AIROBOT: CONVERSATIONAL-BASED USER INTERFACE DRIVING INDUSTRY & AGRICULTURE

Roxana IRIMIA[1], Sînică ALBOAIE[1], Dan DONOSĂ[2], Monica RUSU[2]

e-mail: cocamonica.pfa@gmail.com

Abstract

In this article, we present the results of our research project, AIRobot, an innovative approach to integrating Artificial Intelligence (AI) and Natural Language Processing (NLP) across various sectors, including agriculture, industry, and services. The primary objective of our study is to streamline human-software interactions using natural language, thereby enhancing productivity and simplifying operations. In this article, we introduced a robust architecture and the essential components for implementing AIRobot, serving as a solid foundation for future applications in the agriculture, industry, and services sectors.

Keywords: NLP, conversational AI, personal assistant

The overarching vision of AIRobot is to create personal assistants (personal robots) with the expectation that conversational interfaces will be applicable in nearly every application, serving as an alternative to Command Line Interfaces (CLI) or Graphical User Interfaces (GUI). Essentially, there is a well-founded hope that user interfaces will become more conversational, human-like, and closer to the natural way people interact with one another.

Personal assistants (Sun C.C., 2021) are structured as a novel proposition put forth by AIRobot. They are organised within collaborative spaces, allowing for text, audio, and video interactions between the robot assistants and users.

Furthermore, the robot can mediate interactions among users, who can work independently but are always assisted by the robot.

Naturally, we envision it capable of enhancing these interactions to the highest level possible. Another crucial aspect is incorporating security by design (Jo H., 2022). Our architectural decisions prioritise privacy, ensuring that interactions remain restricted to limited groups, untraceable by providers or even by us. This is made possible through encryption and concepts such as Digital Wallets or Decentralized Identities (DIDs) (Dib O., Toumi K., 2020).

We rely on an architecture that supports a professional Key Management System (KMS) based on Hardware Security Modules (HSMs) and Trusted Execution Environments (TEEs). The vision entails using the same code and interaction style on both mobile and desktop platforms, similar to existing communication applications that allow mobile and desktop uploads with minimal differences. While the desktop version utilises a larger screen, the interaction remains consistent across both platforms.

Building upon this vision, AIRobot aims to deliver a seamless, user-friendly experience that fosters efficient communication and collaboration. By leveraging advancements in AI

[1] Axiologic SaaS, Iași, Romania

[2] Iasi University of Life Sciences, Romania

and NLP, personal assistants can understand and respond to user queries, offering accurate and contextually relevant information. As a result, users can focus on their core tasks, increasing productivity and reducing the learning curve associated with traditional interfaces.

In addition to focusing on security and privacy, the AIRobot architecture emphasises scalability and adaptability. This allows for integrating new features, applications, and technologies as they emerge, ensuring that the personal assistants remain up-to-date and continue to meet users' evolving needs. The modular design also promotes customisation, enabling users and organisations to tailor the personal assistants according to their specific requirements.

MATERIAL AND METHOD

AI assistants developed using the paradigm of conversational interfaces will be able to incorporate elements of communication, audio, video, emotion detection, and other functionalities. Of course, during the platform's development, we realised that numerous applications could be built on this foundation. We will analyse, test, research, and improve the platform to support other types of applications beyond the three we described in the project. The companies within our group are collaborating with other interested partners to utilise the AIRobot platform for various application types. The pluggable components of the AIRobot Core are outlined in the figure below.

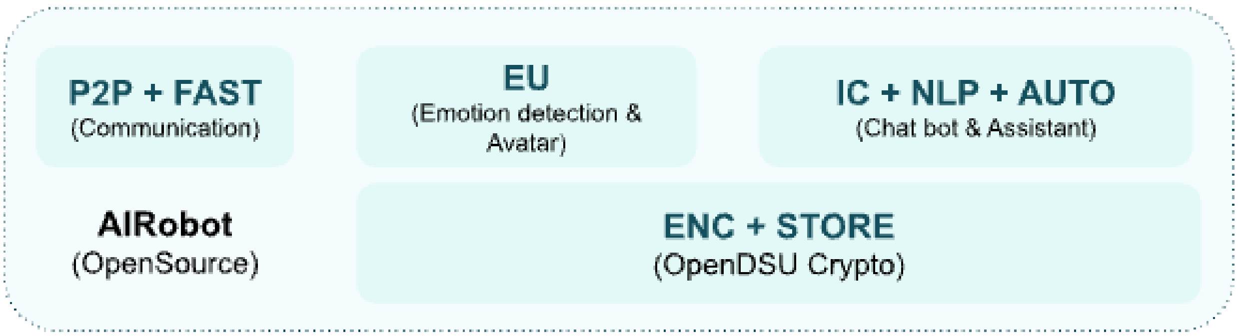

Figure 1 **Pluggable AIRobot Core Components**

As the initial project description mentioned, we envisioned the platform containing a series of plugins or "pluggable" components that provide generic functionalities across all applications we develop. These plugins are grouped according to their roles. For audio, video, and text communication, we have the P2P (Peer-to-Peer) plugin that ensures direct and efficient communication between nodes. Additionally, we have the FAST plugin, which (using Artificial Intelligence techniques and other methods related to emotion detection and significant communication event detection) can enable efficient collaboration without transmitting the entire audio-video stream, using sound and video image reconstruction techniques.

Regarding the EU part, the abbreviation comes from the phrase "from Emoție Umană in the Romanian language but also “Self”). We aim to replace some audio-video communication, which can be costly, slow, and may not work well on weaker networks. However, to maintain human-like communication, transmitting emotions during interactions (Murthy, A.R., Anil Kumar, K.M., 2021), we have this plugin/module that allows emotion detection and displaying an avatar to replace the audio-video stream, making communication more interactive, engaging, and emotionally interaction rich.

We have three plugins designed to work together on natural language processing and conversational interfaces. We have the IC (Conversational Interface) plugin, which we call CI (Conversational Interface), the NLP plugin, and the AUTO plugin. The latter combines the two plugins and enables the learning and programming of rules and intelligent reaction methods for the assistant (for this robot) or ChatBot (a commonly used term).

For security and storage, we have these two plugins, ENC and STORE, which will use OpenDSU technologies (Ursache et al, 2022) to perform encryption in data transit and for stored data encryption.

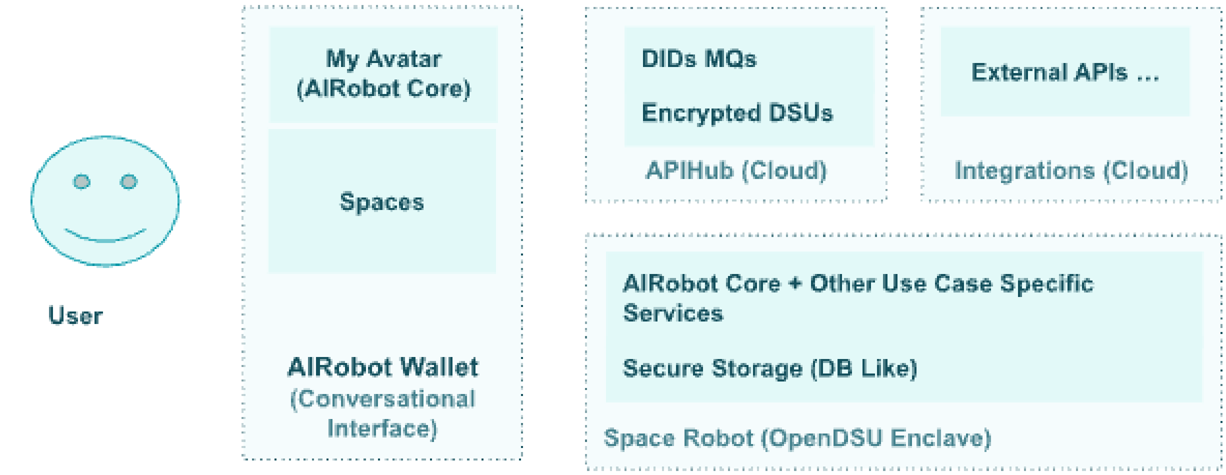

Figure 2 **AIRobot Platform Architecture**

Figure 2 above provides a detailed view of the Conversational Interface (CI), Natural Language Processing (NLP), and Automation (AUTO) components. Simultaneously, the architectural components are presented below.

The user interacts with a Digital Wallet through conversational interface methods. This Wallet consistently features an Avatar, which instantiates the AIRobot Core. Consequently, all listed plugins have implementations related to NLP and the automation of reactions. Each application will have its assistant capable of storing interactions, knowing rules, and processing natural language.

On the other hand, if these assistants or robots represent companies or individuals requiring assistance, remote instantiation occurs in the Cloud, and we deal with a different type of robot. We playfully refer to it as the Space Robot, as interactions occur within certain collaborative spaces for work, organisation, and various forms of collaboration. Therefore, the Space Robot can be assimilated and highlighted as the personal assistant of a collaborative space.

We have these applications, as well as Cloud instances, developed through a technology called OpenDSU Enclaves. This technology enables secure and plugin-enabled storage in different environments, such as HSMs or databases. Of course, this Space Robot will also instantiate an AIRobot Core and can be programmed to react to events, user input, and external events highlighted by these integrations.

Consequently, personal assistants or robots can interact with external systems through integrations specific to each use case or application. It should be noted that no general solution applies to every case or application.

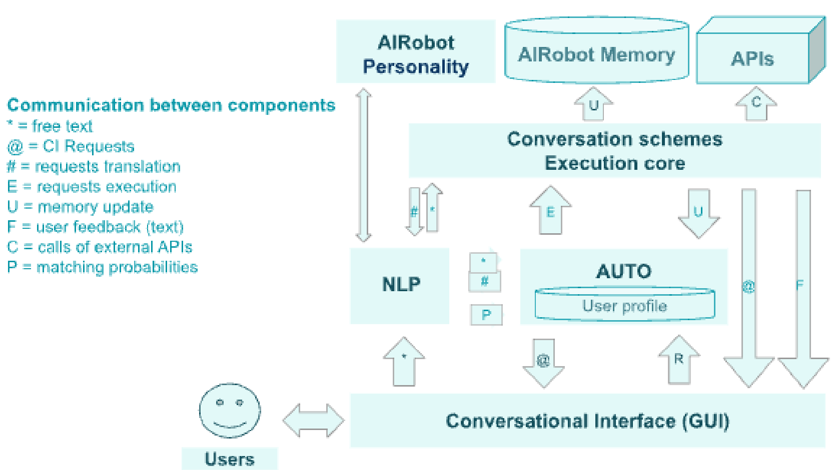

Figure 3 **AIRobot Core Architecture**

As mentioned earlier, AIRobot Core is an abstraction that can be instantiated in Wallets and what we call Space Robot. In the following, we will detail how this interaction is currently planned within an AIRobot. The figure above shows that the reading is done from bottom to top, starting with the Conversational Interface (CI), a GUI (Graphical User Interface), somewhat imaginable, similar to popular chatting applications. From the GUI, inputs are made to the NLP plugin, which starts to make sense and interacts with the AUTO plugin. The latter interacts with what we call an Execution Engine of a Conversational Schema (CS), updates the memory or storage of the robot that learns about the user, the outside world, can interact through APIs (Application Programming Interfaces) with external integrations, etc. The AIRobot (Agent) personality is a collection of CS configured with a specific NLP and AUTO plugin. When we refer to CIRequests (Conversational Interface Requests), the code of a CS can request input from the CI. The types of inputs would be as follows: choose an option from a menu (options), choose an option from a slideshow (cards that can be selected), request values as input data (date, duration, object name, person name, dimensions, area, geographic location).

Each CS is a JSON file containing the following fields: triggers, parameters, validation, and code. Triggers are a series of natural languages that can trigger the Conversational Schema. Parameters are a series of values (types) required to execute the code. The AUTO plugin could link/combine those values from its context or trigger CIRequests to the CI to fill in those values. Validation allows the parameters' values to be checked before executing the code and requests incorrect parameters again (generate CIRequests to CI).

Regarding the JS code, it is executed in the agent's current environment. This code could trigger calls to external APIs and update the agent's current state (AIRobotMemory) or the AUTO plugin's state correlated with the current user. AIRobotMemory is a secure storage space (OpenDSU enclave providing APIs for tables and key values) that the Agent uses as a database.

RESULTS AND DISCUSSIONS

Our natural language processing system is a complex and dynamic process that relies on the interdependent modules of Tokenization, Entity Matching, and Personality modules, which are orchestrated by a central component. The first module, tokenisation (Rai A., Borah S., 2020), is the initial step that extracts essential nouns and verbs from the user's input. This is done by analysing the sentence structure and syntax to identify the most relevant information. The extracted information is then used to determine the targeted entities and actions that must be executed.

The entity matching (Yu J. et al, 2020) module is responsible for correlating the extracted entities and actions with predefined triggers for user flows registered in the application. This is done by matching the extracted information with a predefined set of rules and triggers, which are defined based on the application's intended functionality. The entity matching module is critical to the NLP system since it ensures it executes the correct actions based on user input.

The personality module serves two critical functions in the NLP system. Firstly, it stores and manages the conversational schemas for each application, which provide a consistent user experience. Secondly, it stores user-related information and patterns, allowing the AIRobot to understand the user's behaviour better as they interact with the application. This understanding is critical to improving the user's experience since it allows the system to tailor responses and actions to their specific needs and preferences.

Our NLP system's integration with the Microsoft Semantic Kernel (Maeda J., Chaki E., 2023) is a fundamental aspect of our approach, allowing our system to handle complex user inputs with conversational intent and requests that extend beyond the predefined application steps. This integration gives our system access to an extensive language model and APIs, enabling it to communicate effectively with other external tools, ultimately enhancing the user experience.

One key advantage of the Semantic Kernel is its ability to store conversation context memory. This feature is particularly useful in scenarios where the user requires additional information, clarification, or a summary of the previous interaction. By storing context memory, our system can understand the user's intent and respond appropriately, leading to more productive and meaningful conversations; but also works with the Personality module and stores semantic user-related information, which is a key part of this module.

Another valuable feature of the Semantic Kernel (Bisson S., 2023) is its capability to define skills to be called during flow execution. These skills enable our system to perform complex tasks more efficiently, providing a more user-friendly experience. As an example, we have implemented a defined SK skill to serialise a JSON file containing a daily report into a more easily understandable summary, allowing the user to access the information they need quickly and even ask for clarifications.

Moreover, the Connectors functionality provided by the Semantic Kernel SDK is a critical feature we have integrated into our flow execution. This functionality enables our system to execute actions and retrieve information from external sources, such as adding or retrieving data from the user's calendar or sending emails. The Connectors functionality expands the capabilities of our NLP system, allowing it to interact with other applications and services seamlessly.

Overall, our NLP system is a sophisticated and dynamic approach that relies on the interdependent modules of tokenisation, entity matching, and personality modules. Integrating our NLP system with the Microsoft Semantic Kernel is a crucial aspect of our approach, providing numerous benefits, including the storage of conversation context memory, the ability to define skills, and the Connectors functionality. This integration enables us to provide a more comprehensive and effective NLP system that can handle complex user inputs and enhance the user experience.

CONCLUSIONS

The AIRobot project will catalyse further innovation within AI, NLP, and conversational interfaces as it progresses. By breaking down barriers between users and technology, AIRobot has the potential to revolutionise the way we interact with software systems across various sectors. Ultimately, our vision is to empower individuals and organisations to harness the full potential of AI-driven personal assistants, creating a more connected, efficient, and intuitive future. Our research showcases the transformative power of AIRobot in fostering a new wave of AI and NLP-enabled solutions across multiple industries. By simplifying human-software interactions and harnessing the latest advancements in AI and NLP, AIRobot offers a promising approach to revolutionising the global landscape of technology applications.

Our findings reveal the significant impact of AIRobot's generic architecture on the efficiency and accessibility of software systems in the target domains. We demonstrate how conversational AI interfaces can revolutionise how users interact with technology, leading to substantial improvements in operational processes and overall user experience.

Moreover, we discuss the challenges and adaptive strategies encountered during the project's development. The results highlight the potential of AIRobot to serve as a model for future AI and NLP-driven projects, paving the way for innovative applications and sustainable growth in partnering firms.

ACKNOWLEGMENTS

This research is co-financed by the European Fund for Regional Development through the Competitiveness Operational Program 2014 – 2020, project “Establishment and implementation of partnerships for the transfer of knowledge between the Iasi Research Institute for Agriculture and Environment and the agricultural business environment”, acronym “AGRIECOTEC”, SMIS code 119611.

REFERENCES

Bisson, S., 2023 - *Semantic Kernel: A bridge between large language models and your code.* Retrieved April 20, 2023, available at: https://www.infoworld.com/article/3693310/semantic-kernel-a-bridge-between-large-language-models-and-your-code.html.

Ursache, C., Sammeth, M., Alboaie, S., 2022 - *OpenDSU: Digital Sovereignty in PharmaLedger.* arXiv:2209.14879

Dib, O., & Toumi, K., 2020 - *Decentralized identity systems: Architecture, challenges, solutions and future directions.* Annals of Emerging Technologies in Computing (AETiC), Print ISSN, 2516-0281.

Friedrich, F., Mendling, J., & Puhlmann, F., 2011 - *Process Model Generation from Natural Language* Text. Notes on Numerical Fluid Mechanics and Multidisciplinary Design, 482–496. https://doi.org/10.1007/978-3-642-21640-4_36

Yu, J., Bohnet, B., Poesior, M., 2020 - *Named Entity Recognition as Dependency Parsing*. arXiv:2005.07150

Jo, H., 2022 - *Impact of Information Security on Continuance Intention of Artificial Intelligence Assistant.* Procedia Computer Science, 204, 768–774. https://doi.org/10.1016/j.procs.2022.08.093

Jofche, N., Mishev, K., Stojanov, R., Jovanovik, M., Zdravevski, E., & Trajanov, D., 2022 - *Named Entity Recognition and Knowledge Extraction from Pharmaceutical Texts using Transfer Learning.* Procedia Computer Science 721–726, available at: https://doi.org/10.1016/j.procs.2022.07.107

Barbieri, L., Madeira, E.R., Stroeh, K., van der Aalst, W.M.P., 2022 - T*owards a Natural Language Conversational Interface for Process Mining.* Part of the Lecture Notes in Business Information Processing book series (LNBIP,volume 433)

Maeda, J., Chaki, E., 2023 - *What is Semantic Kernel?* Microsoft. Retrieved April 20, 2023, available at https://learn.microsoft.com/en-us/semantic-kernel/whatissk.

Murthy, A. R., & Anil Kumar, K. M., 2021 - *A Review of Different Approaches for Detecting Emotion from Text.* IOP Conference Series: Materials Science and Engineering, available at: https://doi.org/10.1088/1757-899x/1110/1/012009.

Rai, A., & Borah, S., 2020 - *Study of Various Methods for Tokenization. Applications of Internet of Things*, 193–200, available at: https://doi.org/10.1007/978-981-15-6198-6_18

Sun, C. C. 2021 - *Analyzing Determinants for Adoption of Intelligent Personal Assistant: An Empirical Study,* available at: https://doi.org/10.3390/app112210618.