Support of Migrant Reception, Integration, and Social Inclusion by Intelligent Technologies

WANNER, Leo, BOWEN, Daniel, BURGOS, Marta, CARRASCO, Ester, ČERNOCKÝ, Jan, CODINA, Toni, DANILINS, Jevgenijs, DAVEY, Steffi, DE LARA, Joan, DIMOPOULOU, Eleni, EGOROVA, Ekaterina, GEBHARD, Christine, GRIVOLLA, Jens, JARAMILLO-ROJAS, Elena, KLUSCH, Matthias, MAVROPOULOS, Athanasios, MOUDATSOU, Maria, NIKOLAIDOU, Artemisia, NTIOUDIS, Dimos, RODRÍGUEZ, Irene, ROSGOVA, Mirela, SHEKHAWAT, Yash, SHVETS, Alexander, SOBKO, Oleksandr, TZIONIS, Grigoris and VROCHIDIS, Stefanos

Citation:

Copyright and re-use policy

Support of Migrant Reception, Integration, and Social Inclusion by Intelligent Technologies

Leo Wanner [1,2,*], Daniel Bowen [3], Marta Burgos [4], Ester Carrasco [5], Jan Černocký [6], Toni Codina [7], Jevgenijs Danilins [8], Steffi Davey [9], Joan de Lara [5], Eleni Dimopoulou [10], Ekaterina Egorova [6], Christine Gebhard [11], Jens Grivolla [2], Elena Jaramillo-Rojas [12], Matthias Klusch [12], Athanasios Mavropoulos [13], Maria Moudatsou [10], Artemisia Nikolaidou [14], Dimos Ntioudis [13], Irene Rodríguez [7], Mirela Rosgova [14], Yash Shekhawat [8], Alexander Shvets [2], Oleksandr Sobko [3], Grigoris Tzionis [13] and Stefanos Vrochidis [13]

1. Catalan Institute for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain
2. Department of Engineering, Pompeu Fabra University, 08002 Barcelona, Spain; jens.grivolla@upf.edu (J.G.); alexander.shvets@upf.edu (A.S.)
3. NTT Data, 08005 Barcelona, Spain; daniel.antonio.bowen.vera@nttdata.com (D.B.); oleksandr.sobko@nttdata.com (O.S.)
4. Método Estudios Consultores, 36206 Vigo, Spain; martaburgos@metodogrupo.com
5. Generalitat de Catalunya, 08038 Barcelona, Spain; ester.carrasco@gencat.cat (E.C.)
6. Department of Computer Graphics and Multimedia, Technical University Brno, 612 00 Brno, Czech Republic; cernocky@fit.vutbr.cz (J.Č.); ekaterina.egorova@firma.seznam.cz (E.E.)
7. ISocial Foundation, 08018 Barcelona, Spain; tcodina@isocial.cat (T.C.); irodriguez@isocial.cat (I.R.)
8. Nurogames GmbH, 50676 Köln, Germany; jevgenijs.danilins@nurogames.com (J.D.); yash.shekhawat@nurogames.com (Y.S.)
9. CENTRIC, Sheffield Hallam University, Sheffield S1 1WB, UK; steffi.davey@shu.ac.uk
10. PRAKSIS, 10437 Athens, Greece; e.dimopoulou@praksis.gr (E.D.); m.moudatsou@praksis.gr (M.M.)
11. Caritas Hamm, 59065 Hamm, Germany; gebhard@caritas-hamm.de
12. Deutsches Forschungszentrum für Künstliche Intelligenz, 66123 Saarbrücken, Germany; elena.jaramillo@dfki.de (E.J.-R.); matthias.klusch@dfki.de (M.K.)
13. Center for Research & Technology Hellas, 57001 Thermi, Greece; mavrathan@iti.gr (A.M.); ntdimos@iti.gr (D.N.); gtzionis@iti.gr (G.T.); stefanos@iti.gr (S.V.)
14. Center for Security Studies, KEMEA, 11527 Athens, Greece; a.nikolaidou@kemea-research.gr (A.N.); m.rosgova@kemea-research.gr (M.R.)
* Correspondence: leo.wanner@upf.edu

Abstract: Apart from being an economic struggle, migration is first of all a societal challenge; most migrants come from different cultural and social contexts, do not speak the language of the host country, and are not familiar with its societal, administrative, and labour market infrastructure. This leaves them in need of dedicated personal assistance during their reception and integration. However, due to the continuously high number of people in need of attendance, public administrations and non-governmental organizations are often overstrained by this task. The objective of the Welcome Platform is to address the most pressing needs of migrants. The Platform incorporates advanced Embodied Conversational Agent and Virtual Reality technologies to support migrants in the context of reception, integration, and social inclusion in the host country. It has been successfully evaluated in trials with migrants in three European countries in view of potentially deviating needs at the municipal, regional, and national levels, respectively: the City of Hamm in Germany, Catalonia in Spain, and Greece. The results show that intelligent technologies can be a valuable supplementary tool for reducing the workload of personnel involved in migrant reception, integration, and inclusion.

Keywords: human-centered artificial intelligence; intelligent technologies for migrant support; conversational artificial intelligence; virtual reality; agent-based information delivery; agent-based coaching; agent-based training

1. Introduction

Migration is a constant on the world's political agenda, especially migration provoked by wars, civil unrest, and economic crises that lead to millions of individuals being forced to flee their homes and become refugees. This migration poses a societal challenge in the host countries. Prominent examples include the massive refugee flows from Syria and Iraq in 2015, from Ukraine in 2022, and the enduring flows from the African continent to Europe and from Latin America to the US. In these cases, most of the migrants come from different cultural and social contexts, do not speak the language, and are not familiar with the societal, administrative, and labour market infrastructure of the host country. This leaves them in need of dedicated personal assistance during their reception and integration. However, due to the continuously high number of people in need of attendance, public administration and non-governmental organizations are often overstrained with this task. The consequences are delays in legal registration, cancelled or shortened language and/or integration courses, difficulties surrounding incorporation into the labour market, and more, with all this leading to frustration, parallel societies, and poverty on the migrant side and lack of tolerance and acceptance on the side of the host country's society. For these reasons, support by intelligent assistance technologies is in high demand.

AI-based technologies (e.g., in educational, healthcare, and social or basic care applications) have already been successfully used for a long time. While some of these technologies have also already been used for migrant assistance [1–3], there is a significant difference between the functionality of the overwhelming majority of such applications and the complexity of the requirements towards an application that serves the needs of migrants and refugees, henceforth referred to as "Third-Country Nationals", TCNs (by using the term "Third-Country Nationals", we follow the official terminology of the European Commission). While these applications usually cover one specific task or serve one specific target group, the needs of TCNs are manifold; furthermore, the group of TCNs is very heterogeneous, such that any one-for-all solution will not work. Moreover, the available applications usually use one type of technology, such as personal assistants, virtual reality, or decision support. However, all of these are needed in order to address the needs of TCNs, and they need to be integrated so as to complement each other.

In what follows, we introduce the Welcome Platform, developed as part of a European Commission-funded project, aimed at providing an advanced solution to support TCNs in addressing some of their most pressing needs in Europe. The preliminary design of the Welcome Platform has been presented in [4]. The platform combines the technologies of intelligent Embodied Conversational Agents (ECAs) and Virtual Reality (VR) with traditional ways of information delivery such as FAQs and links to online material in order to serve TCNs best. Users can access services through a mobile device-based application in the language of their preference. The VR and the ECA-running mobile device applications are both based on the same technology, namely *Unity* (https://unity.com/), which facilitates their complementarity and integration.

The Welcome Platform has been successfully evaluated in trials with migrants in three European countries. This evaluation shows that intelligent technologies are a valuable supplementary tool for reducing the workload of personnel involved in migrant reception, integration, and inclusion.

The distinctive and innovative features of the Welcome Platform, compared to the available systems, are as follows:

- It aims to serve TCNs at all stages of their lives after arrival in the host country, namely, reception, integration, and social inclusion.
- It uses the most suitable type of technology for each of the addressed needs: ECAs, where verbal guidance is required, for example in the creation of a CV for job applications; VR, where visual experience and action is of use, such as in the context of language learning or acquaintance with cultural facilities; and a Frequently Asked Questions (FAQs) section, where standard information is available to address common concerns.

- It takes into account that the needs of migrants that must be addressed by municipal, regional, and state organizations may be different. For illustration, we show its use in the city of Hamm, Germany, the region of Catalonia in Spain, and at the national scale in Greece. At the same time, modular extension allows the Welcome Platform to cover other needs in other locations.

The remainder of this article is structured as follows: the next section presents the background of our work, i.e., the identified needs of TCNs and their alignment with intelligent technologies that address them best. Section 3 places our work in the context of related applications. In Section 4, we outline the architecture of the Welcome Platform and its individual modules. Section 5 illustrates the application of Welcome ECAs, while Section 6 describes the application of Welcome VR technologies. Section 7 summarizes the evaluation of the Welcome Platform in user trials. Finally, Section 8 draws some conclusions from the presented work. In view of the numerous abbreviations used in the paper, we provide a table with the full names of the Platform's different modules along with their corresponding abbreviations in an annex.

2. Background

In what follows, we first briefly outline the certified needs of the TCNs, and then discuss what types of technologies are best suited to address each of these needs.

2.1. Concerns of TCNs

The three central concerns related to migration in host countries are reception, integration, and social inclusion (https://reliefweb.int/report/world/story-journey-across-europe-first-reception-integration-migrants, accessed on 21 May 2024). Each of these concerns is a collaborative problem-solving task in which public administration officers and NGO employees (henceforth, "stakeholders") are much more than mere passive information service providers. They need to engage with the TCNs in order to acquire information or provide information, guide them on how to accomplish a specific task, or complete a task themselves, with information either acquired or at hand. The range of matters that a TCN has to deal with (and that stakeholders have to support them in) in the context of reception, integration, and social inclusion are manifold, and depend on the administrative structure and the legal framework of the host country. One of the consequences of this dependency is the different interpretation of what reception, integration, and inclusion imply. For instance, in Catalonia, Spain, reception already implies language training, introduction to the labour market and to the Catalan society and coaching, along with the recognition of academic qualifications and previous professional experience (https://ec.europa.eu/migrant-integration/integration-practice/first-reception-service_en, accessed on 21 May 2024). In Germany, however, all of this is addressed as an integration task (https://www.bamf.de/EN/Themen/Integration/ZugewanderteTeilnehmende/Integrationskurse/InhaltAblauf/inhaltablauf-node.html, accessed on 21 May 2024). Therefore, it is important, first, to briefly define how these three terms are used in our work.

Reception (https://www.unhcr.org/media/reception-asylum-seekers-including-standards-treatment-context-individual-asylum-systems, accessed on 21 May 2024) implies at least registration and orientation on the basic rights of TCNs in the given host country. While registration procedures differ from host country to host country, all of them share commonalities in requirements such as coordinating a visit to public administration, filling out registration forms, delivery of specific documentation, etc. Thus, TCNs require some guidance on where to find the corresponding public administration office, which information to introduce in what part of the form, what types of documentation are needed and in what time frames, etc. Orientation should at least cover information on the rights to seek asylum, receive medical assistance, enjoy temporary resident status, etc. – in short, human rights information.

Integration (https://www.oecd.org/regional/regional-policy/Migration-Flyer-FINAL.pdf, accessed on 21 May 2024) is likely to be the broadest and most heterogeneous topic. It

subsumes, e.g., language learning, incorporation into the labour market, learning about social and health services and the education system. In particular, language learning is central; even more than in conventional Second Language Learning setups, topic-oriented vocabulary acquisition is crucial. As a matter of fact, the ability to name things one refers to in different contexts, whether in interaction with the authorities or with fellow citizens, is of extraordinarily high importance.

Social inclusion (https://www.oecd.org/regional/Local-inclusion-Migrants-and-Refugees.pdf, accessed on 21 May 2024) is not equal to integration, as a person can be integrated into a host country in the sense that they speak the language and know and use all of the available services offered by the state, but may still be socially marginalized, i.e., not be able to participate in the social life of the host country. Participation in guided tours through social installations such as, e.g., public libraries or other cultural facilities, learning about social event calendars, and learning about administrative or societal details of the host region or country are only some of the possibilities that can foster better social inclusion.

Figure 1 summarizes the concerns of TCNs that have been identified as high-priority needs in the contexts of reception, integration, and inclusion by the stakeholders involved in the development of the Welcome Platform.
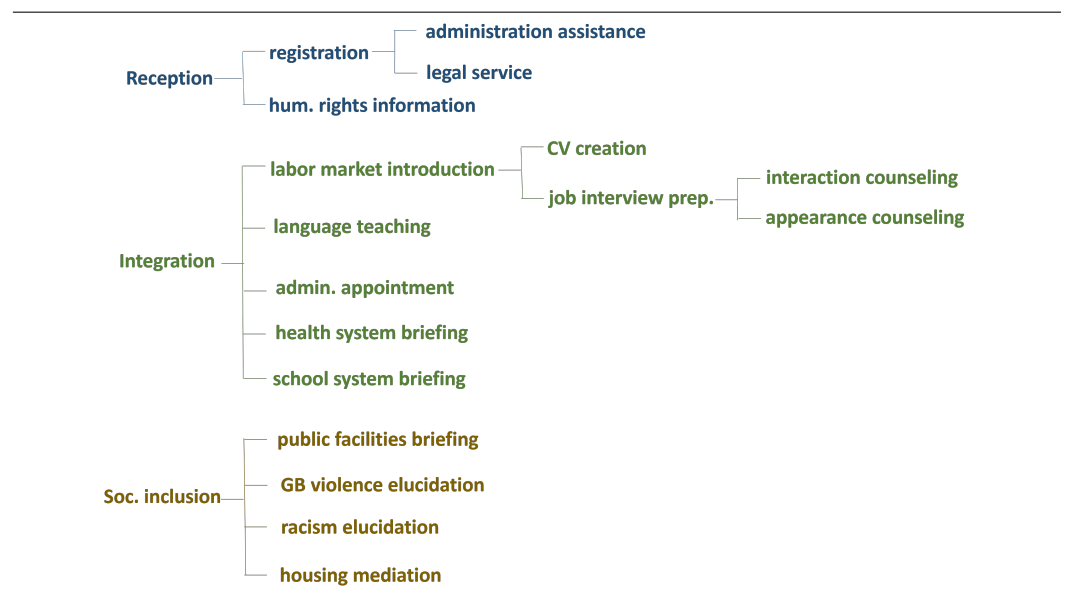

Figure 1. High-priority concerns of TCNs as captured in the leaves of the hierarchies related to reception, integration, and social inclusion.

2.2. How Can Intelligent Technologies Help?

The concerns of TCNs displayed in Figure 1 can be classified as needs for information provision, coaching, or training. For instance, an explanation of how a public library functions or what is needed to arrange an appointment with a lawyer is classified as information provision, while help with registration forms, CV creation, or advice on the apparel for a job interview is considered coaching, and improving skills for a job interview is considered training. As can be observed in the literature (see Section 3 below), each of the tasks of information provision, coaching, and training suggests specific types of technologies for its realization.

Information provision has most often been approached as a question-answering problem, which is then solved either as a simple static FAQ list, if standard general information is to be provided (https://www.searchenginejournal.com/best-faq-page-examples/267709/#close, accessed on 9 May 2024), or as a dynamic open domain model based on state-of-the-art natural language processing techniques [5]. In an interactive information request-delivery context with possible follow up clarification or detailing requests, a conver-

sational agent setup is more appropriate [6]. This is also the case for information delivery to TCNs.

Coaching is by nature an interactive setup in which a knowledgeable coach provides guidance and feedback that is needed for achieving a specific goal or task to someone with little or no experience. The task of coaching often involves an Embodied Conversational Agent acting as a virtual coach [7]. The use of Virtual Reality in which both the coach and the coached individual are immersed is also common [8]. The virtual coach can also be embedded into a Virtual Reality setup [9].

Training is occasionally mentioned in the context of coaching, especially when ECAs act as trainers [10,11]. However, it is, in principle, a different task, as it implies "the organized procedure by which people learn knowledge and/or skill for a definite purpose" [12], which is a priori not a given in the case of coaching. This procedure can be supervised and guided by a trainer, in which case it implies an aspect of coaching, or be self-guided; in the case of TCNs, it is the former. The most prominent training application for TCNs is language learning. Computer-Assisted Language Learning (CALL) is a very active line of research and development. An entire range of commercial and experimental programs is available, e.g., Babbel (https://uk.babbel.com/), Rosetta Stone (https://rosettastone.com), Lingo Pie (http://lingopie.com/), and more. The use of ECAs [13,14] and VR [15–18] for this task has also been widely explored. Indeed, both offer significant advantages. Conversational agents allow for personalized reactions to the observed mistakes of the learner along with explanatory dialogues and natural guidance to support the learner through exercises. VR allows objects to be visualized and facilitates actions for which the learners are supposed to learn the vocabulary. It also supports multiple-learner setups. This has been shown to have a great potential for improving learning outcomes [19].

Which types of technologies are most suitable for the implementation of information provision, coaching, or training depends on the characteristics of the information provided in respect of the topic being taught, coached, or trained. Table 1 shows the alignment followed in the Welcome Platform.

Table 1. Alignment of services with technologies. 'VPA': Virtual Personal Assistant"; 'VR': Virtual Reality; 'FAQ': Frequently Asked Questions; 'info': Information Provision; 'coach': Coaching; 'train': Training; 'GB': Gender-based (Violence).

		Info	**Coach**	**Train**
reception	registration		VPA	
	TCN rights	FAQ		
integration	CV creation		VPA	
	job interview interaction			VR
	job interview appearance		VR	
	language teaching			VR / VPA
	adm. appointment		VPA	
	health system info	FAQ / VPA		
	schooling system info	FAQ / VPA		
soc. inclusion	public facilities intro	VR		VR
	GB violence & racism	FAQ / VR		VR
	gender bias & discrim.	FAQ / VR		VR
	host country geography	VR		VR
	education	FAQ / VR		VR
	housing mediation		VPA	

In the context of reception, TCNs are coached through the registration procedures by a Virtual Personal Assistant (VPA) and are offered information they need to know about their rights through FAQs. During integration, services related to CV creation, recommendations on job interview appearance, language teaching, and arrangement of appointments with public administration also imply coaching by a VPA. As pointed out above, VR is a very useful technology in this context for language teaching or learning. Basic information on the health and schooling systems in the host country is best communicated by FAQs, while how to act during job interviews is best trained through VR. VR is also the optimal instrument for the majority of social inclusion services related to information provision, e.g., showing a TCN around in public facilities such as libraries, public gyms, or swimming pools or explaining what is meant by gender-based violence or how to fight racism. A convenient means for provision of some generic information on gender-based violence and measures against racism are the FAQs, while VPA is the most suitable means to mediate between TCNs who are looking for shared housing possibilities. As acknowledged by experts and as shown in [20], social inclusion services are ideally suited for language learning activities; therefore, all of the social inclusion services covered in the Welcome Platform include a language teaching component.

3. Related Work

Thanks to neural machine learning techniques, the recent advances in research areas related to the individual Welcome Platform technologies are extensive, and it would definitely go beyond the scope of this section to attempt to delve into detail of each of them. Therefore, we refer the reader to recent surveys on conversational agent technologies [21], dialogue management [22], speech recognition [23], language understanding [24], controllable natural language generation [25], computer-assisted language learning [26], and the use of VR technologies for learning [27] and social interaction [28]. In what follows, we focus on the related work on ICT specifically for migrant reception, integration, and inclusion.

Several studies have assessed and acknowledged the capability of ICT and AI to support the reception, integration, and social inclusion of migrants; see, e.g., [20,29–33]. Others have evaluated the potential of specific technologies such as, e.g., mobile devices or VR for integration [34–36] and language learning [37], respectively.

An increasing number of applications targets one or several tasks related to migrant reception, integration, and social inclusion. Thus, [38] presents a prototype and [39] a more mature version of the smartphone-based *MApp* application for language learning and social inclusion, including language lessons designed to assist informal learning in everyday life, with a focus on situational language needs and a social forum for peer support, cultural information, comments, and practice. Informal learning is personalized by assessing the history of the user's interaction with the application. In [40], the authors describe a mobile application for informal language learning using gamification techniques, while [41] focus on the co-design aspect of an embodied social chatbot that can provide answers to questions on six topics (language learning, internship application, vocational training, school, university application, and student finance) related to social integration. Interaction with the user is implemented using the *ProtoPie* tool (https://www.protopie.io/, accessed on 18 May 2024). *MyMigrationBot*, presented in [42], uses another off-the-shelf dialogue manager, namely, *Twilio* (https://www.twilio.com/, accessed on 18 May 2024). It interacts with migrants through a Facebook Messenger based-interface to acquire personality traits such as extroversion, emotional stability, and agreeableness, along with the competencies required for their current job. *NADINE-bot* [1] retrieves answers to the questions of migrants from repositories of FAQs on the administrative procedures of different EU countries. The focus is on segmentation and matching of the retrieved information with the user's inquiry to deliver the most relevant chunk, while a chat module is added to simulate conversational capacity. The *IMMERSE* platform [2] offers a search engine for different

types of relevant information and decision support, e.g., in the context of a job application where no verbal interaction is foreseen.

Despite all these advances, we can conclude in view of the multiple heterogeneous concerns of TCNs listed in Section 2.1, which require different types of technologies to address them (cf. Section 2.2), that none of the reviewed related works offers an integrated solution that incorporates FAQs, advanced conversational agent technologies, and VR. The Welcome Platform aims to provide such a solution.

4. The Welcome Platform

4.1. Overview of the Welcome Platform

As mentioned above, the Welcome Platform is an integrated solution of technologies for FAQ answering, Virtual Personal Assistants, and VR that provides information provision, coaching, and training services to TCNs. Apart from the needs of the TCNs, the Welcome Platform also caters for the needs of stakeholders, i.e., public administrators, NGOs and language teaching personnel, by providing decision support technologies in terms of a Visual Analytics Component (VAC) and a Teacher Panel (TP). However, in this article, we focus on the applications that aim to address the needs of TCNs.

The integration provided by the Welcome Platform goes beyond a common interface that facilitates access to different types of technologies. Thus, the Personal Assistants may refer to specific FAQs during interaction with TCNs, and VR can incorporate avatars. Figure 2 outlines the high-level architecture of the Welcome Platform.
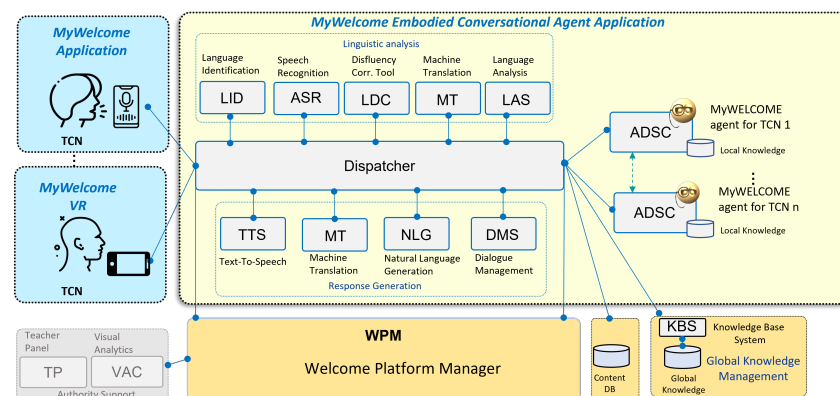

Figure 2. High-level architecture of the Welcome Platform. 'LID': Language Identification, 'ASR': Automatic Speech Recognition, 'LDC': Language Disfluency Correction, 'MT': Machine Translation, 'LAS': Language Analysis, 'TTS': Text-to-Speech, 'NLG': Natural Language Generation, 'DMS': Dialogue Management Service, 'ADSC': Agent-Driven Service Coordination, 'KBS': Knowledge Base System, 'WPM': Welcome Platform Manager, 'TP': Teacher Panel, 'VAC': Visual Analytics Component.

The Platform is managed through the Welcome Platform Manager (WPM), which offers software maintenance features such as displaying the status of the individual modules in the platform, the logs produced by the agents and other components as a result of their interaction, etc. Furthermore, administrators (public administration officers or NGO employees) can use the WPM to register users and manage their profiles. Upon registration, each user is assigned their personal instance of the VPA (referred to as *MyWelcome Agent*), which is active during interaction with its "Master" and with other agents and dormant otherwise.

A registered user (in our case, a TCN) can access the Platform via one of its two front-ends: MyWelcome Application and MyWelcome VR. The MyWelcome Application is installed on the user's mobile device. It provides easy and centralised access to several services, including, e.g., profile management, contact with public administration, links to relevant external online information, personalization of the appearance of the personal assistant, written and oral interaction with the personal assistant, display of personal

documentation created during the interaction with the ECA (e.g., filled out registration or first reception forms or the CV), download of the PDF version of the created documentation, and more. In addition, the Application runs the Frequently Asked Questions section and Vocabulary Learning Exercises services (cf. Figure 3 for a screenshot).

The MyWelcome VR serves as an engaging means for information delivery, coaching, and training. Figure 4 shows a game setup for word spelling in the context of language learning.
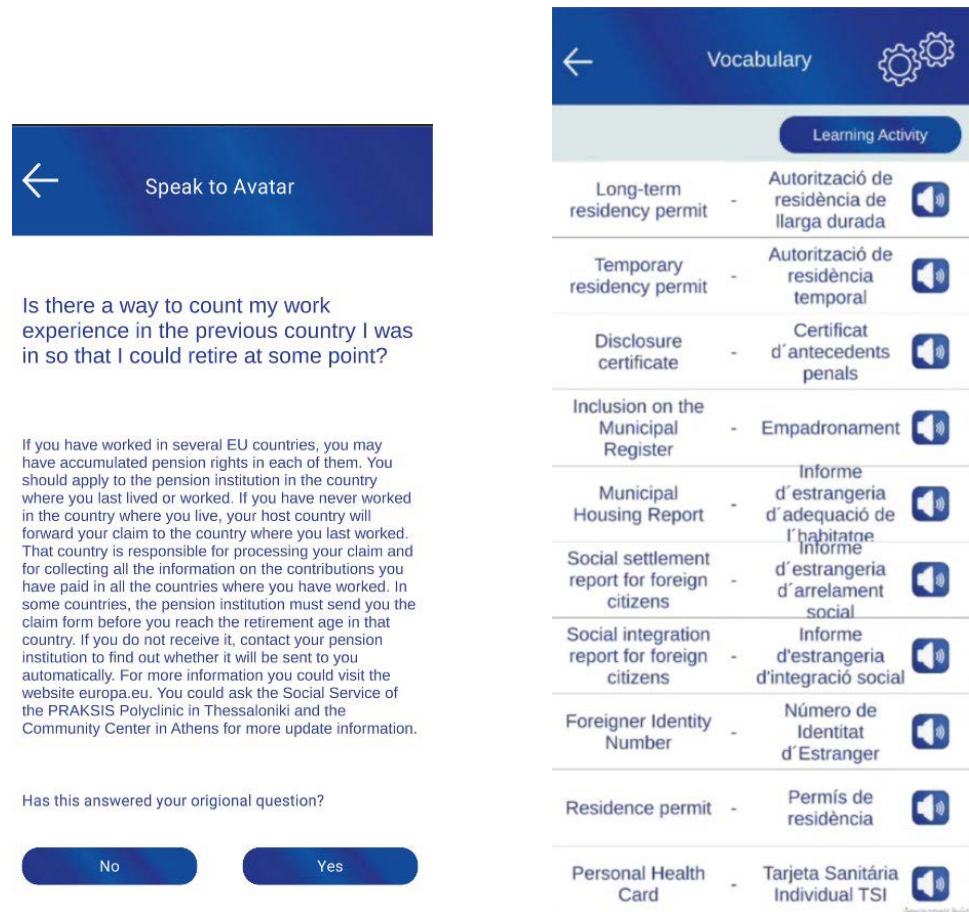

Figure 3. FAQs and language learning exercise featured by the MyWelcome Application (MWA).

Figure 4. Word spelling VR setup.

Both the MyWelcome Application and the MyWelcome VR are implemented in Unity (https://unity.com), which facilitates the shared use of common libraries and features and prepares the ground for the integration of avatars of the ECAs into VR (the current release of the Welcome Platform does not feature this integration). Both are linked to the Dispatcher, which mediates the partially asynchronous data requests and delivery between the front-ends, the modules that compose the Platform, and the data and knowledge stores. The modules include technologies for the analysis of spoken language analysis: Language Identification (LID) and Automatic Speech Recognition (ASR); for translation and analysis of the obtained transcripts and written language user interventions: Machine Translation (MT) and Language Analysis (LAS); and for generation of written and spoken agent interventions: Natural Language Generation (NLG) and Text-to-Speech (TTS). The actions and interventions of the personal agent of a user are planned by the Agent-Driven Service Coordination (ADSC) instance deployed for this specific user and by the Dialogue Management Service (DMS).

4.2. The Global Data and Knowledge Stores

The data and knowledge stores shown in Figure 2 are global, i.e., accessed by the WPM or by any of the ADSC instances. In addition, each ADSC instance disposes of a Local Agent Knowledge Repository (LAKR), which is controlled by its Knowledge Management Service (KMS) and which contains the agent's factual knowledge on the TCN profile and domain knowledge according to the global Welcome ontology in the Welcome Domain Knowledge Repository (WDKR) hosted by the KBS.

In this section, we describe the global stores; for the presentation of the LAKR and KMS, see Section 4.4.1, in which ADSC is described in more detail. The global data and knowledge stores are composed of the Content Database (CDB) and the Global Knowledge Base (GKB), which is managed by the Knowledge Base System. The CDB contains information that is communicated literally to the TCN (e.g., legally sensitive statements for which alteration is not foreseen) and auxiliary information for the technologies (e.g., sentence templates for Natural Language Generation). Its main components are a MySQL DB, Network File Storage, and Solr for rapid full text search. It can be accessed via RESTful APIs and a Web UI.

The GKB contains a number of different ontology subrepositories, including, e.g., (i) the Domain Knowledge Repository, which manages information that can be of use to all agents (i.e., all TCNs), such as the address of the immigration office in a city or the structure of a CV; (ii) the Agent Template Repository, which contains the characteristics of the MyWelcome agents needed for their dynamic creation upon the registration of a user; (iii) the Agent Repository containing the list of agents that have been deployed on the Platform and their current status (active or dormant); (iv) the FAQs, which are modeled in the ontology because they can be referred to during the dialogue of the ECA with the TCN; and (v) user-related data such as the TCN or organization profile definitions, etc.

All GKB subrepositories are encoded in OWL2 (https://www.w3.org/TR/owl2-overview/, accessed on 7 June 2024) and stored in RDF4J triple stores (https://rdf4j.org/, accessed on 7 June 2024). Figure 5 shows the definition of the TCN and organization profiles and the FAQs, respectively; the blue rounded rectangles are classes, while the green parallelograms indicate properties.

Management of the GKB (insertion, retrieval, updating, and removal of information via SPARQL queries) is realized by the Knowledge Base Service (KBS). The KBS also creates the Local Agent Repositories for newly initiated agents (see below).
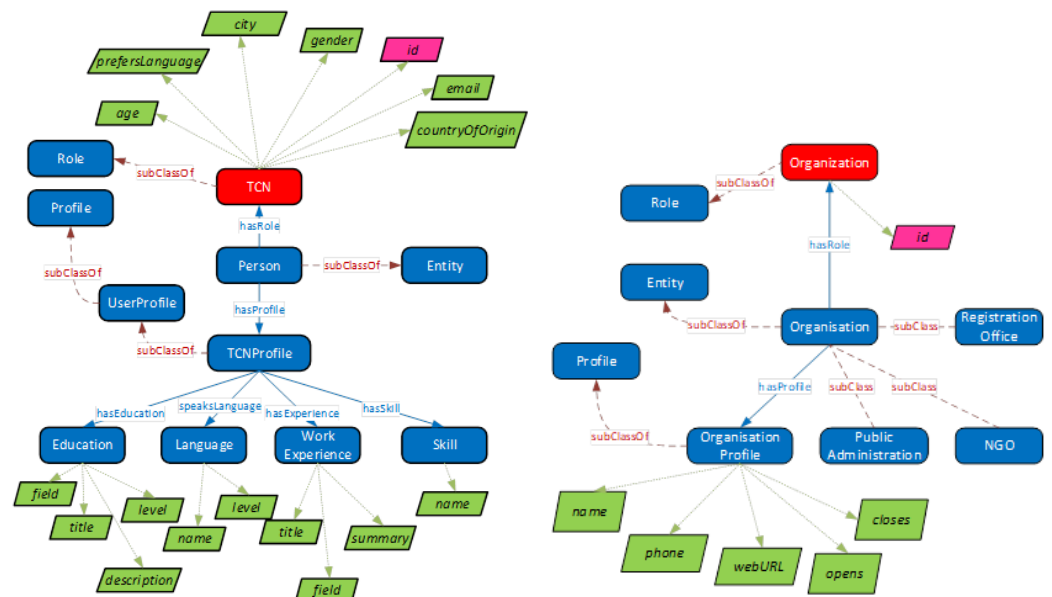
Figure 5. The definition of the TCN (**left**) and organization (**right**) profiles in the WELCOME ontology.

4.3. Language Technologies

As indicated in Figure 2, the Platform incorporates an entire range of language technologies required for spoken and written communication with the user in the language of their preference. Thus far, the Welcome Platform covers Moroccan (Darija) and Levantine Arabic, Catalan, English, German, Greek, and Spanish. Modern Standard Arabic is used to react to interventions of TCNs in Darija or Levantine Arabic due to the lack of reliable text-to-speech technologies for Darija and Levantine Arabic.

It is important to note that the Welcome Platform's language technologies are not system-specific, and can be integrated via the Dispatcher into any other application or used as stand-alone modules. In the current release, they are mainly used by the deployed MyWelcome Agents, while the MyWelcome Application and the MyWelcome VR front-ends also use language identification, speech recognition, machine translation, and text-to-speech synthesis. Next, we briefly introduce each of these technologies.

4.3.1. Language Identification (LID)

The LID model in the Welcome Platform is a Gaussian Linear classifier model [43] trained on sentence embeddings (or *i-vectors* [44]), which are obtained from 80-dimensional acoustic *Stacked Bottle-Neck* (SBN) features [45]. Publicly available corpora (NIST's Language Recognition Evaluation challenges, Babel, Fisher, Common Voice, etc.) have been used for training all three models (the Gaussian classier, the i-vector extractor, and the SBN feature extractor). The LID accuracy on held-out data from these corpora is between 0.94 (for Spanish) and 0.97 (for Catalan and German). Welcome Platform domain-specific native speaker monologues and/or dialogues were recorded for Catalan, Greek, and Moroccan and Levantine Arabic. The LID accuracy on these recorded data ranged between 0.4 for Catalan and 0.96 for Greek.

4.3.2. Automatic Speech Recognition (ASR)

For each of the seven languages covered in the Welcome Platform, an ASR model was trained separately on the corresponding language data using the *Speechbrain* toolkit (https://speechbrain.github.io/, accessed on 14 June 2024) in a character-based or word fragment-based end-to-end setup and the *wav2vec* pretrained model [46]. The output of the ASR module is a "one-best" text transcription of an utterance, with separate word confidences that are used by the Knowledge Management and the Dialogue Management Services to decide whether a clarification of what language is being spoken is needed.

The Word Error Rate (WER) on in-domain data ranges between 10.4% for Catalan and 19.4% for Spanish. On user-collected data, the WER is considerably higher (e.g., 39.6% for Catalan). This is likely due to the different genres of data used for training and the poorer acoustic quality of the user data.

4.3.3. Machine Translation (MT)

MT plays an important role in the multilingual setup of the Welcome Platform. Due to the complexity of the deep semantic analysis of low-resourced languages such as Moroccan and Levantine Arabic, English serves as interlingua. The ASR transcriptions are translated into English before analysis. The verbal reactions of the agent are also generated in English, then translated into the language spoken by the TCN.

Two Multilingual Neural Translation (MNT) models have been trained: one for translation between English and the other covered Indo-European languages (Catalan, French, German, Greek, and Spanish), and another one for translation between Arabic (Modern Standard Arabic and Moroccan and Levantine Arabic) and the Welcome Platform's Indo-European languages (including English). Both MNT models are based on the encoder–decoder transformer architecture [47] and trained using the MarianNMT framework (https://marian-nmt.github.io/, accessed on 12 September 2024). For training, the OPUS7 collection of open-source parallel corpora (https://opus.nlpl.eu/, accessed on 12 September 2024) and data collected as part of the project (850 M sentence pairs for the English ↔ Indo-European model and 730 M sentence pairs for the Arabic ↔ Indo-European model) were used. Assessment of the MT models in terms of the BLEU and COMET metrics showed that the models largely achieve state-of-the-art quality, such as NLLB 1.3B and NLLB 0.6B (https://ai.meta.com/research/no-language-left-behind/, accessed on 12 September 2024).

4.3.4. Language Disfluency Correction (LDC)

Spontaneous speech often contains hesitation or disfluency markers, such as fill words (*hmm*, *uh*, *you know*, …), corrections, repetitions, etc. [48]. As such markers pose a challenge, especially for syntactic language analysis, the three most common disfluency markers (correction, repetition, and restart, possibly preceded by a fill word) are eliminated before the (potentially translated) verbal intervention of the user is passed on for language analysis. For this purpose, the model from[49] is integrated for detection, classification, extraction, and correction of these three types of disfluency.

4.3.5. Language Analysis Service (LAS)

The LAS is composed of two submodules: the Surface Language Analysis (SLA) submodule and the Deep (or Semantic) Language Analysis (DLA) submodule. SLA employs the UDPipe2 (https://github.com/ufal/udpipe, accessed on 14 September 2024) pipeline, to obtain the universal dependency tree (https://universaldependencies.org/, accessed on 14 September 2024) of a given statement after its tokenization, PoS-tagging, and lemmatization. In DLA, the obtained *surface* dependency tree is further enriched by outputs of Named Entity Recognition (NER), Concept Extraction (CE), Word Sense Disambiguation (WSD), geolocation, and speech act detection. For NER, Spacy's off-the-shelf models (https://spacy.io/, accessed on 4 July 2024) are used. For geolocation, the data from OSM (https://www.openstreetmap.org/, accessed on 5 July 2024) and GeoNames (https://www.geonames.org/, accessed on 5 July 2024) are converted into search indices for a basic search engine realized in the Welcome Platform. Concept extraction is performed by the pointer generator network model from [50].

The enriched universal dependency tree is then projected via an intermediate deep syntactic structure onto a semantic structure using a faster and more efficient reimplementation of the MATE [51] graph transducer grammar. The semantic structure, in which the nodes are labeled with concepts and the edges with predicate–argument relation labels, is

mapped onto an ontological representation and incorporated into the Local Knowledge Repository of the user's agent; see Figure 6 for the corresponding ontology schema.
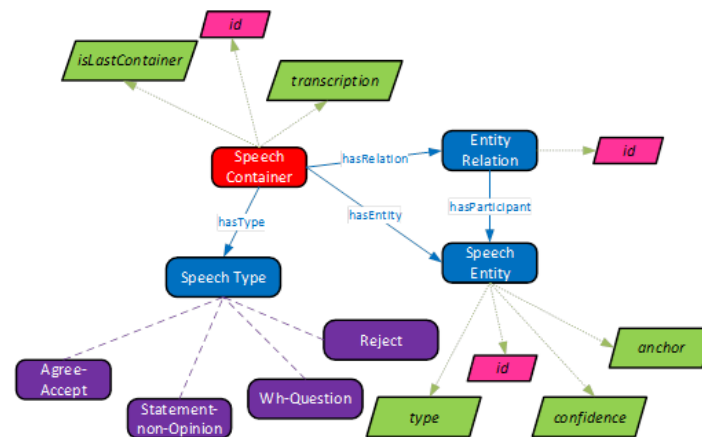

Figure 6. Ontology schema for the representation of the language analysis output.

4.3.6. Multilingual Natural Language Generation (NLG)

Many of the state-of-the-art task-oriented dialogue applications use sentence template-based Natural Language Generation (NLG) [52–54]. However, template-based generation falls short in the context of the Welcome Platform. Thus, although a significant amount of information is standardized, and as such can be handled well by templates (e.g., "*Your nearest registration office is located in* <ADDRESS>"), a lot of information is encoded in the ontology and calls for flexible realization. In this case, full-fledged sentence generation is thus more appropriate [55–57]. As a consequence, we use a combination of template-based and full-fledged sentence generation. Both are realized using an extended version of the graph transducer-based FORGe generator [58], which is grounded in the multistratal linguistic model of Meaning–Text Theory [59]. The generator receives a semantic structure from the Dialogue Management Service as input. If a standard statement is to be generated, corresponding sentence template(s) are retrieved from the Content Database and filled with the information from the structure; otherwise, FORGe's language-specific generation grammars are used.

4.3.7. Text-to-Speech Synthesis

For text-to-speech (TTS) synthesis of the generated agent turns in Catalan, English, French, German, Greek, Modern Standard Arabic, and Spanish, the Open TTS (https://github.com/synesthesiam/opentts, accessed on 23 April 2024) and Google TTS (https://cloud.google.com/text-to-speech, accessed on 23 April 2024) have been integrated. The voice matches the gender of the avatar chosen by the user as embodiment of the agent.

4.4. Personal Assistant Realization

Unlike the majority of state-of-the-art ECA applications [22,60], the Welcome Platform's personal assistant needs to be able not only to conduct a meaningful dialogue with a human user but also to retrieve external information (e.g., the opening hours of an office) and to reason (e.g., determine the office nearest to the residence of the TCN). This is achieved by a two-module configuration of the central part of the Conversational Agent (CA): the Agent-Driven Service Coordination (ADSC) and the Dialogue Management Service (DMS). The two modules configuration allows us to handle both non-verbal and verbal actions in depth. As a rule, non-verbal actions imply reasoning based on the user's input along with their user profile, interaction history, background knowledge, etc., for retrieval of information via external services or to update data repositories. This type of action is very different in nature from planning for the realization of a specific verbal action in context, which is truly a "dialogue management" task.

Based on the outcome of the analysis of the current move of the user, their interaction history, user profile, and background knowledge on the conversation topic, the ADSC reasons to select and compose "services" that constitute the appropriate reaction of the agent, which can be of verbal or non-verbal nature. TCN registration, delivery of health system information, and collaborative creation of a CV for job application are examples of such services (see Table 1).

The execution of a service is modelled by one or more Behaviour Trees (BTs) of the CA. Realization of those parts of the BT that imply verbal interaction with the user is assumed by the DMS based on the input it receives from the ADSC (via the Dispatcher) in terms of a Dialogue Input Package (DIP). A DIP is a frame structure that includes a number of slots. Each slot refers to a particular aspect or topic of an ongoing conversation that is relevant for providing the requested (composite) service to the TCN, such that by selecting specific slots for realization in the next agent move, the DMS determines the flow of the dialogue. Figure 7 illustrates the interplay between the ADSC and the DMS.
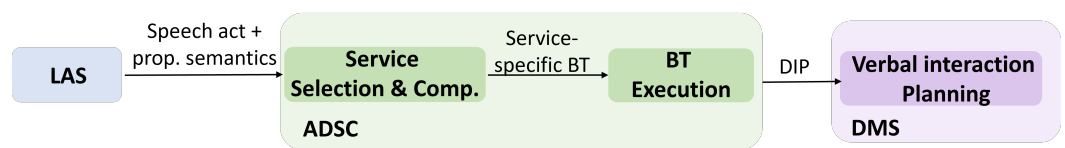

Figure 7. Agent-Driven Service Coordination (ADSC)–Dialogue Management Service (DMS) interplay. 'LAS': Language Analysis Service, 'BT': Behaviour Tree, 'DIP': Dialogue Input Package.

4.4.1. Agent-Driven Service Coordination (ADSC)

Each MyWelcome agent (recall that each TCN is attended by their own personalized agent) has a proprietary ADSC instance assigned to it. The main component of an ADSC instance is the agent deployment and execution framework Access Java Agent Nucleus (AJAN) server, which is equipped with HTTP interfaces for external interaction [61]. The AJAN Agent (also referred to as "Agent Core", AC) deployed in an ADSC instance accesses knowledge in three local RDF repositories: the Local Agent Knowledge Repository (LAKR), the Local Service Repository (LSR), and the Local Agent Repository (LAR), and is interfaced with its Semantic Service Computing (SSC) submodule for semantic selection and composition of relevant services for the TCN as well as with the Knowledge Management Service (KMS) for management of the local knowledge repository and semantic reasoning of the given Agent Core. Figure 8 shows the internal configuration of an ADSC instance.
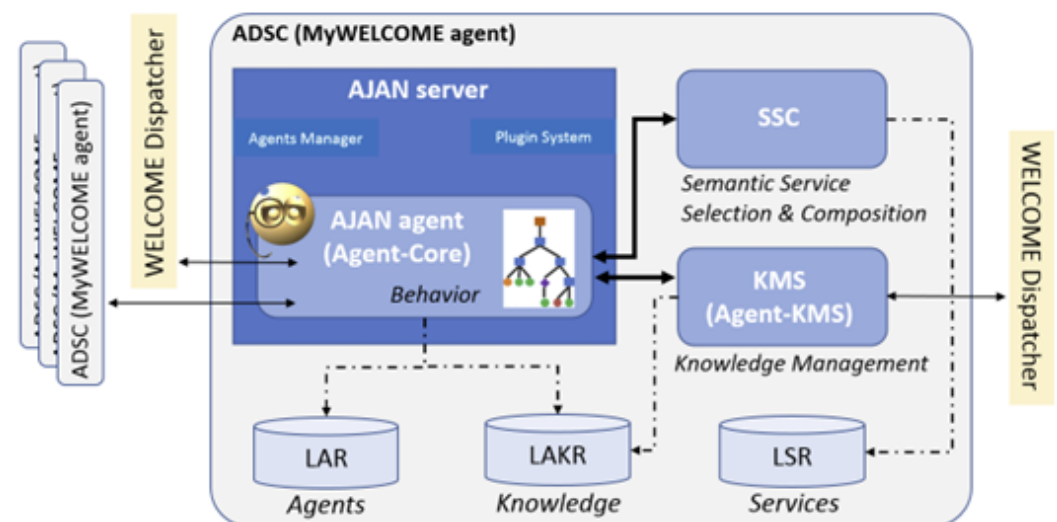

Figure 8. The architecture for Agent-Driven Service Coordination (ADSC). 'AJAN': Access Java Agent Nucleus, 'SSC': Semantic Service Computing, 'KMS': Knowledge Management Service, 'LAR': Local Agent Repository, 'LAKR': Local Agent Knowledge Repository, 'LSR': Local Service Repository.

4.4.2. Agent Core (AC)

As mentioned above, the behavior of a MyWelcome Agent (and consequently of its Agent Core) is modelled using Behavior Trees (BTs) that are conditioned by the facts stored in the agent's LAKR (obtained by the Language Analysis Service from the moves of the TCN and fed into the LAKR by the Knowledge Management Service; see Figure 9 for illustration). Facts are accessed via SPARQL queries attached to the nodes of the tree and evaluated with respect to demands of social services. To react to a demand, the AC invokes its internal Semantic Service Computing (SSC), which identifies the corresponding service in the Local Service Repository (LSR). If no relevant service is found in the LSR, the AC invokes the SSC to call its semantic service composition planner to satisfy the given request of a service and provide the corresponding BT back to the AC. If no service composition matches the goal, the agent asks the DMS to inform the user that the service request cannot be satisfied and proposes to check the information available in the Frequently Asked Questions provided in the MyWelcome Application.
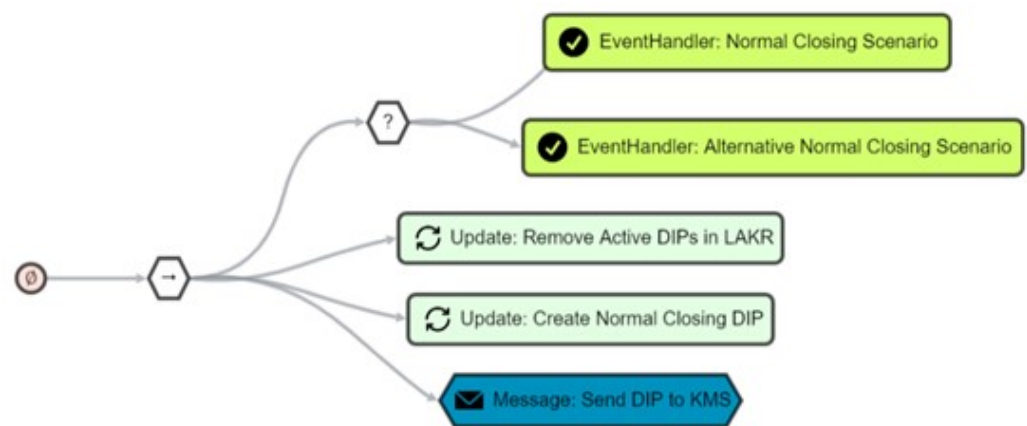

Figure 9. A sample compressed behavior tree; the actions in the tree create a farewell DIP (to be realized by the DMS) and reset the Local Knowledge Repository of the agent. 'DIP': Dialogue Input Package, 'LAKR': Local Agent Knowledge Repository, 'KMS': Knowledge Management Service.

The BTs of the service(s) selected or dynamically composed by the SSC are executed by the AC. This task involves interaction of the agent with the TCN to provide or request information. For this purpose, the Agent Core communicates with the DMS, which is responsible for determining a targeted dialogue strategy based on the information encoded in the DIP by the AC.

4.4.3. Local Service Repository (LSR)

The Local Service Repository (LSR) is a copy of the Welcome Service Repository (WSR), and as such contains the list of registered semantic services from which the SSC of an agent can make a selection in the course of the interaction between the agent and the TCN.

A local copy of the services is maintained for each agent in order to ensure private service execution bookkeeping. For initialization, a pull mechanism is executed by the AC. The AC sends a message (via the Dispatcher) to a specific endpoint of the KBS to retrieve the complete content of the WSR for initializing the LSR; this communication is asynchronous. As already the LAKR, the LSR encodes the representation of the services in an RDF4J triple store with a SPARQL endpoint.

4.4.4. Local Agent Repository (LAR)

The Local Agent Repository (LAR) of an agent contains the list of active agents with which the agent can communicate during a multiple agent interaction scenario. During initialization of the local repositories, the AC populates the LAR with information from the Welcome Agent Repository (WAR) through a pull mechanism. In particular, the AC sends a message to the KBS via the Dispatcher to retrieve the WAR content. Upon receipt of this

request, the KBS component of the platform retrieves from the WAR only agents with the status "active".

As the other local knowledge repositories, LAR is realized in terms of an RDF4J triple store.

4.4.5. Agent Knowledge Management Service (KMS)

The Knowledge Management Service (KMS) of an agent stores, queries, and updates the knowledge in its LAKR. As shown in Figure 10, the KMS is composed of dedicated components for Knowledge Base Population, Dynamic Ontology Extension, and the Semantic Reasoning Framework. Knowledge Base Population is responsible for translating received data (e.g., the MyWelcome Application, from the Language Analysis Service or from the Agent Core) in various formats into an RDF-based representation that adheres to the schema of the Welcome Platform Ontologies as well as for introducing the formatted data into the LAKR. Dynamic Ontology Extension facilitates the integration of information from an external multilingual semantic network.
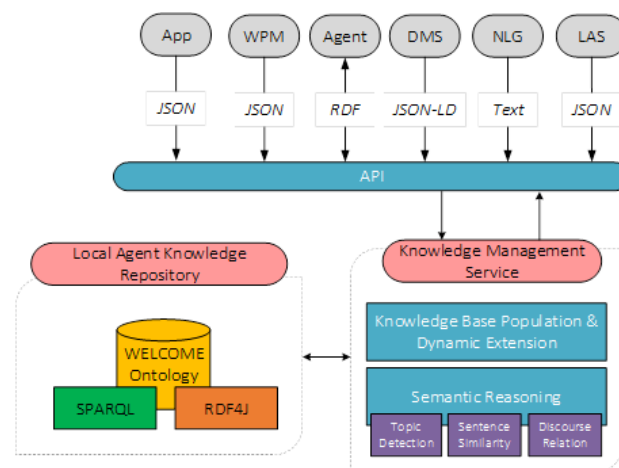

Figure 10. Architecture of the Knowledge Management Service (KMS). 'App': Application, 'WPM': Welcome Platform Manager, 'DMS': Dialogue Management Service, 'NLG': Natural Language Generation, 'LAS': Language Analysis Service.

Each time the KMS receives an input from the Language Analysis Service (LAS), ontology extension is triggered to update and extend the LAKR with entities found in the BabelNet [62] lexical resource as being semantically related to the entities in the utterance of the user. The Semantic Reasoning Framework implements a reasoning framework, combining native OWL2 reasoning and SPARQL rules to evaluate the received input in support of the semantic service selection performed by the AC to determine which slots of the DIP to be passed to the DMS are filled and with what content, as well as to assess whether the user should be referred to the FAQ for the answer(s) to their concern. It is a framework implemented by a combination of Java code, SPARQL rule sets, and machine learning algorithms. In particular, machine learning algorithms are used to classify and categorise text data into predefined categories or labels (i.e., topic detection) and assess sentence similarity for the identification of the FAQ relevance.

4.4.6. Semantic Service Computing (SSC)

Semantic Service Computing (SSC) is invoked by the AC. The AC launches a service request that contains all relevant information (speech act and facts) from the user's move (passed by the LAS to the KMS of the agent). The request is taken up by the iSeM matchmaker of the SSC [63], which retrieves the top-k relevant services from the LSR and passes them to the AC. If no relevant services have been identified, the SSC's service composition planner aims to dynamically compose a service. The planner works as an offline state-based

action planner [64]. Its initial state is a set of facts in OWL extracted by the KMS from the LAKR (its fact base). The initial state describes the current state of the interaction with the user. The goal state of the planner, created by the KMS, is a set of facts in OWL that persist after the execution of the action plan. Both the initial and the goal states are then converted into an action planning problem in the Planning Domain Definition Language (PDDL) by the OWL2PDDL converter of the service planner in the SSC called by the AC. The planner produces a list of sequential action plans, which are then assembled into a service plan and returned to the AC.

4.4.7. Dialogue Management Service (DMS)

As stated above, the Dialogue Management Service (DMS) is one of the two central modules that support interaction with the user. Specifically, the DMS operates on the slots of the Dialogue Input Package (DIP), which it receives from an agent according to the service that is to be executed. A DIP reflects the dialogue and system status up to the last user utterance for a specific service (cf. the diagram on the left in Figure 11 for the overall structure of the DIP). Each of its slots represents information to be requested from the user or provided to the user in the context of the service in question. For instance, *'informIntroductionEducation'*, *'obtainDegreeTitle'*, *'obtainDegreeCertificate'*, *'obtainSchool'*, *'obtainYear'*, *'obtainSchoolCompleted'*, *'obtainGrade'*, *'obtainIfAdditionalEducation'*, *and 'confirmEducationInformation'* are the slots of the DIP *'EducationInformation'* in the CV creation service. The diagram on the right in Figure 11 shows the rich meta-information properties of the individual slots of the DIPs.
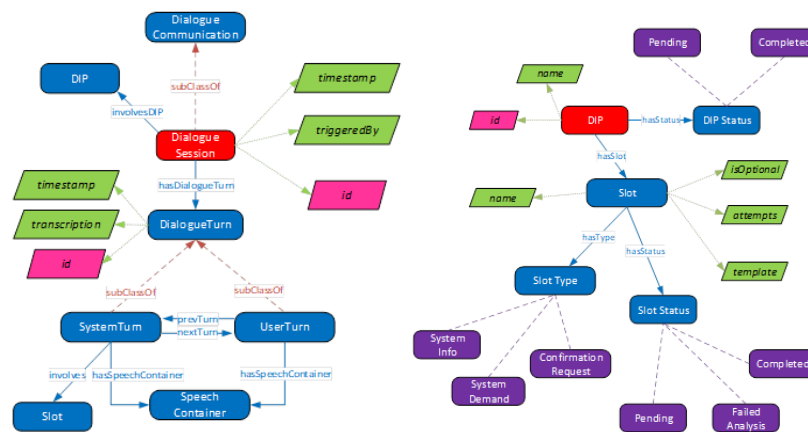

Figure 11. The structure of a DIP (**left**) and the properties of the slots of a DIP (**right**). 'DIP': Dialogue Input Package.

The central element of the DMS is the *policy* for mapping states to dialogue actions by which the DMS decides on the next communicative move of the agent. In this context, a dialogue action consists of selecting one or more slots of the current DIP and determining speech acts for their verbalization. The selected slots, their content as specified in the LAKR of the agent, and the determined speech acts are passed to the Natural Language Generation (NLG) module for verbalization. In addition to the semantic content encoded in terms of RDF triples, the DMS output includes references to pre-rendered sentence templates in the Content DB, a template ID for retrieving a template associated with a slot to be filled with RDF triples, and lexicalisations of entities included in the triples in various languages as provided by KMS. Figures 12 and 13 provide examples of a policy application for the same DIP in different states. Both the input and output of the DMS are serialized as JSON-LD messages.

Figure 14 displays the architecture of the DMS. A number of rule-driven service policies have been defined for DIP slot selection, pertinent LAKR content identification, and speech act determination. Each of the policies is encoded by a distinct Behavior Tree

that selects slots with the status 'pending' in the order that best fits the task-oriented dialogues within the designed services. In practical terms, this means that TCNs first receive an introductory statement that puts them into context (as, e.g., *"Now, let's proceed with your education. I will ask you about your school career. Please, start from the recent stage and then we can continue chronologically backwards item by item"*). Then they are requested to provide some information or confirm (in reaction to a yes/no question) the statement of the agent. For this purpose, the DMS selects the 'SystemInfo' slot, which is optionally followed by a 'SystemDemand' or 'ConfirmationRequest' slot.
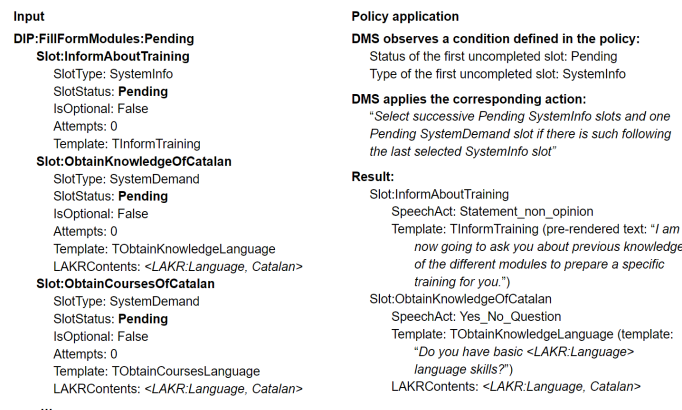

Figure 12. An example of the policy application for slots of the "Pending" status. 'DIP': Dialogue Input Package, 'DMS': Dialogue Management Service.
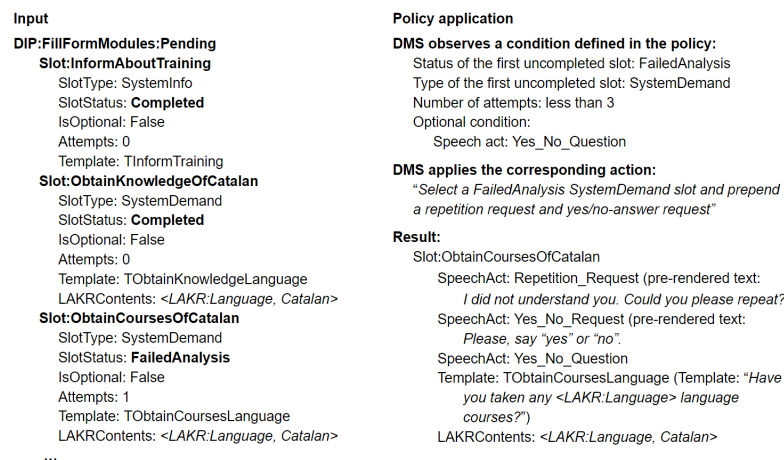

Figure 13. An example of the policy application for slots of the "Failed Analysis" status. 'DIP': Dialogue Input Package, 'DMS': Dialogue Management Service.
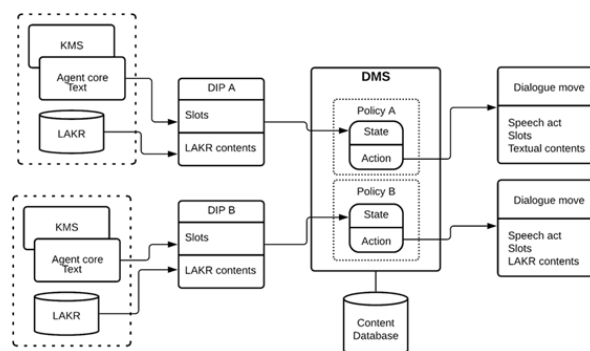

Figure 14. Architecture of the Dialogue Management Service (DMS). 'KMS': Knowledge Management Service, 'LAKR': Local Agent Knowledge Repository, 'DIP': Dialogue Input Package.

Predefined strategies are adopted to react to less foreseen situations; for instance, in the case of a failed analysis of a user statement by the ASR or LAS, the DMS selects a request for its repetition. After a maximum number of failed repetitions (as a rule, the maximum is set to three), the DMS selects the NLG template that informs the user about the interaction problem and advises them to consult with the relevant public administration or NGO personnel.

4.5. Cohorts of Personal Assistants

Some of the services covered by the Welcome Platform imply the interaction of several agents on behalf of their "Masters". Examples of such services are *Language Course Coordination* and *Co-habitation Search Coordination*. In the former, the agents of TCNs attending a language course collaborate with each other to help form groups of learners with similar language proficiency, while in the latter the agents of TCNs looking for shared housing spaces collaborate to find cohabitants with mutually matching expectations and preferences.

To accommodate this functionality, the AJAN of an agent (see Section 4.4.1) is complemented by a Multi(M)AJAN plugin module (in addition to SSC, KMS, etc.), which takes over the tasks of multiple agent coordination and communication. The MAJAN module consists of two submodules, MAJAN Plugin and MAJAN Web. MAJAN Plugin provides generic template SPARQL-BTs with nodes and SPARQL queries that are useful for three specific types of Multiagent Coordination protocols, namely, FIPA-Request-Interaction, Coalition-Structure-Generation, and Clustering, while MAJAN Web provides monitoring and evaluation functions; see [61] for further details.

5. MyWelcome Agents in Use

Figure 15 illustrates the respective pipelines of the modules involved in interaction between a TCN and their MyWelcome Agent via the MyWelcome Application frontend. The dark green bars mark the Dispatcher. In the case of a spoken user intervention, the MyWelcome Application sends an audio (*wav*) file to the Language Identification (LID) module via the Dispatcher, which provides confidence scores for the possible languages of the intervention. In extreme cases (e.g., a very noisy environment or highly accented speech), it is possible for all confidence scores to be zero. In this case, the Knowledge Management Service (KMS) of the agent receives an error code which activates the Dialogue Management Service (DMS). The DMS creates the representation of a request to repeat the intervention, which is verbalized by the Natural Language Generator (NLG). If the confidence scores are low, a request to confirm that the intervention is in the language with the highest score is generated; otherwise, the language tag of the language with the highest confidence score is assigned to the file and passed to the Dispatcher such that the latter can forward it to the corresponding Automatic Speech Recognition (ASR) module. In the case of a textual intervention, the MyWelcome Application identifies the language of the intervention. If the language is different from English, the transcription of the intervention is forwarded by the Dispatcher to the Machine Translation (MT) module. The English original/English translation is processed by the Language Disfluency Correction (LDC) module. The obtained grammatically fluid intervention is processed by the Language Analysis Service (LAS). The LAS produces a representation that is incorporated by the KMS of the agent into its Local Knowledge Repository. The agent creates a DIP to react to the move of the user. This DIP is filled by the KMS with information from the LAKR of the Agent and passed to the DMS. The DMS plans the verbalization of the reaction of the Agent. The output of the DMS is a semantic structure which is passed to the NLG for verbalization. If the NLG opts for the use of sentence templates instead of full-fledged generation, the corresponding templates are retrieved from the Content DB via the Content Management Service (CMS). If the conversation is not in English, the generated verbal move is translated; otherwise, it is passed directly to the TTS.
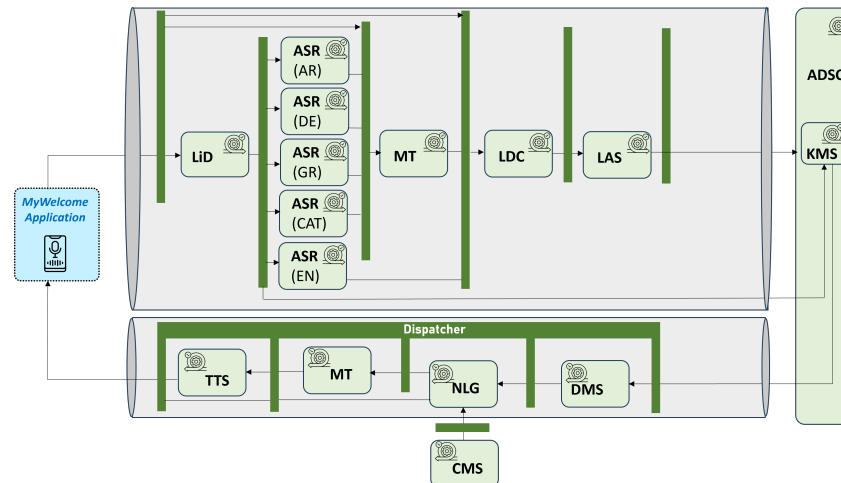
Figure 15. The pipeline of interaction processing in the Welcome Platform. 'LID': Language Identification, 'ASR': Automatic Speech Recognition, 'MT': Machine Translation, 'LDC': Language Disfluency Correction, 'LAS': Language Analysis Service, 'ADSC': Agent-Driven Service Coordination, 'KMS': Knowledge Management Service, 'DMS': Dialogue Management Service, 'NLG': Natural Language Generation, 'TTS': Text to Speech, 'CMS': Content Management Service, 'AR': Arabic, 'DE': German, 'GR': Greek, 'CAT': Catalan, 'EN': English.

If the agent takes the initiative and initiates a conversation, the flow starts with the creation of a DIP by the agent.

The two general pipelines (as sketched above) are applied in all three main tasks related to the support of TCNs: reception, integration, and social inclusion (cf. Table 1). In what follows, we present examples for each of the three tasks in terms of interaction patterns, as is common in agent-based interaction frameworks.

5.1. MyWelcome Agent-Supported Reception

MyWelcome Agent supported reception focuses on guiding the TCNs during the registration procedures that TCNs usually need to go through. This can be, e.g., First Reception upon entry into the host country (as in Catalonia) or the registration required for legal consultation (as in Greece). Because registration procedures are generally very similar as far as the actions of the agent are concerned, in what follows we focus on the First Reception Service.

The goal of the First Reception Service is to provide a general description of the reception procedure foreseen for TCNs in Catalonia, to inform TCNs that they must be registered in order to apply for the service, and to help registered TCNs fill out the reception form which allows them to apply for the service. To do this, the agent acquires the information that is usually collected by the reception officer through dialogue, then presents it to the TCN in terms of a form that can be edited and sent as a PDF file to a registered email account. Figure 16 shows the corresponding generalized interaction pattern between the agent and the TCN. The continuous (light blue) boxes denote the actions of the agent, while the dashed (light green) ones represent the actions of the TCN. Each action of the TCN may imply a subsequent clarification dialogue (marked by an arrow circle) initiated by the agent in case of a low confidence ASR output.
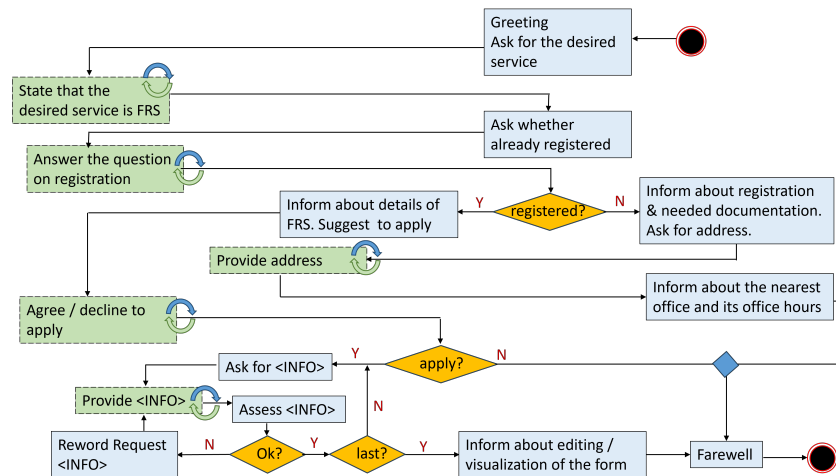
Figure 16. Generalized Agent–TCN interaction pattern during the First Reception Service (FRS). Here, '<INFO>' stands for personal information of the TCN, including biographical data, residence address, etc.

5.2. MyWelcome Agent-Supported Integration

Integration of TCNs into the host country is where the MyWelcome Agent comes most into play for information provision and coaching. For example, the information provided by the agent might concern the healthcare or schooling systems in a host country. In addition, any further topic that cannot be easily covered by the FAQs, for instance, because it is likely to lead to follow-up questions, is a candidate for coverage. Each topic can contain several more specific subtopics. For instance, in the context of schooling, TCNs may be interested in knowing about the holidays during the school year, the way the performance of students is graded, etc. According to experts, the most pressing question is the recognition of foreign diplomas in the host country. Figure 17 shows a fragment of the interaction pattern for information about school diploma recognition. This pattern forms part of a more generic pattern that involves interactions regarding the recognition of any kind of education diploma. The beginning of this interaction is skipped in the figure; we show only a part of the interaction pattern here due to the complexity of the full pattern.
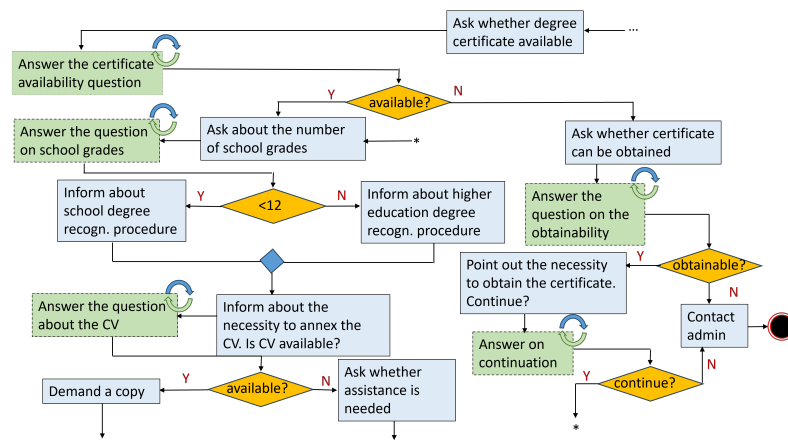

Figure 17. Fragment of the Agent–TCN interaction pattern for briefing on school diploma recognition in the host country.

The coaching of TCNs by the MyWelcome Agent in the context of integration currently focuses on the creation of the CV for job application and making appointments either with a job center or with legal services. A complete CV usually consists of several sections, including biographical data, education, languages spoken, job experience, etc. Thus,

extensive interaction patterns have been designed for completion of each of these sections. Figure 18 displays a general overview of these patterns.
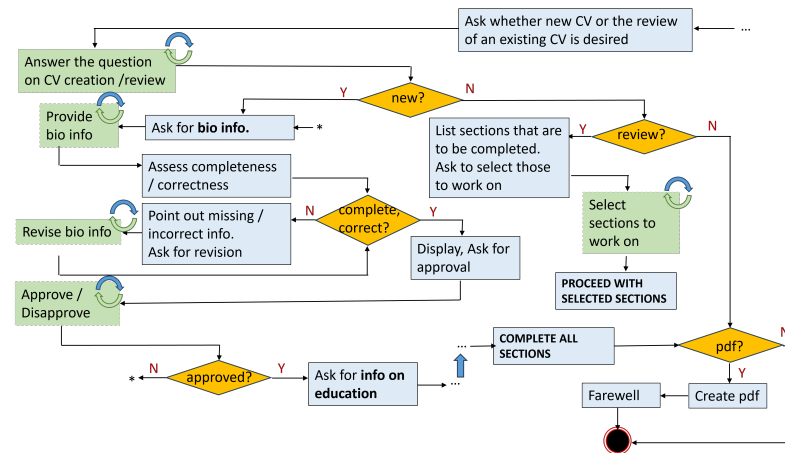

Figure 18. Agent–TCN interaction pattern for CV creation.

5.3. MyWelcome Agent-Supported Social Inclusion

A prominent example of the involvement of the MyWelcome Agent in social inclusion of TCNs is mediation in the context of the search for shared housing. In this case, collaboration among multiple MyWelcome Agents comes into play, as mentioned in Section 4.5.

Figure 19 displays the details of this coordination. A TCN who decides to initiate a search for cohabitants communicates his/her location preferences via the MyWelcome Application interface. The preferences are noted by the KMS of the TCN's agent and passed to the KBS, i.e., the global Knowledge Base System, which identifies the agents of other individuals who are interested in cohabiting in one or several of the same locations. The list of relevant agents is passed to the KMS of the TCN's agent, which solicits the WPM to wake them up. In the course of this interaction between the TCN's agent and the woken-up agents, which is managed by the MAJAN plug-in of the respective agents, personal data are shared and the users' compatibility in terms of the characteristics specified by the searching TCN of the "Masters" is determined. The list of compatible users is then passed to the MWA by the TCN's agent for inspection by the "Master".
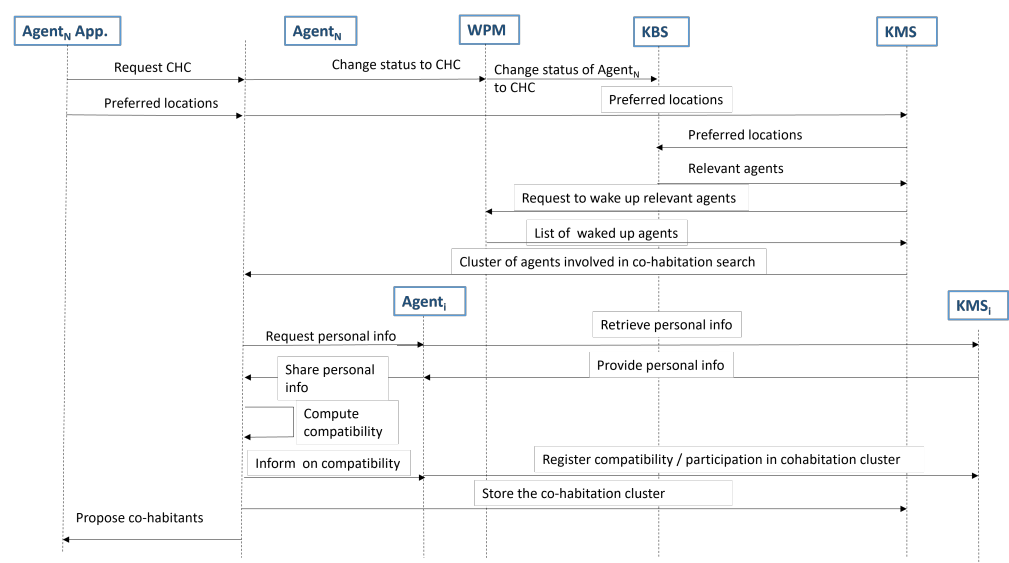

Figure 19. Cohabitation coordination between multiple agents. 'CHC': Cohabitation Coordination, 'WPM': Welcome Platform Management, 'KBS': Knowledge Base System, 'KMS': Knowledge Management Service.

6. MyWelcome VR in Use

As in the case of MyWelcome Application, MyWelcome VR draws upon several modules of the Welcome Platform. Figure 20 illustrates the pipeline of the modules involved in interaction between a TCN and the MyWelcome VR. As can be observed, this process concerns a number of the Platform's language technologies.
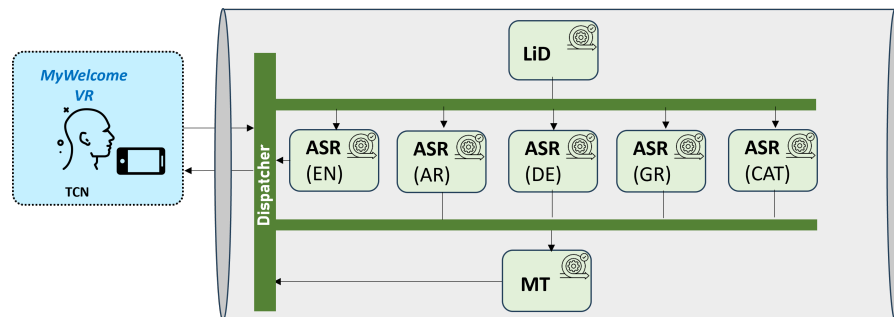

Figure 20. Pipeline of interaction with VR in the Welcome Platform. 'LID': Language Identification, 'ASR': Automatic Speech Recognition, 'MT': Machine Translation, 'EN': English, 'AR': Arabic, 'DE': German, 'GR': Greek, 'CAT': Catalan.

As indicated in Table 1, MyWelcome VR targets, first of all, TCN integration and social inclusion, supporting information provision, coaching, and training in accordance with the insight that active interaction with visually displayed information facilitates its comprehension [65].

6.1. VR-Supported Integration

VR-supported integration of TCNs in the Welcome Platform focuses on language learning and assistance with the incorporation of the TCN into the labour market. It is undisputed that VR can effectively support language learning [18]. Therefore, the Welcome Platform includes language learning aspects in its integration and social inclusion setups. One of these setups is a word spelling game in which a canon shoots balls with a letter written on each of them; the user must place the balls in the correct position in the sockets on a table or throw the ball into the bin to correctly spell a given word. Another is a vocabulary learning game in which the user is shown words in the host country language, the word translations into their mother tongue, and images, then is asked to align all of them. The language learning setup repertoire is complemented by training for interaction in formal language, as would be expected in a job interview, and social inclusion-related vocabulary learning (see below).

The need to support the incorporation of TCNs into the labour market is manifold, as finding employment is one of their primary concerns. TCNs may wonder, for instance, what clothes should be worn for a job interview, what the interviewee is expected to know about the company, how to act during the interview, and how formally one is expected to speak, all of which are concerns that vary from country to country.

MyWelcome VR turned out to be an ideal instrument for coaching TCNs on these matters. Below, we illustrate how this is realized using two example situations: a simulated job interview, and choosing the right outfit to wear for a job interview. In the context of interview simulation, the user appears in front of a virtual panel of interviewers (cf. Figure 21). The questions asked by the interviewers and the expected answers are retrieved from the professional online source "Your 2023 Guide to the Most Common Interview Questions and Answers" https://www.themuse.com/advice/interview-questions-and-answers (accessed on 12 February 2024) and enriched by TCN-relevant question–answer pairs. The answers provided by the TCNs are recorded to allow for later analysis and comparison with the expected answers by a human coach.

Figure 21. Simulated job interview setup.

To coach TCNs with respect to appropriate clothes for a job interview, the user is asked to choose items from the offered outfits (cf. Figure 22 for two examples) and dress a mannequin with the chosen items. The mannequin is inspected and feedback on its appearance is provided.

Figure 22. Selecting the right clothing for a job interview.

6.2. VR-Supported Social Inclusion

The Welcome Platform focuses on three broad aspects of social inclusion: presentation of public facilities such as public libraries and gyms, introduction to the regional geography of the host country and its most emblematic historical or cultural monuments, and transmission of cultural and social values in relation to gender-based violence and racism, gender discrimination, and education. Figure 23 illustrates the virtual setup for exploring public facilities.
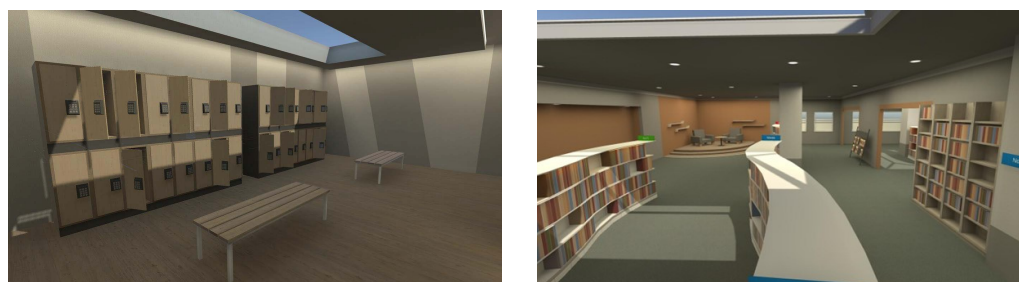

Figure 23. Presentation of public facilities: the locker room of a gym (**left**) and library reception (**right**).

Figure 24 features a regional geography VR mini-game for Germany as the host country. The user is asked to assign names to regions on the displayed map and locate the monuments.

As mentioned above, the Welcome Platform also uses VR to train TCNs on the local vocabulary in the host country's language. This can take place in a virtual *inclusion room* where the user can interact with panels located on the walls of the room, in dedicated language learning rooms, or in separate setups that feature the respective public facility (in the current release, library and public gym).

Figure 24. Regional geography game for learning the locations and names of the states in Germany.

7. User-Oriented Evaluation

In addition to the technical evaluation of the individual modules of the Welcome Platform, the MyWelcome Application and MyWelcome VR have been evaluated by the users after the release of the first, second, and third prototypes of the Platform. In what follows, we describe the setup of the evaluation, its results, and the lessons drawn from the results.

7.1. Evaluation Setup

The evaluation was carried out in trials with TCNs in the city of Hamm, Germany, in Greece, and in Catalonia, Spain. In Germany, the evaluation took place on the premises of Caritas; in Greece, it took place on the premises of PRAKSIS; and in the case of Catalonia, it took place on the premises of three Catalan municipality authorities with a high proportion of TCNs.

The participants in the trials were recruited by personal invitation. The invitations stressed that rejection or acceptance did not imply any administrative or legal consequences. Upon agreeing to participate in the trials, the participants were asked to sign a consent form and fill out a questionnaire with some basic personal information (e.g., age, education, marital status). While this information was not considered for the assessment or for judging the quality of the Welcome Platform services, it told us a lot about the profiles of the users who are interested in our current technologies and for whom we should seek to make them more attractive (see Section 7.3). Table 2 displays the number of participants at each site and for each prototype. Note that not all participants tested all of the offered services; due to the severe contact restrictions during the peak of the COVID-19 pandemic, the second prototype was tested by fewer TCNs.

The table highlights two features of the participants that appear to be especially relevant in this context, namely, average age and educational level.

Table 2. Number of participants in the evaluation trials of the first, second and third prototypes of the Welcome Platform services (in bold), their average age ('age'), and the percentage with university/college education ('he').

	# Participants		
	1st	**2nd**	**3rd**
Catalonia, Spain	**27** (age: 43; he: –)	**21** (age: 37; he: 90%	**30** (age: 33; he:62%)
Greece	**17** (age: 29; he: 41.8%)	**6** (age: 41; he: 50%	**27** (age: 40; he: 67%)
Hamm, Germany	**10** (age: 28.8; he: 30%)	**5** (age: 29; he: 75%)	**9** (age: 33; he: 45%)

During the trials, the participants were provided with a laptop or tablet to navigate in the Welcome Platform applications as well as with VR glasses. Assistance by personnel

was offered whenever a participant asked for it. After the trial, all participants filled out an evaluation form with a number of statements concerning central aspects of the services. For each of the statements, the users could mark 'strongly disagree', 'tend to disagree', 'neither agree nor disagree', 'tend to agree', or 'strongly agree'. For a more convenient presentation, we mapped this scale to the five-value Likert scale, with 'strongly disagree' as '1' and 'strongly agree' as '5'.

7.2. Evaluation Results

Figures 25 and 26 respectively show the progression of user satisfaction with the general features of the increasingly mature MyWelcome Application and MyWelcome VR across the three countries.
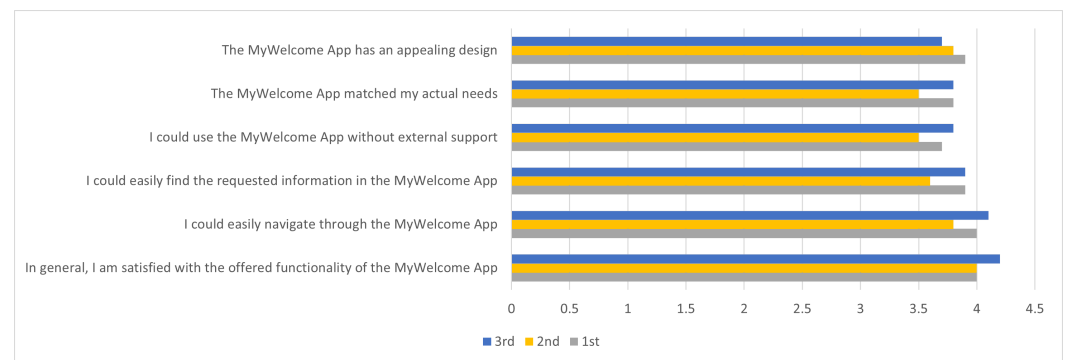

Figure 25. User satisfaction during evaluation of the increasingly mature MyWelcome Application. On the X axis, '0–4.5' represents the five-value Likert scale, with '1' being the worst and '5' the best; '3rd' stands for the third (and final) prototype of the Application, '2nd' for the second prototype, and '1st' for the first prototype.
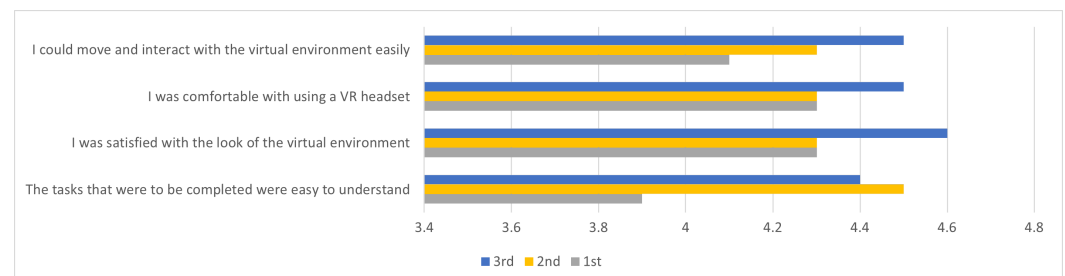

Figure 26. User satisfaction during evaluation of the increasingly mature MyWelcome VR. On the X axis, '3.4–4.8' represents the five-value Likert scale, with '1' being the worst and '5' the best; '3rd' stands for the third (and final) prototype of the Application, '2nd' for the second prototype, and '1st' for the first prototype.

It is interesting to observe that the users considered the design of the first prototype of the MyWelcome Application more appealing than that of the second prototype and the final release. Considering that the overall design did not change significantly, it can be hypothesized that the users were more permissive in their scoring of the first prototype due to the novelty of the application during first prototype trials. The second prototype of the application was in general evaluated somewhat more negatively. This can be explained by the deployment of additional unstable services compared to the first prototype, along with higher expectations of users who also participated in the first prototype trials. Overall, the third prototype of the MyWelcome Application was rated close to 4 on the five-value Likert scale, which we consider to be a good outcome.

In the case of the MyWelcome VR, a clear increase in user satisfaction can be seen with the third prototype compared to the first and second prototypes, with the exception of the clarity of the tasks. On this item, the second prototype was evaluated somewhat better

(4.5, as compared to 4.4 for the third prototype). This is most likely due to the significant increase in the services that the users had to evaluate and their complexity. The average rating of the third prototype was around 4.5 out of 5, with 5 being the best possible score.

In order to provide the reader with more detailed insight into the acceptance of the final versions of the MyWelcome Application and MyWelcome VR by the users, we present, in what follows, the evaluation of the third prototype in more detail.

The services in the context of which the MyWelcome App was evaluated included, among others, FAQs, domain-specific vocabulary lookup, and a number of agent-driven services, e.g., briefing on the healthcare system of the host country, training on booking an appointment with public administration, collaborative CV creation, etc.

Because FAQs are the most common service across migrant support applications (see Section 3), and as CV creation is the most complex and challenging service in which a single personal assistant is involved, we first display the evaluation outcome for these two services. With respect to the FAQs, the majority of users agreed ('tend to agree' or 'strongly agree') that the FAQ section provides useful information and that the provided information is sufficiently detailed (cf. Figure 27). In addition, all 25 users who rated this statement chose either 'tend to agree' (20) or 'strongly agree' (5) for the question on whether all relevant topics were covered. This assessment is especially valuable and encouraging to us since users of the Welcome Platform heavily depend on this information.
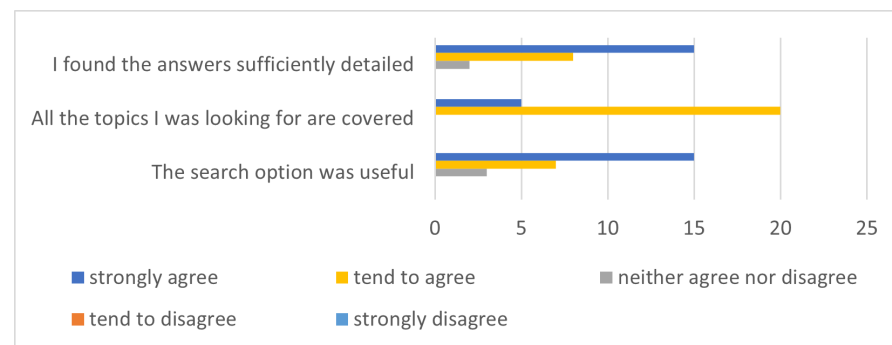

Figure 27. User evaluation of the FAQ section in the MyWelcome Application; on the X axis, '0–25' specifies the number of users.

Figure 28 displays the cumulative figures across the three countries obtained from the questionnaires on the collaborative CV creation. It shows that the users generally tended to agree that the agent provided by the MyWelcome Application is a usable and useful instrument for CV creation. The only characteristic of the agent that the majority of users did not appreciate was its slow speed, which also negatively influenced the entire process of CV creation. One user was also not satisfied with the resulting CV, while 12 users chose either 'tend to agree' or 'strongly agree' for the question on whether the created CV was to their satisfaction. However, it should be noted that 6 out of the 19 users who rated this statement selected 'neither agree nor disagree'. Further work with users is needed in order to identify and address the reason for this neutral stance.

MyWelcome VR was similarly evaluated in a series of different services that addressed vocabulary learning (in a number of setups), learning the host country's geography, job interview training, and job interview appearance coaching, among others. The outcomes of the evaluations of one of the vocabulary learning setups (vocabulary spelling) and of the job interview training (cf. Figure 21) are considered as representative for the VR-driven services.

The VR-based word spelling game was evaluated in Greece and Germany. Figure 29 displays the outcome of the evaluations. It can be seen that the word spelling service was very well received by the users.

The job interview training service was tested in Catalonia, Spain and in the city of Hamm, Germany. Figure 30 displays the outcome of the evaluation. The figure shows that this service was not rated as well as the word spelling service, but better than the

MyWelcome Application-based CV creation coaching service. One of the ten users felt that the service was not easy to use, and the same user felt nervous during the exercise. More detailed feedback would be needed in order to address these issues.
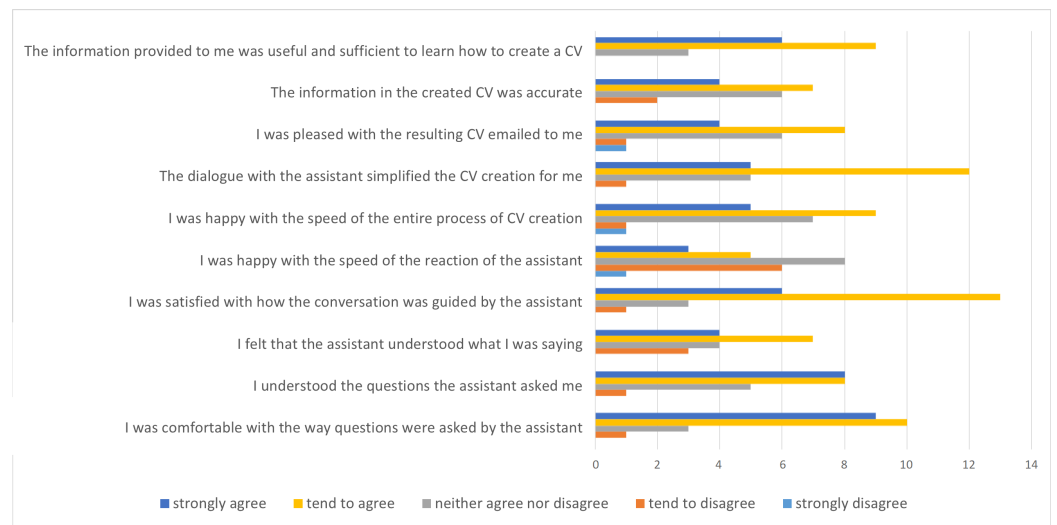

Figure 28. User evaluation of agent-supported CV creation via the MyWelcome Application; on the X axis, '0–14' indicates the number of users.
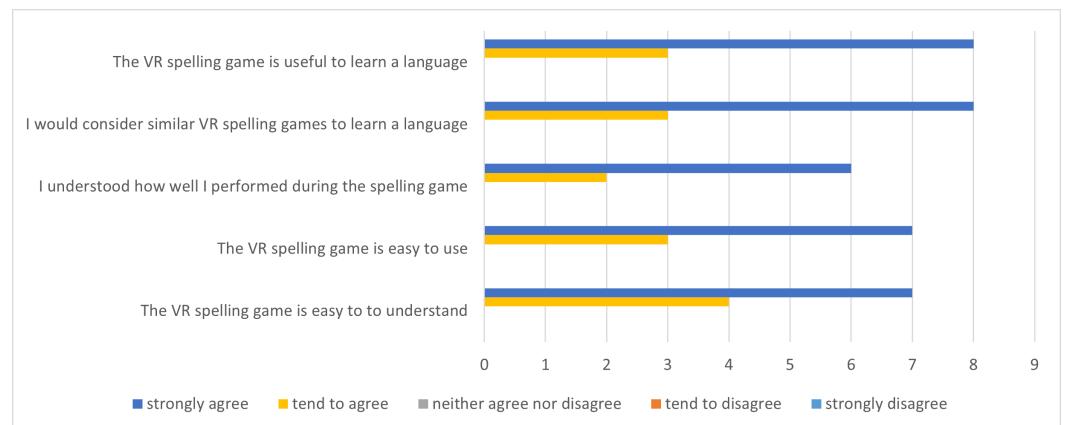

Figure 29. User evaluation of the word spelling exercise in VR; on the X axis, '0–9' marks the number of users.
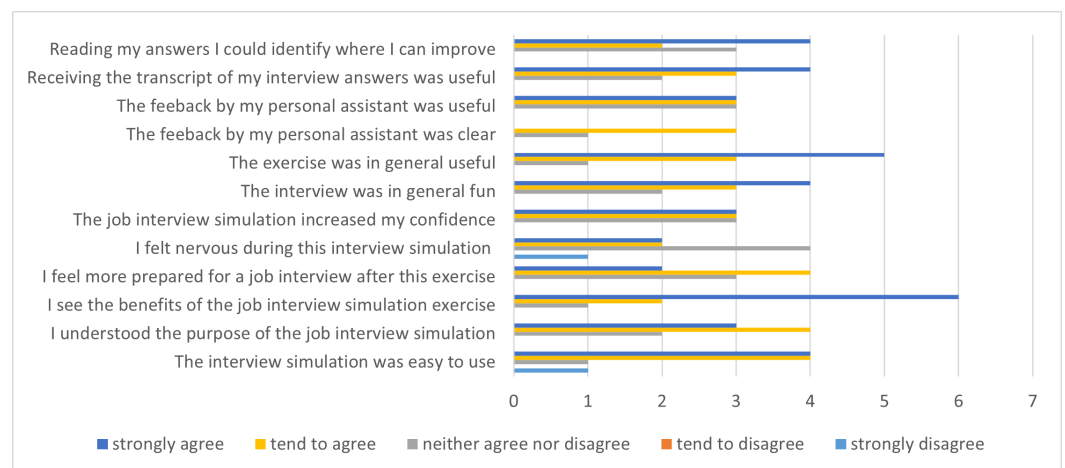

Figure 30. User evaluation of job interview training in VR; on the X axis, '0–7' denotes the number of users.

7.3. Discussion

The trials demonstrated that the services offered by the Welcome Platform are useful and were well-received by the TCNs. The results show that the Platform addresses the concerns that TCNs rated to be of the highest priority, including assistance in job-seeking, education (especially host country language learning), and information about health and schooling services in the host country.

In particular, the results highlight the success of VR-based techniques among all participants of the trials independently of their age and education level, with the VR-based services oriented towards language learning and social integration enjoying especially great popularity. This confirms the outcome of recent studies arguing that VR is apt to increase the motivation and performance of users (cf. [66,67], among others). No participants complained about nausea, dizziness, or other conditions that have been observed in connection with the use of VR headsets. Nonetheless, it must be acknowledged that these conditions are rather common, which limits the range of potential users of VR-based services. Another potential drawback of VR-based services is that due to the relatively high price of the headsets, they can be expected to be used only on the premises of NGOs or migrant integration authorities.

Apart from the speed of the Application, a number of services should be further improved or extended, among them, the job interview training and CV creation processes. In particular, job interview training would be considerably more effective if the user could receive direct feedback on their interaction with the interviewer panel, e.g., how an interviewer was addressed, how a question was answered, etc. However, this would require substantial advances in state-of-the-art natural language processing technologies, which is beyond the scope of the present work.

The CV creation service proved to be very useful, but was also complex and time-consuming, and should be redesigned to make it more attractive to users. Further studies are needed to obtain an optimal design.

In addition to the improvement and extension of some individual services, attention should also be paid to a more basic fact. As can be observed in Table 2, a significant number of participants in the evaluation trials had a higher education degree, and many of them were below 30 years of age (the average age in all three trials oscillated between 35.8 and 36.5). This means that while the use of new technologies was perceived as attractive by the younger TCNs, our assessment did not reach TCNs beyond 45 years of age on a large scale. Because these TCNs equally frequently use mobile devices, and as such are potential users of Welcome Platform services, there is a challenge in making the Platform's services appealing to them as well.

Overall, the significantly diverging number of participants in the three host countries, their limited age range, and the dominance of participants with higher education can be considered as limitations of the evaluation trials carried out thus far. However, even with this limitation, the feedback received during the trials leaves no doubt as to the Welcome Platform's usefulness.

8. Conclusions

The continuously high number of migrants and refugees is increasingly overstraining the services provided by public administrators and NGOs in host countries. One way to address this challenge is the use of intelligent technologies that can assist Third-Country Nationals (TCNs) with their needs. A number of proposals have been made in the literature to facilitate more efficient language learning by TCNs. More recently, social integration and support in administrative procedures have also been addressed. However, TCNs need assistance at all stages of their life in their new home country. In this paper, we present the prototype of a platform that integrates an intelligent personal assistant and VR technology to assist TCNs during first reception procedures, integration, and everyday challenges in selected host countries (Catalonia in Spain, Germany, and Greece). The personal assistant technologies are grounded in advanced knowledge-based agent planning techniques.

Unlike previous proposals, in which interaction with the user is dealt with entirely through a dialogue management module, we separate the knowledge-driven and language-agnostic tasks involved in interaction from the language discourse planning tasks. The language technologies that form part of the personal assistant are also used for VR. This makes VR accessible by users with different linguistic backgrounds.

The Welcome Platform uses state-of-the-art deep machine learning models for language translation and language analysis. In preliminary experiments that we carried out with deep learning-based models in the context of the Welcome Platform, we noticed hallucinations such as providing the incorrect address and opening hours of a police office for registration. To avoid any risk of hallucination, from which deep learning-based agent and dialogue models still suffer, we relied on classical machine learning rather than on neural machine learning techniques for these tasks. It will be a matter of future work to explore the use of deep learning for dynamic acquisition of interaction patterns and behavior trees and for dynamic construction of DIPs to ensure more flexible discourse planning by the dialogue management module. Another important topic for future research is cross-dialogue act reference resolution, which would allow both the user and the agent to make references to what the other party has said in the past.

For use in practice, the issues identified in Section 7.3 should be addressed. The individual modules of the Platform need to mature further, and if required extended to cover additional languages and additional needs of TCNs. Thanks to the modularity of the Platform, these adaptations can be realized with limited effort.

While the Welcome Platform focuses on the areas we identified as high-priority needs of the TCNs in our three host countries, it can be extended by services that address other needs, such as information on child care, traditions and holidays in the host country, cultural event calendars, etc. Furthermore, it can be extended to cover the needs of TCNs in other host countries.

Finally, the Welcome Platform could also be adapted to address the needs of other target groups beyond TCNs, such as elderly or physically impaired people.

Author Contributions: L.W.: Conceptual design of the Platform, coordination of the design and implementation of the Platform, coordination of the design and realization of user trials and system evaluation, writing of the paper, funding acquisition; D.B. and O.S.: Conceptual design of the Platform, development and integration of the Platform; M.B.: Design of the language learning services; E.C., J.d.L., E.D., I.F., C.G. and M.M.: Design of the user services, design of the evaluation trials, realization of the evaluation trials; J.Č. and E.E.: Design and realization of the automatic speech recognition and language identification components; T.C. and I.R.: Realization of evaluation trials; J.D. and Y.S.: Realization of the virtual reality services of the Platform; S.D.: Design and development of the MyWelcome Application and MyWelcome VR front-ends, coordination of the development of user services, realization of the language learning services; J.G. and A.S.: Conceptual design of the Platform, design and realization of the dialogue management, language analysis, language generation, and text-to-speech modules; E.J.-R. and M.K.: Conceptual design of the Platform, design and realization of the agent-related modules; A.M., D.N., G.T. and S.V.: Conceptual design of the Platform, design and realization of the knowledge acquisition, management, and expansion services, design and realization of the language disfluency and correction module; A.N. and M.R.: Design and coordination of user trials and their evaluation. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The complete anonymized evaluation questionnaires are available upon request. Personal data of the users and the datasets collected during the Project are subject to data privacy restrictions.

Abbreviations

The following abbreviations are used in this manuscript:

AC:	Agent Core
ADSC:	Agent-Driven Service Coordination
ASR:	Automatic Speech Recognition (module)
BT:	Behavior Tree
CA:	Conversational Agent
CALL:	Computer-Assisted Language Learning
CDB:	Content Database
CHC	Cohabitation Coordination
CMS:	Content Management Service
KMS:	Knowledge Management Service
DIP:	Dialogue Input Package
DLA:	Deep Language Analysis
DMS:	Dialogue Management Service
ECA:	Embodied Conversational Agent
FAQ:	Frequently-Asked Question
GB:	Gender-Based (violence)
GKB:	Global Knowledge Base
KMS:	Knowledge Management Service
KBS:	Knowledge Base System
LAR:	Local Agent Repository
LAKR:	Local Agent Knowledge Repository
LSR:	Local Service Repository
LAS:	Language Analysis Service
LDC:	Language Disfluency Correction (module)
LID:	Language Identification (module)
NMT:	Neural Machine Translation
MT:	Machine Translation (module)
MWA:	MyWelcome Application
MWVR:	MyWelcome Virtual Reality
NER:	Named Entity Recognition
NLG:	Natural Language Generation (module)
SLA:	Surface Language Analysis
SSC:	Semantic Service Computing
TCN:	Third-Country National
TP:	Teacher Panel
TTS:	Text-to-Speech (module)
UI:	User Interface
VAC:	Visual Analytics Component
VPA:	Virtual Personal Assistant
VR:	Virtual Reality
WAR:	Welcome Agent Repository
WDKR:	Welcome Domain Knowledge Repository
WER:	Word Error Rate
WPM:	Welcome Platform Manager
WSR:	Welcome Service Repository
WSD:	Word–Sense Disambiguation

References

1. Lelis, A.; Vretos, N.; Daras, P. NADINE-BOT: An Open Domain Migrant Integration Administrative Agent. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW), London, UK, 6–10 July 2020.
2. Ntioudis, D.; Kamateri, E.; Meditskos, G.; Karakostas, A.; Hubery, F.; Bratskaz, R.; Vrochidis, S.; Akhgar, B.; Kompatsiaris, I. IMMERSE: A Personalized System Addressing the Challenges of Migrant Integration. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW), London, UK, 6–10 July 2020.
3. Noennig, J.R.; Cserpes, B.; Ceola, F.; Barski, J.; Brandenburger, K.M.; Malchow, M. The Migrant Integration Platform MICADO—A Tool for Social Integration and Cohesion. In Proceedings of the International Forum on Digital and Democracy, Rome, Italy, 17–18 November 2022.
4. Wanner, L.; Klusch, M.; Mavropoulos, A.; Jamin, E.; Puchades, V.M.; Casamayor, G.; Çernocký, J.; Davey, S.; Domínguez, M.; Egorova, E.; et al. Towards a Versatile Intelligent Conversational Agent as Personal Assistant for Migrants. In *Proceedings of the Advances in Practical Applications of Agents, Multi-Agent Systems, and Social Good*; The PAAMS Collection; PAAMS 2021; Lecture Notes in Computer Science; Dignum, F., Corchado, J., De La Prieta, F., Eds.; Springer: Cham, Switzerland, 2021; Volume 12946, pp. 316–327.
5. Zhang, Q.; Chen, S.; Xu, D.; Cao, Q.; Chen, X.; Cohn, T.; Fang, M. A Survey for Efficient Open Domain Question Answering. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, ON, Canada, 9–14 July 2023; pp. 284–295.
6. Zaib, M.; Zhang, W.E.; Sheng, Q.Z.; Mahmood, A.; Zhang, Y. Conversational question answering: A survey. *Knowl. Inf. Syst.* **2022**, *64*, 3151–3195. [CrossRef]
7. Passmore, J.; Tee, D. Can Chatbots like GPT-4 replace human coaches: Issues and dilemmas for the coaching profession, coaching clients and for organisations. *Coach. Psychol.* **2023**, *19*, 47–54.
8. Babar, P.P.; Barry, M.; Peiris, R. Understanding Job Coaches' Perspectives on Using Virtual Reality as a Job Training Tool for Training People with Intellectual Disabilities. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–7.
9. Norouzi, N.; Kim, K.; Bruder, G.; Erickson, A.; Choudhary, Z.; Li, Y.; Welch, G. A Systematic Literature Review of Embodied Augmented Reality Agents in Head-Mounted Display Environments. In *Proceedings of the ICAT-EGVE 2020—International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, 2–4 December 2020*; Argelaguet, F., McMahan, R., Sugimoto, M., Eds.; The Eurographics Association: Eindhoven, The Netherlands, 2020. [CrossRef]
10. Provoost, S.; Lau, H.M.; Ruwaard, J.; Riper, H. Embodied Conversational Agents in Clinical Psychology: A Scoping Review. *J. Med. Internet Res.* **2017**, *19*, e151. [CrossRef] [PubMed]
11. Schouten, D.G.; Deneka, A.A.; Theune, M.; Neerincx, M.A.; Cremers, A.H. An embodied conversational agent coach to support societal participation learning by low-literate users. *Univ. Access Inf. Soc.* **2023**, *22*, 1215–1241. [CrossRef]
12. Beach, D.S. *Personnel: The Management of People at Work*; Macmillan: Toronto, ON, Canada, 1985.
13. Petrović, J.; Jovanović, M. Conversational Agents for Learning Foreign Languages—A Survey. In Proceedings of the Sinteza 2020—International Scientific Conference on Information Technology and Data Related Research, Belgrade, Serbia, 17 October 2020.
14. Xiao, F.; Zhao, P.; Sha, H.; Yang, D.; Warschauer, M. Conversational agents in language learning. *J. China Comput. Assist. Lang. Learn.* **2023**, *3*. [CrossRef]
15. Lin, T.J.; Lan, Y.J. Language learning in virtual reality environments: Past, present, and future. *Educ. Technol. Soc.* **2015**, *18*, 486–497.
16. Parmaxi, A. Virtual reality in language learning: A systematic review and implications for research and practice. *Virtual Interact. Learn. Environ.* **2020**, *31*, 172–184. [CrossRef]
17. Peixoto, B.; Pinto, R.; Melo, M. andCabral, L.; Bessa, M. Immersive Virtual Reality for Foreign Language Education: A PRISMA systematic review. *IEEE Access* **2021**, *9*, 48952–48962. [CrossRef]
18. Hua, C.; Wang, J. Virtual reality-assisted language learning: A follow-up review (2018–2022). *Front. Psychol.* **2023**, *14*. [CrossRef]
19. Chen, B.; Wang, Y.; Wang, L. The Effects of Virtual Reality-Assisted Language Learning:A Meta-Analysis. *Sustainability* **2022**, *14*, 3147. [CrossRef]
20. Díaz Andrade, A.; Doolin, B. Information and Communication Technology and the Social Inclusion of Refugees. *MIS Q.* **2016**, *40*, 405–416. [CrossRef]
21. Nagbal, R.; Fatty, S.; Brizan, D. A Survey of Conversational Styles and Systems. In Proceedings of the 16th International Conference on Human System Interaction (HSI), Paris, France, 8–11 July 2024.
22. Reimann, M.M.; Kunneman, F.A.; Oertel, C.; Hindriks, K.V. A Survey on Dialogue Management in Human-robot Interaction. *J. Hum.-Robot Interact.* **2024**, *13*, 22. [CrossRef]
23. Kheddar, H.; Hemis, M.; Himeur, Y. Automatic speech recognition using advanced deep learning approaches: A survey. *Inf. Fusion* **2024**, *109*, 102422. [CrossRef]
24. Lenci, A. Understanding natural language understanding systems. *Sist. Intell.* **2023**, *32*, 277–302.
25. Wang, J.; Zhang, C.; Zhang, D.; Tong, H.; Yan, C.; Jiang, C. A recent survey on controllable text generation: A causal perspective. *Fundam. Res.* **2024**, *in press*. : 10.1016/j.fmre.2024.01.001 [CrossRef]
26. Mohsen, M.A.; Althebi, S.; Alsagour, R.; Alsalem, A.; Almudawi, A.; Alshahrani, A. Forty-two years of computer-assisted language learning research: A scientometric study of hotspot research and trending issues. *ReCALL* **2024**, *36*, 230–249. [CrossRef]

27. Dunmoye, I.D.; Rukangu, A.; May, D.; Das, R.P. An exploratory study of social presence and cognitive engagement association in a collaborative virtual reality learning environment. *Comput. Educ. X Real.* **2024**, *4*, 100054. : 10.1016/j.cexr.2024.100054 [CrossRef]

28. Han, E.; Bailenson, J.N. Social Interaction in VR. In *Oxford Research Encyclopedias, Communication*; Oxford University Press: Oxford, UK, 2024.

29. Codagnone, C.; Kluzer, S. *ICT for the Social and Economic Integration of Migrants into Europe*; Technical Report; Joint Research Centre, Institute for Prospective Technological Studies: Seville, Spain, 2011.

30. Reichel, D.; Siegel, M.; Andreo, J.C. *ICT for the Employability and Integration of Immigrants in the European Union*; Technical Report; Joint Research Centre, Institute for Prospective Technological Studies: Seville, Spain, 2015.

31. Leligou, H.C.; Anastasopoulos, D.; Vretos, N.; Solachidis, V.; Kantor, E.; Plašilová, I.; Girardet, E.; Montagna, A.; Vlahaki, F.; et al. Experiences and Lessons Learnt from the Evaluation of ICT Tools for and with Migrants. *Soc. Sci.* **2021**, *10*, 344. [CrossRef]

32. Regina, P.; De Capitani, E. Digital Innovation and Migrants' Integration: Notes on EU Institutional and Legal Perspectives and Criticalities. *Soc. Sci.* **2022**, *11*, 144. [CrossRef]

33. Akhgar, B.; Hough, K.L.; Samad, Y.A.; Bayerl, P.S.; Karakostas, A. *Information and Communications Technology in Support of Migration*; Springer: Berlin/Heidelberg, Germany, 2022.

34. Drydakis, N. Mobile applications aiming to facilitate immigrants' societal integration and overall level of integration, health and mental health. Does artificial intelligence enhance outcomes? *Comput. Hum. Behav.* **2021**, *117*, 106661. [CrossRef]

35. Bradley, L.; Berbyuk Lindström, N.; Sofkova Hashemi, S. Integration and Language Learning of Newly Arrived Migrants Using Mobile Technology. *J. Interact. Media Educ.* **2017**, *1*, 1–9. [CrossRef]

36. Kirya, M.; Debattista, K.; Chalmers, A. Using virtual environments to facilitate refugee integration in third countries. *Virtual Real.* **2023**, *27*, 97–107. [CrossRef]

37. Kukulska-Hulme, A.; Gaved, M.; Jones, A.; Norris, L.; Peasgood, A. Mobile language learning experiences for migrants beyond the classroom. In *He Linguistic Integration of Adult Migrants/L'Intégration Linguistique des Migrants Adultes: Some Lessons from Research/Les Enseignements de la Recherche*; Beacco, J.C., Krumm, H.J., Little, D., Thalgott, P., Eds.; De Gruyter Mouton: Berlin, Germany; Boston, MA, USA, 2017; pp. 219–224.

38. Kukulska-Hulme, A.; Gaved, M.; Paletta, L.; Scanlon, E.; Jones, A.; Brasher, A. Mobile Incidental Learning to Support the Inclusion of Recent Immigrants. *Ubiquitous Learn. Int. J.* **2015**, *7*, 9–21. [CrossRef]

39. Jones, A.; Kukulska-Hulme, A.; Norris, L.; Gaved, M.; Scanlon, E.; Jones, J.; Brasher, A. Supporting immigrant language learning on smartphones: A field trial. *Stud. Educ. Adults* **2017**, *49*, 228–252. [CrossRef]

40. Ngan, H.; Lifanova, A.; Jarke, J.; Broer, J. Refugees Welcome: Supporting Informal Language Learning and Integration with a Gamified Mobile Application. In *Proceedings of the Adaptive and Adaptable Learning*; EC-TEL 2016; Lecture Notes in Computer Science; Verbert, K., Sharples, M., Klobučar, T., Eds.; Springer: Cham, Switzerland, 2016; Volume 9891.

41. Chen, Z.; Lu, Y.; Nieminen, M.P.; Lucero, A. Creating a Chatbot for and with Migrants: Chatbot Personality Drives Co-Design Activities. In Proceedings of the ACM Designing Interactive Systems Conference, Eindhoven, The Netherlands, 6–10 July 2020; pp. 219–230.

42. Chlasta, K.; Sochaczewski, P.; Grabowska, I.; Jastrzębowska, A. MyMigrationBot: A Cloud-based Facebook Social Chatbot for Migrant Populations. In Proceedings of the 1st Workshop on Personalization and Recommender Systems. Co-located with the 17th Conference on Computer Science and Intelligence Systems, Seattle, WA, USA, 18–23 July 2022, pp. 51–59.

43. Martınez, D.; Plchot, O.; Burget, L.; Glembek, O.; Matějka, P. Language Recognition in iVectors Space. In Proceedings of the INTERSPEECH, Florence, Italy, 28–31 August 2011.

44. Dehak, N.; Kenny, P.; Dehak, R.; Dumouchel, P.; Ouellet, P. Front–end factor analysis for speaker verification. *IEEE Trans. Audio Speech Lang. Process.* **2011**, *19*, 788–798. [CrossRef]

45. Fér, R.; Matějka, P.; Grezl, F.; Plchot, O.; Veselý, K.; Černocký, J. Multilingually Trained Bottleneck Features in Spoken Language Recognition. *Comput. Speech Lang.* **2017**, *46*, 252–267. [CrossRef]

46. Schneider, S.; Baevski, A.; Collobert, R.; Auli, M. wav2vec: Unsupervised Pre-training for Speech Recognition. In Proceedings of the INTERSPEECH, Graz, Austria, 15–19 September 2019.

47. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017, pp. 6000–6010.

48. Stubbe, M.; Holmes, J. You Know, eh and Other "Exasperating Expressions": An Analysis of Social and Stylistic Variation in the Use of Pragmatic Devices in a Sample of New Zealand English. *Lang. Commun.* **1995**, *15*, 63–88. [CrossRef]

49. Passali, T.; Mavropoulos, T.; Tsoumakas, G.; Meditskos, G.; Vrochidis, S. LARD: Large-scale Artificial Disfluency Generation. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, Marseille, France, 20–25 June 2022; pp. 2327–2336.

50. Shvets, A.; Wanner, L. Concept Extraction Using Pointer-Generator Networks and Distant Supervision for Data Augmentation. In Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management (EKAW 2020), Bolzano, Italy, 26–29 September 2020; pp. 120–135.

51. Bohnet, B.; Wanner, L. Open Soucre Graph Transducer Interpreter and Grammar Development Environment. In Proceedings of the Language Resources and Evaluation Conference (LREC), Valletta, Malta, 17–23 May 2010; pp. 211–218.

52. Kale, M.; Rastogi, A. Template Guided Text Generation for Task-Oriented Dialogue. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Virtual, 16–20 November 2020; pp. 6505–6520. [CrossRef]

53. Liang, Z.; Hu, H.; Xu, C.; Miao, J.; He, Y.; Chen, Y.; Geng, X.; Liang, F.; Jiang, D. Learning Neural Templates for Recommender Dialogue System. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*; Moens, M.F., Huang, X., Specia, L., Yih, S.W.t., Eds.; Online and Punta Cana, Dominican Republic; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 7821–7833. [CrossRef]
54. Sun, Q.; Xu, C.; Hu, H.; Wang, Y.; Miao, J.; Geng, X.; Chen, Y.; Xu, F.; Jiang, D. Stylized Knowledge-Grounded Dialogue Generation via Disentangled Template Rewriting. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Seattle, WA, USA, 10–15 July 2022; pp. 3304–3318. [CrossRef]
55. Wang, H.; Cui, M.; Zhou, Z.; Wong, K.F. TopicRefine: Joint Topic Prediction and Dialogue Response Generation for Multi-turn End-to-End Dialogue System. In Proceedings of the 5th International Conference on Natural Language and Speech Processing (ICNLSP 2022), Trento, Italy, 16–17 December 2022; pp. 19–29.
56. Sun, Q.; Wang, Y.; Xu, C.; Zheng, K.; Yang, Y.; Hu, H.; Xu, F.; Zhang, J.; Geng, X.; Jiang, D. Multimodal Dialogue Response Generation. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland, 22–27 May 2022; pp. 2854–2866. [CrossRef]
57. Ahmad, Z.; Ekbal, A.; Sengupta, S.; Bhattacharyya, P. Neural response generation for task completion using conversational knowledge graph. *PLoS ONE* **2023**, *198*, e0269856. [CrossRef]
58. Mille, S.; Dasiopoulou, S.; Wanner, L. A Portable Grammar-Based NLG System for Verbalization of Structured Data. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, Limassol, Cyprus, 8–12 April 2019; pp. 1054–1056. [CrossRef]
59. Mel'čuk, I. *Dependency Syntax: Theory and Practice*; SUNY Press: Albany, NY, USA, 1988.
60. Dai, Y.; Yu, H.; Jiang, Y.; Tang, C.; Li, Y.; Sun, J. A Survey on Dialog Management: Recent Advances and Challenges. *arXiv* **2021**, arXiv:2005.02233.
61. Antakli, A.; Kazimov, A.; Spieldenner, D.; Jaramillo Rojas, G.; Zinnikus, I.; Klusch, M. AJAN: An Engineering Framework for Semantic Web-Enabled Agents and Multi-Agent Systems. In Proceedings of the PAAMS Conference, Guimarães, Portugal, 12–14 July 2023; pp. 15–27.
62. Navigli, R.; Ponzetto, S.P. BabelNet: Building a very large multilingual semantic network. In Proceedings of the 48th annual meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; pp. 216–225.
63. Klusch, M.; Kapahnke, P. The iSeM Matchmaker: A Flexible Approach for Adaptive Hybrid Semantic Service Selection. *Web Semant.* **2012**, *15*, 1–14. [CrossRef]
64. Helmert, M.; Röger, G.; Karpas, E. Fast Downward Stone Soup: A Baseline for Building Planning Portfolios. In Proceedings of the ICAPS WS on Planning and Learning, Freiburg, Germany, 11–16 June 2011; pp. 28–35.
65. Allman, S.A.; Cordy, J.; Hall, J.P.; Kleanthous, V.; Lander, E.R. Exploring the Perception of Additional Information Content in 360° 3D VR Video for Teaching and Learning. *Virtual Worlds* **2022**, *1*, 1–17. [CrossRef]
66. Ozgun, O.; Sadık, O. Implementation of VR Technologies in Language Learning Settings: A Systematic Literature Review. *Educ. Policy Anal. Strateg. Res.* **2023**, *18*, 32–61. [CrossRef]
67. Dick, E. *Current and Potential Uses of AR/VR for Equity and Inclusion*; Technical Report; Information Technology & Innovation Foundation: Washington, DC, USA, 2021.