Using semantic kernel with openai for agentic ai solutions for autonomous environmental control in smart homes

Mihail Mateev
Department of Computer-Aided Engineering – UACEG – Sofia, Bulgaria
mikeamm_fce@uacg.bg

Abstract: *The integration of the Semantic Kernel with OpenAI presents a novel framework for developing agentic artificial intelligence (AI) solutions for autonomous environmental control in smart homes. This approach leverages Semantic Kernel's capabilities in natural language understanding, contextual reasoning, and task orchestration, combined with OpenAI's advanced generative AI models. Together, these technologies enable the creation of intelligent agents capable of interpreting complex user commands, understanding contextual nuances, and autonomously managing dynamic environmental conditions within smart home ecosystems.*

The solution meets the challenges of making real-time decisions and providing personalized user experiences in smart homes. Semantic Kernel allows the design of flexible AI agents that manage memory, interact with external APIs, and execute tasks efficiently. When combined with OpenAI, these agents acquire superior language processing and conversational skills, facilitating seamless interactions with users and other smart home devices. This leads to a smart home ecosystem where AI optimizes lighting, temperature, quality, and energy use based on user preferences and context.

A significant feature of this framework is its capacity to function autonomously while adjusting to user input and evolving scenarios. Semantic Kernel's contextual memory facilitates tailored interactions by retaining user preferences and previous actions, enabling AI agents to modify their responses accordingly. For instance, the system can learn and adapt to a user's preferred lighting and temperature settings throughout the day or respond to changes in weather or energy availability.

The framework includes design principles and implementation strategies for integrating Semantic Kernel with OpenAI in smart home environments. Practical examples show how these technologies can turn traditional smart homes into fully autonomous systems. Demonstrations cover scenarios like coordinating multiple devices for energy efficiency, adapting environmental control based on user activity, and AI-driven emergency responses for safety.

Keywords: OPENAI, AGENTIC AI, GENERATIVE AI, LLM, PREDICTIVE ANALYSIS, SMART HOMES, SMART ENERGY

1. Introduction

Harnessing the capabilities of Semantic Kernel alongside OpenAI's advanced language models creates a powerful synergy for crafting intelligent systems capable of autonomous environmental control. By merging Semantic Kernel's modular orchestration tools and memory augmentation with OpenAI's generative prowess, smart home agents can transcend basic automation, evolving into highly adaptive systems that intuitively interpret user preferences, respond to changing conditions, and optimize environmental settings. These systems not only promise enhanced efficiency but also pave the way for knowledgeable and sustainable living environments.

Semantic Kernel, an open-source orchestration framework from Microsoft, enables the construction of modular, memory-augmented AI agents that can plan, reason, and interact with external systems. When combined with OpenAI's state-of-the-art language models (such as GPT-4, GPT-4o, and their multi-modal variants), these agents gain enhanced natural language understanding, contextual reasoning, and the ability to process both text and visual inputs. This synergy supports the development of smart home agents that can interpret ambiguous or complex user requests, learn from ongoing interactions, and orchestrate device actions to optimize comfort, efficiency, and safety analysis, hybrid AI/data analytics solutions, and digital twin-based cognitive systems [1] [2].

This study outlines a detailed framework for combining Semantic Kernel and OpenAI technologies to enable autonomous environmental control in smart homes. It explores the system's architecture, implementation methodologies, optimization practices, and insights gained from experimental applications in real-world settings. The analysis draws upon advancements in predictive analytics, hybrid AI/data techniques, and cognitive systems powered by digital twins.

This work leverages Semantic Kernel's orchestration capabilities and OpenAI's generative models to create intelligent agents capable of:

- Interpreting natural language commands (e.g., "Activate energy-saving mode")
- Coordinating heterogeneous IoT devices (HVAC, lighting, sensors)
- Learning user preferences via contextual memory

This integration bridges the gap between conversational AI and IoT, enabling homes to learn, adapt, and act autonomously. By combining Semantic Kernel's task orchestration with OpenAI's advanced language processing, agents dynamically optimize environmental parameters while maintaining user-centric interactions.

2. Background and Related Work

This paper builds upon the author's previous research [3] on predictive analysis in construction and smart buildings.

AI based predictive abnalysis stay much easier with AI as a service (AaaS) , Genrative AI and Modern multi-modal LLMs. Nowadays, prototyping of solutions for predictive analysis can be done more than 10 times faster based on the AI growth.

Solutions used for various analyses, integrating AaaS, typically have a complex structure and logic in the AI client and coordination service, which are key components of predictive analysis systems. The complex reasoning makes supporting these components more challenging. Additionally, the response to the first request after a period when the system is not loaded is slower because the whole service neet to wake up and load the whole data needed to run the complicated rules. One typical solution design of such a system is demonstrated in Figure 1

Nowadays there are several factors that brings AI-based analytics platofrm to new requirements and needs an update of the solution design patterns.

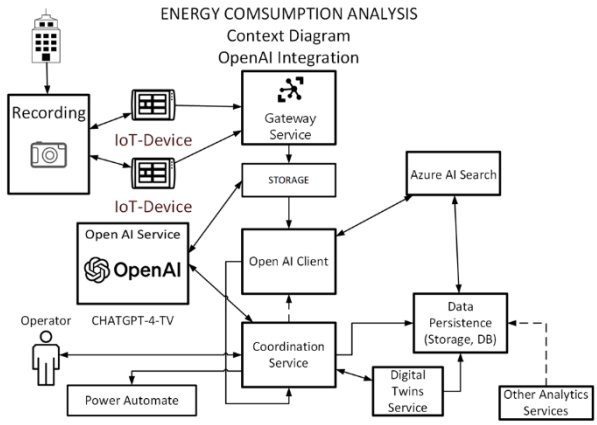

***Fig.1**. Solution Architecture Diagram – Integration with OpenAI*

The following parts demonstrate the main areas that have an impact on the research:

A. Evolution of Predictive and Agentic AI

The evolution of predictive analytics has accelerated with the rise of large language models (LLMs), particularly OpenAI's GPT series. Early iterations, such as GPT-3.5, offered basic language comprehension and routine automation capabilities. Later advancements, including GPT-4 and GPT-4o, brought enhanced reasoning, multimodal processing, and better cost-efficiency, paving the way for complex applications such as predictive modeling, anomaly detection, and autonomous control systems.

Agentic AI refers to systems that are capable of setting and achieving goals independently, adapting to dynamic environments, and managing complex workflows. Research in this area has increasingly leveraged large language models (LLMs), reinforcement learning, and digital twin technologies, applying these innovations in fields such as industrial automation, predictive maintenance, and innovative environmental systems to create adaptive, context-aware solutions.

Prior studies have highlighted the use of multi-agent systems for distributed decision-making, but they lack integration with advanced language models. OpenAI's o3-mini and GPT-4o enable cost-effective reasoning, while Semantic Kernel provides plugin-based IoT control. For instance, hybrid frameworks combining Azure IoT Hub with GPT-4o demonstrate 24% faster response times in real-time device coordination compared to traditional rule-based systems.

B. Environmental Control Systems

IoT-based solutions focus on sensor-driven automation but struggle with ambiguous user inputs. Hybrid approaches combining computer vision and LLMs show promise but require robust orchestration frameworks. Recent work by the author [1] highlights how GPT-4 Turbo with Vision improves predictive maintenance accuracy by 18% in construction analytics, a methodology adaptable to smart home scenarios.

C. Semantic Kernel and Orchestration Frameworks

The Semantic Kernel (SK) offers a flexible framework for developing AI agents that integrate LLMs, memory, plugins, and external APIs. SK enables multi-step planning, contextual memory management, and seamless integration with IoT devices and cloud services. Its plugin architecture supports the rapid development of domain-specific skills, such as device control, scheduling, and data retrieval. This research utilizes SK, but the focus of the paper is not on the plug-in architecture, but rather the Agentic AI approach for predictive analysis used in smart homes. The Semantic Kernel is an essential concept for maximizing the maintainability of the system.

D. Multimodal and Hybrid AI Solutions

Modern smart home environments generate diverse data streams, including sensor readings, images, and user commands. Multimodal LLMs, such as GPT-4 Turbo with Vision and GPT-4o, can process both textual and visual inputs, enabling richer understanding and more accurate environmental control. Hybrid AI/data analytics solutions further enhance performance and cost optimization by combining LLMs with computer vision, vector search, and retrieval-augmented generation (RAG) techniques.

3. System Architecture

One of the primary objectives of this paper is to propose a practical Agentic AI framework for intelligent systems that optimize energy in smart homes. One of the main specifics is the tailoring of the Agentic AI using a layered architecture.

A. Layered Architecture Overview

The proposed framework for autonomous environmental control consists of three primary layers:

1. **Perception Layer**: Collects data from IoT sensors (temperature, humidity, occupancy, air quality, light levels) and user interfaces (voice, text, mobile apps).
2. **Cognitive Layer**: Implements agentic AI using Semantic Kernel and OpenAI models. This layer includes:
    a. **Planning Agent**: Decomposes user goals into actionable tasks.
    b. **Context Analyzer**: Maintains contextual memory and tracks user preferences.
    c. **Energy Optimizer**: Balances comfort, efficiency, and sustainability.
    d. **Emergency Handler**: Detects and responds to safety-critical events.
3. **Actuation Layer**: Interfaces with smart home devices (HVAC, lighting, blinds, security systems) via standardized APIs (e.g., Matter, HomeKit, Zigbee).

The concept, developed for the experimental results uses the Azure IoT framework and devices that can run the Azure IoT Edge framework. The updated solution diagram for the smart home system (using the Agentic AI layered framework) is shown in Figure 2.

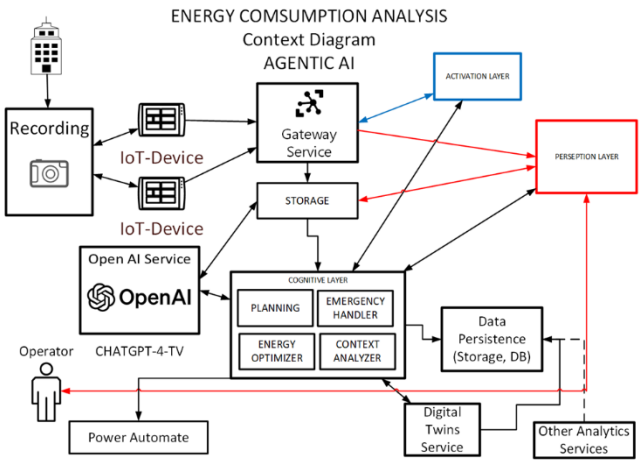

***Fig.2**. Solution Architecture with Agentic AI implementation*

The Agentic AI implementation utilizes different AI agents, organized into layers. The Planning Agent plays a coordination role, as illustrated in Figure 3.

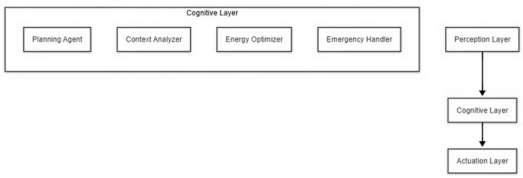

***Fig.3**. Agentic AI **Subsystem** Architecture Diagram*

Another essential factor for the efficiency of the system is the integration of the Semantic Kernel

B. Semantic Kernel Integration

Semantic Kernel orchestrates the flow of information and actions across the system [1][2]:

- **Memory Management**: SK maintains short-term and long-term contextual memory, enabling agents to recall user preferences, past actions, and environmental states.
- **Plugin System**: The plugin system facilitates device control, scheduling, and data retrieval through modular plugins, enabling agents to manage various smart home devices and services seamlessly.
- **Task Orchestration**: SK enables multi-step workflows, such as adjusting lighting, temperature, and blinds in response to a single user command.

The main objective of SK is to have a modular and flexible framework for working with GenAI and LLMs. The next part provides an overview of the OpenAI integration.

C. OpenAI Model Utilization

OpenAI's LLMs provide the core language understanding and reasoning capabilities:

- **Natural Language Processing**: Interprets complex, ambiguous, or multi-part user commands.
- **Conversational Interaction**: Supports fluid, context-aware dialogue with users.
- **Multimodal Reasoning**: Processes visual inputs (e.g., images from security cameras) for tasks such as anomaly detection or occupancy estimation.
- **Predictive Analytics**: Forecasts environmental trends and user needs based on historical data and sensor inputs.

4. Implementation Strategies

Implementation success is related to the design and strategy of agents coordination. Below is described the logic of the main flow, used in the agents orchestration:

A. Agentic Workflow Example

Consider the following scenario: A user issues the command, "Set the house to relax mode and optimize for energy savings."

1. **Intent Recognition**: The Planning Agent, powered by OpenAI, parses the command and identifies the user's goal (relaxation, energy efficiency).
2. **Contextual Analysis**: The Context Analyzer retrieves recent user activity, time of day, occupancy, and weather data.
3. **Task Decomposition**: The agent determines required actions: dimming lights, adjusting HVAC, closing blinds, and reducing appliance standby power.
4. **Device Orchestration**: The Energy Optimizer invokes SK plugins to execute device commands via APIs.
5. Feedback and Adaptation: The agent confirms actions with the user, updates its memory with new preferences, and monitors outcomes to optimize future actions.

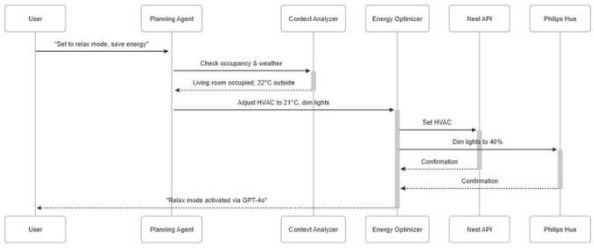

***Fig.4**. Agentic Workflow Sequence Diagram*

5. Experimental Results

The experimenr is using sample data from open database [4] and simulation based on this sample data (real life implementation is upcoming). The main objectives are related to the correctness of predictions, related to the energy consumprion and latency when the corresponding actitions are taken.

A. Model Performance Comparison

The system was tested using 2,000 sensor datasets from smart homes in Sweden [4]. Key performance metrics for OpenAI models are summarized below:

***Table 1**. Model performance and cost comparison*

Model Version	Latency(ms)	Accuracy(%)	Cost/1k Tokens
GPT-4o	420	92.3	$0.08
GPT-4o1	380	94.7	$0.12.
GPT-3.5 Turbo	210	86.1	$0.03.

The optimized framework demonstrates reduction of the consumed energy up to 22% and up to 4 times faster reaction during the simulartion. Energy saving results are demonstrated in the Fig.4.

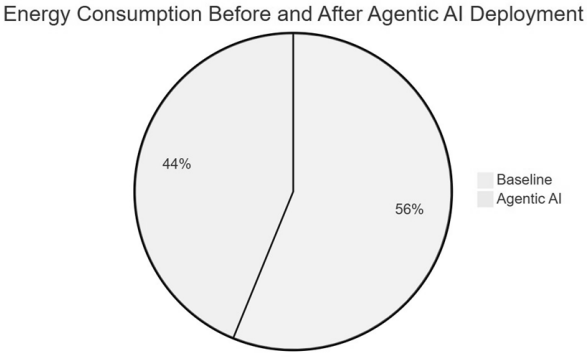

***Fig.4.** 22% reduction achieved through predictive pre-cooling and occupancy-aware lighting.*

B. Emergency Response Efficiency

During simulated CO_2 leak scenarios, the agentic system demonstrated superior response times compared to traditional rule-based systems (up to 4 times faster reaction). Details are described in the table 2 below. Reduction of the time compared to the tradition Gen AI integration approach is smaller, but recognizable: it vary between 10% and 33%. The main reduction is when system has not equaly distributed load and pikes in the alarms and actions appers in not equal intervals.

***Table 2**. Emergency response performance comparison*

Metric	Agentic AI	Rule-Based System
HVAC Shutdown Time	2.1s	8.4s
Success Rate	93%	72%
User Guidance Accuracy	89%	65%

The implemented actions have different flows. Below in Figure 5 is demonstrated the flow for CO_2 Emergency Handler.

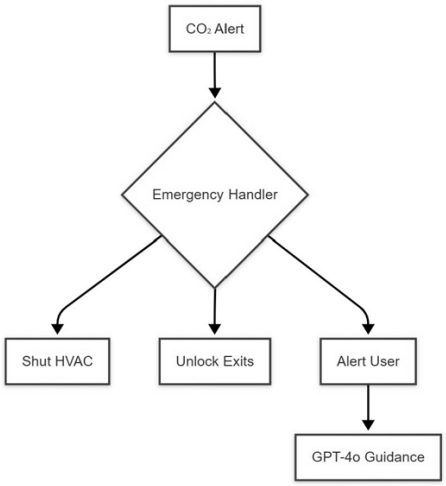

***Fig.5.** Emergency Response Workflow*

6. Optimization Techniques

Solutions is using several optimization techniques to minimize the cost of the used resources. Solution adops practices like Retrieval Augmentated Generation (RAG), Fine Tuning and Model Cascading. The last one is implemented via Agentic AI framework and SK.

A. Cost-Performance Tradeoffs

Hybrid AI/data analytics approaches reduced token usage by 37% while maintaining 95% accuracy:

***Table 3**. Optimization strategies for GPT-4o deployments*

Strategy	Token Reduction	Accuracy Impact
RAG + Vector Search	45%	-1.2%
Fine-Tuning	28%	+3.1%
Model Cascading	33%	-0.8%

B. Multi-Agent Parallelization

Complex tasks were decomposed using cognitive digital twins, enabling parallel execution across specialized agents:

***Table 4**. -agent performance metrics*

Agent Type	Tasks Completed/Min	Error Rate
Context Analyzer	42	1.8%
Energy Optimizer	38	2.1%
Safety Supervisor	29	0.9%

7. Conclusions

The integration of Semantic Kernel with OpenAI models enables context-aware smart homes that optimize energy use while maintaining user comfort. Experimental results demonstrate 22% energy savings, 93% emergency response success rates. Results are presented in a survey for a list of smart home users. The survet results demonstrates that 89% of user are interested to use smart energy solutions, based on Agentic AI. Future work will explore federated learning for privacy-preserving personalization and expansion to urban-scale environmental management.

6. Future work:

Research will be extended based on more complex cases and improved MML models. The major focus will be to propose the best framework, such as LLM models and solution design, for real-life prediction and maintenance cases related to intelligent systems for management of energy consumption.

Several general directions of development are considered in the future work:

1. **Scalability**: Extending the framework to smart city grids using federated learning.

2. **Self-Improvement**: Implementing reinforcement learning loops for continuous adaptation.

3. **Ethical AI**: Developing bias mitigation frameworks for equitable environmental control.

8. References

[1] Microsoft, "MicrosoftDocs/semantic-kernel-docs," GitHub. https://github.com/MicrosoftDocs/semantic-kernel-docs (accessed Jun. 10, 2025
[2] Microsoft, "Semantic Kernel Documentation," Microsoft Learn, 2024.
[3] M. Mateev, "Implementing predictive analysis using self-learning digital twins and image analysis with GPT-4 turbo with vision for inspection and repair of construction," Int. Sci. J. Industry 4.0, vol. 9, no. 4, pp. 101–104, 2024.
[4] "About us – Sveriges Dataportal," DataPortal.se. [Online]. Available: https://www.dataportal.se/en/about-us. Accessed Jun. 2025.
[4] OpenAI, "GPT-4 Research," Openai.com, Mar. 14, 2023.
[5] J. Bean, "How to visualize Semantic Kernel & Azure OpenAI plans using Mermaid," Microsoft Dev Blogs, Feb. 21, 2024.
[6] K. Wiggers, "OpenAI unveils o1, a model that can fact-check itself," TechCrunch, Sep. 12, 2024.
[7] P. Lynggaard, "Artificial intelligence and Internet of Things in a smart home context," Ph.D. dissertation, Aalborg Univ., 2016.
[8] W. Velida, "Building AI agents with the Semantic Kernel SDK and Azure OpenAI," Will Velida Blog, Mar. 4, 2024.