha, \beta \in [0, 1]$, available grid-cells V, prompts $Prompt_1$, $Prompt_{reflection}$.
2: **Output:** Grid-cell allocation S_{limit}.
3: (Init) $S_0 \leftarrow Greedy(b, V)$
4: (Init) Update $Prompt_1$ with S_0, $h(S_0)$ and the other inputs.
5: **for** $i = 1$ to limit **do**
6:     (LLM) $d \leftarrow \text{LLM}(\text{Prompt}_i)$
7:     (Solve) $S_i \leftarrow GuidedGreedy(\alpha, \beta, b, V, \emptyset, d)$.
8:     (Compare) Compute differences Δf, Δh (Eq. 5) and update $Prompt_{reflection}$.
9:     (Verbal Feedback) Query LLM to generate verbal reflection.
10:     (Optimize Prompt) Update Prompt_{i+1}, based on Eq. (6), (7).
11: **Return:** S_{limit}

Remark 2 (A heuristic improvement for $\beta = 1.0$). *When $\beta = 1.0$, we use the following heuristic to improve the performance of Algorithm 1: before line 6, we run the Greedy algorithm to allocate $\lceil \alpha \cdot b \rceil$ facilities, and provide this allocation to the LLM as a part of the prompt in line 6 (the greedy allocation is serving as a minimum district allocation, like we do in line 6 of Algorithm 3). We then skip line 7 and continue with line 8. This approach allows the LLM to leverage the greedy selection while preserving the same theoretical guarantee.*

Algorithm 2 GuidedGreedy(α, β, b, V, d)

1: **Input:** Parameters $\alpha, \beta \in [0, 1]$; budget $b \in \mathbb{N}_{>0}$; available cells V;existing district allocation S, district budgets $d \in \mathbb{N}^r$.
2: **Output:** Grid-cell allocation $\hat{S}$.
3: $\hat{S} \leftarrow S$
4: **while** $|\hat{S}| < b$ **do**
5:     $c \leftarrow \arg\max_{v \in V} \left(f(\hat{S} \cup \{v\}) - f(\hat{S})\right)$
6:     $c_d \leftarrow \arg\max_{v \in \bigcup_{j:d_j>0} T_j} \left(f(\hat{S} \cup \{v\}) - f(\hat{S})\right)$
7:     Let j_d be the district such that $c_d \in T_{j_d}$
8:     **if** $|\hat{S}| > \lceil \alpha b \rceil$ **or** $f(\hat{S} \cup \{c_d\}) - f(\hat{S}) \geq \beta \cdot (f(\hat{S} \cup \{c\}) - f(\hat{S}))$ **then**
9:         $\hat{S} \leftarrow \hat{S} \cup \{c_d\}$
10:         $d_{j_d} \leftarrow d_{j_d} - 1$
11:     **else**
12:         $\hat{S} \leftarrow \hat{S} \cup \{c\}$
13: **return** $\hat{S}$

4.2 Theoretical Guarantee

Since the coverage function f is monotone and submodular, the standard greedy algorithm achieves a $(1 - 1/e)$ approximation ratio [20]. We extend this result to our hybrid framework, which combines greedy selection with LLM-guided refinement.

Theorem 4.1 generalizes the classic result of [12]. While that work establishes approximation guarantees for allocations closely following a purely greedy strategy, we extend the analysis to two dimensions: (i) arbitrary monotone submodular functions (beyond coverage-specific objectives), and (ii) partial selection settings where fewer than k cells are chosen. This latter generalization is crucial in our setup, as our iterative framework allows the LLM to adjust or supplement a subset of the greedy-selected locations.

Theorem 4.1. *Let $b \in \mathbb{N}_{>0}$ and $\alpha, \beta \in [0, 1]$ be the balance parameters provided as inputs to Algorithm 1. Let $OPT_b \subseteq V$ denote an optimal subset of size b maximizing f. Denote by S_{limit} the allocation returned by Algorithm 1. Then*

$$f(S_{limit}) \geq \left(1 - (1 - \tfrac{\beta}{b})^{\alpha b}\right) f(OPT_b) \geq \left(1 - e^{-\alpha\beta}\right) f(OPT_b).$$

This result guarantees that even when part of the selection process is delegated to the LLM—subject to the α–β trade-off parameters—the final allocation still retains an exponential coverage bound comparable to the classical greedy algorithm. The parameters α and β are provided to Ethiopian stakeholders as tools to determine the optimal balance between offering guarantees and adhering to expert advice. We anticipate that different regions in Ethiopia will select varying guarantee parameters, depending on their specific need to balance the influence of current human planners with coverage considerations. The parameter β is particularly useful when the coverage values of different cells are similar. In such cases, it prevents the greedy algorithm from selecting a cell solely because it offers a marginally better coverage gain despite much worse alignment with the advice. The parameter α allows the LLMs some

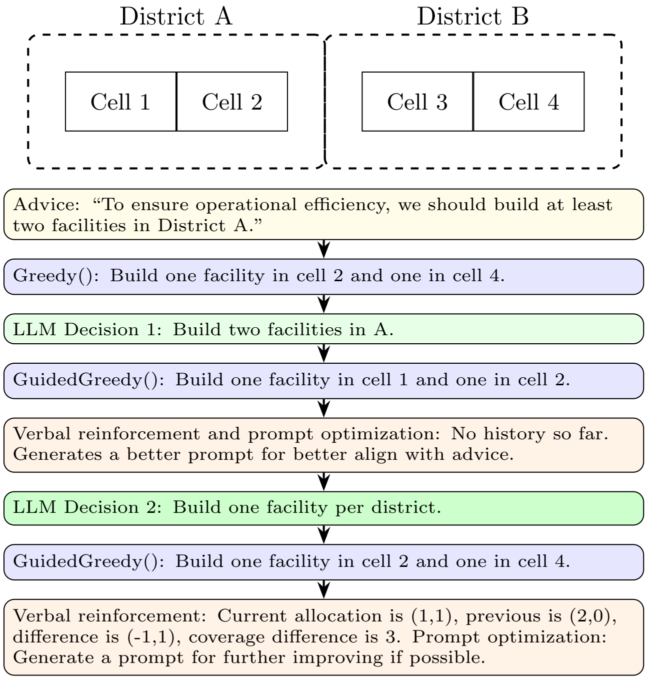

Figure 4: Toy running example with two districts and four cells; For GuidedGreedy, we use the parameters $\alpha = 0, \beta = 1$ allowing the LLM to fully determine the district allocation. Greedy() denotes the greedy algorithm of [20].

flexibility to follow the advice even when it recommends cells that provide very low coverage improvement.

An Example. In Figure 4, we present a toy running example of our approach. The example considers a region consisting of two districts, each containing two grid-cells. Each grid-cell represents a 1 km × 1 km area on a map of Ethiopia, and our approach focuses on selecting districts and cells for building health facilities. Given an advice sentence, we first apply the algorithm of [20] to obtain a grid-cell allocation. We then invoke the LLM to produce a district allocation, incorporating this grid-cell allocation among other factors. Based on this district allocation, we run the GuidedGreedy algorithm to compute a refined grid-cell allocation. Verbal reinforcement and prompt optimization are used to improve the prompt, and the updated prompt is supplied to the LLM in the subsequent iteration. This procedure repeats until a final grid-cell allocation is obtained. Note that typical regions in Ethiopia generally have many more districts.

4.3 Supporting Online Budget

An important extension of our framework considers the setting where budgets arrive sequentially over multiple rounds [5], for example, one allocation round per fiscal year. We assume that all candidate grid-cells remain available throughout the planning horizon and that new budget increments become available at the beginning of each round. To adapt Algorithm 1 to this setting, we modify its update rule to incorporate the historical allocations from previous rounds. Specifically, in line 6 of Algorithm 3, the prompt to the LLM now includes a *minimum district allocation vector* representing facilities already selected in earlier years. This ensures that the LLM respects all previously established allocations and builds upon them when optimizing the new round.

Let b_r denote the available budget in round r, and suppose the process has reached round t. The cumulative number of facilities planned up to this point is therefore $b = \sum_{r=1}^{t} b_r$. The LLM prompt for round t requests an updated allocation consistent with this cumulative budget and the existing district allocation. The multi-round variant of our approach is summarized in Algorithm 3.

Algorithm 3 Multiple Rounds Facility Location

1: **Input:** Budgets $b_1, ..., b_t$, advice set $\mathbf{A}$, balance parameters $\alpha, \beta \in [0, 1]$, available grid cells V, existing facilities S_e, prompts $Prompt_0$, $Prompt_{reflection}$, time horizon t.
2: **Output:** Allocation S
3: $S \leftarrow \emptyset$
4: **for** $r = 1$ to t **do**
5:     (Init) $S_0 \leftarrow Greedy(b, V)$
6:     (Init) Update $Prompt_0$ with S_0, $d(S_0)$, minimum district allocation $h(S)$, and the other inputs.
7:     **for** $i = 1$ to limit **do**
8:         (LLM) $d \leftarrow$ LLM(Prompt$_i$).
9:         (Solve) $S_i \leftarrow GuidedGreedy(\alpha, \beta, b_r, V, S, d)$.
10:         (Compare) Compute differences Δf, Δh (Eq. 5) and update $Prompt_{reflection}$.
11:         (Verbal Feedback) Query LLM
12:             to generate verbal reflection.
13:         (Optimize Prompt) Update Prompt$_i$,
14:             based on Eq. (6), (7).
15:     $S \leftarrow S_{limit}$
16: **Return:** S

The following theorem expands the results of Theorem 4.1 to the online settings.

THEOREM 4.2. *Let $b_1, \ldots, b_t$ denote the budgets, and let $\alpha, \beta \in [0, 1]$ be the balance parameters provided as inputs to Algorithm 3. Let $b = \sum_{i=1}^{t} b_i$ and let $OPT_b \subseteq V$ denote an optimal subset of size b maximizing f. Denote by S the allocation returned by Algorithm 3. Then*

$$f(S) \geq \left(1 - (1 - \frac{\beta}{b})^{\alpha b}\right) f(OPT_b) \geq \left(1 - e^{-\alpha\beta}\right) f(OPT_b).$$

REMARK 3. *This online-budget formulation allows the model to handle dynamically arriving resources while maintaining continuity across years. In practice, this property is essential for national-scale public health planning, where upgrades are deployed incrementally based on evolving financial and logistical constraints. The framework thus generalizes naturally from a single-shot allocation problem to an online, multi-horizon setting, preserving the same theoretical guarantees on coverage at each stage.*

5 EXPERIMENTS

5.1 Dataset

We evaluate the proposed framework using Ethiopia's projected 2026 population data [34]. Walking accessibility is computed following [5] with the global friction map [33], assuming a maximum two-hour walking distance. The experiments cover three regions—Afar, Somali, and Benishangul Gumuz—characterized by diverse terrain and population densities. To emulate multi-stakeholder guidance, we queried Gemini-2.5-Pro to generate 20 expert advice sentences grouped into four sets (five per expert) with intentional contradictions to mimic real decision environments. For quantitative evaluation, an alignment function was automatically synthesized by Gemini-2.5-Pro to score allocations between 0 and 1 according to the textual advice. All iterative allocation and feedback steps were executed using Gemini-2.5-Flash as shown in Figure 3. The used prompts are available in the extended version of the paper [28].

5.2 Experiment 1: Is the Verbal Feedback Useful?

In Figures 5,6,7 we present the results for two variants of our approach: i) **Quantitative feedback**, where alignment signals are numeric; and ii) **Verbal feedback**, where updates are guided by natural-language reflections from the LLM. Across all three regions, the verbal-feedback variant achieved notably higher advice alignment, confirming that language-based reasoning effectively internalizes qualitative goals. The quantitative variant, in contrast, yielded slightly higher coverage values, reflecting its direct optimization focus. Importantly, both variants consistently satisfied the theoretical $1 - e^{-\alpha\beta}$ coverage guarantee (even if we assume that the optimal selection, OPT_b, covers the entire population). Among the three regions, Benishangul Gumuz exhibited the largest gap between the two modes, suggesting stronger contextual heterogeneity where language input provides added value.

Remark 4. *Note that in all regions, the coverage decreases as the alignment with advice increases. We justify this decrease by the additional knowledge that the human expert provides. For example, we may prioritize improving health outcomes in districts with poor existing health status, even if it results in a lower overall coverage.*

5.3 Experiment 2: Change in Actual Coverage and Alignment as the Parameter α Increases

In Figure 8, we measured changes in coverage and alignment to advice as a function of the parameter α, while keeping $\beta = 1$. As expected, larger values of α lead to higher coverage improvement but lower advice alignment. One exception is that in the Afar region, the coverage in $\alpha = 0.0$ is worse than in $\alpha = 0.25$, indicating that, in practice, setting $\alpha = 0.25$ offers little benefit for coverage, since the LLM iterative approach achieves similar results without enforcing greedy selections. One thing to note is that as the greedy algorithm is deterministic, there are no variance bars for $\alpha = 1$. These results validate the interpretability of α: small α values promote qualitative conformity, while larger α values recover the classical submodular-optimization behavior.

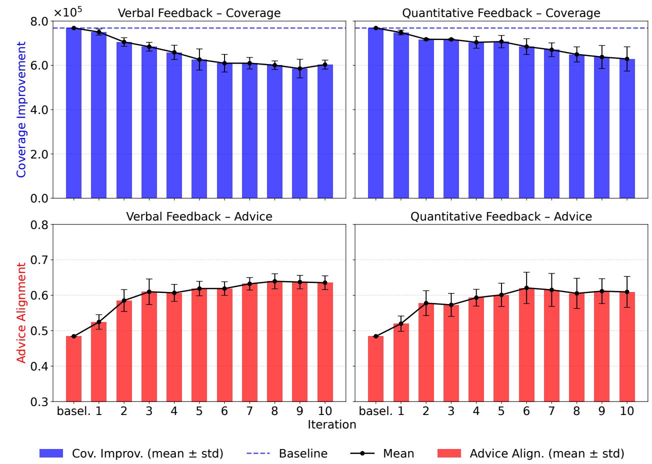

Figure 5: Verbal vs. quantitative feedback for Afar; $\alpha = 0.0$, $\beta = 1.0$.

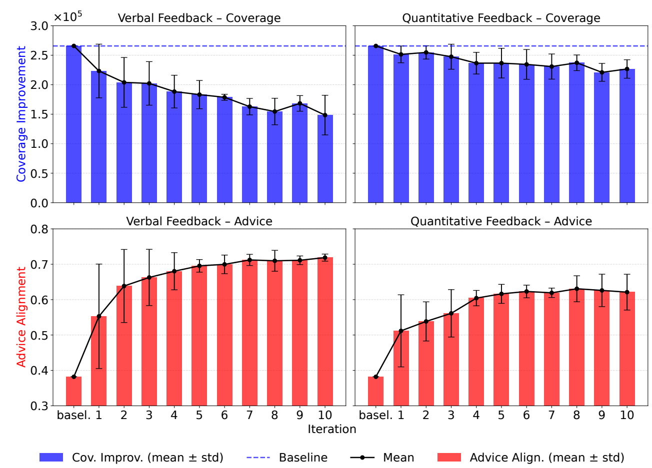

Figure 6: Verbal vs. quantitative feedback for Benishangul Gumuz; $\alpha = 0.0$, $\beta = 1.0$.

5.4 Experiment 3: Is It Useful to Have a Longer Time Window?

In the extended version [28], we examine whether incorporating a longer feedback history improves the performance of verbal refinement. Specifically, we compare two variants: one that includes information from the past three iterations (*3-step window*) and another that considers only the immediately preceding iteration (*1-step window*). In the *no-window* variant, the feedback prompt contained the current and previous allocations but excluded their differences in allocation and total coverage. Overall, after ten iterations, the observed differences between the two best variants (1-step and 3-step windows) were minimal in both population coverage and advice alignment. This suggests that the LLM effectively internalizes historical trends even with limited contextual information, making explicit multi-step computation largely redundant.

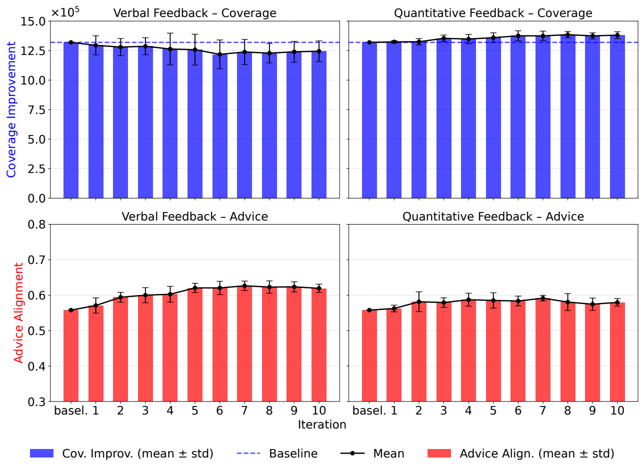

Figure 7: Verbal vs. quantitative feedback for Somali; $\alpha = 0.0$, $\beta = 1.0$.

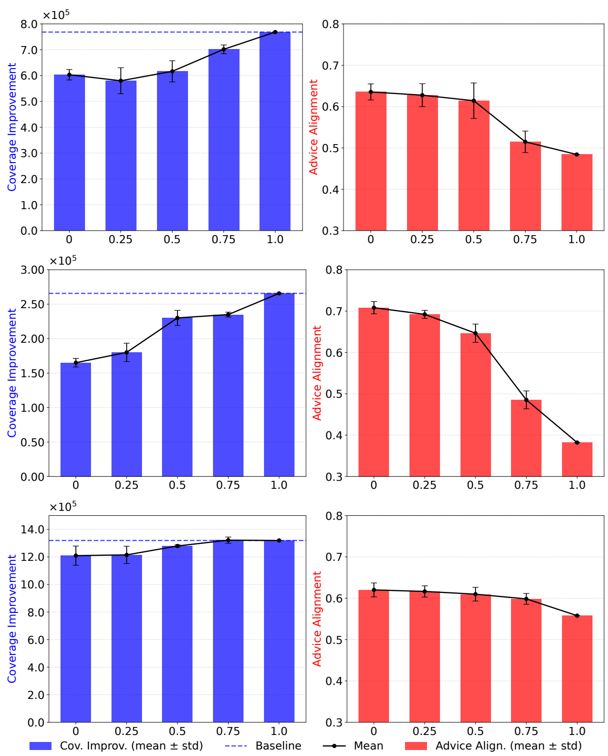

Figure 8: Coverage and alignment for different values of α in Afar (top), Benishangul Gumuz (middle), and Somali (bottom) regions; $\beta = 1.0$.

Furthermore, extending the feedback window to three iterations did not yield measurable gains in either metric. Consequently, we adopt the 1-step window setting as our default configuration, as it achieves comparable performance while being slightly more stable and computationally efficient.

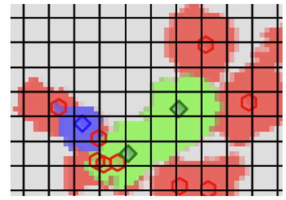

Figure 9: Red hexagons indicate existing facilities. Green rhombuses show locations selected by the algorithm with $\alpha = 0.5$, while blue rhombuses correspond to $\alpha = 0$ (β is always 1.0). Colored areas represent the accessibility areas of the facilities. Each grid cell covers 5 km × 5 km.

5.5 Analyzing Visually the Different Allocation Algorithms

Figure 9 presents the resulting facility allocations in a subregion of the Afar region, generated under two parameter settings: $\alpha = 0, \beta = 1$ and $\alpha = 0.5, \beta = 1$. This visualization highlights how different trade-off configurations lead to distinct spatial allocation patterns. While the two solutions exhibit substantial overlap in high-priority areas—indicating agreement on core population centers—the overlap is not complete. The $\alpha = 0.5$ configuration, which emphasizes coverage optimization, tends to expand toward denser clusters, whereas the $\alpha = 0$ configuration better reflects qualitative expert preferences. Together, these contrasting patterns illustrate the interpretability and tunability of the proposed framework in balancing quantitative coverage with qualitative alignment.

5.6 Summary of Findings

Overall, the experimental results demonstrate that incorporating language-based feedback substantially enhances interpretability without compromising the theoretical coverage guarantees of the underlying optimization framework. The trade-off parameter α provides a clear and intuitive mechanism for balancing quantitative coverage performance with qualitative expert alignment, enabling flexible policy tuning across regions. Moreover, the analysis shows that short feedback windows are sufficient for stable convergence, indicating that the LLM can effectively capture temporal patterns without extensive historical context. Finally, the visual comparisons reveal that different parameter settings produce distinct yet interpretable allocation patterns, offering planners and policymakers valuable insights into the trade-offs between efficiency and fairness in real-world deployment.

6 CONCLUSIONS

In this paper, we present a novel framework for optimizing bi-objective problem in complex environment. Our LEG framework employs an adaptive greedy algorithm integrated within an LLM-based refinement loop, ensuring coverage guarantees while jointly optimizing advice alignment and coverage. We evaluated the framework on three real-world datasets, demonstrating its effectiveness. Combined with existing literature, our work brings us closer to addressing pressing real-world challenges.

ACKNOWLEDGMENTS

We are very grateful to Kasahun Sime and Wondesen Nigatu of the Ministry of Health, Ethiopia, for their continuous support and valuable advice. The findings and conclusions in this report are those of the authors and do not represent any recommendations or official positions.

REFERENCES

[1] Pieter Abbeel and Andrew Y Ng. 2004. Apprenticeship learning via inverse reinforcement learning. In *International conference on Machine learning*.

[2] Amir Ahmadi-Javid, Pardis Seyedi, and Siddhartha S Syam. 2017. A survey of healthcare facility location. *Computers & Operations Research* 79 (2017), 223–263.

[3] Sisay Mulugeta Alemu, Gerd Weitkamp, Abera Kenay Tura, Jelle Stekelenburg, and Regien Biesma. 2025. Optimizing maternal health facility locations in Ethiopia: A geospatial approach for upgrading to emergency obstetric and newborn care. *Journal of Transport Geography* 125 (2025), 104208.

[4] Nikhil Behari, Edwin Zhang, Yunfan Zhao, Aparna Taneja, Dheeraj Nagaraj, and Milind Tambe. 2024. A decision-language model (dlm) for dynamic restless multi-armed bandit tasks in public health. *Advances in Neural Information Processing Systems* 37 (2024), 3964–4002.

[5] Davin Choo, Yohai Trabelsi, Fentabil Getnet, Samson Warkaye Lamma, Wondesen Nigatu, Kasahun Sime, Lisa Matay, Milind Tambe, and Stéphane Verguet. 2025. Optimizing Health Coverage in Ethiopia: A Learning-augmented Approach and Persistent Proportionality Under an Online Budget. In *Proceedings of the AAAI Conference on Artificial Intelligence*. AISI Track, to appear; also available as arXiv preprint arXiv:2509.00135.

[6] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems* 30 (2017).

[7] Lutfi Eren Erdogan, Nicholas Lee, Sehoon Kim, Suhong Moon, Hiroki Furuta, Gopala Anumanchipalli, Kurt Keutzer, and Amir Gholami. 2025. Plan-and-act: Improving planning of agents for long-horizon tasks. In *International Conference on Machine Learning*. Also available as arXiv preprint arXiv:2503.09572.

[8] Lorenzo Jaime Yu Flores, Ramon Rafael Tonato, Gabrielle Ann dela Paz, and Valerie Gilbert Ulep. 2021. Optimizing health facility location for universal health care: A case study from the Philippines. *PloS one* 16, 9 (2021), e0256821.

[9] Abraham Haileamlak and Israel Ataro. 2023. The Ethiopian health extension program (HEP) is still relevant after 15 years of implementation although major transformation is essential to sustain its gains and relevance. *Ethiopian journal of health sciences* 33, Spec Iss 1 (2023), 1.

[10] Conor F Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M Zintgraf, Richard Dazeley, Fredrik Heintz, et al. 2022. A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems* 36, 1 (2022), 26.

[11] Nathaniel Hendrix, Samson Warkaye, Latera Tesfaye, Mesfin Agachew Woldekidan, Asrat Arja, Ryoko Sato, Solomon Tessema Memirie, Alemnesh H Mirkuzie, Fentabil Getnet, and Stéphane Verguet. 2023. Estimated travel time and staffing constraints to accessing the Ethiopian health care system: A two-step floating catchment area analysis. *Journal of Global Health* 13 (2023), 04008.

[12] Dorit S Hochbaum and Anu Pathria. 1998. Analysis of the greedy approach in problems of maximum k-coverage. *Naval Research Logistics (NRL)* 45, 6 (1998), 615–627.

[13] Zican Hu, Wei Liu, Xiaoye Qu, Xiangyu Yue, Chunlin Chen, Zhi Wang, and Yu Cheng. 2025. Divide and Conquer: Grounding LLMs as Efficient Decision-Making Agents via Offline Hierarchical Reinforcement Learning. In *International Conference on Machine Learning*. Also available as arXiv preprint arXiv:2505.19761.

[14] Jiaming Ji, Tianyi Qiu, Boyuan Chen, Borong Zhang, Hantao Lou, Kaile Wang, Yawen Duan, Zhonghao He, Jiayi Zhou, Zhaowei Zhang, et al. 2023. Ai alignment: A comprehensive survey. *arXiv preprint arXiv:2310.19852* (2023).

[15] Cheol Woo Kim, Jai Moondra, Shresth Verma, Madeleine Pollack, Lingkai Kong, Milind Tambe, and Swati Gupta. 2025. Navigating the Social Welfare Frontier: Portfolios for Multi-objective Reinforcement Learning. In *International Conference on Machine Learning*.

[16] Cheol Woo Kim, Shresth Verma, Mauricio Tec, and Milind Tambe. 2026. Preference Robustness for DPO with Applications to Public Health. In *Proceedings of the AAAI Conference on Artificial Intelligence (to appear)*. Also available as arXiv preprint arXiv:2509.02709.

[17] Minae Kwon, Sang Michael Xie, Kalesha Bullard, and Dorsa Sadigh. 2023. Reward Design with Language Models. In *International Conference on Learning Representations*. Also available as arXiv preprint arXiv:2303.00001.

[18] Yecheng Jason Ma, William Liang, Guanzhi Wang, De-An Huang, Osbert Bastani, Dinesh Jayaraman, Yuke Zhu, Linxi Fan, and Anima Anandkumar. 2024. Eureka: Human-Level Reward Design via Coding Large Language Models. In *International Conference on Learning Representations*. Also available as arXiv preprint arXiv:2310.12931.

[19] Kaisa Miettinen. 1999. *Nonlinear multiobjective optimization*. Vol. 12. Springer Science & Business Media.

[20] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. 1978. An analysis of approximations for maximizing submodular set functions—I. *Mathematical programming* 14, 1 (1978), 265–294.

[21] Andrew Y Ng, Stuart Russell, et al. 2000. Algorithms for inverse reinforcement learning. In *International Conference on Machine Learning*.

[22] Ministry of Health [Ethiopia]. 2020. Realizing Universal Health Coverage through Primary Healthcare: A Roadmap for Optimizing the Ethiopian Health Extension Program (2020–2035).

[23] Sebastian Peitz and Sedjro Salomon Hotegni. 2025. Multi-objective deep learning: Taxonomy and survey of the state of the art. *Machine Learning with Applications* (2025), 100700.

[24] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. 2013. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research* 48 (2013), 67–113.

[25] SS Radiah Shariff, Noor Hasnah Moin, and Mohd Omar. 2012. Location allocation modeling for healthcare facility planning in Malaysia. *Computers & Industrial Engineering* 62, 4 (2012), 1000–1010.

[26] Hua Shen, Tiffany Knearem, Reshmi Ghosh, Kenan Alkiek, Kundan Krishna, Yachuan Liu, Ziqiao Ma, Savvas Petridis, Yi-Hao Peng, Li Qiwei, et al. 2024. Towards bidirectional human-ai alignment: A systematic review for clarifications, framework, and future directions. *arXiv preprint arXiv:2406.09264* (2024).

[27] Haotian Sun, Yuchen Zhuang, Lingkai Kong, Bo Dai, and Chao Zhang. 2023. Adaplanner: Adaptive planning from feedback with language models. *Advances in neural information processing systems* 36 (2023), 58202–58245.

[28] Yohai Trabelsi, Guojun Xiong, Fentabil Getnet, Stéphane Verguet, and Milind Tambe. 2026. Health Facility Location in Ethiopia: Leveraging LLMs to Integrate Expert Knowledge into Algorithmic Planning. *arXiv preprint arXiv:2601.11479* (2026).

[29] Shresth Verma, Niclas Boehmer, Lingkai Kong, and Milind Tambe. 2025. Balancing act: prioritization strategies for llm-designed restless bandit rewards. In *International Conference on Game Theory and AI for Security*. Springer, 376–394.

[30] Hao Wang, Peng Luo, and Yimeng Wu. 2023. Research on the location decision-making method of emergency medical facilities based on WSR. *Scientific Reports* 13, 1 (2023), 18011.

[31] Huihui Wang, Roman Tesfaye, Gandham N. V. Ramana, and Chala Tesfaye Chekagn. 2016. *Ethiopia Health Extension Program: An Institutionalized Community Approach for Universal Health Coverage*. World Bank.

[32] Yufei Wang, Wanjun Zhong, Liangyou Li, Fei Mi, Xingshan Zeng, Wenyong Huang, Lifeng Shang, Xin Jiang, and Qun Liu. 2023. Aligning large language models with human: A survey. *arXiv preprint arXiv:2307.12966* (2023).

[33] D. J. Weiss, A. Nelson, C. A. Vargas-Ruiz, K. Gligorić, S. Bavadekar, E. Gabrilovich, A. Bertozzi-Villa, J. Rozier, H. S. Gibson, and et al. Shekel, T. 2020. Global maps of travel time to healthcare facilities. *Nature Medicine* 26 (2020), 1835–1838.

[34] WorldPop. 2025. WorldPop Population Counts 2015–2030: Ethiopia — Spatial Distribution of Population. https://data.humdata.org/dataset/worldpop-population-counts-2015-2030-eth. Constrained estimates of total population per 100m grid cell in Ethiopia for 2015–2030, Geotiff format, CC BY4.0 license.

[35] Guojun Xiong and Milind Tambe. 2026. VORTEX: Aligning Task Utility and Human Preferences through LLM-Guided Reward Shaping. (2026). To appear; also available as arXiv preprint arXiv:2509.16399.

[36] Ke Yang, Yao Liu, Sapana Chaudhary, Rasool Fakoor, Pratik Chaudhari, George Karypis, and Huzefa Rangwala. 2025. Agentoccam: A simple yet strong baseline for llm-based web agents. In *International Conference on Learning Representations*. Also available as arXiv preprint arXiv:2410.13825.

[37] Yao Zhang, Zijian Ma, Yunpu Ma, Zhen Han, Yu Wu, and Volker Tresp. 2025. Webpilot: A versatile and autonomous multi-agent system for web task execution with strategic exploration. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 23378–23386.