But which skills?
Natural Language Processing tools and the identification of high-demand skills in online job advertisements

Gianni Anelli

Gianni Anelli is a PhD student in the Institute of Employment Studies (IER) at the University of Warwick, UK.

ABSTRACT
Skills assessment is essential for today’s labour market. There are many factors that change the requirements for the workplace. More than ever, it is important to monitor which skills are in high demand so that workers stay employed and companies do not lose productivity. This research discusses the relevance of data from online job portals for this task. It then uses a skill extractor in online job advertisements from Chile to identify and extract the skills employers place in their online job advertisements through skills dictionaries. The study shows modest results when using the European Skills, Competences and Occupations (ESCO) dictionary but an enhanced and much-improved result when adding an inductively constructed dictionary of the national labour market. Using this method would allow a new input of information to be incorporated into labour market information systems that would enable better decisions to be made by the various actors in the labour market.

KEY WORDS
high-demand skills, online job advertisements, skills dictionaries, skill extractor, bottom-up approach

Introduction
Results of the second wave of the European Skills and Jobs Survey (ESJS) in 2021 indicated that 45% of respondents are concerned about technological skills obsolescence and the need to acquire new knowledge and skills’ (Cedefop, 2022:16). Moreover, in more recent findings, 81% of employees express a desire to participate in
training programmes to keep their skills up to date (ManpowerGroup, 2023). Similar uncertainty is shared by employers. Based on the 2018 Manpower Group survey, Cárdenas et al. (2020) highlighted that 45% of the world’s employers cannot find the skills they need in the labour market. These are signs of the striking changes sweeping across the world of work and the increasing uncertainty for all actors within the labour market. There are clear indications that workers will need to re-skill or up-skill to stay employable because of the emergence and restructuring of new tasks and the transition to potential new jobs (Kanders et al., 2020).

The problem, nevertheless, seems to be that the change in job skills is faster than the capacity of educational institutions and the working population in general to adapt to those changes. For the workforce, this represents a risk, as workers lack updated information that would enable them to invest in their education or training to keep themselves employed. Ultimately, this leads to skills imbalances with social and economic costs for individuals, firms and national economies (McGuinness, Poliakas & Redmond, 2018; Brunello & Wruuck, 2019; Gal et al., 2019). This situation suggests there is a need for better and more detailed flows of information on what specific skills are in demand in the workplace. These would make it possible to generate re-skilling and up-skilling programmes for workers as well as make adjustments to educational institution curriculums to ensure that people can meet the demands of quickly changing workplaces (Cedefop, 2019b).

In this context, online job portals have gained relevance in the study of the labour market in more depth. Their use by employers and job seekers has increased year after year, creating exponential growth in the amount of information stored on these websites about the labour market (Amato et al., 2015). Globally, in 2018 the job applications made through these job portals accounted for a fifth of all hires (Cedefop et al., 2021). With this rising accumulation of data from online portals, a different stream of valuable information has opened for assessing trends in demand for skills in the labour market (Orlik et al., 2020). As a result, analysis which makes use of data science techniques (Big Data) applied to sources such as online job portals (OJPs) has expanded as a method of study (Eurofound, 2021). Still, there is scope to test and expand the limits of current lines of study. In addition to monitoring vacancies posted online as a supplement to national statistics, online labour market data has further advantages to deliver, especially about which skills are in high demand. Therefore, the question behind this study is to inquire how to achieve effective results in detecting high-demand skills within online job advertisements.

Advances in Natural Language Processing (NLP) and text mining tools make it possible to extract more and richer information about skills requirements within online job advertisements. The research presented here constructed a skill extractor for capturing those skills within job advertisements. For the construction of a skill extractor, this study initially used the European Skills, Competences, and Occupations (ESCO) Skill Dictionary. Although this is a dictionary legitimised by the work and prestige of ESCO, the results indicate only a 30% effectiveness in finding skills. This relatively low effectiveness can be attributed to the fact that the ESCO dictionary uses labels that are too wordy and too explanatory compared with those normally used in online job advertisements to refer to skills. For instance, while ESCO’s dictionary
contains the skill ‘using spreadsheet software’, it does not have a skill named ‘Excel’, the label most advertisements use when referring to the most common spreadsheet package, Microsoft Excel.1

The study showed considerable differences when the ESCO dictionary is complemented by a manually constructed dictionary that takes into consideration the vocabulary actually used by employers in local labour markets. Therefore, using an inductive search on how skills are mentioned within job advertisements in online job portals – in this case, the country of Chile – this dictionary enhances the preliminary results obtained using only the ESCO skills dictionary. When combined with a skills dictionary from a national labour market, the effectiveness in detecting skills increases.

Approaches to the measurement of skills
Traditionally, occupations and their skills have been assessed and updated using survey data, employers’ interviews or experts’ opinions (Beblavý et al., 2018). However, those methods are based on data, most times outdated, which minimises the ability to make proper decisions by labour market actors. These methods not only require considerable financial resources but also take a long time to collect and process (Chen, 2021). This is why the European Commission has declared that it is essential, as part of the objectives for the evaluation of current trends in the labour market, to ‘strengthen skills intelligence, highlighting the need for online “real-time” information on skills demand’ (Eurofound, 2021:4).

As a result, the spotlight has turned to the data within OJPs. What at first were websites with utility as a search engine to promote and find jobs, over time, have become tools for researching trends in the labour market (Cárdenas et al., 2020). Compared to annual or biennial surveys, they bring advantages such as data collection in (quasi) real-time, lower costs for extracting information and improvements in the accuracy and level of detail of the analysis of skills (MAC, 2017; Cárdenas et al., 2020; Cedefop et al., 2021). In the last decade, the use of data from OJPs has been incorporated as a supplementary input to traditional sources for the estimation of vacancy rates (Askitas et al., 2018; Tijdens, Beblavý and Thum-Thysen, 2018; MAC, 2019; Eurofound, 2021). However, the value of information contained in online job advertisements allows for richer lines of research.

‘Skills’, the mixture of abilities and knowledge to adequately perform a task (Rodrigues, Fernández-Macías & Sostero, 2021), provide a more in-depth level of job requirements. For many years, educational qualifications were the most common proxy for talking about skills due to the complexity of getting data at this level of detail (OECD, 2017). Data from online job advertisements can have that level of depth. Working with ads from OJPs, it is possible to assess the skills that employers deliberately place in their job advertisements and not just the ones selected to be in a

---

1 ESCO Skills Dictionary has a ‘preferred’ label and several ‘alternative’ labels. The one analysed and used in this research is the preferred label. Alternatives are not used because in many cases there are no labels. When they do exist, they are all part of the same string without any separator between the alternative labels. This makes it highly improbable to generate an extractor that follows some kind of pattern to differentiate between them. For instance, in the case of the skill ‘using spreadsheet software’, one of the alternative labels is ‘Microsoft Office Excel’. The problem is that it is one label among 28 others without punctuation and can only be recognised manually when one ends and the other begins.
questionnaire (Cedefop et al., 2021). This can be considered as a bottom-up approach, where the information emerges without prior restrictions, guided by a data-driven approach (Colombo, Mercorio & Mezzanzanica, 2018).

**Different approaches to the skills extraction problem**

To extract the information, it is first necessary to pre-process and clean the data. Raw data from OJPs will not provide valuable information unless cleaned. When working with text strings, each researcher must calibrate their text cleaning process according to their research purposes.

Text mining refers to all the semi-automatic techniques necessary to identify, clean and extract information from unstructured text (Januzaj et al., 2019; Fareri et al., 2020). Its goal is to convert text-based content into a more convenient format, eliminating all that does not add value for research purposes. Separation of text into words (tokenising), case sensitivity, removal of unnecessary words, punctuation, symbols and blank spaces are important steps for pre-processing text (Beblavý et al., 2018; Lovaglio et al., 2018; Chernova, 2020; Vladimirovna & Ibrahim, 2020; Lunn, Zhu & Ross, 2020).

The extraction of skills is a problem belonging to the subworld of information extraction. In this case, it seeks to identify those parts of a text that are related to the demand for skills (Sharma, 2019). The general framework for this type of problem is known as the emerging NLP tasks (Chernova, 2020; Fareri et al., 2020; Lunn, Zhu & Ross, 2020). NLP refers to teaching machines to understand and process human language to perform specific actions to differentiate the meaning that certain words have within the rules of the specific language, a meaning that often depends on its context. In other words, it is teaching machines to deal with the ambiguity of human semantics (Ates, Bostanci & Serdar, 2021). To achieve this goal, it is necessary to build a pipeline with several steps. The first step is data cleaning with the text mining techniques applied so far. As the specificity of the task to be assigned requires greater precision, NLP models become more complex (Chernova, 2020; Lunn, Zhu & Ross, 2020). These range from Bag of Words models, through rule models, text classification models to more complex ones such as language transformation models (Akhtyamova, 2020).

Techniques associated with NLP tasks can be classified into machine learning supervised and unsupervised models (Cobb et al., 2018). Supervised models are those where people train the model with the expected response in training data. The algorithm, therefore, learns what it should then look for in new data. By contrast, unsupervised models are those where there is no human intervention. The algorithm learns relationships and patterns from the data itself, to produce an outcome. In the case of skill extraction, the same distinction exists. Within the applied techniques, most approaches to this type of problem are based on supervised models (Sharma, 2019; Chernova, 2020; Wings, Nanda & Adebayo, 2021).²

² Despite the results obtained with unsupervised models, the problem with these approaches is that, because job advertisements have information on multiple topics and not just skills and tasks, they extract too much noise. Sometimes they leave out relevant skills and include companies’ propaganda (Chernova, 2020).
Many studies have created a training dataset, trained a model on a specific technique and then applied it to unseen data for testing. Techniques such as Word2vec or FastText as non-contextual word embeddings and long short-term memory techniques or BERT as contextual word embeddings combined with Part of Speech Tagging (POS) or Name Entity Recognition (NER) have been used in this area (Chernova, 2020; Bhola et al., 2021; Luoma & Pyysalo, 2021; Wings, Nanda & Adebayo, 2021; Vermeer et al., 2022).

Despite the progress in the computational processing of text and the immediate attention of machine learning techniques, the research presented here uses a rule-based model which is a useful and effective technique for information extraction. By using dictionaries with predetermined skills terms, skills extraction can be solved by matching those mentioned in job descriptions with a skills dictionary. Each term or phrase identified as a skill is converted to a n-gram and then looked into the strings of the text so the NLP algorithm functions as a search engine (Appadoo, Soonnoo & Mungloo-Dilmohamud, 2020; Brancatelli, Marguerie and Brodman, 2020).

The question then is where to find a dictionary of skills against which to check job advertisements. The first dictionary of skills used in our research was from ESCO Taxonomy. ESCO is the European Union’s occupational, and skills classifier constructed through the labour markets and educational institutions of its members. It is for this reason that this skills dictionary was selected. This skills dictionary contains a diverse set of countries with varying levels of development. It is, therefore, more likely to have skills used in the Chilean labour market than the skills in O*NET, based on the US economy. Public and private actors use this multilingual classifier dictionary of occupations and skills as the reference language for employment and education (Asenitou, 2015; European Commission, 2021). Accordingly, the dictionary is available in 26 languages, including Spanish, which facilitates matching with Chilean labour advertisements. In its latest version in Spanish, the dictionary contains 13,891 skills categorised into cross-sector, occupation-specific, sector-specific, and transversal skills. However, compared with its US counterpart, ESCO has three times the number of job profiles and six times the number of skills of O*NET among its records (Rentzsch & Stanina, 2020; Fareri et al., 2021).

Construction of the skill extractor

The job advertisements used in this research were provided by the System for the Analysis of Employment Portals (SABE according to its Spanish acronym) Project in Chile. The SABE Project seeks to collect and standardise information from different job portals in Chile. Different random batches of data were used for the year 2022.

---

3 Phrases were built in the form of n-grams. N grams correspond to sets of n consecutive words in the form of smaller chunks turning sentences to a better suited form for analysis and annotations (Wings, Nanda and Adebayo, 2021). In other words, n-grams are phrases with two or more words together where n represent how many they are. They are composed essentially by nouns, verbs, adverbs, or adjectives. For instance, an example of a 2-grams would be ‘critical thinking’, 3-grams would be ‘work under pressure’ and so on.

4 Austria, Belgium – Luxembourg, Bulgaria, Czech Republic, Croatia, Denmark, Estonia, Germany, Hungary, Italy, Latvia, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden.

5 The SABE Project in Chile has working agreements with several OJPs in the country. More information is available at https://sabe.wic.cl/
First, in order to compare the entries in the dictionary of skills against the text strings of job advertisements, the skill extractor was built using the library spaCy on Python. Among the many advantages of spaCy, the library performs a numerical vectorisation of the words for any text operation, that is, a numerical representation where words with similar meaning and context appear closer together. Accordingly, spaCy performs tokenisation in which each word is assigned a number according to a pretraining sample loaded into the library which is called vectorisation.

It is on this vectorisation that the skill extractor works and the search becomes more efficient, operating much better with larger amounts of data. Typical search engines run by multiplying the size of the text string by the size of the list of terms the user is looking for to find matches. That means that the complexity rapidly increases with large amounts of data and a longer list of terms to look for in the text. Because the skill extractor was constructed on spaCy, it has a softer level of complexity compared to other available tools. The function runs by multiplying the size of the text data by the logarithm of the size of the list of terms looked for in the text (NewsCatcher, 2022).

Yet, the major challenge with the ESCO skills dictionary – or any institutional skill dictionary – for this purpose is that it was not designed to be matched with skills in job advertisements but, instead, is intended to describe its listed skills. This means that the way they are written is somewhat lengthy when measured by the total number of words for each skill. This represents a problem because it makes it more difficult to match against the terms used in online job advertisements. The more words, the more synonyms for each word used are possible. In contrast, advertisements have limited space for a job description, so they need to ‘economise’ the language used to place their information, not only on the skills and task requirements of the job but also on other aspects of the recruitment process, or advertising hooks to attract more applicants.

To face this problem, every skills label in ESCO’s dictionary was pre-processed and cleaned as well as the online job advertisements. Table 1 compares the first results of the skill extractor using the online job advertisements and dictionary before and after being filtered. Accordingly, the first extractor using unfiltered data shows modest results in finding skills in job advertisements. The extractor was able to find skills in only 30% of all job ads. In contrast, the second extractor used the filtered ESCO skills dictionary on filtered job descriptions and the results increased up to 46% of all job advertisements with at least one skill on their job description.

Table 1: Results of extractors with ESCO skills dictionary

<table>
<thead>
<tr>
<th>Extractor</th>
<th>Job description</th>
<th>Skills dictionary</th>
<th>Total Nº skill matches</th>
<th>Nº of ads with one or more matches</th>
<th>% of ads with one or more matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unfiltered</td>
<td>Unfiltered</td>
<td>2,616</td>
<td>1,466 /4,950</td>
<td>30%</td>
</tr>
<tr>
<td>2</td>
<td>Filtered</td>
<td>Filtered</td>
<td>5,197</td>
<td>2,292 /4,950</td>
<td>46%</td>
</tr>
</tbody>
</table>

Source: own elaboration.
Crucial at this stage was the manual validation of the job ads with and without skill matches. There was an iterative revision of results: which advertisements were matched and how many skills they supposedly had. In this examination, it became clear that ESCO’s skills dictionary was not detecting many skills declared by employers. Several online job advertisements contained multiple skills mentioned while the skill extractor did not identify even one. This happened in most cases for one of three reasons: the dictionary did not have all the skills demanded in job advertisements, the skills labels in the dictionary had too many words of length, or the dictionary simply used a more academic language than that used by employers when describing the skills they need.

Assembling a Chilean dictionary
Consequently, a dictionary with inductively selected terms, keywords and phrases used in job advertisements was constructed for this research. A list of skills was created by reviewing over 4,000 randomly selected job advertisements for the Chilean Labour Market from the SABE Project’s data. As the assessment of job advertisements progressed, the use of words and phrases mentioning skills became more consistent, and patterns emerged. A similar structure on how skills appeared in job descriptions was acquired. This knowledge was applied to build phrases and construct the skills dictionary as in other similar studies (Sharma, 2019; Brancatelli et al., 2020).

According to Rentzsch and Staneva (2020), the main advantage of combining dictionaries is that skills terms and phrases with a more macroscopic and long-term view of the labour market can be retrieved from the expert dictionary. On the other hand, the specific Chilean dictionary allows for a near real-time mapping of micro trends in demand for skills and the terms used.

Accordingly, the rules for constructing the dictionary of Chilean skills were as follows. First, avoid one-word skills so as not to generate false matches with other parts of the text on job descriptions. For example, the ESCO dictionary skill 6,407 is ‘values’. Whenever the word values appeared in a job description, it appeared as a skill in the extractor results. Cases like this were avoided in the construction of the Chilean dictionary. The only exceptions were for specific terms associated with a software, for example, Excel, Python or Java or proper nouns that by themselves are understood as skills such as ‘plumbing’. Second, because the texts do not have stop words, two or more n-gram skills are based on the combination of verbs and nouns, for example, ‘taxable financial knowledge’, ‘preventive maintenance’ and ‘welding structures’. Third, keywords or phrases must be mutually exclusive. There cannot be skill phrases that contain each other in a similar skill and generate double matches. For instance, there cannot be ‘review financial statements’ and ‘review financial statements services’. As a result, currently the Chilean dictionary has 2,286 labels to search for skills.

6 Additionally, ‘values’ was not considered in the ESCO dictionary for the skill extractor.
Results combining ESCO and Chilean skills dictionaries

The combined ESCO and Chilean skills dictionaries were loaded into the skill extractor. When it searched for skills in different samples of data, it got similar results as shown in Table 2. The accuracy of the model was consistent throughout all samples.

The results in Table 2 show the number of skill matches for the total sample and a breakdown of skill matches per ad. The second column indicates the number of total skills matches for all advertisements in the sample. The third and fourth columns are the number and percentage of how many job advertisements have at least one skill match on their job description against the skills dictionaries. The fifth column indicates the number of ads by the skill matches they have. For instance, in the first sample when it says '3 skills – 679' it means there were 679 advertisements in the sample that mentioned three skills in their job description.

A review of the results indicates that the majority of job advertisements state at least one skill per job. In contrast, between 16% and 18% of advertisements make no mention of any skill in their job description. To ensure the accuracy of the results, a visual check on those job advertisements was put in place, and this confirmed they did not mention any skill in their description. In fact, these job advertisements corresponded to low-skilled jobs, where detailed information about the skills required was not usually provided. This could mean that employers take for granted the skills needed for those types of positions. By contrast, a review of the job advertisements with the highest number of skills mentioned showed that they tended to be high-skilled jobs. Nevertheless, these are general results for all job advertisements, without distinction by economic sector or occupation.

The results obtained by the skill extractor are shown in more detail in Figure 1. This shows the 15 skills that were most in demand in the sample of 34,605 job advertisements from the SABE Project. The skills most frequently mentioned by employers are 'customer services', 'responsibility and commitment' and 'Excel'. These may reflect the important weight of the Commerce sector within job advertisements, and the importance of customer service in making sales. The results also highlight the importance of commitment to work as well as the importance of being able to use essential digital tools for everyday tasks. Once the advertisements are classified by occupational groups, the skills will provide more detailed insights into different types of jobs.

Conclusions and limitations

In today's society, there are constant changes that demand new skills for the workplace. This is why it is vital to assess continuously the skills in demand to keep the workforce competent in the skills wanted by employers and thus remain employable.

Job advertisement data in OJPs are a step forward for monitoring these trends in the labour market. They have a high granularity of information regarding the requirements of employers for the positions to which job seekers apply, and provide this information with the advantage of being financially less expensive and quicker to gather than surveys (Turrell et al., 2018; Cedefop et al., 2021). This study shows an effective way to achieve results with detailed and timely insights into those skills demanded in online job advertisements. By complementing the skills of the ESCO Dictionary with a
bottom-up inductive dictionary on a specific labour market, a useful new input with valuable information can be incorporated into a country’s Labour Information System.

Even so, this approach to the extraction of information on skills in high demand has its limitations that must be considered. This method requires a more significant human effort to prepare and process raw and unstructured data than traditional sources (Vladimirovna and Ibrahim, 2020; Cedefop et al., 2021). Similarly, there is a latent risk that the information, instead of being representative, could be biased toward specific economic sectors or occupations (MAC, 2017). Taking all these things into account, the information extracted from OJPs should not be considered as a substitute for data obtained by traditional methods’ data but as a complement to what already exists (Bosch et al., 2018; Cedefop, 2019a).

<table>
<thead>
<tr>
<th>Data sample</th>
<th>Total Nº skill matches on sample</th>
<th>Nº ads with at least one match</th>
<th>% ads with at least one match</th>
<th>Quantity of skills – Number of ads matches</th>
</tr>
</thead>
</table>
| 4,950       | 15,382                          | 4,120/4,950                   | 83%                           | 0 Skills – 830
|             |                                 |                               |                              | 1 Skills – 885 |
|             |                                 |                               |                              | 2 Skills – 846 |
|             |                                 |                               |                              | 3 Skills – 679 |
|             |                                 |                               |                              | 4 Skills – 510 |
|             |                                 |                               |                              | 5 Skills – 350 |
|             |                                 |                               |                              | => 6 Skills – 850 |
| 5,215       | 16,281                          | 4,291/5,215                   | 82%                           | 0 Skills – 924 |
|             |                                 |                               |                              | 1 Skills – 940 |
|             |                                 |                               |                              | 2 Skills – 875 |
|             |                                 |                               |                              | 3 Skills – 648 |
|             |                                 |                               |                              | 4 Skills – 557 |
|             |                                 |                               |                              | 5 Skills – 369 |
|             |                                 |                               |                              | => 6 Skills – 901 |
| 5,250       | 16,557                          | 4,381/5,250                   | 83%                           | 0 Skills – 869 |
|             |                                 |                               |                              | 1 Skills – 944 |
|             |                                 |                               |                              | 2 Skills – 827 |
|             |                                 |                               |                              | 3 Skills – 687 |
|             |                                 |                               |                              | 4 Skills – 650 |
|             |                                 |                               |                              | 5 Skills – 437 |
|             |                                 |                               |                              | => 6 skills – 833 |
| 34,605      | 97,836                          | 29,060/34,605                 | 84%                           | 0 Skills – 5,545 |
|             |                                 |                               |                              | 1 Skills – 7,273 |
|             |                                 |                               |                              | 2 Skills – 6,438 |
|             |                                 |                               |                              | 3 Skills – 5,199 |
|             |                                 |                               |                              | 4 Skills – 3,359 |
|             |                                 |                               |                              | 5 Skills – 2,021 |
|             |                                 |                               |                              | => 6 Skills – 4,770 |

* Results are filtered by counting unique skill matches for each ad. This avoids overcounting matches if the same skill is repeatedly mentioned in the job description.

Source: own elaboration.
A further caution to consider relates to the assumptions behind skills present in job advertisements. The assumption is that for each job advertised, employers select the most critical or fundamental skills they need for the vacancy, and those skills are the ones that are explicitly placed in the advertisements (Rios et al., 2020). However, each vacancy also requires other skills that are not mentioned but are implicitly necessary for the job. These are likely to be understated in the educational level requirements. Therefore, the skills employers include in the advertisements are those that, in the employer’s experience, are essential but there are others that are implicit and necessary, which must not be ignored.

Concerning the skill extractor itself, it produces consistent results, but there are limitations. Because it works with dictionaries of keywords and phrases, there might always be wordings that employers use to refer to a skill that are different from those in the dictionaries. One example is ‘know, understand, comply with and enforce the requirements established in the risk prevention policies’, which is quite extensive in the length of words for a job advertisement. Although employers tend to economise in this respect, it is still possible to find such expressions. In the same sense, many other employers use ‘implement risk prevention policies’. When employers are too descriptive in their language and excessively wordy, it might be that the skill extractor does not pick out that skill. In other words, this approach needs to be constantly updated and adapted to the particularities of how language is used by employers in labour markets to describe skills (Gugnani & Misra, 2020).

Finally, it is relevant to mention that this line of research opens doors for future research on more profound implications of the swift changes within the labour market. There is a growing concern about whether the workforce can keep up with technological changes. The methodology and results of this study would allow, on the one hand, to observe which skills are most in demand, for example, those related to...
the use of new technologies. On the other hand, it makes it possible to explore in future research the extent to which those technical skills are associated with improvements in wages and working conditions offered by employers. This would make it possible to assess whether higher skills are associated with better working conditions and which skills generate the greatest changes in these conditions. Calderón-Gómez et al. mention how, in the Spanish labour market, there are a large number of jobs that include tasks related to ICT and other digital skills, while another group of workers do not use them and experience precarious conditions with few changes in occupational progression (Calderón-Gómez et al., 2020). In order to build on such findings and investigate their social implications in greater depth, this research opens up promising new avenues. For example, it could provide a tool to enable policymakers to know precisely which skills are associated with higher employability. Due to its low cost of implementation, this methodology makes it possible to provide updated information on the skills that are most employable on a permanent basis, enabling the implementation of training programmes, reducing digital gaps, and contributing to social mobility.

Additionally, by categorising jobs into occupations using classifiers such as O*NET or ISCO, it would be possible to generate valuable information regarding more specific occupational groups, for example, Software Developers, Electrical Engineering Technicians or Nursing Professionals, among others. By this means, analysing the shared variation of skills in high demand over time could provide insights into emerging skills, those that are becoming obsolete or skill mismatches in the labour market.

© Gianni Anelli, 2023

ACKNOWLEDGEMENTS
I would like to express my gratitude for the support of the SABE Project and the National Labour Observatory of Chile and SENCE for providing me with the data on online job advertisements that made this study possible.

REFERENCES


YM-D3pNKWA


ManpowerGroup (2023) Demand for Skilled Talent Persists for Q1 Despite Global Headwinds


