Job Role Description and Skill Matching in a Rapidly Changing Labor Market Using Knowledge Engineering

George Stalidis and Selini Kyriazidou

Abstract The job market is continuously evolving and suffers from skill mismatch, while digital platforms for job seeking, human resource management and training planning are widely adopting intelligent matching engines. Efficient job description and successful skill matching require standard terminologies for job positions and their requirements/qualifications, such as the 'European Skills, Competences, Qualifications and Occupations' (ESCO). The aim of this paper is to search whether ESCO—as the most representative job-related information model—has been adopted in the current rapidly evolving job market and the degree in which standardized job roles and their related skillsets are in line with the content found in the current job ads. Additionally, we intend to identify possible missing elements of this framework, towards its wider adoption and advanced skill-matching recommendation systems. As a representative case, the study was focused on selected IT professions in the Greek labor market. To this end, we applied a text mining process to 400 job ads, in order to capture the skillsets required by recruiting companies. The identified requirements for the selected job roles were used to model part of the Greek IT labor market. This model was then compared with the suggested requirements of the ESCO framework. It was found that the degree of matching between the skills in ads and the skills suggested by ESCO, is notably small and that the skills frequently requested in current IT job ads that were not included in ESCO, were mostly related to recently developed technologies and to soft skills.

Keywords Job description · Skill matching · Text mining · Knowledge modeling

1 Introduction and Background of Research

The job market is continuously evolving due to major advances in economy, society and technology, and seems to suffer from problems like skill mismatch, skill shortage, over qualification and issues with talent management [1]. At the same time, modern

G. Stalidis (✉) · S. Kyriazidou
International Hellenic University, 17th Km Thessaloniki-Sindos, Thessaloniki, Greece
e-mail: stalidgi@ihu.gr

A. Kavoura et al. (eds.), *Strategic Innovative Marketing and Tourism*, Springer Proceedings in Business and Economics, https://doi.org/10.1007/978-3-031-51038-0_21

digital platforms for job seeking, human resource management and training planning are widely adopting intelligent matching engines and data mining methods to automatically monitor the job market and to produce successful recommendations [2]. A generally accepted need for efficient job description and for successful skill matching is the global use of a standard terminology, in which job positions and their associated tasks, knowledge and skill requirements are unambiguously described. In their review, Rentzsch and Staneva [3] present the most important skills taxonomies, which are available today for this purpose, together with representative labor market-related applications. The 'European Skills, Competences, Qualifications and Occupations' (ESCO) (https://ec.europa.eu/esco/portal/home) is one of the most notable internationally influential job-related standardizations in digital form. It is an active European project which develops a dynamically refined multilingual classification of occupations, skills, competences and qualifications, that ensures a consistent understanding across Europe of these three dimensions and captures the links among them.

ESCO has already considerable impact and has been adopted as a basis by a considerable number of research and application development projects in the areas of labor market analysis, human resource management, skill-matching and job seeking services [3]. However, it is argued that companies do not make full use of standard terminologies, neither within their HR activities nor in their recruiting processes, leading to proprietary and ambiguous job advertisements. As an effort to investigate the reasons behind this, the question is raised whether the ESCO framework is expressive enough and up-to-date, in order to cover the needs of rapidly evolving disciplines such as Information and Communication Technology (ICT).

The aim of this paper is to search whether ESCO—as the most representative existing job-related information model—is applicable in the current job market and the degree in which its content (i.e. job roles and related skillsets) is in line with the content found in the current job postings. Additionally, we intend to determine if ESCO is up to date in a continuously evolving labor market. As a representative case of a rapidly changing and demanding area, we applied our research to the Information and Communication Technologies (ICT) field in the contemporary Greek labor market (post-Covid and post-economic crisis). More specifically, the research questions are:

- Which are the latest trends in the knowledge, skills and competencies that Greek companies in the ICT sector require?
- Which technologies are more frequently asked in job postings?
- In which degree do job descriptions in job ads and job seeking platforms make use of standard terminology?
- Is ESCO aligned with the current findings in job postings as regards the required skills in the ICT sector in Greece?

Within the work presented in this paper, we firstly applied a text mining process to a large number of job ads in the IT sector in Greece, in order to capture a "live" picture of the skillsets required by recruiting companies. The selected job roles and

extracted requirement terms (knowledge, skills and soft skills) were used as a basic model, which was then compared with the suggested requirements of ESCO.

2 Literature Review

2.1 The Contemporary Labor Market, Skill-Matching and the Need for Standardized Vocabularies

In the contemporary dynamic, globalized and digitalized labor market, important challenges are related to skill matching, i.e. the successful matching between employers and job seekers. Considering the enormous amount of positions and candidates, and the complexity of the elements to be considered, it is clear that there is high demand for intelligent and efficient computerized methodologies on behalf of companies and job search services. On behalf of the individual job seekers, modern services can analyze their professional profile and recommend opportunities, as well as indicate the knowledge gaps and skills they are lacking to succeed in their career [4]. The wide adoption of search engines and intelligent matching systems, demanded a standard terminology to be used globally as a solid basis for information exchange. The same is true for digital platforms for human resource management and training planning, where it is necessary to unambiguously describe job roles and to link them to skills, knowledge and qualifications in a consistent and machine understandable form.

On the level of public organizations, a crucial sector is skills Intelligence i.e. the monitoring and analysis of labor market, forecasting and identification of misalignment between job offer and demand. The size and complexity of the skill mismatch problem [1], gave an additional boost to efforts for standardization of terminology, consistency and semantic clarity [1, 5, 6]. Standard terminologies should be able to accommodate the changing requirements of knowledge and skills and be expressive enough to capture critical elements of job roles. At the same time, job role description should be consistent across different countries and cultures.

The current research work is focused on the labor market in Greece and, in particular, on the sector of Information Technology (IT). The specific field has been selected as a representative case of a rapidly evolving field, which is particularly active in job offering/seeking and with significant challenges related to talent management and skills shortage. The labor market of IT in Greece has similar structure with the corresponding international market. Although a relatively new sector, it has already been rapidly developed [7]. In addition to domestic companies, large international companies have established local departments, employing large numbers of employees with qualifications similar with those in the international market. It is also notable that the recent pandemic has introduced the element of tele-working, it has affected the working relations and influenced the mix of required skills [8]. According to a recent

survey by Mariani et al. [9], the most frequent requirements found in job openings in IT are: communication, involvement and responsibility, problem solving and analysis, team working, and the formal qualifications: previous work experience, knowledge of the English language, and education degree in related field.

2.2 Job Description Taxonomies and Standards

The standardization of job descriptions started several decades ago, as they were important for statistical studies, forecasts and policy formulation. The "Standard Classification of Occupations" (ISCO) (https://www.ilo.org/public/english/bureau/stat/isco/) of the International Labor Organization (ILO) is an important such classification still in use (latest version ISCO-08). The 'European Skills, Competences, Qualifications and Occupations' (ESCO) (https://ec.europa.eu/esco/portal/home) and the Occupational Information Network (O*NET) (https://www.onetonline.org/) are the most internationally influential job-related knowledge bases. From the one side they describe jobs (occupations) organized in hierarchies, from the other side they describe skills/knowledge elements and qualifications, and also provide links that inform which occupations are related to which Skills & Competences and Knowledge.

In this work, we focus on the European standard ESCO, rather than the American O*NET. ESCO is an active project which supports a dynamically refined multilingual classification that ensures a consistent understanding. It describes around 3000 occupations, 13,000 skill and competencies, and the relations among them in 27 languages. De Smedt et al. [5] describe the structure of ESCO, which covers three different pillars: (i) occupations, (ii) knowledge, skills and competences, and (iii) qualifications. In each one, concepts and terms are structured in hierarchies and given unique codes and identifiers on the semantic web, so that essential knowledge relevant to the labor market is represented in a standardized and machine readable form using the Linked Open Data (LOD) framework. In Fig. 1, the knowledge, skills and competencies associated to the occupation of ICT application developer are presented in tabular form. These terms are standardized and readable by computerized engines and additional information is provided e.g. for which other occupations is a skill essential.

ESCO has been used by multiple projects and researchers as a basis for job related terminology and as a knowledge base of the European labor market [3]. It has been used e.g. by Europass as a tool to record the skills and professional experience of third-country nationals in a standardised way, by the Organisation for Economic Cooperation and Development (OECD) as a basis for analyzing skills supply and demand, and by the openSKIMR project (https://openskimr.eu) [4] to inform their algorithms for skill profile analysis and individual career planning.

In their research work, given the rapid technological change, Chiarello et al. [10] emphasize the importance of the timely updating of ESCO. In order to investigate whether it is in line with the latest needs of industry 4.0, they measure the degree of

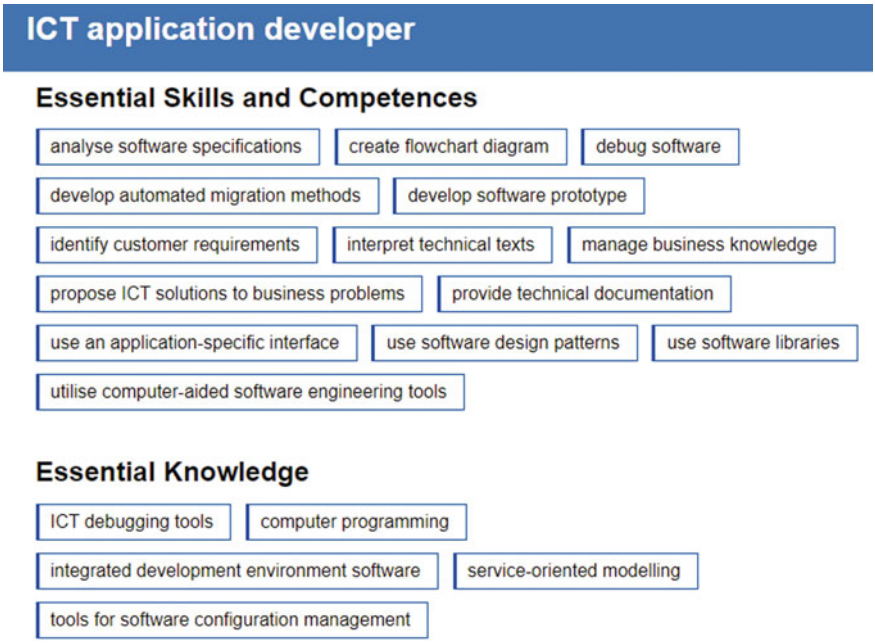

Fig. 1 The knowledge, skills and competencies associated to the occupation of ICT application developer (extract from ESCO page)

alignment between the ESCO taxonomy and the latest technological trends, as found in scientific literature. They concluded that most of the top-80 relevant technologies found in papers are also included in ESCO.

2.3 Technologies for Taxonomy Building and Skill Extraction

Ontologies are used widely as knowledge modeling tool to organize and describe job related knowledge. Their purpose is to allow advanced computerized matching based on semantics rather than keywords. Khobreh et al. [6] adapted the generic Neon ontology development methodology to an ontology-based framework for systematically expressing the knowledge, skills, and abilities which are prerequisites for jobs, further extending it to links with vocational education and training.

As regards the development of job framing standards, the two major approaches are the expert-driven and the data-driven ones. ESCO and O*NET are developed and maintained by domain experts, organized in working groups within the framework of large projects. Since this approach is highly demanding in human effort, considerable research is performed in automatic knowledge extraction from data, employing text mining and other machine learning methods [3]. The data-based approach has

been adopted by large companies in the field of labor market, such as Google and Linkedin, which create data-driven ontologies by mining large volumes of advertisements. The source in this case is the unstructured text from job ads, listings of job offers, comments in social media or even reports and literature. A knowledge engineering process identifies terms, recognizes their semantics and structures the resulting information by building appropriate knowledge bases. These constructed proprietary knowledge bases are usually components of a more complex knowledge model, which is linked to existing standard ontologies, such as the case of Cloud Talent Solution, developed by Google and used inside the job search feature of the Google search engine (https://cloud.google.com/blog/products/gcp/cloud-jobs-api-machine-learning-goes-to-work-on-job-search-and-discovery). Their approach was to link the O*Net ontology with their own proprietary one, which is built by extracting occupation descriptions and skills from hundreds of thousands of company websites. The methods reported in the literature for automatic extraction of job related terms and of building semantic-based knowledge bases vary from simple text mining processes to machine learning methods based on deep learning. Representative methodologies are presented in [2, 11], while a review of semantic mining methods is provided in [12].

3 Analysis Methods

The analysis performed within this work, included 3 parts: (a) Automatic extraction from job postings of job titles and the associated knowledge, skills and qualifications using text mining techniques, in order to capture the current labor market trends. (b) Retrieval from ESCO of the corresponding occupation title and related skills, expressed in standard terminology. (c) Comparison between the automatically captured picture and the content of ESCO, in order to investigate the degree of alignment.

The input data were postings published in the largest and most widely used job finding sites active in the Greek market, like indeed.gr, carrerjet.gr and jobfind.gr. The data collection was performed within the period June to October 2022, using a web crawling application. The keywords used were software application developer and programmer (in Greek). Filtering was applied to remove duplicates, resulting to a final number of analyzed postings of 403. The text mining process was implemented in RapidMiner Studio (https://rapidminer.com/) using the Text Processing extension. The process included tokenization (splitting the text into a sequence of tokens), decapitalization, filtering (removal of webpage tags, symbols and stopwords e.g. articles, and etc.), creation of n-grams (sets of more than one word), and frequency counting. The parametrization was optimized after some experimentation with the size of n-grams, the minimum frequency and the stopwords.

In parallel, we retrieved from ESCO the suggested knowledge and skills & competencies associated with the occupations of programmer and ICT application developer. These were retrieved in the standard coded terminology, as offered by ESCO, in

machine understandable form (JSON file). Both the essential and the optional knowledge and skills & competencies were retrieved. The comparison between ads and ESCO was performed by a software component (in Java) using the Sørensen–Dice method with the similarity coefficient adjusted to 70%. The results were observed and a limited manual curation was applied to correct some missed matches.

4 Results

The most frequent requirements for programmers and software developers in Greece were technical knowledge related to web development, mobile computing and databases. Considerable requests for soft skills were also found, such as team spirit, consistency, analysis and responsibility, as well as formal qualifications such as English language and higher education degree. The comparison between the skillsets included in ESCO and the ones found in ads showed that the degree of matching is notably small. Only 53% of the ESCO terms appear in ads, and many—even essential—ones do not appear at all (e.g. content development processes, compliance on cloud, Saas model, and others). Most of the matching terms appear in ESCO as optional (see Table 1). Furthermore, it was interesting that the skills missing from ESCO were mostly related to recently developed technologies and to soft skills.

Table 1 Frequency of ESCO terms in job ads

Knowledge term in ESCO	Type of knowledge/skill	Matching term in job ads	Frequency
Php	Optional	Php	234
Sql	Optional	Sql	224
Javascript framework	Optional	JavaScript framework	209
Css	Optional	Css	189
Java	Optional	Java	86
Python	Optional	Python	65
Software application environment	Optional	Software development	46
Android	Optional	Android	32
Ajax	Optional	Ajax	20
Ios	Optional	Ios	20
Cyberspace security	Essential	Cybersecurity	11

5 Discussion and Conclusions

An important finding from this study was that, although ESCO was widely used in research and by large platforms, proprietary free language was used in the majority of job ads posted in the IT area in Greece. The matching between the skillsets suggested by ESCO and the requirements frequently found in ads was smaller than expected. This misalignment was partly explained by the ambiguity of the free-style descriptions, unprecise terms and the mix of English-Greek language used in the Greek job postings. The implemented methods for analyzing job postings showed limitations and called for tuning and some manual intervention. Although more sophisticated methods have been reported [2, 11], it was clear that automatic skill extraction from free-style descriptions poses challenges and allows errors. This was a strong argument in favor of conformance to standard taxonomies. Considering the value of job description clarity for both companies and job seekers, it is suggested to employers and human resource managers to fully benefit from existing standards like ESCO.

Another finding was that the above misalignment was also attributed to substantial differences between the elaborate work of expert groups and "real time" data-driven updates. Thus, despite the dynamic nature of ESCO, an additional updating mechanism is indeed essential in rapidly evolving fields. It was therefore confirmed that all stakeholders involved in skill matching and labor market analysis can greatly benefit from a component for knowledge extraction from data and dynamic model building, in order to keep up with rapid changes.

References

1. Brunello G, Wruuck P (2021) Skill shortages and skill mismatch: A review of the literature. Journal of Economic Surveys, 35:1145–1167. https://doi.org/10.1111/joes.12424
2. Vankevich A, Kalinouskaya I. (2021) Better understanding of the labour market using Big Data. Ekonomia i prawo. Economics and Law 20:677–692. https://doi.org/10.12775/EiP.2021.040
3. Rentzsch R, Staneva M. (2020) Skills-Matching and Skills Intelligence through curated and data-driven ontologies. In: DELFI Workshops 46–57. https://doi.org/10.18420/delfi2020-ws-106
4. Mirski P, Bernsteiner R, Radi D. (2017) Analytics in human resource management the Open-SKIMR approach. Procedia Computer Science. 122:727–734. https://doi.org/10.1016/j.procs.2017.11.430
5. De Smedt J, le Vrang M, Papantoniou A. (2015) ESCO: Towards a Semantic Web for the European Labor Market. In: WWW2015 Workshop: Linked Data on the Web (LDOW2015)
6. Khobreh M, Ansari F, Fathi M, Vas R, Mol ST, Berkers HA, Varga K. (2015) An ontology-based approach for the semantic representation of job knowledge. IEEE Transactions on Emerging Topics in Computing 4(3):462–473. https://doi.org/10.1109/TETC.2015.2449662
7. Karamanis K, Beneki C, Ioakimidis M. Greek (2018) Labour market: The evaluation of minimum wage and unemployment during the period 2000–2017. Journal of International Studies 11(4):93–105. https://doi.org/10.14254/2071-8330.2018/11-4/7

8. Pouliakas K (2020) Working at Home in Greece: Unexplored Potential at Times of Social Distancing? IZA Discussion Paper No. 13408. https://doi.org/10.2139/ssrn.3636637
9. Mariani P, Marletta A, Mussini M, Zenga M. (2021) Professional profiles and job requirements in labour demand: An analysis of the Italian information technology sector. Social Indicators Research 156:651–667. https://doi.org/10.1007/s11205-020-02394-9
10. Chiarello F, Fantoni G, Hogarth T, Giordano V, Baltina L, Spada I. (2021) Towards ESCO 4.0–Is the European classification of skills in line with Industry 4.0? A text mining approach. Technological Forecasting and Social Change 173:121177. https://doi.org/10.1016/j.techfore.2021.121177
11. Konstantinidis I, Maragoudakis M, Magnisalis I, Berberidis C, Peristeras V. (2022) Knowledge-driven Unsupervised Skills Extraction for Graph-based Talent Matching. In: 12th Hellenic Conference on Artificial Intelligence pp 1–7. https://doi.org/10.1145/3549737.3549769
12. Ristoski P, Paulheim H. (2016) Semantic Web in data mining and knowledge discovery: A comprehensive survey. Journal of Web Semantics 36:1–22. https://doi.org/10.1016/j.websem.2016.01.001