Integrating Knowledge into Conversational Agents for Worker Upskilling

Rubén Alonso, Danilo Dessí, Antonello Meloni, Marco Murgia, and Reforgiato Recupero Diego

1 Introduction

The supply and demand for labor, wherein employees give the supply and employers provide the demand, are referred to as the labor market or job market. They are a crucial part of every economy and are closely connected to the markets for capital, goods, and services. According to a recent International Labour Organization (ILO)

R. Alonso (✉)
R2M Solution s.r.l., Pavia, Italy

Programa de Doctorado, Centro de Automática y Robótica, Universidad Politécnica de Madrid-CSIC, Madrid, Spain
e-mail: ruben.alonso@r2msolution.com

D. Dessí
Knowledge Technologies for the Social Sciences Department, GESIS – Leibniz Institute for the Social Sciences, Cologne, Germany
e-mail: danilo.dessi@gesis.org

A. Meloni · M. Murgia
Mathematics and Computer Science Department, University of Cagliari, Cagliari, Italy
e-mail: antonello.meloni@unica.it; m.murgia98@studenti.unica.it

R. R. Diego
Mathematics and Computer Science Department, University of Cagliari, Cagliari, Italy

R2M Solution s.r.l., Pavia, Italy
e-mail: diego.reforgiato@unica.it; diego.reforgiato@r2msolution.com

J. Soldatos (ed.), *Artificial Intelligence in Manufacturing*,
https://doi.org/10.1007/978-3-031-46452-2_15

assessment[1], the present global economic slump is likely to drive more employees to take poorer quality, poorly paid employment that lack job security and social safety, exasperating disparities that have already been made worse by the COVID-19 crisis. According to the same report, only 1.0% of the world's employment is expected to expand in 2023, which is less than half of the rate in 2022. The number of unemployed people worldwide is projected to increase by almost 3 million to 208 million in 2023, representing a 5.8% global unemployment rate. The limited labor supply in high-income nations is partly to blame for the moderate extent of this predicted increase. This would reverse the drop in unemployment seen globally between 2020 and 2022.

The COVID-19 epidemic has, therefore, recently had a significant effect on the world labor market. Additionally, new difficulties have arisen that are also negatively affecting the workplace, such as rapid increases in inflation, disruptions in the supply chain, and the conflict in Ukraine.[2] It is vital to take action to advance social justice by tackling issues like young employment, informality, productivity, and gender parity. To create a lasting and inclusive recovery that leaves no one behind, policymakers, companies, employees, and civil society must collaborate. This entails spending money on education and training, enhancing social safety nets, encouraging good work, and moving forward.

The industrial sector, and in particular the manufacturing sector, is not immune to this situation, since it is a very competitive labor market in which it is difficult to recruit top talents that are experienced with new technologies such as robotics or AI. The sector also suffers the fast pace of innovative technologies, making it difficult for both companies and employees, without adequate training and practice, to be up-to-date and competitive. The manufacturing sector needs to retain talent and adapt to new needs, through professional development activities and investment in employee growth, and at the same time be able to find quick solutions, such as cross-training, to be able to cover leaves of absence or production peaks. This calls for new training, retention, and recruitment strategies.

Since 2021, several industries and jobs have been predicted to grow by CareerBuilder,[3] a 1995-founded employment website with operations in 23 countries and a presence in over 60 markets. CareerBuilder offers labor market information, talent management software, and other recruitment-related services. The listed employment spans a variety of industries, including leisure, arts, renewable energy, personal services, healthcare, and information technology. It is possible to upload and build a resume and look for one of many jobs posted by different organizations.

[1] https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_865332.pdf

[2] https://ilostat.ilo.org/assessing-the-current-state-of-the-global-labour-market-implications-for-achieving-the-global-goals/.

[3] https://www.careerbuilder.com/.

Other systems providing similar services are Monster.com,[4] OfferZen,[5] LinkedIn,[6] Glassdoor,[7] JobStreet,[8] ZipRecruiter,[9] Dice,[10] G2 Deals,[11] Indeed Hiring Platform,[12] Hired,[13] Cord,[14] Circa,[15] Naukri,[16] to name a few.

Using these platforms to hunt for a job is simple, but getting chosen for the top chances is often a challenge. For instance, if the position requires a computer science specialist with fund-raising skills, a candidate with only computer science skills is unlikely to be chosen because they lack crucial qualifications like proposal writing experience or start-up development. For the aforementioned example, one should be aware that writing innovation proposals is a crucial skill for fulfilling fund-raising needs. It can be not easy to find answers to queries like these and information on how to develop the necessary skills. One possibility that has recently been exploited is the use of conversational agents [15]. Impressive outcomes have been seen in chats across a variety of areas since ChatGPT's release. Although ChatGPT claims to respond to follow-up queries, acknowledge mistakes, challenge false premises, and reject unsuitable requests, it frequently generates responses that are false and may deceive the user who interacts with it. Authors in [12] have further demonstrated this issue by pointing out related ChatGPT problems in the academic community. Even in this situation, ChatGPT partially made up some of the response's components, and the user was left in the dark as to which components are correct and which are incorrect. Therefore, even though ChatGPT occasionally offers accurate responses, it cannot be relied entirely upon and a different technique should be investigated.

The use of external domain knowledge can be the solution to the aforementioned issues. There are several lexical databases or online taxonomies that have been collected and designed by different organizations that can be relied on when asking for information in such a sense. As such, in this chapter, we will first list the most important existing conversational agents which can be leveraged within the labor domain for worker upskilling. Then, we illustrate and describe all the resources related to the labor domain. We will also provide a solution that integrates conversational agents with the information provided by one of such resources showing the benefits and impact that can be attained. The idea is to leverage such

[4] https://www.monster.com/.
[5] https://www.offerzen.com.
[6] https://www.linkedin.com/.
[7] https://www.glassdoor.com/.
[8] https://www.jobstreet.com/.
[9] https://www.ziprecruiter.co.uk.
[10] https://www.dice.com/.
[11] https://www.g2.com/deals.
[12] https://www.indeed.com/employers/hiring-platform.
[13] https://hired.com/.
[14] https://cord.co/.
[15] https://circaworks.com/.
[16] https://www.naukri.com/.

resources to boost existing conversational agents when asking questions in the underlying domain and overcoming their limitations.

The remainder of this chapter is organized as follows. Section 2 lists the scientific papers that have been published related to conversational agent technologies for the labor domain and worker upskilling. Section 3 will describe in depth the identified conversational agent systems and will illustrate the details of their functionalities and their main limitations, whereas lexical resources within the labor market domain that can be used to provide reliable domain knowledge are included in Sect. 4. A possible solution of a conversational agent integrated with one of such resources is presented in Sect. 5. Expected benefits and impact of the proposed solution are reported in Sect. 6. Section 7 ends the paper with conclusions and outlines future directions.

2 Related Work

Conversational agents can be found in a variety of domains such as mental health [9], lifestyle [5], and customer service [17] to support users of online platforms to cope with daily tasks and challenges. They make use of natural language as a personalization enabler, i.e., they allow a user to interact with complex systems using his/her own language rather than using a limited number of predefined options (e.g., command line, buttons in a web interface, etc.) [7]; this makes the interaction with the conversational agents similar to person-to-person interaction.

Recently, conversational agents have been released for the labor domain to support employers to find new employees as well as candidates who are looking for a job and are interested to improve their skills to be eligible. The reason is that the use of conversational agents can make the recruitment process and interviews more inclusive and efficient, and interviewees seem to be willing to trust the conversational agents in confiding information and listening to their words [18]. Conversational agents can support people during each stage of their career path starting from their early studies (e.g., in choosing a university or study path), through the search for a job (e.g., conversational agents that perform preliminary interviews), and during their employment (e.g., a conversational agent that guides employees in learning a new technology). For example, the conversational agent built on top of the IBM Watson suite[17] proposed in [10] presents a system to support new college students in the decision-making process to choose the right major on the basis of necessary skills and employment opportunities. Another recent conversational agent relevant to the labor domain is *GuApp*. It supports candidates to find a job in specific geographical areas using the content of the *Gazzetta Ufficiale*, the Italian public sector's official journal of records, and a knowledge graph built on top of the

[17] IBM Watson: https://www.itechbm.com/watson.

ISTAT website,[18] which provides a taxonomy of professions organized by sector, and DBpedia.[19]

In line with these advancements, the research community looked into the use of knowledge resources to inject knowledge into such systems and provide tools to support candidates in self-evaluating their CVs, and workers to improve their skills. For example, these novel tools can identify which skills a candidate should acquire to become eligible for a specific job. An example of such technologies is described in [16], where the authors looked into which tech skills are taught in colleges and which skills are truly required by the job market. They used Natural Language Processing (NLP) for detecting entities and keywords from computer science courses and job descriptions and provide recommendations for tech courses or topics for tech career development. In [2] a novel method to match the content of CVs with the O*NET database, a large database of work and worker characteristics and skill requirements, is proposed. The authors used advanced NLP technologies to automatically parse the content of CVs, state-of-the-art transformer models based on the *SentenceTransformers* framework [14] to encode the extracted information into latent representations, and a semantic layer to match the information from CVs to the O*NET database. This novel technology is provided to the public through a demo [3]. The same technology is also employed by the same authors to provide *STAR-BOT* [1], a conversational agent that can help explore the O*NET database using a novel grammar and transformer model to understand the user request and deliver information about jobs. In addition to this, *STAR-BOT* can also suggest educational online courses that can help the upskilling of workers if successfully passed.

3 Existing Conversational Agents

In this section, we will describe some state-of-the-art conversational agents that can be employed, among others, also within the labor domain.

ChatGPT Released in November 2022, ChatGPT[20] is an artificial intelligence (AI) chatbot created by OpenAI. It is a form of generative AI that allows users to enter prompts to receive humanlike images, text, and videos created by the AI. It is constructed on top of the foundational large language models (LLMs) GPT-3.5 and GPT-4 and has been fine-tuned utilizing both supervised and reinforcement learning strategies. OpenAI claims that "Reinforcement Learning from Human Feedback" (RLHF) was used to train ChatGPT. The model initially underwent supervised fine-tuning, in which OpenAI trainers acted as both a human user and an AI bot. To

[18] ISTAT: https://www.istat.it/en/.

[19] https://www.dbpedia.org/.

[20] https://openai.com/blog/chatgpt.

fine-tune the model for conversational usage, the trainers used this to develop a dialogue sequence that mimicked how people converse in real life. Later, ChatGPT was enhanced by developing a reward model to be used for the following phase of reinforcing learning. In order to produce responses, this entailed AI trainers engaged with the tool. Afterward, the responses were ranked according to their quality. With this information, there was a further fine-tuning phase called Proximal Policy Optimization. ChatGPT employs deep learning algorithms to assess incoming text prompts and produce responses based on patterns in the data it has been trained on. It can comprehend subtle language differences and deliver responses of the highest caliber because it has been trained on a huge corpus of literature, including books, papers, and websites. Users can give the chatbot feedback by clicking the "thumbs up" or "thumbs down" icons next to each response in order to help it learn. Users can also offer more textual comments to enhance and perfect upcoming dialogue. Users can ask a wide range of topics on ChatGPT, from straightforward inquiries to more difficult ones like, "What is the meaning of life?". ChatGPT is skilled in STEM fields and has the ability to troubleshoot or write code. There is no restriction on the kinds of queries that can be asked to ChatGPT. ChatGPT uses data only through 2021; therefore, it is unaware of events and data after that point. Additionally, because it is a conversational chatbot, users can request additional details or ask that it try again when producing content.

A list of limitations of ChatGPT is described in the following. It does not properly comprehend how intricate human language is. Words are generated using ChatGPT based on input. As a result, comments could come off as superficial and lacking in profundity. Moreover, ChatGPT could respond incorrectly if it does not fully comprehend the question. On top of that, responses may come off as artificial and robotic. The training data covers up to 2021 and, therefore, ChatGPT has no knowledge of what happened later. As ChatGPT is still being trained, providing comments when a response is erroneous is advised. ChatGPT may misuse terms like "the" or "and." Due to this, information must still be reviewed and edited by humans in order to make it read more naturally and sound more like human writing. ChatGPT does not cite sources and does not offer interpretation or analysis of any data or statistics. It is unable to comprehend irony and sarcasm. It can focus on the incorrect portion of a question and be unable to shift. For instance, if we ask, "Does a horse make a good pet based on its size?" and then we ask "What about a cat?". Instead of providing information regarding keeping the animal as a pet, ChatGPT might only concentrate on the animal's size. Because ChatGPT is not divergent, it cannot shift its response to address several questions in one.

Bing It is a chatbot created by Microsoft. Bing is integrated with Microsoft's Bing search engine and is made to respond to users' inquiries in a way that is specific to their search history and preferences.[21] The Bing AI chatbot may help with a variety of tasks, including question answering, providing recommendations,

[21] https://www.bing.com/?/ai.

presenting pertinent search results, and having casual conversations. Similar to ChatGPT, the new AI-powered application on Bing responds to user requests using a selection of words determined by an algorithm that has learned from scanning billions of text documents on the Internet. Bing AI indexes the entire web in order to produce a response. Because of this, the chatbot has access to the most recent news, information, and studies at the time the user submitted his/her query. As an AI model capable of error, OpenAI has already acknowledged that ChatGPT is prone to hallucinations and incorrect responses. Bing AI has made some efforts to address this problem. When the user asks a question in Bing AI, it will respond with footnotes that will take the user directly to the original source of the answer. Bing works on top of GPT4, the newest version of OpenAI's language model systems. It has also been integrated into Skype chat. Through a straightforward chat interface, this new experience is intended to give the user access to a rich and imaginative source of knowledge, inspiration, and solutions to user queries. Microsoft published in a blog post[22] that Bing was prone to getting off track, especially after "extended chat sessions" of 15 or more inquiries, but claimed that user feedback was helping it to make the chat tool better and safer.

Bard Google created the conversational generative AI chatbot and is still working on its release. It was created to compete with ChatGPT and other language models. Bard is based on the LaMDA (Language Model for Dialogue Applications) family of large language models. It was built on top of Google's Transformer neural network architecture, which was also the basis for ChatGPT's GPT-3.5 language model. Eventually, it will be employed to augment Google's own search tools as well as provide support and interaction for businesses. Differently from ChatGPT, Google Bard can use up-to-date information for its responses. Bard may occasionally provide inaccurate, misleading, or false information while presenting it confidently. For instance, Bard might provide developers with working code that is incomplete or fails to produce the correct output. Google Bard is already accessible via a waitlist, but it is difficult to predict when it will be accessible to everyone. However, Google's CEO stated that Google Bard would soon be used to improve Google Search, thus Bard might become more widely accessible shortly.

Chatsonic Chatsonic[23] is a GPT-4 based chatbot that tries to solve the main limitations of ChatGPT. Chatsonic uses Google to grasp information about the latest events. This makes this chatbot able to support users with timely answers and information about events that take place at any moment. It also provides a few sources harvested from the Internet used to generate the answer which can have a relevant impact on the trust that a user can give to the conversational agent. Chatsonic can generate images and can be integrated within the Chrome browser for efficiently working with everyday tools and web platforms such as Gmail, LinkedIn,

[22] https://blogs.bing.com/search/february-2023/The-new-Bing-Edge-%E2%80%93-Learning-from-our-first-week.

[23] https://writesonic.com/chat.

Twitter, etc. In the context of worker upskilling, it can be used to find online courses that can be attended to learn or improve specific skills. One of its main limitations is that each user has a budget of $10k$ words, and when finished, a premium account needs to be activated.

Copilot Copilot[24] is a language model that slightly differs from the others because it targets computer scientists and developers and their daily task of writing new source code. In fact, this model is trained on billions of lines of code. It allows one to write code faster, suggests more lines of code, proposes implementations of methods, and enables developers to focus on the overall approach instead of dispersing energies on trivial tasks. However, this tool is not perfect and it has several limitations; it often delivers code and method implementations that do not make sense and, additionally, it does not test the code that is written. Therefore, the developer has to verify that the automatically generated code is functional, as well as the quality of the delivered results. This is a crucial aspect for the accountability of the work, and for ensuring the quality of the developed software to customers. Nevertheless, Copilot is an interesting tool for the labor domain considering computer science-related areas and might be a plus in a large variety of companies in which software is mostly used for simple tasks.

ELSA-Speak English Language Speech Assistant[25] (ELSA) is an application based on artificial intelligence and speech recognition technologies designed to support users in learning English. It can be used to improve English pronunciation and provides lessons and exercises. The application provides real-time feedback to the users to enhance their speaking skills. ELSA-Speak can be employed within the labor market to prepare for job interviews (for example, candidates might be more confident to have an interview in English), to help non-native English speakers in their work, and to improve their English skills in order to increase opportunities for career advancement.

4 Skills, Competences, and Occupations

For some time now, different initiatives have been presented to help students, the workforce, and companies to meet the needs of the labor market. All of them have in common that they are initiatives supported by government agencies, are updated periodically, and offer the data in a public and open way. In this section, we will focus on three of the most relevant databases and information systems on occupations and competencies.

[24] https://github.com/features/copilot.

[25] https://elsaspeak.com/en/.

O*NET O*NET[26](Occupational Information Network) is a program sponsored by the U.S. Department of Labor/Employment and Training Administration, which brings together occupational information to understand how the changing labor environment impacts the workforce. One of the main outputs of the program is the O*NET database, a collection of hundreds of standardized descriptions related to more than 1000 occupations found in the USA. This collection is updated periodically through worker surveys supplemented by information from occupation experts. Each occupation in the O*NET database is associated with a set of skills, abilities, and knowledge and is linked to a variety of tasks and activities. The version 26.2 (February 2023) includes 35 `Skills`, 33 `Knowledge`, 52 `Abilities`, 41 `Work Activities`, 4127 `Tools Used`, 17,975 `Task Ratings`, and 8761 `Technology Skills`. Each of these entities allows for the categorization of each occupation, its compatibility, and the specific detection of the needs of the occupation. O*NET is based on a conceptual model that provides a framework to identify the most relevant information about occupations even across jobs and industries. This is useful to link titles referring to the same occupation or to search for intersectoral relationships.

ESCO ESCO[27] (European Skills, Competences, and Occupations) is the European reference terminology describing and classifying occupations and relates them to the most relevant skills and competencies. In addition to being freely and openly available, it is translated into 27 languages (including the 24 official EU languages) and connects to several international frameworks and classifications, such as the International Standard Classification of Occupations (ISCO)[28] or the European Qualifications Framework. ESCO v1.1.1 includes more than 3000 occupations, categorized into 10 groups, mapped with ISCO, and more than 13000 skills, subclassified into 4 types of concepts: knowledge, skills, attitudes and values, and language skills and knowledge. This hierarchy is partially based on the entities and elements of O*NET and NOC (another occupational resource described in the next section). This partial link with other databases, the mapping with ISCO, the connection with qualification levels, and above all, its multilingual support make it particularly attractive for applications that require adaptation to the languages that ESCO supports and especially for linking students and job seekers with employers and employment services.

NOC NOC[29](National Occupation Classification) is the national system for the description of occupations, in French and English, and aligned with the Canadian labor market situation. This open database is updated annually with major revisions every 10 years and is based on data from censuses, employment services, labor regulations, and public consultations among others, which are then analyzed and

[26] O*NET OnLine, National Center for O*NET Development, www.onetonline.org/.

[27] https://esco.ec.europa.eu/.

[28] https://www.ilo.org/public/english/bureau/stat/isco/isco08/.

[29] https://noc.esdc.gc.ca/.

processed, and discussed by working groups. Like the aforementioned resources, NOC also includes a hierarchy and structure of occupations and is linked to the Occupational and Skills Information System (OaSIS),[30] the Canadian database with detailed information on more than 900 occupations. It includes, among other information, the main duties of each occupation, together with the skills, abilities, and even the most common interests among the people who hold those occupations. Among the most interesting functionalities of NOC and OaSIS are the comparison between occupations, the link with Training, Education, Experience, and Responsibilities (TEER)[31] to understand the training and education required to perform the main duties of an occupation, and the search occupations by interests (e.g., investigative, artistic, or social).

In addition to the above, it is important to mention other international and national classifications that are either of reference or are applicable in other countries and are inspired by or refer to the above. The International Standard Classification of Occupations [6] (ISCO) is an international classification developed by the International Labor Organization of the United Nations. This classification and, in particular, its publications ISCO-88 and ISCO-08 define groups and titles of occupations associated with tasks. In general, these database and associated documents are less comprehensive than the three previous ones, but in certain occupational databases, connections with ISCO allow occupations to be linked between databases in different countries.

One of the most common criticisms of these classifications is that they are created and adapted in high-income countries. For this reason, initiatives to map competencies in other low/middle-income countries are also being undertaken. Two examples of these initiatives are STEP (Skills towards Employment and Productivity) and PIAAC (Programme for the International Assessment of Adult Competencies). STEP is a World Bank program to improve the understanding of skills (cognitive, non-cognitive, and technical) in the labor market and relates skills to education, through household surveys and employer-based surveys. It is a program of the Organisation for Economic Cooperation and Development (OECD) that analyzes socio-labor and educational characteristics and links them to competencies at different cognitive levels such as literacy, numeracy, and problem-solving. The first cycle of PIAAC started in 2011 analyzing 42 countries and its second cycle will run until at least 2029 covering 30 countries.

All these initiatives and databases provide a stable and updated background, based on reliable sources and in many cases that can be linked together, which allows us to propose ideas and proofs of concept such as the one mentioned in the following section.

[30] https://noc.esdc.gc.ca/Oasis/OasisWelcome.

[31] https://www.statcan.gc.ca/en/subjects/standard/noc/2021/introductionV1.

5 Proposed Solution

The goal of the solution we have proposed in recent papers [1–3] is to make available to users, who are not always technologically prepared, the information contained within any of the resources described in Sect. 4 in the form of natural language and to help them analyze which skills and personal knowledge are most useful for their application in any job.

One of the worker upskilling resources covered in Sect. 4, the O*NET database, contains many types of information. Our system uses the titles of the principal job categories that cover all the jobs in the North American job market, their descriptions, and the alternative titles for each one. Abilities, Skills, Knowledge, Work Activities, Tasks, Technology Skills, and Tools are the information that characterizes each job category. The number of elements for each job is constant for the first four categories and variable for the last three. The items in the first five present a score relative to their importance to the job they refer to, while the last two are lists of tools or technologies commonly used in each job category.

More in detail, the system we propose performs the following tasks:

1. Determine how much the user's resume is appropriate for the chosen job category.

   The system uses NLP techniques to analyze the user's resume and to extract its most important information. The database entities and the extracted content from the CV are compared using semantic similarity. The resume score for a job is the sum of the scores of the discovered elements divided by the maximum score obtainable for the job itself and a corrective factor required to avoid penalizing jobs with few items in the entities with a variable number of elements. The system draws attention to the lacking aspects of the CV and suggests ways for users to increase their knowledge of each one.
2. Determine the user's suitability for the selected job.

   The system prompts the user to enter his/her self-assessed level (none, little, medium, good, excellent) in terms of Abilities, Knowledge, Skills, Work Activities, and Tasks entities taken from the database and associated with the selected job. The system returns a percentage score, where 60% is the suitability threshold and a list of lacking knowledge or abilities that the user might improve to raise his/her score.
3. Determine what occupation the user is best suited for.

   The system prompts the user to enter his/her self-assessed level (none, little, medium, good, excellent) in terms of Abilities, Knowledge, Skills, and Work Activities entities taken from the database and returns a list with five suggested job categories.
4. Perform conversation with the user.

   To integrate the knowledge provided by O-NET into the conversational agent, we first chose a reduced set of questions that it must be able to answer. The base question templates are:

- `Alternate names of jobs`
- [*Abilities* | *Knowledge* | *Skills* | *Tasks* | *Tech-Skills* | *Tools used* | *Work activities*] `that are important or necessary in any job`
- `Description of a job`
- `The similarities between a couple of jobs`
- `Which jobs require a specific` [*Ability* | *Knowledge* | *Skill* | *Task* | *Tech-Skill* | *Tool* | *Work activity*]
- `Recommendations on how to improve one` [*Ability* | *Knowledge* | *Skill* | *Task* | *Tech-Skill* | *Tool* | *Work activity*]

Then we implemented the query models necessary to retrieve the data to generate the answers.

When the conversational agent recognizes a question in the user's input that is part of the set of known questions, it requests the database for the data necessary to prepare the answer in natural language.

The system uses semantic similarity to match user input with one of the predefined question templates. If the similarity value is not high enough, the system will answer without the database knowledge, returning a default answer, or using a generative pre-trained model. More details of our system can be found in [1–3].

6 Expected Challenges, Benefits, and Impact

In terms of challenges, employee engagement in training activities and work commitments is one of the key elements. Knowing the information about their job and potential growth paths, while reducing doubts about the stability of their job, allows their adaptation and evolution, and thus engagement. Interactive chatbots and conversational tools can simplify the task of obtaining this information, for example, through informal interactions and by simply relying on the chatbot as a companion that offers relevant and appropriate information, focused on the characteristics of each occupation.

As far as technical challenges are concerned, it is still necessary to reach an adequate balance between the use of large language models (e.g., ChatGPT) and models trained with concrete content based on validated or accepted reference information. One of the main problems of LLMs is the lack of determinism in the results [11] and the potential failures in the recommendations [13]. Both can be compensated by relying on databases such as the ones mentioned above and prioritizing the results based on the information in those databases.

It is important to also mention and consider the challenges of conversational agents related to privacy. While many of the ideas mentioned in this article do not require storing worker information, and in our use cases we have demonstrated this, there are several cases, such as group data discovery or candidate analysis, where the handling of worker data could end up in the processing of private data, aggregation

of personal data and in some cases in profiling. The topic related to the concerns about user data's inappropriate use in user-chatbot interactions is being researched and surveyed [4, 8], and still there are several debates about it at the international level (e.g., in April 2023 Italy was the first Western country to block ChatGPT[32]). This calls for a critical analysis of the different uses of chatbots for upskilling and, above all, an alignment with the initiatives for the maintenance of the privacy of workers and users, from the General Data Protection Regulation to the Artificial Intelligence Act.

Regarding benefits and impact, at the scientific and technical level, and as can be seen in the state of the art, advances in conversational agents and language models are clear. All these advances have an impact on the development of solutions such as the one we present, for example, by providing better interaction capabilities or even the possibility of answering open-domain questions. At the economic level, as also discussed in the introduction, there is a demand for workers and a demand for knowledge for specific activities, and upskilling agents can facilitate this task. At the social level, it impacts workers and the unemployed, in terms of analyzing their strengths, detecting training or internship needs, and adapting their CVs to fit specific jobs.

From the worker's perspective, conversational AI agents based on standardized and reliable databases of occupational information have direct impacts and benefits.

On the one hand, these conversational solutions have a direct impact on workers' knowledge of their work activity. The workers themselves can at any time obtain reliable information about their job and learn about skills, activities, and knowledge relevant to the task they are carrying out. Moreover, the worker can ask the conversational agent which are the usual technical skills for a certain occupation or activity to imagine where they might be oriented.

On the other hand, it is beneficial for their training and career. The employees can detect points in which their profile would benefit from training and at the same time detect training needs and create training plans to adapt to the needs of their position, or to another position they plan to move to. This is one of the other benefits of these conversational agents since the workers can compare themselves with professions and positions in which they would be interested in changing or being promoted.

From a more organizational and business point of view, there are also a couple of areas where AI-based conversational agents such as the ones mentioned in this article can be beneficial. The part of assistance in hiring processes is a point in which these agents will have a direct impact: from providing candidates with information about the position, to pre-screening or determining the suitability of a candidate by using conversational agents to compare the candidate's CV with the baseline of the occupational reference database. Conversational agents are also beneficial in the detection of training needs for specific positions or the creation of group training plans, based on the needs of multiple workers.

[32] https://www.bbc.com/news/technology-65139406.

Therefore, the impact on the HR area is clear: talent management, discovery of training paths, HR resources, growth programs, etc. All of them benefit in the short term from conversational agents like the ones mentioned above. Proofs of concept and use cases conducted with conversational agents, based in our proposed solution, acting as virtual interviewers, skills gap detectors or as supporters for the selection of better training, demonstrate the potential of these technologies. Not only at the level of large technology companies but the aforementioned solutions can also make a difference at the level of SMEs in industrial and manufacturing environments. Where they may not have large HR departments but need reliable, objective, and consistent feedback for resource upskilling and solutions based on standardized information, starting from the same baseline and with clear and repeatable criteria.

7 Conclusions

This chapter explores the potential uses, benefits, and limitations of conversational agents within the labor market. The chapter guides the reader through the state-of-the-art research on this topic and presents which technologies are currently employed. Furthermore, the chapter analyzes conversational agents that are disrupting the labor market, where AI tools are expected to be employed for a large variety of tasks such as learning new skills, supporting document writing, code development, language skills improvement, and so on. Along with these conversational agents, we also introduce the reader to valuable sources of information which can be leveraged to build AI systems on top. Finally, we present how potential solutions can be used for specific use cases and discuss how all these technologies can impact the labor market.

As main lessons, we would like to mention the need for both the workforce and companies to adapt to the rapidly changing world of new technologies and how this can be beneficial, for example, by taking advantage of it to better understand occupations, retain talent, hire, gain job knowledge, or adapt to new activities. It is also important to consider the need to efficiently integrate large language models and validated and reliable databases. Finally, it is essential to consider and analyze the privacy implications of these solutions. It is possible to develop privacy-aware conversational systems for the upskilling of workers, but there is a possibility that certain solutions may abuse data. For this reason, it is critical to provide these systems with consideration of Human Centricity aspects and all the factors that compose it, including privacy.

In the medium term, it is expected that these technologies will be integrated into the regular workflow of employees and will have a positive impact on employee engagement and professional development. For this reason, we expect that this chapter will foster several visions to the reader about the role of AI-based conversational agents in the labor domain and worker upskilling which might bring unprecedented development in the field.

Acknowledgments This work has been partly supported by EU H2020 research and innovation programme project STAR—Safe and Trusted Human Centric Artificial Intelligence in Future Manufacturing Lines (Grant n. 956573).

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5—Call for tender No.3277 published on December 30, 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union—NextGenerationEU. Project Code ECS0000038—Project Title eINS Ecosystem of Innovation for Next Generation Sardinia—CUP F53C22000430001—Grant Assignment Decree No. 1056 adopted on June 23, 2022 by the Italian Ministry of University and Research (MUR)

References

1. Alonso, R., Dessí, D., Meloni, A., Recupero, D.R.: Incorporating knowledge about employability into conversational agent. Appl. Intelligence submitted (2023)
2. Alonso, R., Dessí, D., Meloni, A., Recupero, D.R.: A novel approach for job matching and skill recommendation using transformers and the o*net database. Big Data Res. submitted (2023)
3. Alonso, R., Dessí, D., Meloni, A., Reforgiato Recupero, D.: A general and NLP-based architecture to perform recommendation: A use case for online job search and skills acquisition. In: Proceedings of the 38th ACM/SIGAPP Symposium On Applied Computing. ACM Special Interest Group on Applied Computing (SIGAPP) (2023)
4. Belen Saglam, R., Nurse, J.R.C., Hodges, D.: Privacy concerns in chatbot interactions: When to trust and when to worry. In: Stephanidis, C., Antona, M., Ntoa, S. (eds.) HCI International 2021 - Posters, pp. 391–399. Springer International Publishing, Cham (2021)
5. Fadhil, A., Gabrielli, S.: Addressing challenges in promoting healthy lifestyles: the al-chatbot approach. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare, pp. 261–265 (2017)
6. Hoffmann, E.: International Statistical Comparisons of Occupational and Social Structures, pp. 137–158. Springer US, Boston, MA (2003). https://doi.org/10.1007/978-1-4419-9186-7_8
7. Hussain, S., Ameri Sianaki, O., Ababneh, N.: A survey on conversational agents/chatbots classification and design techniques. In: Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019) 33, pp. 946–956. Springer (2019)
8. Ischen, C., Araujo, T., Voorveld, H., van Noort, G., Smit, E.: Privacy concerns in chatbot interactions. In: Følstad, A., Araujo, T., Papadopoulos, S., Law, E.L.C., Granmo, O.C., Luger, E., Brandtzaeg, P.B. (eds.) Chatbot Research and Design, pp. 34–48. Springer International Publishing, Cham (2020)
9. Lee, M., Ackermans, S., Van As, N., Chang, H., Lucas, E., IJsselsteijn, W.: Caring for Vincent: a chatbot for self-compassion. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
10. Lee, T., Zhu, T., Liu, S., Trac, L., Huang, Z., Chen, Y.: CASExplorer: A conversational academic and career advisor for college students. In: The Ninth International Symposium of Chinese CHI, pp. 112–116 (2021)
11. Maddigan, P., Susnjak, T.: Chat2vis: Generating data visualisations via natural language using ChatGPT, codex and gpt-3 large language models. IEEE Access (2023)
12. Meloni, A., Angioni, S., Salatino, A., Osborne, F., Reforgiato Recupero, D., Motta, E.: Integrating conversational agents and knowledge graphs within the scholarly domain. IEEE Access **11**, 22468–22489 (2023). https://doi.org/10.1109/ACCESS.2023.3253388
13. Oviedo-Trespalacios, O., Peden, A.E., Cole-Hunter, T., Costantini, A., Haghani, M., Kelly, S., Torkamaan, H., Tariq, A., Newton, J.D.A., Gallagher, T., et al.: The risks of using ChatGPT to obtain common safety-related information and advice. Available at SSRN 4346827 (2023)

14. Reimers, N., Gurevych, I.: Sentence-Bert: Sentence embeddings using Siamese Bert-networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics (2019). https://arxiv.org/abs/1908.10084
15. Singh, S., Beniwal, H.: A survey on near-human conversational agents. Journal of King Saud University - Computer and Information Sciences **34**(10, Part A), 8852–8866 (2022). https://doi.org/10.1016/j.jksuci.2021.10.013. https://www.sciencedirect.com/science/article/pii/S1319157821003001
16. Vo, N.N., Vu, Q.T., Vu, N.H., Vu, T.A., Mach, B.D., Xu, G.: Domain-specific NLP system to support learning path and curriculum design at tech universities. Comput. Educ. Artif. Intell. **3**, 100042 (2022)
17. Xu, A., Liu, Z., Guo, Y., Sinha, V., Akkiraju, R.: A new chatbot for customer service on social media. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 3506–3510 (2017)
18. Zhou, M.X., Mark, G., Li, J., Yang, H.: Trusting virtual agents: The effect of personality. ACM Trans. Interact. Intell. Syst. (TiiS) **9**(2-3), 1–36 (2019)