# GiDi: A Virtual Assistant for Screening Protocols at Home

Andrés Piñeiro-Martín[1,2], Carmen García-Mateo[1], Laura Docío-Fernández[1], María del Carmen López-Pérez[2], and Ignacio Novo-Veleiro[3]

[1]GTM research group, AtlanTTic Research Center, University of Vigo, Vigo, Spain
[2]Balidea Consulting & Programming S.L., Santiago de Compostela, Spain
[3]Home Hospitalization Unit, University Hospital of Santiago de Compostela, Spain

## Abstract

Home hospitalisation is emerging as a key pillar in the evolution of medical care, providing effective and safe hospital care for those patients for whom hospitalisation at home is the best option. However, its applicability is limited by its reliance on healthcare professionals physically travelling to patients' homes. This paper presents GiDi, a virtual assistant designed for screening patients with acute heart failure during home hospitalisation, providing healthcare professionals with information about the patient's condition, and allowing them to prioritise and focus on patients who really need attention. GiDi, fluent in Galician and Spanish, overcomes the challenge of bilingual environments. Developed with state-of-the-art open-source technology, it adheres to stringent healthcare data governance. This work and demonstration showcases our experience in integrating GiDi's components and AI modules, in close collaboration with medical professionals, and presents a robust industrial solution tailored to the Galician-Spanish context.

## 1 Introduction

Home hospitalisation (Hospitalización a Domicilio, HADO (Hermida-Porto et al., 2015)) is a key factor in the evolution of healthcare, particularly within the framework of the Servicio Gallego de Salud (SERGAS). As healthcare adapts to meet the increasing needs of an ageing population with chronic conditions, HADO embodies a patient-centred approach, providing compassionate care in the familiarity of the patient's home. However, there are practical limitations to its scalability - the need for healthcare professionals to physically visit patients at home, while crucial for personalised care, presents logistical challenges. To overcome these challenges and unlock the full potential of home hospitalisation, a key opportunity lies in exploring how emerging technologies, such as virtual assistants, can enhance and streamline patient care.

GiDi[1], a virtual assistant designed for monitoring acute heart failure (Farmakis et al., 2015) patients at home, addresses the practical challenges of HADO while improving the efficiency of healthcare professionals. By providing real-time patient information and prioritisation capabilities, GiDi enables healthcare teams to focus where needed, ultimately improving the quality and responsiveness of home care. In this context, GiDi's role goes beyond the use of technology to redefine what is possible within the home hospitalisation model.

Developed using an Ethics by Design methodology, GiDi is the result of collaboration with medical professionals from the IDIS (Instituto de Investigación Sanitaria de Santiago de Compostela) Foundation, as well as stakeholders and end-users. It incorporates state-of-the-art language technology to create a functional pipeline in Galician and Spanish, demonstrating the feasibility of building industrial solutions in bilingual and low-resource language environments using open source technology and robust data governance. This paper outlines the design of GiDi, provides detailed descriptions of its modules and explains its demonstration.

## 2 GiDi Description

### 2.1 Origin, design and pilot

The GiDi project (June 2022 - October 2024) is the result of a collaboration between Balidea S.L.[2], the GTM Department of the University of Vigo[3] and the IDIS Foundation[4], and is part of the Eurostars-3 programme, co-funded by the CDTI and supported by the Horizon Europe Research and Innovation Framework Programme of the European Union.

[1]GiDi takes its name from the phonetic transcription of the initials of "GrandDaughter".
[2]https://balidea.com/
[3]http://gtm.uvigo.es/en/
[4]https://www.idisantiago.es/

The assistant is designed to adapt a daily screening protocol for acute heart failure, based on an alert system designed by experts at the IDIS Foundation. This system, developed according to the Ethics by Design methodology (Piñeiro-Martín et al., 2022), involves GiDi asking patients questions by voice about basic measurements (sleep, weight, urine, oxygen saturation, blood pressure, etc.) that trigger alerts. Using natural language understanding, GiDi then communicates the screening results to healthcare professionals, providing a comprehensive and technologically advanced approach to healthcare screening.

The project is in its final phase of testing and development. In the coming months (June 2024), a pilot phase will start to evaluate the screening and the prototype. The pilot will be conducted with end users, in collaboration with the HADO team and with nursing home companies as stakeholders.

### 2.2 ASR multilingual system

GiDi is designed to comprehend spoken Galician and Spanish, and to achieve this, advanced multilingual models and multilingual strategies have been employed within the Automatic Speech Recognition (ASR) module. Specifically, we have performed multilingual fine-tuning of the new wav2vec bert 2.0 speech encoder (Barrault et al., 2023) and of the popular pre-trained XLS-R model (Babu et al., 2022) using balanced data from the Librispeech (Panayotov et al., 2015), Common Voice (Ardila et al., 2020) and audios from the FalAI dataset (Piñeiro-Martín et al., 2023; Piñeiro-Martín et al., 2024).

### 2.3 Virtual assistant text-based solution

To develop our text-based conversational solution, we used Rasa (Bocklisch et al., 2017), a collection of open-source Python libraries designed for developing conversational software. Rasa allows the development of Natural Language Understanding (NLU), Dialogue Management (DM) and Natural Language Generation (NLG) modules in both Galician and Spanish. This is achieved through bilingual/language-agnostic pipelines and language models, such as LaBSE (Language-agnostic BERT Sentence Embedding) (Feng et al., 2022), incorporating state-of-the-art Natural Language Processing (NLP) techniques, including the use of Transformers architectures (Bunk et al., 2020).

To create the decision tree and implement the screening protocol, we use a pipeline that integrates static rules and a model that learns from examples of conversations (Vlasov et al., 2019). Finally, due to the nature of the assistant's responses, they are based on templates and are slightly modified depending on the context of the conversation.

### 2.4 Bilingual text-to-speech

Given the bilingual nature of the assistant and the user's ability to switch between Galician and Spanish during the conversation, GiDi must deliver speech with consistent performance and characteristics in both languages. It is essential that GiDi adapts its voice to the language chosen by the user.

Due to the limited availability of suitable options, it was decided to use the synthetic voice available in Galician and to adapt the messages in Spanish to be compatible with this voice. For Galician, a language with relatively low resources, the only open-source text-to-speech (TTS) system available was COTOVÍA's voice (Banga et al., 2008; Díaz et al.), which unfortunately fell short of contemporary standards in terms of quality. However, a recent breakthrough has been achieved with the release of the first neural synthetic voice model for Galician through the Proxecto NÓS (de Dios-Flores et al., 2022). This newly developed voice model is now seamlessly integrated into GiDi.

### 2.5 Demonstration Plan

To showcase the capabilities of GiDi, we have developed a comprehensive demonstration plan that highlights its key features. Users can access the virtual assistant through any web browser[5], enabling them to explore and experience the following functionalities in Galician and Spanish:

- **Screening Protocol Simulation:** Provide the user with the ability to initiate and simulate the screening protocol, report the measures and generate the appropriate alarms. Access to the assistant's decision tree and screening protocol is provided for this simulation. In addition, the patient can follow the protocol by scheduling and continuing, pausing and resuming, proactively skipping and initiating, correcting and requesting measurements.

- **Language Handling:** GiDi supports bilingual conversations in Galician and Spanish. GiDi is able to recognise the language of the

[5]The test assistant will be accessible via the following link: https://gidibot.balidea.com/

speaker and switch accordingly. To achieve this, we have developed a multilingual communication strategy where the assistant only assumes the language until it clearly recognises the user's language or the user indicates which language they want to communicate in. This strategy involves the use of a multilingual ASR model, ASR models for each language once a language has been unambiguously recognised, and a text-based language identification (LID) model based on FastText (Bojanowski et al., 2017).

- **Advanced Intent Classification through Language Understanding:** Improve intent classification by using sentence embeddings from language agnostic models such as LaBSE and by searching for keywords in the user's message.

- **Contextual Understanding:** Analyzing both user messages using self-attention mechanisms over the sequence of dialogue turns and the bot's memory (slots) for nuanced contextual comprehension.

- **Enhanced Named Entity Recognition (NER):** Identifying and transcribing spoken numbers and measurements to ensure protocol adherence. For this purpose, we fine-tune the BERT model (Devlin et al., 2019) to perform NER.

- **Contextual Response Replay:** Users can effortlessly revisit the context of the conversation by requesting a repetition of the last response, ensuring clarity and maintaining a smooth interactive experience.

The demonstration will allow access to a test prototype to explore the functionalities presented above, but full access to screening will be limited as this is not the objective of this demonstration, and user identification functionalities and integration with the management platforms developed for this project will be disabled. Figure 1 shows an example of a conversation in the demo chat widget.

As part of the demonstration, a guide will be provided to help identify how to test the main functionalities of the assistant, indicating the measures included in the screening protocol, or examples of how to communicate with the assistant.

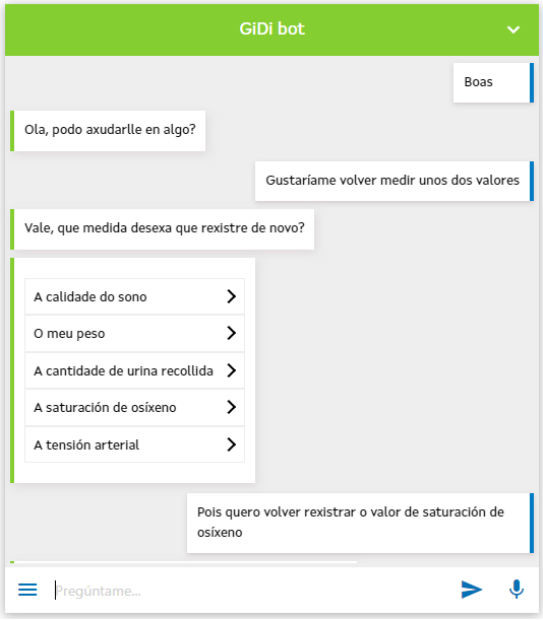


Figure 1: Conversation example in the GiDi chat widget.

## 3 Conclusions

In summary, GiDi, our bilingual virtual assistant for home hospitalisation, demonstrates the successful integration of open source technologies to build a robust pipeline. Developed using an Ethics by Design approach, GiDi not only skilfully navigates bilingual conversations, but also streamlines health screening protocols for acute heart failure patients. This work highlights the feasibility and effectiveness of building end-to-end solutions using open source tools, and offers a look at the potential transformative impact of virtual assistants in home healthcare.

## Acknowledgements

This research has been supported by the Galician Innovation Agency through the program "Doutoramento Industrial" and by the Xunta de Galicia through the grants: "Centro singular de investigación de Galicia accreditation 2019-2022" and GPC ED431B 2021/24.

## References

Rosana Ardila, Megan Branson, Kelly Davis, Michael Kohler, Josh Meyer, Michael Henretty, Reuben Morais, Lindsay Saunders, Francis Tyers, and Gregor Weber. 2020. Common voice: A massively-multilingual speech corpus. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4218–4222.

Arun Babu, Changhan Wang, Andros Tjandra, Kushal Lakhotia, Qiantong Xu, Naman Goyal, Kritika Singh, Patrick von Platen, Yatharth Saraf, Juan Pino, Alexei Baevski, Alexis Conneau, and Michael Auli. 2022. XLS-R: Self-supervised Cross-lingual Speech Representation Learning at Scale. In *Proc. Interspeech 2022*, pages 2278–2282.

Eduardo R Banga, Francisco Méndez, Francisco Campillo, Gonzalo Iglesias, and Laura Docío. 2008. Descripción del sintetizador de voz Cotovía para la evaluación Albayzin TTS.

Loïc Barrault, Yu-An Chung, Mariano Coria Meglioli, David Dale, Ning Dong, Mark Duppenthaler, Paul-Ambroise Duquenne, Brian Ellis, Hady Elsahar, Justin Haaheim, et al. 2023. Seamless: Multilingual expressive and streaming speech translation. *arXiv preprint arXiv:2312.05187*.

Tom Bocklisch, Joey Faulkner, Nick Pawlowski, and Alan Nichol. 2017. Rasa: Open source language understanding and dialogue management. *arXiv preprint arXiv:1712.05181*.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146.

Tanja Bunk, Daksh Varshneya, Vladimir Vlasov, and Alan Nichol. 2020. Diet: Lightweight language understanding for dialogue systems. *arXiv preprint arXiv:2004.09936*.

Iria de Dios-Flores, Carmen Magarinos, Adina Ioana Vladu, John E Ortega, José Ramom Pichel Campos, Marcos Garcia, Pablo Gamallo, Elisa Fernández Rei, Alberto Bugarín Diz, Manuel González González, et al. 2022. The nós project: Opening routes for the galician language in the field of language technologies. In *Proceedings of the workshop towards digital language equality within the 13th language resources and evaluation conference*, pages 52–61.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Francisco L Campillo Díaz, Francisco J Méndez Pazó, and Eduardo Rodríguez Banga. Estado actual y líneas futuras del sistema de conversión texto-voz gallego-castellano Cotovía.

Dimitrios Farmakis, John Parissis, John Lekakis, and Gerasimos Filippatos. 2015. Acute heart failure: epidemiology, risk factors, and prevention. *Revista Española de Cardiología (English Edition)*, 68(3):245–248.

Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2022. Language-agnostic BERT sentence embedding. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 878–891, Dublin, Ireland. Association for Computational Linguistics.

Leticia Hermida-Porto, LM Dopico-Santamariña, Fernando Lamelo-Alfonsín, B Aldamiz-Echevarría Iraurgui, MA Silva-César, and Luciano Vidán-Martínez. 2015. Hospitalización a domicilio en hospitales públicos gallegos. *Galicia Clínica*, 76(1):7–12.

Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. 2015. Librispeech: an ASR corpus based on public domain audio books. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 5206–5210. IEEE.

Andrés Piñeiro-Martín, Carmen García-Mateo, Laura Docío-Fernández, María del Carmen López-Pérez, and José Gandarela-Rodríguez. 2024. FalAI: A Dataset for End-to-end Spoken Language Understanding in a Low-Resource Scenario. In *Proceedings of the LREC-COLING 2024 Conference*. LREC-COLING.

Andrés Piñeiro-Martín, Carmen García-Mateo, Laura Docío-Fernández, and María del Carmen López-Pérez. 2022. Ethics Guidelines for the Development of Virtual Assistants for e-Health . In *Proc. IberSPEECH 2022*, pages 121–125.

Andrés Piñeiro-Martín, Carmen García-Mateo, Laura Docío-Fernández, and María del Carmen López-Pérez. 2023. FSTP Project Report - BEST Assistants - Building E2E Spoken-Language Understanding Systems for Virtual Assistants in Low-Resources Scenarios.

Vladimir Vlasov, Johannes EM Mosig, and Alan Nichol. 2019. Dialogue transformers. *arXiv preprint arXiv:1910.00486*.