ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.834

GROWIA: Claude Sonnet 4 LLM-Based Web Application for Promoting Healthy Food Production Through Urban Agriculture in Lima, Peru

Tate-Carhuaricra, Kristy Emma[1]; Mendoza-Meza, David Fabrizio[2]; Copaja-Cornejo, Richard Nivaldo[3]
[1,2,3]Universidad Peruana de Ciencias Aplicadas, Lima, Perú, u20211c227@upc.edu.pe, u202110957@upc.edu.pe, pcsircop@upc.edu.pe

Abstract*– Urban agriculture (UA) represents a critical solution for food security in urban environments; however, its widespread adoption is limited by the lack of specialized and personalized guidance that adapts to residents' space constraints and microclimate conditions. This study presents GROWIA, a web application leveraging Claude Sonnet 4 Large Language Model (LLM) through specialized prompt engineering to democratize agricultural knowledge and facilitate the production of nutritious foods in Metropolitan Lima. Unlike traditional rule-based systems requiring manual rule encoding or fine-tuning approaches demanding significant computational resources, GROWIA employs a carefully crafted prompt engineering methodology that transforms a general-purpose LLM into a domain-specific agricultural advisor. Implemented with a three-tier architecture and specialized prompt engineering, the application processes contextual user information to generate personalized cultivation plans through natural language interaction. Dual validation included technical evaluation by 8 agronomist experts on 19 cultivation plans and usability testing with 14 end users without experience. Results demonstrate a Correct Recommendations Rate of 89.5% (p=0.032), a System Usability Scale score of 85.2, a 55.4% reduction in perceived barriers, and 71.4% effective adoption. GROWIA demonstrates that domain-specialized LLMs through prompt engineering can effectively democratize agricultural expertise, offering a scalable, replicable solution for addressing urban food security challenges across Latin American metropolitan contexts.

Keywords-- **Prompt engineering, large Language Models, urban agriculture, food security, AI-powered recommendations.**

I. Introduction

Access to nutritious and safe food represents a critical challenge in Peru, especially in urban areas with high population density. According to the report published by the Food and Agriculture Organization of the United Nations (FAO) in 2024, during 2023, 51.7% of the Peruvian population experienced moderate or severe food insecurity, equivalent to approximately 17.6 million people, of whom 6.9 million faced severe food insecurity conditions [1].

Faced with this situation, various Latin American and Caribbean countries have implemented multidimensional strategies to improve access to healthy foods, addressing nutritional, monetary and community aspects [2]. In Peru, some institutions have adopted similar approaches; however, these initiatives do not integrate technological solutions that allow scaling, personalizing, and automating assistance in urban food production.

In the academic field, innovative proposals have been developed that integrate artificial intelligence (AI) into agricultural systems. For example, the Adaptive-RAG (Adaptive Retrieval-augmented generation) framework has been applied to optimize greenhouse management through AI [3], while in China a language model specialized in agricultural pest and disease management has been designed [4]. However, scientific literature focused on the use of intelligent technologies to facilitate direct access to nutritious foods in urban environments remains scarce. This gap between available technological capabilities and the needs of urban consumers motivates the development of specific solutions that democratize specialized agricultural knowledge.

The motivation of this research is to enable residents of Metropolitan Lima without agricultural experience to produce nutritious foods through personalized urban agriculture advisory. This work aspires to contribute to the reduction of food insecurity in Lima through a scalable technological solution that facilitates the self-supply of healthy foods in urban environments with limited spaces.

This work proposes the design and implementation of GROWIA, a web application leveraging Claude Sonnet 4 LLM through prompt engineering to provide personalized urban agriculture advisory via natural language interaction. The application interprets queries, processes contextual user information (available space, district climate conditions, nutritional objectives, previous experience), and generates specific cultivation plans leveraging the language model's reasoning capability.

The paper is organized as follows: Section II reviews the related work. Section III describes the architecture and components of the proposed system. Section IV details the experimental methodology. Section V discusses the results. Finally, conclusions are outlined in Section VI.

II. Related works

In the field of artificial intelligence applied to agriculture, nine articles were identified that present approaches closely related to this research. These studies represent recent advances in applying AI technologies to diverse agricultural contexts.

TABLE I
Summary of Approaches and Results from Related Research

Approach	Contribution	Results	Year	Source

Expert AI model	xAI system for detecting climate risks in European agriculture, combining expert knowledge and ML.	Good performance in temperature events, acceptable in droughts and rainfall deficits.	2025	[5]
Deep learning-based AI model	DB-CGANNet model for disease detection in rice leaves.	Accuracy of 98.9% and 99.08%, better than previous CNN, RNN, and hybrid models.	2025	[6]
Interpretable AI model	ARAG framework for greenhouse management with interpretable AI and natural language explanations.	Better coherence and utility compared to basic systems.	2025	[7]
AI integration with IoT	Integration of AI with dynamic and adaptive IoT irrigation systems.	The system achieved IWUE = 1.37kg/m³ in wheat, with potential for improvement.	2025	[8]
Specialized LLM AI model	Specialized LLM (IPM-AgriGPT) for pest and disease management.	Outperforms ChatGLM3-6B and Chinese-LLaMA-7B; comparable to GPT-3.5-Turbo.	2025	[9]
Crop recommendation system with explainable AI (XAI)	Combines decision trees with LIME for interpretable recommendations.	R^2 = 0.9415 and errors less than 1.0 in yield prediction.	2024	[10]
Smart hydroponic system with IoT sensors and ML	Integrates ESP32, Bluetooth, and five algorithms for real-time recommendations.	97.5% accuracy with Random Forest and 90% efficiency in water conservation.	2024	[11]
Validation of IoT and AI tool effectiveness in urban agriculture	Mixed methodology: literature analysis and surveys of urban farmers.	70% reported better operational efficiency and 30% reduction in water use.	2023	[12]
Agricultural chatbot based on Artificial Neural Network	NLP architecture with ANN for intent and entity classification.	94.2% accuracy in classification and 1.3s average response time.	2022	[13]

The work [5] presents an explainable artificial intelligence (xAI) model that integrates agroclimatic expert knowledge with an ensemble of 1000 XGBoost models for multi-risk detection of adverse climate events in European agriculture. The approach stands out for its capability to generate probabilistic results with uncertainty estimation and interpretable explanations based on SHAP metrics, demonstrating effectiveness in managing risks associated with heat waves, cold spells, droughts, and water deficits.

The study [6] develops a hybrid neural network that combines graph convolutions and attention modules for detection and classification of diseases in rice leaves. Its method incorporates advanced preprocessing and bio-inspired feature selection, achieving an accuracy rate of 98.9%, superior to previous models.

On the other hand, the work [7] proposes an interpretable artificial intelligence framework for greenhouse management that translates complex predictive control (MPC) decisions into natural language explanations adapted to the user's technical level. It integrates semantic techniques and GPT-4o-based language generation, improving the accessibility and utility of the system for farmers.

The study [8] integrates AI with automated IoT irrigation systems in Eastern Europe, evolving from static models to dynamic real-time decisions to optimize water use. The system combines environmental sensors, microcontrollers, and OpenAI models to generate contextualized recommendations, achieving a water efficiency index of 1.37kg/m³, a competitive result that demonstrates the scalability and replicability of the solution in the face of climate challenges.

The work [9] develops a large language model (LLM) specialized in agricultural pest and disease management. It uses advanced fine-tuning techniques with LoRA and reasoning distillation (CoT) to adapt a generalist model to a vertical domain, improving accuracy, safety, and reduction of hallucinations in responses, surpassing ChatGLM3-6B and Chinese-LLaMA-7B, and being comparable to GPT-3.5 Turbo.

The research [10] proposes XAI-CROP, a crop recommendation system based on decision tree algorithms integrated with explainable artificial intelligence techniques. This system uses Local Interpretable Model-agnostic Explanations (LIME) to provide understandable justifications for agricultural recommendations, achieving a coefficient of determination of 0.94152 and prediction errors consistently below the critical threshold of 1.0.

The study [11] presents an intelligent hydroponic system based on AIoT that integrates specialized sensors (NPK, pH, humidity) with machine learning algorithms to automate crop recommendations and real-time nutritional monitoring. This approach employs five algorithms trained with data from the Indian Chamber of Food and Agriculture, where Random Forest achieves 97.5% accuracy. Experimental validation with lettuce cultivation demonstrates 90% efficiency in water conservation compared to traditional methods.

The research [12] proposes a technology adoption framework in urban agriculture through a mixed research methodology that includes literature analysis and farmer surveys. Results show that the integration of IoT and AI improves operational efficiency, where 70% of farmers report measurable improvements and automated irrigation systems achieve a 30% reduction in water use, ensuring sustainability in urban contexts.

Finally, the study [13] develops AgroBot, an agricultural assistant chatbot based on Artificial Neural Networks (ANN) to address information gaps in the agricultural sector through natural language processing. The ANN architecture compares queries against a structured agricultural knowledge base, achieving 94.2% accuracy in intent classification and an average response time of 1.3 seconds.

While these studies demonstrate significant advances in AI applications for agriculture, they focus primarily on commercial settings or require fine-tuning/specialized knowledge bases [9], [13], increasing implementation complexity. GROWIA addresses these gaps by leveraging prompt engineering for domain specialization, enabling rapid deployment in resource-constrained Latin American contexts without computational-intensive retraining.

III. Contribution

A. *Methodology and Proposed Contribution*

Limited access to nutritious foods in Metropolitan Lima represents a challenge for the food security of the urban population. Although urban agriculture (UA) offers a viable solution for local food production, residents face significant barriers including uncertainty about the appropriate crops for constrained urban spaces, lack of personalized technical guidance adapting to individual microclimates, and absence of accessible expertise for maintenance and troubleshooting.

Traditional agricultural advisory approaches—government extension services, printed manuals, generic online resources—fail to provide personalized, interactive guidance for urban residents without experience. These solutions face scalability limitations and cannot adapt to individual contexts (specific space dimensions, sun exposure patterns, time availability).

1) Natural Language Interaction: Users describe their situation conversationally without requiring technical terminology or structured forms. The system understands colloquial descriptions such as "I have a small balcony with morning sun" and translates these into actionable agricultural parameters.

2) Context-Aware Personalization: The LLM processes multiple contextual factors simultaneously—available space geometry, sun exposure patterns, user time availability, district climate conditions, tool availability, and experience level—to generate recommendations specifically adapted to each user's unique constraints. This multi-dimensional personalization cannot be achieved through simple decision trees or static databases.

3) Scalable Knowledge Democratization: The system provides expert-level agricultural guidance without requiring human expert availability for each consultation. Through specialized prompt engineering, the LLM maintains agronomic accuracy while adapting its communication style to novice users, effectively democratizing specialized knowledge that would otherwise require years of agricultural training or costly professional consultation.

The proposed approach advances beyond rule-based expert systems that require manual encoding of decision rules and cannot adapt to unforeseen combinations of constraints. It also differs from traditional recommender systems that rely on collaborative filtering or content-based algorithms operating on static crop databases. Instead, GROWIA employs the LLM's reasoning capabilities to generate adaptive, contextual knowledge for diverse urban farming scenarios, offering a methodology replicable across different metropolitan environments and cultural contexts in Latin America.

B. *Application Architecture Model*

The technological solution follows a three-tier architecture designed for scalability, maintainability, and real-time responsiveness, as shown in Fig. 1.

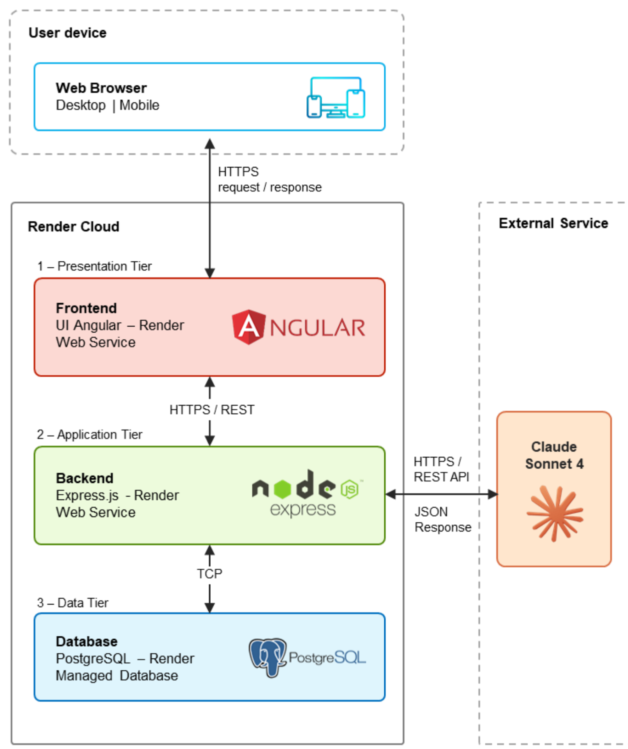

Fig. 1 Physical Architecture Diagram.

1) Presentation layer: Implemented using Angular 18, providing a responsive user interface optimized for both desktop and mobile devices. Users interact through a standard web browser via HTTPS protocol.

2) Application layer: Developed with Express.js, an API Gateway is deployed that manages RESTful endpoints for authentication, user administration, and processing of requests directed to the language model. This layer incorporates core components such as a prompt engineering engine, responsible for constructing structured queries for the Claude API that integrate the user's specific context.

3) Data layer: Implemented in PostgreSQL, it serves as the central repository for system information management. This layer stores user profiles, registered crops, and interaction history with the application.

Additionally, there is an external APIs layer where consumption of the Claude API stands out, providing advanced natural language processing services through Claude Sonnet 4. This integration follows an HTTPS communication protocol with data exchange in REST/JSON format, where the backend consumes artificial intelligence services as a specialized external component.

C. *User-Application Interaction Flow*

GROWIA implements a five-phase conversational workflow that transforms novice users into autonomous urban farmers through progressive AI-guided support, as illustrated in Fig. 2.

The authentication phase (Phase 1) loads user profiles from PostgreSQL, including interaction history and registered crops. The initial query phase (Phase 2) begins when users access the conversational assistant, which requests essential contextual information: (a) available space, (b) watering schedules, (c) gardening tools, (d) climate conditions, and (e) sun exposure. This structured data collection enables subsequent personalized recommendations.

In recommendation generation (Phase 3), the backend constructs specialized prompts integrating user context with system constraints (Lima-adapted species, agricultural calendar, urban viability). Claude Sonnet 4 processes requests and generates three recommended crops ordered by viability, each including species, suitability rationale, difficulty estimate, and harvest time.

The detailed plan generation phase (Phase 4) activates when users select a crop. The system requests exhaustive cultivation plans including: (1) substrate preparation, (2) planting instructions with measurements, (3) watering schedules, (4) growth timeline with milestones, and (5) nutritional benefits. Plans use conversational format with visual markers and accessible language for beginners.

Finally, in follow-up (Phase 5), users register crops in the "My Crops" module, functioning as a digital journal for observations and problems. The system incorporates these entries as context in future interactions, enabling specific advice on detected issues (pests, slow growth, wilting).

This five-phase flow ensures that GROWIA not only provides technically correct recommendations but also maintains continuous support adapted to the real evolution of the user's crops, differentiating itself from traditional static recommendation systems.

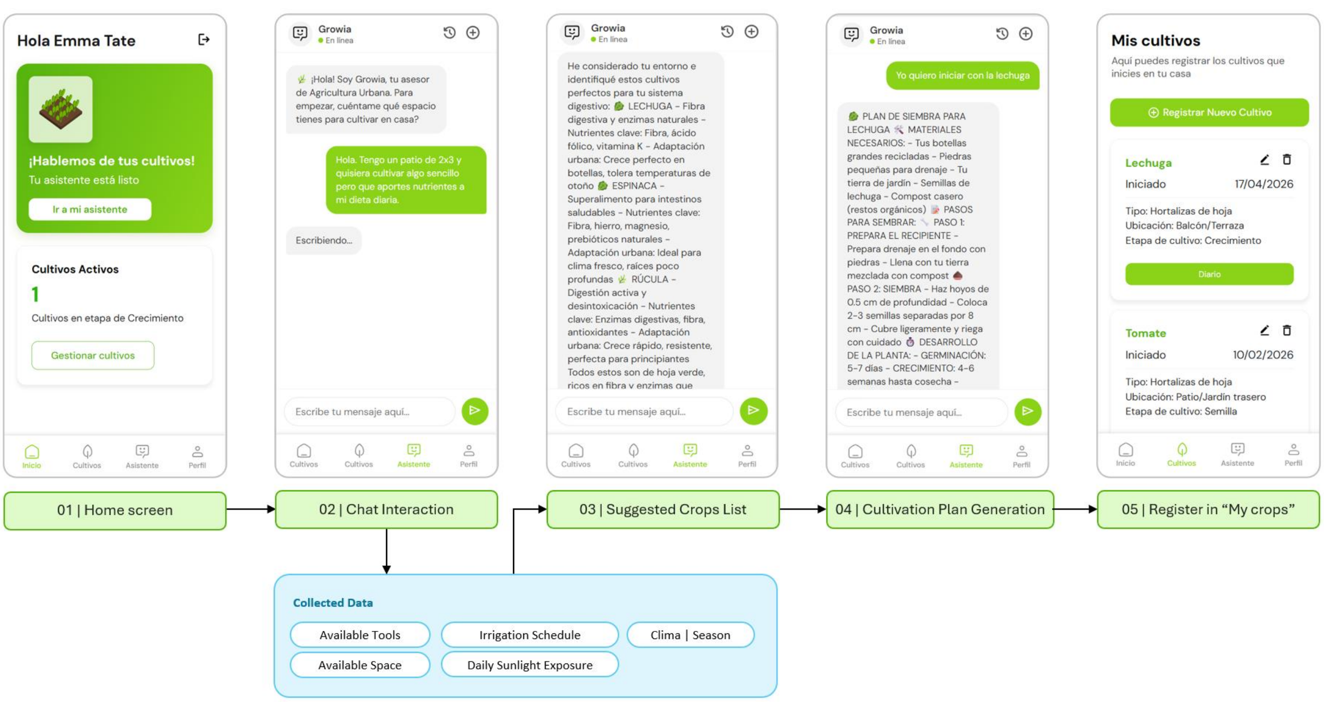

Fig. 2 Flow diagram of the interaction in GROWIA's main functionality.

D. Prompt Engineering Methodology

The core innovation of GROWIA lies in its specialized prompt engineering approach that transforms a general-purpose LLM into a domain-specific agricultural advisor. Unlike traditional chatbot implementations that rely on fine-tuning or retrieval-augmented generation (RAG), GROWIA employs a carefully crafted system instruction that constrains the model's behavior and knowledge domain to edible urban agriculture.

1) System Instruction Design: The prompt engineering strategy implements three critical constraints that ensure agronomic validity and user safety:

Domain Restriction: The system instruction explicitly limits recommendations to edible crops only, preventing the LLM from suggesting ornamental plants or potentially toxic species. This constraint is enforced through the directive: "You are a personalized Urban Agriculture advisor. Only recommend edible crops."

Contextual Data Requirements: Before generating recommendations, the system mandates collection of four essential parameters: (a) available space (specified in square meters or descriptive terms such as "medium balcony"), (b) irrigation schedules (morning, afternoon, or evening availability), (c) available gardening tools, and (d) local climate conditions. This structured data collection ensures that subsequent recommendations are contextually appropriate and feasible for the user's specific constraints.

Progressive Interaction Protocol: The prompt implements a multi-stage conversation flow that mimics expert agricultural consultation: (a) initial data gathering phase with clarifying questions, (b) confirmation phase where the system summarizes collected information and requests user validation ("Do you want to add more details?"), (c) recommendation generation trigger, and (d) detailed cultivation plan delivery once the user selects a specific crop.

2) Technical Implementation: The system instruction is injected into the Claude Sonnet 4 request as a system message, separate from the conversation history stored in the messages array. This architectural decision ensures that the behavioral constraints persist across all interactions while allowing the conversational context to evolve naturally. The implementation uses the Claude API with the following configuration:

Model: claude-sonnet-4-20250514

Max tokens: 2000 (allowing detailed cultivation plans)

Temperature: Default (0.5 for balanced creativity and consistency)

System message: Specialized agricultural advisor instruction

Messages: Dynamic conversation history with user context

3) Response Structure Optimization: The prompt engineering approach guides the LLM to generate cultivation plans with specific structural elements essential for novice success: (a) required tools specification, (b) step-by-step planting instructions, (c) watering schedules with frequency and optimal timing, (d) growth milestones with expected timeframes, and (e) conversational language using accessible terminology and encouraging tone.

This prompt engineering methodology differs from traditional approaches in three key aspects: First, it achieves domain specialization without requiring model fine-tuning or external knowledge bases, reducing computational costs and implementation complexity. Second, it maintains conversational flexibility while ensuring structured information collection, balancing user experience with technical requirements. Third, it leverages the LLM's inherent reasoning capabilities to adapt recommendations to diverse urban contexts, rather than relying on static rule-based systems that require manual updates for new scenarios.

IV. EXPERIMENTS

To validate the effectiveness of our web application based on Claude Sonnet 4, we implemented a dual evaluation study that combines technical validation by experts and usability evaluation with end users. This dual approach ensures both the agronomic accuracy of the AI-generated recommendations and the system's effectiveness in facilitating food access through urban agriculture for people without prior experience.

A. Expert Validation

In the validation of the system, four agricultural engineers and four coordinators of urban agriculture (UA) programs in Metropolitan Lima were contacted to assess the technical accuracy and feasibility of the system-generated recommendations.

1) Evaluation Protocol: The validation followed a structured three-phase protocol designed to simulate realistic usage scenarios while maintaining evaluation rigor:

Phase 1 - System Introduction: Experts received a standardized presentation explaining GROWIA's objectives, target users (urban residents without agricultural experience), and intended use cases (food self-sufficiency in space-constrained environments). This contextualization ensured experts evaluated recommendations against appropriate standards for novice urban farmers rather than commercial agricultural criteria.

Phase 2 - Interactive Simulation: Each expert independently interacted with GROWIA simulating diverse user profiles. Experts were instructed to vary key parameters across simulations: (a) space constraints, (b) sun exposure patterns (full sun, partial shade, limited light), (c) time availability, and (d) seasonal contexts (current season and hypothetical future seasons). This variation ensured comprehensive evaluation across the system's intended operating range. Each expert generated a minimum of 2 cultivation plans for different species, resulting in 19 total evaluated plans.

Phase 3 - Structured Evaluation: Experts completed a specialized survey instrument designed to assess five critical dimensions of recommendation quality through closed questions with 5-point Likert scales (1=completely inadequate, 5=excellent) and binary checklists for information completeness.

2) Evaluation Instrument: The validation survey (Table II) assessed five technical dimensions critical for agricultural recommendation systems. These questions were specifically designed to evaluate the effectiveness of GROWIA's recommendations and ensure the collection of accurate information about critical aspects of the system.

TABLE II
APPLICATION VALIDATION INSTRUMENT QUESTIONS

ID	Question	Evaluation Method
Q1	As an expert, evaluate whether the recommended crop was correct for the context you described to the assistant.	Likert Scale (1-5)
Q2	Is the recommended crop appropriate for planting in the current season in Lima?	Likert Scale (1-5)
Q3	Select only the elements that are present in the generated plan.	Multiple checklist
Q4	Considering the space you indicated to Growia, is the crop viable to reach its mature size in the available space without restrictions?	Likert Scale (1-5)
Q5	Imagine that a person without agricultural experience receives this plan. Could they implement it using only this information?	Likert Scale (1-5)

Item Q1 is formulated to collect baseline data for calculating the Correct Recommendations Rate (CRR), determining how many of the generated plans are agronomically accurate for the simulated cultivation context. Complementarily, question Q2 focuses on collecting the information necessary for the Temporal Accuracy (TA) metric, identifying whether the recommended crop is appropriate for the current season and agricultural calendar.

On the other hand, Q3 is used to validate the Completeness Index (CI), verifying through a checklist he presence of all essential elements in the generated plan. Question Q4 seeks to collect information for the Spatial Adequacy (SA) indicator, evaluating whether the crop is viable and without restrictions for the limited space that the expert or simulated user indicated to the system. Finally, item Q5 is designed to feed the calculation of the Self-Sufficiency Index (SSI), a crucial aspect, by determining whether the generated plan is sufficiently clear and complete to be implemented by a novice user without the need for external consultation.

3) Metrics and Statistical Analysis: We defined two primary metrics representing fundamental technical validity:

Correct Recommendations Rate (CRR): Measures the proportion of plans that meet minimum agronomic standards. A plan is considered agronomically correct if it obtains a score ≥ 4 on question Q1 (Agronomic Accuracy). This metric validates that the LLM generates technically sound and contextually appropriate content. The CRR is calculated as:

$$CRR = \left(\frac{N_{correct}}{N_{total}}\right) \times 100\% \qquad (1)$$

Where:

- $N_{correct}$: number of plans with score ≥ 4 on Q1
- N_{total} : total evaluated plans (*n* = 19)

The success threshold was defined as CRR ≥ 80%, representing the minimum acceptable accuracy for agricultural advisory systems. This threshold was established based on expert consensus that 1 in 5 incorrect recommendations could undermine user trust and potentially cause crop failure leading to discouragement.

Self-Sufficiency Index (SSI): Measures the proportion of plans that can be implemented by beginners without the need for additional expert knowledge. A plan is considered self-sufficient if it obtains a score ≥ 4 in question Q5 (Technical Self-Sufficiency). This metric validates that the recommendations are actionable by the target audience (people without UA experience). The SSI is calculated as:

$$SSI = \left(\frac{N_{self-sufficiency}}{N_{total}}\right) \times 100\% \qquad (2)$$

Where:

- $N_{self-sufficiency}$: number of plans with score ≥ 4 on Q5
- N_{total} : total evaluated plans (*n* = 19)

The success threshold was defined as $SSI \geq 70\%$ recognizing that complete self-sufficiency for beginners is a more demanding standard. For statistical analysis, a significance level of $\alpha = 0.05$ was established for all tests. The 95% confidence intervals were calculated using the Wilson method for binomial proportions, appropriate for small sample sizes.

Secondary metrics (Temporal Accuracy, Completeness Index, Spatial Adequacy) were calculated using the same methodology to provide comprehensive evaluation across all quality dimensions, though primary conclusions were based on CRR and SSI as the most critical indicators of system effectiveness.

B. *End User Validation*

In this study, the system's usability, its capacity to reduce perceived barriers, and the adoption intention among end users without agricultural experience were evaluated. The group consisted of 14 participants residing in Metropolitan Lima.

TABLE III
END USER EVALUATION INDICATORS

Identifier	Indicators
PBRI	Perceived Barriers Reduction Index
SUS	System Usability Scale
AAR	Application Adoption Rate
UAPI	Urban Agriculture Practice Intention

The evaluation protocol followed a pre-test, intervention and post-test design. During the pre-test (7 minutes), participants completed a form with the following sections: (1) demographic context, (2) Perceived Barriers for Urban Agriculture (PBUA), and (3) knowledge assessment with 10 multiple-choice questions on essential UA concepts.

The intervention phase was where participants independently interacted with the system completing the following tasks:

- User registration
- Interact with the AI assistant to obtain personalized recommendations
- Generate at least one cultivation plan
- Register the crop to initiate tracking

In the post-test phase (10 minutes), four metrics were defined that capture access facilitation, usability, and behavioral impact. The Perceived Barriers Reduction Index (PBRI) measures the pre-post percentage change in perceived barriers, directly evaluating whether the application facilitates access to UA. The System Usability Scale (SUS) is calculated according to Brooke's standard method, validating that the application is sufficiently usable. The UA Practice Intention (UAPI) measures the percentage of users with positive intention to practice UA, predicting actual adoption probability. The Application Adoption Rate (AAR) extracts the percentage of users who registered at least one crop, validating effective engagement with the main functionality.

V. Results

The findings obtained during the two system validation phases are presented. Results are organized following the experimental design structure, reporting first the expert validation findings, followed by end user results.

A. *Expert Validation Results*

Eight experts evaluated 19 cultivation plans covering diverse species: leafy crops (lettuce, spinach), fruits (tomato, pepper), roots (carrot, radish), and aromatic herbs (basil, rosemary). This diversity allowed evaluating the system's capacity to generate accurate recommendations across different crop categories adapted to the Lima urban context.

The Correct Recommendations Rate (CRR) reached 89.5% (95% CI: 78.2-91.1%), significantly exceeding the 80% target ($p = 0.032$). Of the 19 recommendations evaluated, 17 received scores of 4 or 5 in agronomic accuracy, while 2 obtained scores of 3. One case involved overly ambitious crop selection for limited space, while another recommended fertilization rates more suitable for experienced gardeners.

The Self-Sufficiency Index (SSI) reached 89.5% (95% CI: 70.8-86.1%), exceeding the 70% target ($p = 0.014$). This result indicates that approximately four out of five recommendations are sufficiently detailed for beginners to implement them without external consultation. Experts particularly highlighted: (1) clarity of step-by-step instructions, (ii) the inclusion of solutions to common problems faced by beginners, and (iii) accessible language.

TABLE IV
Performance Metrics and Expert Validation Results ($n = 8$ experts, 19 evaluated plans)

Metric	Target	Result	95 CI%	p-value
CRR	≥ 80%	89.5%	[78.2, 91.1]	0.032∗
TA	≥ 75%	89.5%	[75.4, 89.7]	0.028∗
CI	≥ 75%	88.4%	[87.1, 95.4]	0.001∗∗
SA	≥ 85%	94.7%	[80.1, 93.2]	0.041∗
SSI	≥ 70%	89.5%	[70.8, 86.1]	0.014∗

∗Statistically significant result. ∗∗Highly significant result.

Fig. 3 shows the performance of all five metrics against their respective baseline thresholds, demonstrating that GROWIA exceeds minimum acceptable standards across all dimensions with margins ranging from 9.5% (CRR) to 24.7% (SA), providing robust evidence of system effectiveness.

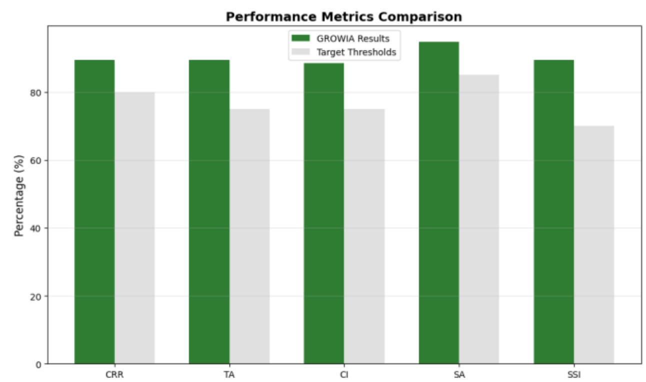

Fig. 3 Performance comparison of GROWIA metrics against baseline thresholds

The detailed expert validation results are presented in Table IV. Among the secondary metrics, Spatial Adequacy (SA) was the highest performing (94.7%), confirming effective adaptation to urban space constraints. Likewise, Temporal Accuracy (TA) (89.5%) validates adequate consideration of Lima’s climate seasons, while the Completeness Index (CI) reached 88.4%, averaging 4.6 of the 5 evaluated informative elements.

Beyond quantitative metrics, experts provided valuable qualitative observations. Agricultural engineers particularly appreciated the system's ability to adapt communication style to novice audiences while maintaining technical accuracy, noting that "the recommendations sound like a patient teacher rather than a technical manual." UA program coordinators highlighted the system's potential for scaling their community education efforts, with one coordinator stating: "This could provide basic guidance to 50-100 families simultaneously, allowing our limited staff to focus on complex troubleshooting and in-person support for families facing unusual challenges."

B. *End User Validation Results*

The 14 participants successfully completed the validation and interaction process with the web application. After post-evaluation, it was observed that all users who started the test managed to complete the assigned tasks.

TABLE V
END USER VALIDATION RESULTS ($n = 14$)

Metric	Target	Result	Status
PBRI	≥ −30%	−55.4%	Exceeded
SUS	≥ 68	85.2	Exceeded
UAPI	≥ 70%	85.7%	Exceeded
AAR	≥ 30%	71.4%	Exceeded

Following analysis of data collected through the web application and associated forms, the following findings were identified:

In the Perceived Barriers Reduction Index (PBRI), a 55.4% decrease in barriers to practicing urban agriculture was recorded, significantly exceeding the 30% objective. Detailed analysis indicates that the application demonstrated particular effectiveness in mitigating barriers associated with technical unfamiliarity and planning complexity.

The System Usability Scale (SUS) evaluation yielded a score of 85.2, placing the system in the 80th percentile of acceptance and classifying it within the range of usability excellence.

Regarding Urban Agriculture Practice Intention (UAPI), 85.7% of participants expressed a positive intention to perform this activity during the next three months.

Finally, the Application Adoption Rate (AAR) revealed that 71.4% of users generated and registered at least one cultivation plan during the sessions, evidencing the system's capacity to promote concrete action and facilitate initial adoption.

VI. Discussion and Implications

The dual validation results demonstrate three critical findings with broader implications for AI deployment in development contexts.

First, Technical viability of prompt engineering: The 89.5% CRR achieved without fine-tuning validates that domain specialization can be accomplished through prompt engineering alone. This methodological contribution reduces computational costs and infrastructure requirements, particularly relevant for Latin American contexts where GPU access and ML expertise are limited. Unlike systems requiring specialized knowledge bases [13] or fine-tuned models [9], GROWIA's approach enables rapid deployment across diverse urban environments without retraining.

Second, Barrier reduction mechanisms: The 55.4% decrease in perceived barriers primarily manifested in three domains: (a) uncertainty about appropriate crop selection reduced from 78.6% to 21.4% of users, (b) concerns about lack of guidance decreased from 85.7% to 28.6%, and (c) time commitment anxiety fell from 64.3% to 35.7%. This pattern suggests that personalized, conversational AI interfaces address psychological barriers more effectively than static information resources.

Third, Scalability for social impact: The 71.4% adoption rate during brief evaluation sessions suggests high conversion from interest to action. From a public health economics perspective, GROWIA's per-user cost (estimated $0.15-0.30 per consultation via API) is negligible compared to human agricultural extension services ($10-20 per consultation), enabling government or NGO deployment at metropolitan scale. This cost structure validates conversational AI as a viable intervention for urban food security in resource-constrained settings.

These findings suggest that GROWIA represents not merely a technical solution but a replicable model for leveraging LLMs to democratize specialized knowledge in domains critical to human wellbeing—with applications beyond agriculture to areas such as primary healthcare guidance, legal aid, or financial literacy.

VII. Conclusion

In this study, we developed GROWIA, an AI-powered web application for urban agriculture advisory. Implemented with Claude Sonnet 4 and a three-tier architecture that enables obtaining user context to provide more suitable crops, GROWIA was validated with 8 agronomists who evaluated 19 plans and 14 end users in Lima, Peru. Key metrics such as Correct Recommendations Rate (CRR), System Usability Scale (SUS), and Perceived Barriers Reduction Index (PBRI) showed 89.5% agronomic accuracy, an excellent SUS score of 85.2, and a 55.4% reduction in perceived barriers, highlighting its potential to facilitate food self-sufficiency in urban environments.

What distinguishes this application is its focus on democratizing agricultural knowledge through conversational AI, significantly reducing barriers for users without prior experience. Future research will focus on strengthening external validity through larger and stratified samples across urban contexts, as well as conducting longitudinal evaluations over complete crop cycles to assess real-world agricultural outcomes such as yield and crop survival. Additionally, the integration of computer vision and IoT-based environmental sensing is expected to enable adaptive, real-time recommendations, evolving GROWIA from a planning tool into a continuous decision-support system, while further deployments in other Latin American cities will allow evaluating the scalability and transferability of the approach.

References

[1] FAO, IFAD, UNICEF, WFP, and WHO. *The State of Food Security and Nutrition in the World 2024 – Financing to End Hunger, Food Insecurity and Malnutrition in All Its Forms*. Rome, 2024. https://doi.org/10.4060/cd1254en

[2] FAO, *"Global Report on Food Crises: Number of people facing acute food insecurity rose to 258 million in 58 countries in 2022,"* FAO News and Media, May. 03, 2023. https://www.fao.org/newsroom/detail/global-report-on-food-crises-GRFC-2023-GNAFC-fao-wfp-unicef-ifpri

[3] J. Huang, M. Wang, Y. Cui, J. Liu, L. Chen, T. Wang, H. Li, and J. Wu, "Layered Query Retrieval: An Adaptive Framework for Retrieval-Augmented Generation in Complex Question Answering for Large Language Models," *Applied Sciences*, vol. 14, no. 23, p. 11014, 2024. https://doi.org/10.3390/app142311014

[4] K. Chatrabhuj, K. Meshram, U. Mishra, and U. Rathnayake, "Application of Artificial Intelligence in Agri-Tech, Environmental and Biodiversity Conservation," *Array*, vol. 26, p. 100412, 2025. https://doi.org/10.1016/j.array.2025.100412

[5] A. H. Essenfelder, A. Toreti, and L. Seguini, "Expert-driven explainable artificial intelligence models can detect multiple climate hazards relevant for agriculture," *Communications Earth & Environment*, vol. 6, no. 1, 2025. https://doi.org/10.1038/s43247-024-01987-3

[6] R. Raman and S. Jayaraman, "Artificial intelligence for sustainable farming with dual branch convolutional graph attention networks in rice leaf disease detection," *Scientific Reports*, vol. 15, no. 1, 2025. https://doi.org/10.1038/s41598-025-94891-5

[7] R. Naagarajan and S. Streif, "Enhancing Greenhouse Management with Interpretable AI: A Natural Language Interface for Advanced and Optimization-based Control," *Smart Agricultural Technology*, vol. 10, p. 101041, 2025. https://doi.org/10.1016/j.atech.2025.101041

[8] N. C. Gaitán, B. I. Batinas, C. Ursu, and F. N. Crainiciuc, "Integrating Artificial Intelligence into an Automated Irrigation System," *Sensors*, vol. 25, no. 4, p. 1199, 2025. https://doi.org/10.3390/s25041199

[9] Y. Zhang, Q. Fan, X. Chen, M. Li, Z. Zhao, F. Li, and L. Guo, "IPM-AgriGPT: A Large Language Model for Pest and Disease Management with a G-EA Framework and Agricultural Contextual Reasoning," *Mathematics*, vol. 13, no. 4, p. 566, 2025. https://doi.org/10.3390/math13040566

[10] M. Y. Shams, S. A. Gamel, and F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: A study on agricultural decision-making," *Neural Computing and Applications*, vol. 36, no. 11, pp. 5695–5714, 2024. https://doi.org/10.1007/s00521-023-09391-2

[11] M. A. Rahman, N. R. Chakraborty, A. Sufiun, S. K. Banshal, and F. R. Tajnin, "An AIoT-based hydroponic system for crop recommendation and nutrient parameter monitorization," *Smart Agricultural Technology*, vol. 8, p. 100472, 2024. https://doi.org/10.1016/j.atech.2024.100472

[12] A. A. AlZubi and K. Galyna, "Artificial Intelligence and Internet of Things for Sustainable Farming and Smart Agriculture," *IEEE Access*, vol. 11, pp. 78686–78692, 2023. https://doi.org/10.1109/ACCESS.2023.3298215

[13] N. Chandolikar, C. Dale, T. Koli, M. Singh, and T. Narkhede, "Agriculture assistant chatbot using artificial neural network," in *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, Coimbatore, India, 2022, pp. 1–5. https://doi.org/10.1109/ICACTA54488.2022.9753433