Emerging approaches to human-robot collaboration in healthcare

Luigi Gargioni
luigi.gargioni@unibs.it
Department of Information Engineering
University of Brescia
Brescia, Italy

ABSTRACT

Collaborative robots can enhance productivity and efficiency in healthcare. This PhD project aims to investigate new methods and tools for effective interaction with these robots, focusing on programming techniques accessible to domain experts without a background in computer science or robotics. Automating repetitive tasks can allow healthcare professionals to dedicate more attention to critical procedures. For instance, this technology can enhance therapy efficiency and personalized medicine preparation, benefiting patient outcomes. The research will investigate the use of Large Language Models to simplify and optimize robot task programming, reducing the need for technical expertise.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing devices*; • **Computer systems organization** → *External interfaces for robotics.*

KEYWORDS

Human-Robot Collaboration, Collaborative Robots, End-User Development, Large Language Models

ACM Reference Format:
Luigi Gargioni. 2024. Emerging approaches to human-robot collaboration in healthcare. In *International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 03–07, 2024, Arenzano, Genoa, Italy*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3656650.3656747

1 INTRODUCTION

The aim of the PhD project is to develop technologies that enhance and refine existing healthcare work procedures.

Collaborative robots are identified as a pivotal tool in achieving this objective. In the medical sector, healthcare professionals often perform repetitive and low-value tasks that can be delegated to collaborative robots. These robots are well-suited to environments like pharmacies due to their versatility in supporting a wide range of tasks. Robots are becoming increasingly advanced and have the potential to revolutionise pharmacy operations, improving patient safety while enhancing efficiency and productivity.

This paper and the related PhD project will explore two use cases in the pharmaceutical sector: the preparation of personalized medicines, also known as galenic formulations, and the preparation of specific therapies for patients.

The main objective of the project is to provide End-User Development (EUD) environments for end users, specifically pharmacists, to program collaborative robots without requiring technical knowledge. To achieve this objective, the use of emerging Large Language Models (LLMs) will be considered. The advantages and disadvantages will be analyzed to establish a reliable, robust and verifiable method of interaction.

2 RELATED WORK

In [19] an overview of automation and robot technologies that can be applied in pharmaceutical laboratories and industries is presented. Human-robot collaboration is often necessary, as robots lack the dexterity required for tasks such as measuring powder, loading samples, and transportation. Collaborative robots are recommended as a solution to provide safety features for shared workspaces. The article discusses different configurations of robotic solutions, including static robots and mobile manipulators with arms and sensors.

In [2], a survey was conducted on end-user robot programming, which presents unique challenges compared to traditional end-user programming. These challenges include the need for programs to interact with physical objects and environments and catering to users with varying backgrounds and technology literacy levels. The goal of end-user robot programming is to enable users without expertise in robotics or programming to navigate the complexities of robot programming.

Current methods for end-user robot programming include visual programming languages, ranging from computer-oriented notations such as flowcharts to those inspired by educational tools like Scratch and Blockly, or custom solutions (an extended literature review of these methods is presented in [6]). These languages allow users to create programs by dragging and dropping blocks that represent programming constructs or domain-specific concepts.

While natural language programming has been suggested as an alternative approach, its implementation encounters difficulties in managing the intricacy and safety-critical aspects of programming tasks. The recent proliferation of LLMs has revitalised this approach. In [14], an approach that exploits LLMs is presented. The research is from the chemical laboratory domain and the aim is to allow end users to create, test, and execute robot programs using natural language inputs. The limitation affecting this work is that end users do not possess knowledge about programming languages and cannot really understand the defined robot programs and assess their correctness. Other tools such as the one described in [18] use LLMs to generate Python or C++ code, with users providing feedback on code quality and safety. Also in this case, the approach requires that users possess programming knowledge. Finally, in [4], I presented a

AVI 2024, June 03–07, 2024, Arenzano, Genoa, Italy

ACM ISBN 979-8-4007-1764-2/24/06
https://doi.org/10.1145/3656650.3656747

first tentative to adopt a hybrid approach that combines natural language interaction supported by ChatGPT and visual programming. In particular, the developed prototype includes a graphical interface based on puzzle-like blocks to visualize the generated programs, accompanied by the option to modify programs directly through graphic manipulation.

3 DESCRIPTION OF THE RESEARCH PROJECT

3.1 Domain analysis and user research

Galenic preparations are medications compounded locally by pharmacists or veterinarians. They offer an alternative when commercially manufactured drugs are not suitable [7]. These preparations are tailored to meet the specific needs of individual patients, particularly children [5], and to adhere to precise medical directives.

As for the use case of customized therapies, these include the preparation of the correct amount of drugs and antibiotics appropriately divided according to the prescribed doses. This is very important to avoid mistakes during intake and to avoid overdosing that, in case of antibiotics, may lead to resistance to them.

Interviews with domain experts and analysis of documents describing the galenic formulation and therapies preparation processes, along with research on users' profiles, allowed identifying areas where collaborative robots could help professionals carry out specific activities.

3.2 Prototype for galenic formulation use case

This PhD project aims to address some of the challenges affecting the manual production process of galenic preparations by using collaborative robots to assist pharmacists with the most precise, error-prone, and time-consuming tasks.

Specifically, an EUD environment to aid pharmacists in programming the tasks to be performed by a collaborative robot is under development. The environment integrates artificial intelligence features based on NLP. At the same time, it is designed to ensure that end users always maintain complete control over the generated output, i.e. a robot program. A first prototype has been developed for this use case, based only on some steps of the galenic preparation process. The prototype workflow starts defining domain elements such as the required mixing action type and the final container for pills. Additionally, a natural language interface has been developed to create the preparation required by the pharmacist. This preparation is then displayed in a specific graphical interface that allows the pharmacist to check its correctness. This preparation will represent a set of actions that constitute a task for the robot.

3.3 NLP engine selection

To determine the most appropriate NLP approach for the described objectives, various tools and LLMs were analysed, highlighting their advantages and disadvantages. Initially, IBM's Watson [12] was evaluated due to its consolidated adoption since 2013 and extensive usage for intent recognition in natural language sentences. It employs artificial intelligence methodologies, comprising natural language analysis, machine learning and data processing, for data analysis, text categorisation, automatic translation, and the like [11]. On the one hand, the IBM web application is a well-established and long-standing tool; on the other hand, it requires the construction of the intent structure via a rigid flow chart interface [13]. This limits the flexibility of the system to a very custom usage.

After excluding IBM's Watson, an analysis of the most powerful and available LLMs was carried out. An LLM is an AI system that employs a vast amount of text and data to comprehend natural language and produce coherent responses or text. Thanks to this technology, it is possible to recognise sentence intents and maintain a consistent conversational flow. The open-source LLM Llama2 [16] [17] developed by Meta was considered. To utilise this tool, it is necessary to download and install the model locally. This would undoubtedly be advantageous in terms of independence from a third-party provider and customization possibilities; however, it requires significant hardware resources to achieve satisfactory performance [10]. Another LLM-based tool considered is Bard by Google [8]. Bard utilises the PaLM2 model [9] [3], also developed by Google. Its Internet access is a considerable advantage as it provides vast amounts of up-to-date information. However, Bard does not provide any APIs for integration into custom applications at this time, and its level of understanding falls below other LLMs [1].

ChatGPT was finally selected due to its computational capability [15], flexibility in training, and easiness of integration and customization.

4 CONCLUSION AND FUTURE WORK

In my first year of PhD research, I identified two use cases and conducted a domain analysis to design ideal workflows. I compared various NLP approaches, including different LLMs. Based on these analyses, I began developing a prototype for the first use case, and initial feedback obtained through an exploratory user study has been positive. This approach aligns with the concept of EUD, where users describe domain concepts and create robot programs in an unwitting manner. As far as the personalized therapies use case is concerned, a workflow hypothesis was formulated to prepare personalized therapies in collaboration with a collaborative robot. The described approach that exploits NLP for the definition of robot tasks and a specific representation for end-user verification may also be applied to this scenario by creating a suitable graphical interface for the visualization phase. Once the advantages and disadvantages of this approach are better understood, a digital twin could be incorporated into the EUD process to provide a more immersive and controlled feedback of the programming activity. Another potential future development is the investigation of using LLMs for retrieving information required to define robot tasks from a specific data source.

ACKNOWLEDGMENTS

This project and the related PhD scholarship are co-funded by the Italian Ministry[1] and the company *Antares Vision*[2].

I thank Prof. Daniela Fogli and Prof. Pietro Baroni for the supervision in this PhD program.

[1]Piano Nazionale di Ripresa e Resilienza (PNRR): https://www.mef.gov.it/focus/Il-Piano-Nazionale-di-Ripresa-e-Resilienza-PNRR/

[2]Antares Vision S.p.A.: https://www.antaresvisiongroup.com/lifescience/it/home/

REFERENCES

[1] Imtiaz Ahmed, Mashrafi Kajol, Uzma Hasan, Partha Protim Datta, Ayon Roy, and Md Rokonuzzaman Reza. 2023. ChatGPT vs. Bard: A Comparative Study. *UMBC Student Collection* (2023).

[2] Gopika Ajaykumar, Maureen Steele, and Chien-Ming Huang. 2021. A Survey on End-User Robot Programming. 54, 8, Article 164 (oct 2021), 36 pages. https://doi.org/10.1145/3466819

[3] Rohan Anil, Andrew M Dai, Orhan Firat, Melvin Johnson, Dmitry Lepikhin, Alexandre Passos, Siamak Shakeri, Emanuel Taropa, Paige Bailey, Zhifeng Chen, et al. 2023. Palm 2 technical report. *arXiv preprint arXiv:2305.10403* (2023).

[4] Giorgio Bimbatti, Daniela Fogli, and Luigi Gargioni. 2023. Can ChatGPT Support End-User Development of Robot Programs?. In *CEUR Workshop Proceedings*, Vol. 3408. https://ceur-ws.org/Vol-3408/short-s2-03.pdf

[5] Francesca Burlo, Davide Zanon, Paola Minghetti, Valentina Taucar, Giulia Benericetti, Giada Bennati, Egidio Barbi, and Lucia De Zen. 2023. Pediatricians' awareness of galenic drugs for children with special needs: a regional survey. *Italian Journal of Pediatrics* 49 (2023), 76.

[6] Daniela Fogli, Luigi Gargioni, Giovanni Guida, and Fabio Tampalini. 2022. A hybrid approach to user-oriented programming of collaborative robots. *Robotics and Computer-Integrated Manufacturing* 73 (2022), 102234.

[7] Nicolas Fortané. 2020. Antimicrobial resistance: preventive approaches to the rescue? Professional expertise and business model of French "industrial" veterinarians. *Review of Agricultural, Food and Environmental Studies* 102, 1 (January 2020). https://doi.org/10.22004/ag.econ.329942

[8] Google. 2024. Google Bard. https://bard.google.com/chat

[9] Google. 2024. Google PaLM 2. https://ai.google/discover/palm2

[10] HardwareCorner. 2024. Computer Hardware Required to Run LLaMA AI Model Locally (GPU, CPU, RAM, SSD). https://www.hardware-corner.net/guides/computer-to-run-llama-ai-model/

[11] Rob High. 2012. The era of cognitive systems: An inside look at IBM Watson and how it works. *IBM Corporation, Redbooks* 1 (2012), 16.

[12] IBM. 2024. IBM Watson. https://www.ibm.com/uk-en/watson

[13] IBM. 2024. IBM Watson Assistant. https://www.ibm.com/products/watsonx\-assistant

[14] Ulas Berk Karli, Juo-Tung Chen, Victor Nikhil Antony, and Chien-Ming Huang. 2024. Alchemist: LLM-Aided End-User Development of Robot Applications. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (<conf-loc>, <city>Boulder</city>, <state>CO</state>, <country>USA</country>, </conf-loc>) *(HRI '24)*. Association for Computing Machinery, New York, NY, USA, 361–370. https://doi.org/10.1145/3610977.3634969

[15] Zhi Wei Lim, Krithi Pushpanathan, Samantha Min Er Yew, Yien Lai, Chen-Hsin Sun, Janice Sing Harn Lam, David Ziyou Chen, Jocelyn Hui Lin Goh, Marcus Chun Jin Tan, Bin Sheng, et al. 2023. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. *EBioMedicine* 95 (2023).

[16] Meta. 2024. Meta Llama 2. https://ai.meta.com/llama/

[17] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288* (2023).

[18] Sai Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor. 2023. *ChatGPT for Robotics: Design Principles and Model Abilities.* Technical Report. Microsoft. https://www.microsoft.com/en-us/research/uploads/prod/\2023/02/ChatGPT_Robotics.pdf

[19] Ádám Wolf, David Wolton, Josef Trapl, Julien Janda, Stefan Romeder-Finger, Thomas Gatternig, Jean-Baptiste Farcet, Péter Galambos, and Károly Széll. 2022. Towards robotic laboratory automation Plug & Play: The "LAPP" framework. *SLAS Technology* 27, 1 (2022), 18–25. https://doi.org/10.1016/j.slast.2021.11.003