ONLINE FIRST ARTICLE
doi.org/10.26398/IJAS.281

Statistica Applicata - Italian Journal of Applied Statistics

Received: 06.10.23/ **Accepted:** 01.03.24
Published online: 17.04.24

THE USE OF ONLINE JOB ADS TO ANALYSE SKILL CHANGES IN ICT OCCUPATIONS

Adham Kahlawi, Laura Grassini and Lucia Buzzigoli
Department of Statistics, Computer Science, Applications, University of Florence, Florence, Italy

Abstract. *Online job ads are a timely and detailed data source that can be used to get in-depth information about occupations requested by employers, together with the related skills, and to monitor the time evolution of skills demand in the different occupations, even at a detailed territorial level. The paper describes the more recent dynamics in skills demand in Italy for two innovation-related groups of occupations, using a measure of skill change between 2019 and 2021 at the regional level.*

Keywords*: Labour Market, Online Job Ad, Occupation, Skill, Skill Change.*

1. INTRODUCTION

Online Job Ads data (OJAs) is gaining popularity in labour market research due to its ability to provide valuable and timely insights into job offers and specific skill requirements across various levels such as territorial or sectoral contexts (Beręsewicz and Pater, 2021).

Of particular significance is the utilisation of OJAs to study and monitor the development of Information and Communication Technology (ICT) occupations and their related job markets, as they provide a channel that aligns well with the interests of potential candidates. Moreover, OJAs provide highly detailed and up-to-date information on specific and innovative skills, such as those in the ICT domain, which are crucial for employers (Aica et al., 2019).

Undoubtedly, ICT has had a profound impact on business processes, tasks, and organisations in various economic activities (López Cobo et al., 2020). The multifaceted roles of ICT professionals, including research, planning, information systems maintenance, and safeguarding data integrity and security, necessitate a wide range of specialised occupational skills. As technological advancements continue to occur at a rapid pace, these skills are becoming increasingly pervasive throughout the economy, making it essential to study the evolution of skills in these occupations. In Italy, the Observatory of Digital Skills confirms this trend and uses Online Job Advertisements (OJAs) to analyse the changing landscape of ICT professions in response to new technological trends. The Observatory produces annual reports to track these changes[1].

Eurostat's European survey on ICT usage and e-commerce in enterprises provides official statistics that shed light on the recruitment of ICT specialists in Italy and the European Union[2]. During 2012-2022, the number of employed ICT specialists in Europe increased by 57.8%, almost seven times higher than the increase (8.8%) in total employment. The estimate of employed ICT specialists in Italy as a percentage of total employment is 3.5 in 2019 and 3.8 in 2021, not far from the EU-27 values (respectively 3.9 and 4.5), but much smaller than in Northern countries, like Sweden (7.0 and 8.0) or Finland (7.6 and 7.4) (Eurostat, 2023a). On the other hand, the percentage of European Union enterprises recruiting or attempting to recruit ICT specialists in 2021 was 9.5% (4.9% in Italy), and 62.8% of these (61.3% in Italy) had difficulties in recruiting them (Eurostat, 2023b). Cedefop (2023) predicts a strong growth in the demand for ICT professionals up to 2035, while the employment shares of technicians are expected to decline.

The market for these occupations is, therefore, very dynamic, and it is expanding rapidly, also following the significant investments in the digital transition envisaged by the National Recovery and Resilience Plan[3] (PNRR) implemented in Italy through NextGenerationEU funds.

This paper uses the Lightcast dataset[4], which collects online job postings describing occupations and related skills with variables referred to official classifications. The aim is to analyse the more recent dynamics in skills demand in Italy for occupations with the highest concentration of ICT specialists and apply a measure of skill change between 2019 and 2021 at the regional level. Other research contributions have studied OJA data for Italy but are not concerned with comparisons of skill change across regions (among others: Lovaglio et al., 2018; Vannini et al., 2019; Kahlawi et al., 2023; Lucarelli and Righi, 2023).

[1] https://www.anitec-assinform.it/pubblicazioni/studi/. Last access December 2023.
[2] According to the EUROSTAT Glossary (last access May 2023) ICT specialists belong to the ISCO-08 occupation groups 133, 25, 35 and to other unit groups that involve the production of ICT goods and services.
[3] https://www.italiadomani.gov.it/it/home.html. Last access September 2023.
[4] Source: Lightcast™ 2022.

The study extends the line of research initiated by Giambona et al. (2021) in two directions.

Firstly, the temporal comparison involves the pre- and post-pandemic situations (2019-2021). As shown in Unioncamere and ANPAL (2022), the COVID-19 pandemic has forced Italian businesses to adopt and use digital technologies at a faster pace. This has increased the awareness of the need to rethink their business models and move towards greater digitalisation. The community and national-level measures to respond to the crisis have further accelerated the adoption of digital technologies. The Next Generation EU program has explicitly prioritised the digital transition. Despite this, some structural difficulties still hinder a broader digital transformation of the Italian production system. These difficulties are primarily related to production specialisation, governance, and company size. However, Italian companies have been progressing in integrating digital technologies into their business processes, especially recently. Therefore, we can expect a not well-defined picture and, consequently, a not well-defined pattern for skill changes.

Secondly, a deeper analysis of the skill change is presented, putting in evidence negative and positive variations at regional and occupation levels and exploring different types of skills. The results also show the different skill dynamics in two ICT occupation groups: *Professionals* and *Technicians*.

The structure of the paper is the following: the next section discusses the characteristics, pros and cons of OJA data for statistical labour market analysis. Section 3 presents a descriptive analysis of the Lightcast data. Section 4 focuses on methodology. Section 5 shows the results and a discussion, while the final section provides concluding remarks.

2. OJA DATA IN LABOUR MARKET INFORMATION SYSTEMS

Over the past decades, technological advancements, digitalisation, globalisation, and environmental concerns have continuously evolved the labour market (Arregui Pabollet et al., 2019).

In this ever-changing context, information that aids policymakers and decision-makers in tracking market trends and facilitating the alignment of labour demand and supply is precious. Online Job Ads data has garnered increasing attention due to its timely availability and detailed insights into job demand, which enrich traditional labour market data sources derived from official statistics. Furthermore, OJAs that include education requirements hold significant potential as supportive tools in higher education policy, facilitating the implementation of effective training programs (OECD, 2020).

OJAs are therefore considered a valuable component of Labour Market Information Systems, encompassing data on the size, composition, functioning, problems, opportunities, and employment-related intentions of the labour market and its participants (ILO, 2001) and employing intelligent approaches to support operational and decision-making activities (Mezzanzanica and Mercorio, 2019).

However, it is well known that the use of OJA data poses challenges related to their representativeness, reliability, and overall quality (Beręsewicz and Pater, 2021; Cammeraat and Squicciarini, 2021; Cedefop, 2019; ILO, 2020; Fabo and Kurekova, 2022; Napierala, 2022). Creating standardised datasets from individual job ads published online faces technical and contextual complexities, impacting the data's potential applications (Giambona et al., 2021). The need for data cleaning further adds to the complexity, while duplications of the same job announcement across multiple portals introduce noise into the final dataset. Moreover, the unavailability of standards for occupational skill descriptions leads to heterogeneity in employer-provided skills listings.

Notably, OJAs do not fully represent all job openings, as some vacancies may not be posted online. Certain countries tend to overrepresent high-skilled occupations and the sectors requiring them, while low-skilled jobs may be underrepresented (Carnevale et al., 2014; Schmidt et al., 2023). The prevalence of other recruitment methods, regional variations in internet usage for job advertisements, and varying employer preferences further contribute to representativeness issues.

Additionally, not all published job ads correspond to actual vacancies, with companies exploring the labour market without necessarily intending to recruit new resources. Comparisons between web vacancy postings and official vacancy statistics have been explored in previous studies (de Pedraza et al., 2017; Lovaglio et al., 2020), with OJAs also complementing classic survey-based approaches for economic statistics (Turrell et al., 2019). However, European Official Job Vacancy Statistics are only available at the national level, and for Italy, only rates are released. In any case, the OJAs remain a valuable source of information for studying specific skills associated with job demand with a granularity that cannot be deduced from official sources. That also justifies the interest Istat (the Italian National Institute of Statistics) recently expressed for this data type (Vannini et al., 2019; Lucarelli and Righi, 2023).

Contextual factors, such as the features of web job portal systems, the role of Public Employment Services, the level of digitalisation, and demographic and economic structures, are other essential issues to consider while analysing OJA data for territorial analysis. These factors may influence the scope and limitations of international and regional comparisons in countries marked by socioeconomic and digital divides (ILO, 2021). For instance, in Italy, the percentage of individuals using the Internet to look for a job or send a job application[5] is 13.6 in 2019 and 13.7 in 2021, in line with the EURO-27 average (respectively 15.63 and 13.42), but much less than the Nothern European countries (e.g., Denmark presents 37.6 and 36.3; Sweden 32.1 and 34.2).

On the other hand, the OJAs are the only available source that provides detailed information regarding the skills required by employers, although not included in official statistics.

Nowadays, there is great interest in the study of skills (see, above all, the European Agenda for Skills, a five-year plan to help individuals and businesses develop more and better skills and put them to use[6]). The anticipation of skills demanded by the labour market and the definition of consequent intervention strategies can help to identify appropriate measures in the fields of work, education, industrial and local development.

3. LIGHTCAST DATA

In this paper, we use the dataset provided for Italy by Lightcast, previously Burning Glass Technologies and Emsi Burning Glass, which collects online job postings deriving from numerous Internet sources (both online job portals and company websites) that have undergone a cleaning process to remove noise, outliers and duplications (Lightcast, 2022). The dataset contains about 70 variables describing the temporal (opening and closure date of publication) and geographical dimension (job location), as well as the economic activity of the company that posted the ad and the educational level required. Finally, the occupation is described by a list of related skills. Most variables refer to official

---

[5] EUROSTAT Data Browser. Online data code: ISOC_CI_AC_I. Last access 23/08/2023.

[6] https ://ec.europa.eu/social/main.jsp?catId=1223. Last access September 2023.

classifications (such as LAU and NUTS for geographical areas, ATECO2007 for economic activities), and skills are traced back to the European Multilingual Classification of Skills, Competences, Qualifications, and Occupations[7] (ESCO) taxonomy and the one used in the Occupational Information Network[8] system (O*NET).

For European countries, ESCO facilitates a more comprehensive and networked representation of the labour market, incorporating skills and knowledge dimensions in the definition of occupations (European Commission, 2019). Regarding occupations, the first four levels of the ESCO hierarchical classification of occupations (up to 4-digit ESCO level) coincide with those of the international standard classification of occupations ISCO-08 (ILO, 2012). Regarding skills, ESCO taxonomy distinguishes between two skill types: skills/competence concepts and knowledge concepts that are defined under the European Qualifications Framework: 'knowledge' means the outcome of the assimilation of information through learning (European Commission, 2019). On the other hand, 'skill' means the ability to apply knowledge and use know-how to complete tasks and solve problems (Council of the European Union, 2017). In ESCO taxonomy, action verbs are used when creating the terms for skills while learning outcomes (nouns) to specify what is known (knowledge). Moreover, each concept can be classified as 'occupation-specific', 'sector-specific', 'cross-sectoral' and 'transversal'[9]. This hierarchy, called skill reusability, refers to how specific a skill/knowledge is within occupations or economic sectors and is a relative concept because it depends on the relation between the skill/knowledge and a particular occupation.

Information in the Lightcast dataset is also organised on an internal taxonomy derived directly from OJAs (Magrini et al., 2023), and that, in the case of digital skills, enriches the ESCO taxonomy with skills tagged from Stackoverflow (in the following called Lightcast skills/knowledge) (O'Kane et al., 2020). The definition of digital and non-digital skills, which is relevant in our case, is not provided. Still, there is a generic reference to skills *from basic digital tools ('Excel') to advanced programming languages ('Python'), to skills and*

---

[7] This publication uses the ESCO classification of the European Commission.
[8] https://www.dol.gov/agencies/eta/onet. Last access September 2023.
[9] https://esco.ec.europa.eu/en/about-esco/escopedia/escopedia/skill-reusability-level. Last access September 2023.

abilities necessary to use these tools ('Computer literacy') and competencies that rely on digital tools to be carried out ('Data analysis') (Magrini et al., 2023).

Table 1 presents simple summaries of the Lightcast data in 2019 and 2021. In the following tables (unless otherwise specified), # skills refers to skill and knowledge concepts. In the two years, job ads have grown by 53.4%. Consequently, the number of skills requested increased considerably as well (64.4%), while the number of 2-digit and 4-digit ESCO occupations and unique skills remained almost unchanged. There are more 2-digit and 4-digit occupations in the ESCO taxonomy (44 and 426, respectively), but the differences are due mainly to the armed forces occupations. 1,245 unique skills occur in both years; 163 are in 2019, not 2021, and 169 are in 2021 and not in 2019.

The number of unique skills occurring in the dataset is much smaller than in the ESCO taxonomy (13,890), like in other countries (for Germany, O'Kane et al., 2020): some ESCO concepts are too broad to be specified in job postings, while others are too specific and are not used in job postings' language.

The average number of skills by job posting is slightly increasing (from 8.9 to 9.5), although the number of unique skills found is almost stable.

The proportion of Lightcast concepts is nearly 27% in both years.

Table 1: General information on Lightcast data. Years 2019 and 2021

			Avg.	ESCO occupation		#unique skills	
Year	#job ads	#skills	#skills	2-digit	4-digit	Total	Lightcast
2019	1,275,236	11,297,327	8.9	40	420	1408	397
2021	1,956,642	18,568,205	9.5	40	421	1414	379

Table 2 presents the number of job ads and requested skills for 2-digit ESCO occupations: *25 - Information and communication technology professionals* and *35 - ICT Technicians*, which are considered the ones with the highest concentration of ICT specialists.

The percentage increase in the number of job ads and the number of skills between 2019 and 2021 is lower than for the whole dataset (38.0% and 45.3% for ESCO25, 47.8% and 52.4% for ESCO 35). The job ads requesting ESCO 25 occupations are 7.4% of the total in 2019 and 6.6% in 2021 and contain many skills (20.9% of the total in 2019 and 18.5% in 2021). The average number of skills by job ad is therefore high (25.2 in 2019 and 26.5 in 2021).

The distribution of job ads and skills for ESCO 35 is much different: in both years, the number of job ads for technicians is about a fifth of those for professionals, while for the total of skills, the proportion is about 13%. Therefore, the average number of skills by job ad is smaller (16.2 and 16.7) but still much higher than the average calculated on the whole dataset. The higher number of skills for the professional segment is a characteristic finding of job advertisements, which also occurs in non-ICT occupations (Cedefop, 2023).

The distribution between the respective 3-digit ESCO occupations is strongly unbalanced in both years, with groups 251 and 351 prevailing on 252 and 352.

Table 2: Descriptive numbers of Lightcast data for 2-digit ESCO occupations 25 and 35. Years 2019 and 2021

2-digit	*25 ICT professionals*			*35 ICT technicians*		
3-digit	251	252	Total	351	352	Total
# job ads						
2019	81,921	11,973	93,894	17,946	561	18,507
2021	107,850	21,723	129,573	26,352	998	27,350
# skills						
2019	2,079,657	284,909	2,364,566	297,731	1,795	299,526
2021	2,912,534	524,271	3,436,805	452,677	3,861	456,538

In Table 3, which reports the names of the 2, 3 and 4-digit occupations, we identify 251 as *SW & applications developers & analysts* and 351 as *ICT operations and user support technicians*.

Table 3 also shows the differences between the internal 2-digit ESCO occupations 25 and 35 subgroups. In every subgroup, again, one 4-digit occupation prevails over the others with more than 50% job ads and total skills (e.g., *2512 SW developers* in group 251 and *3512 - ICT user support* in group 351). Regarding the average number of skills, as expected, all the occupations show values much greater than the general ones in Table 1, except the ones in group 352.

Table 3: Descriptive numbers of Lightcast data for 2-digit ESCO occupations 25 and 35 subgroups. Years 2019 and 2021

ESCO occupations	% job ads		% skills required		Avg # skills	
	2019	2021	2019	2021	2019	2021
2511 Systems analysts	31.2	34.6	23.8	26.6	19.3	20.8
2512 SW developers	54.8	51.0	60.4	56.8	27.9	30.1
2513 Web & multimedia developers	8.8	8.9	11.0	11.3	31.8	34.5
2514 Applications programmers	3.4	4.3	3.1	4.0	22.8	25.1
2519 N.e.c.	1.7	1.2	1.8	1.2	26.0	27.5
Total 251 SW & applications developers & analysts	100.0	100.0	100.0	100.0	25.4	27.0
2521 DB designers and administrators	7.2	3.8	8.4	4.3	27.6	27.3
2522 Systems administrators	62.9	58.3	59.0	53.5	22.3	22.2
2523 Computer network professionals	13.8	15.0	11.3	14.1	19.6	22.7
2529 N.e.c.	16.1	22.9	21.3	28.1	31.5	29.6
Total 252 DB and network professionals	100.0	100.0	100.0	100.0	23.8	24.1
3511 - ICT operations	20.3	20.4	16.5	15.4	13.5	13.0
3512 - ICT user support	60.4	60.9	55.4	55.0	15.2	15.5
3513 - Computer network and systems	6.0	5.5	7.1	7.3	19.6	22.8
3514 - Web technicians	13.2	13.2	21	22.3	26.4	29.0
Total 351 - ICT operations and user support technicians	100.0	100.0	100.0	100.0	16.6	17.2
3521 - Broadcasting and audiovisual	86.5	83.0	84.5	70.1	3.1	3.3
3522 Telecommunications engineering	13.5	17.0	15.5	29.9	3.7	6.8
Total 352 Telecommunications and broadcasting technicians	100.0	100.0	100.0	100.0	3.2	3.9

4. METHODOLOGY

The measure of skill change is the index proposed by Deming and Noray (2020). We applied it to understand if any changes in the required skills occurred between 2019 and 2021. By defining:

$$\Delta_{os} = \left(\frac{\#\ JobAds_{os}}{\#\ JobAds_{o}}\right)_{2021} - \left(\frac{\#\ JobAds_{os}}{\#\ JobAds_{o}}\right)_{2019} \tag{1}$$

the formula of the skill change index (SCI) for the single occupation *o* is:

$$SCI_o = \sum_{s=1}^{S} |\Delta_{os}| \tag{2}$$

where $\#\ JobAds_{os}$ is the number of job ads for occupation *o*, which require the skill *s,* and $\#\ JobAds_{o}$ is the number of job ads required by occupation *o.* In the present analysis, *o* refers to the most detailed level available, the 4-digit ESCO occupation, in groups 25 and 35.

This index measures the net skill change in each occupation. The higher the index value, the higher the skill change. The single addend assumes the value $\left(\frac{\#\ JobAds_{os}}{\#\ JobAds_{o}}\right)_{2019}$ when $\#\ JobAds_{o,2021} = 0$, and the value $\left(\frac{\#\ JobAds_{os}}{\#\ JobAds_{o}}\right)_{2021}$ when $\#\ JobAds_{o,2019} = 0$. As the index only shows the change in absolute value, we will decompose the total value to discover the contribution of the positive and negative addends in the summation.

Due to the peculiarities of the Italian labour market, which is characterised by a notable territorial specialisation, it may be helpful to report the index value by region to investigate any different geographical patterns in the change of skills. To this aim, we define:

$$\Delta_{rs} = \left(\frac{\#\ JobAds_{rs}}{\#\ JobAds_{r}}\right)_{2021} - \left(\frac{\#\ JobAds_{rs}}{\#\ JobAds_{r}}\right)_{2019} \tag{3}$$

and the formula of the skill change index for the region *r* is:

$$SCI_r = \sum_{s=1}^{S} |\Delta_{rs}| \tag{4}$$

where $\#\ JobAds_{rs}$ is the number of job ads in region *r* requiring the skill *s,* and $\#\ JobAds_{r}$ is the number of job ads in region *r*. Null denominators are treated as in formula (1). We use formula (4) to calculate two indexes separately for the 2-digit ESCO subgroups 25 and 35.

In both skill change indexes, there is also a composition effect. Specifically, the value of SCI_o is affected by the change in the distribution across regions of job ads requiring occupations o. Analogously, SCI_r is affected by the change in the distribution of job ads across occupations required by region r.

Finally, we try to express the heterogeneity of skill changes between occupations (still separately for the ESCO 25 and 35) and Italian regions through the variation ranges of the skill change $SCRO_s$ and $SCRR_s$ defined below:

$$SCRO_s = max(\Delta_{os}) - min(\Delta_{os}) \tag{5}$$

$$SCRR_s = max(\Delta_{rs}) - min(\Delta_{rs}) \tag{6}$$

5. FINDINGS AND DISCUSSION

In this application, unlike Kahlawi et al. (2022), we focus only on the unique skills requested for ESCO 25 and 35 occupations.

That said, the number of unique skills identified and for which the skill change index is computed are 869 and 389, respectively, for ESCO 25 and 35 (Table 4).

Table 4: ESCO and Lightcast unique skills/knowledge involved in the analysis. ESCO 25 and ESCO 35 occupations

Skill/	ESCO 25		ESCO 35	
knowledge	#	%	#	%
ESCO	548	63.1	284	53.2
Lightcast	321	36.9	105	46.8
Total	869	100.0	389	100.0

The proportion of unique Lightcast skills/knowledge concepts is higher than in the whole dataset (see Table 1), especially in group ESCO 35. In addition, more than 80% of unique Lightcast concepts are requested by ESCO 25 occupations. Those concepts are not directly traced to the ESCO taxonomy; most refer to specific software and applications labelled with their names (e.g., ABAP, Ubuntu, etc.), and we consider them as knowledge. After that, the number of knowledge among the ESCO own skills is just over 50% (54% in ESCO occupation 25 and 51% in ESCO occupation 35), while among the Lightcast skills/knowledge concepts, excluding the missing occurrences, the percentage is practically 100% in both groups. Therefore, we obtain Table 5, in which the percentage of knowledge type is remarkably high, especially in ESCO 25.

Table 5: Type of skills involved in the analysis. ESCO 25 and ESCO 35 occupations

Skill/	ESCO 25			ESCO 35		
Knowledge	Knowledge	Skill	Total	Knowledge	Skill	Total
Digital	337	71	408	155	52	207
Non-digital	269	180	449	87	88	175
Total*	606	251	857	242	140	382
Total %	70.7	29.3	100.0	63.4	36.6	100.0

* Skill type: 12 missing values for ESCO 25, 7 missing values for ESCO 35

In our opinion, the fact that knowledge is prevalent in the Lightcast taxonomy is not surprising since the job ads analysis algorithm tends to provide more specific information than the ESCO taxonomy, especially in the ICT domain. Here the purpose is to identify the tools and applications used in this area, which are numerous and constantly evolving. For example, the development and spread of cloud computing, data management services such as data protection and cyber security have enlarged the number of knowledge concepts in the digital area.

Figure 1 shows the skill changes SCI_o between 2019 and 2021 for the 4-digit ESCO occupations in groups 25 and 35. Note that each bar is decomposed to show the contribution of the positive and negative values to the skill change index. What is interesting is that, in many cases, the positives and negatives are almost balanced. The proportion of positive values remains between 42% and 65% for all occupations except two. Considering both the numerical contribution to the value of SCI_o and the proportion of positive values, there is no underlying trend towards growth (or decrease) in the demand for skills except for 3522 (*Telecommunications engineering*), which is, however, an occupation requested by a limited number of job ads and, for this reason, sensitive even to minor variations.

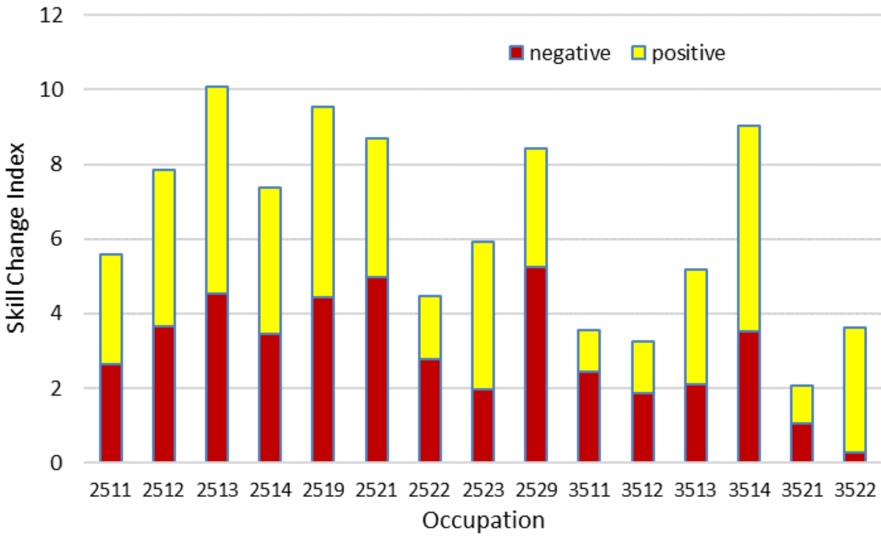

Figure 1: Contribution of negative and positive addends to SCI_o of ESCO 25 and ESCO 35 occupations. Years 2019-2021

We can observe that SCI_o values tend to be smaller for group 35 compared to 25, with one exception (occupation 3514). That likely occurs because of the different number of skills associated, as can be derived from Figure 2, which exhibits a remarkable positive association (correlation coefficient greater than 0.9) between SCI_o values and the average number of skills by job ad (see Table 3).

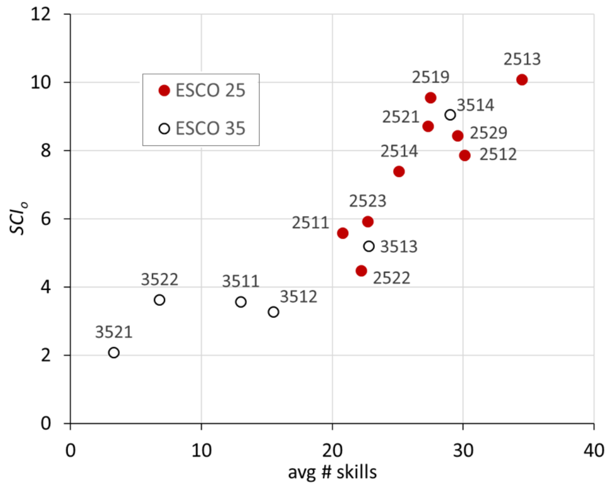

Figure 2: Relationship between *SCI₀* and the average number of skills by job ad for ESCO 25 and ESCO 35 occupations. Years 2019-2021

Figure 3 reports the values of SCI_r at the regional level still separately for occupations ESCO 25 and ESCO 35. Consistently with Figure 1, ESCO group 35 generally presents lower skill change in several Italian regions. Valle d'Aosta (VA) represents a peculiar case: it has the highest value in both occupations generated by only positive addends.

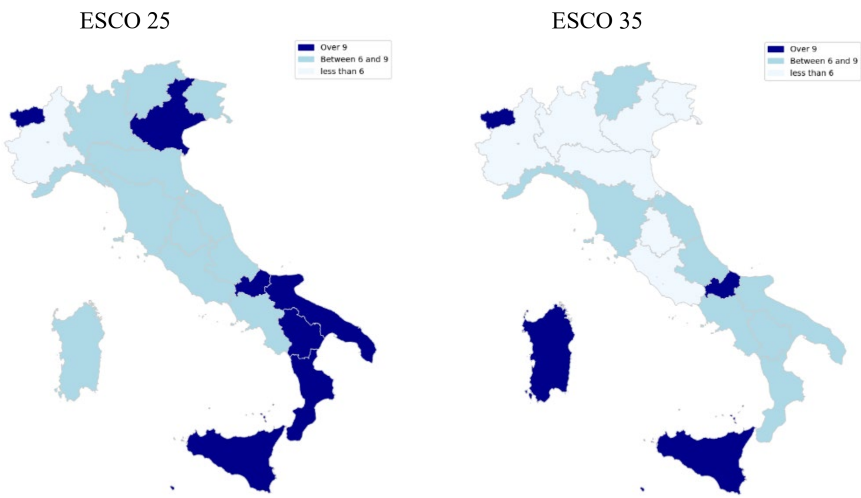

Figure 3: SCI_r for ESCO 25 and ESCO 35 occupations. Years 2019-2021

The well-known territorial divide of Italy in northern, central and southern regions does not occur in the maps: for instance, for ESCO 25, Veneto, with a value greater than nine, belongs to the higher class, like five southern regions, while for ESCO 35, some of the central regions belong to the lower class and others to the intermediate class.

If we also consider the number of unique skills in each region (the ones that are in common in the two years), we find a negative correlation with the SCI_r values for both occupations: -0.64 for ESCO 25 and -0.87 for ESCO 35. This could be due to the different skill weights in the region in relative terms, as in formula (3).

For ESCO 25, there are six northern/central regions where the job ads contain more than 50% of the total unique skills[10]: Lombardia (LOM), Veneto (VEN), Lazio (LAZ), Emilia Romagna (E-R), Piemonte (PIE) and Toscana (TOS). For ESCO 35, the six regions with the higher percentage are the same as before, but the maximum value (for Lombardia) is 26%. These six regions also have the highest total and ESCO 25 job ads (we refer to 2019 data, as in Kahlawi et al. 2022).

The following figures offer a more detailed view of the situation at the territorial level[11], with the decomposition of SCI_r to emphasise positive and negative changes. Even in this case, the negative and positive components of skill change persist in almost all regions. However, some exceptions exist: Valle d'Aosta (VA) shows all positive addends in both occupation groups, while in Veneto (VEN), negative values prevail in the ESCO 25 group.

Overall, there is a slight agreement between ESCO 25 and 35: the correlation coefficient between the two SCI_r series is 0.72.

[10] The percentage is calculated on the average of unique skills in 2019 and 2021 (see Table 1).

[11] The acronyms used to label the Italian regions are the following: PIE (Piemonte), VA (Valle d'Aosta), LOM (Lombardia), TAA (Trentino Alto Adige), FVG (Friuli Venezia Giulia), VEN (Veneto), LIG (Liguria), E-R (Emilia Romagna), TOS (Toscana), UMB (Umbria), MAR (Marche), LAZ (Lazio), ABR (Abruzzo), MOL (Molise), CAM (Campania), PUG (Puglia), BAS (Basilicata), CAL (Calabria), SIC (Sicilia), SAR (Sardegna).

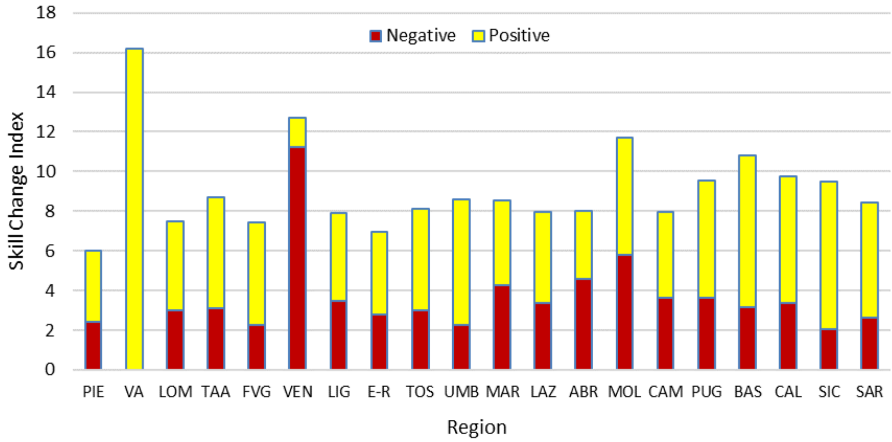

Figure 4: Contribution of negative and positive addends to SCI_r for ESCO 25 occupations. Years 2019-2021

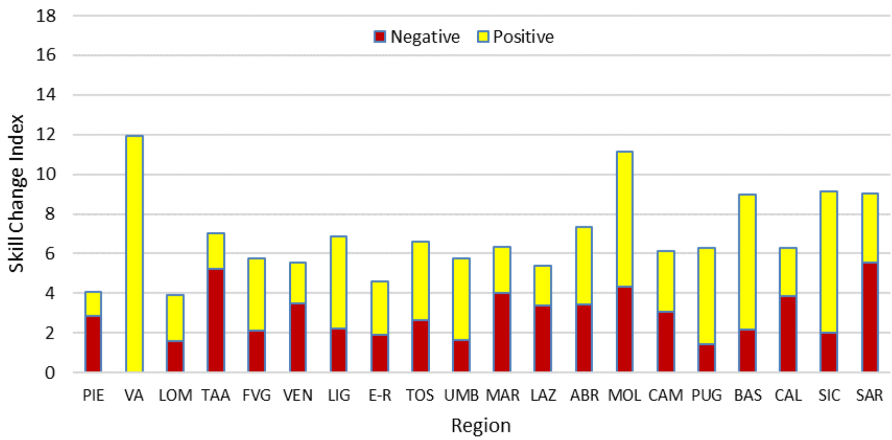

Figure 5: Contribution of negative and positive addends to SCI_r for ESCO 35 occupations. Years 2019-2021

Ultimately, we identify the skills with the highest skill change range across regions and occupations: Figure 6 represents each skill as a point in a Cartesian plane with coordinates $SCRR_s$ and $SCRO_s$ separately for ESCO groups 25 and 35. The lines of the 90th quantile of $SCRR_s$ and the 90th quantile of $SCRO_s$ divide the space into four quadrants. We focus on the skills in the first quadrant, where

$SCRR_s$ and $SCRO_s$ are above the 90th quantile: 49 for ESCO 25 and 25 for ESCO 35.

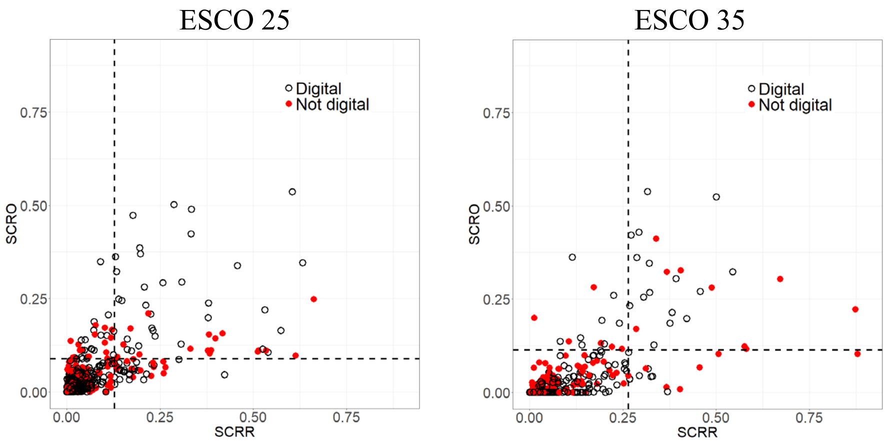

Figure 6: Skill change ranges for ESCO 25 and ESCO 35 occupations. Years 2019-2021

The skills in quadrant I are mostly digital (65% in ESCO 25 and 64% in ESCO 35), while knowledge concepts prevail in group 25 (55%) but not in group 35 (44%). The reusability level of each skill conditioned to the relative occupation is in Table 6: most of the skills are cross-sectoral or sector-specific, and only one is occupation-specific, confirming the relative proportions in groups 25 and 35.

Table 6: Reusability level of skills/knowledge concepts in quadrant I for ESCO 25 and ESCO 35 occupations

	ESCO 25		ESCO 35	
Reusability level	#	%	#	%
Tranversal	9	18.4	7	28.0
Cross-sectoral	15	30.6	8	32.0
Sector-specific	24	49.0	10	40.0
Occupation-specific	1	2.0	0	0.0
Total	49	100.0	25	100.0

To give an idea of the content of quadrant I, Tables 7 and 8 show the ten skills of quadrant I with the highest values for $SCRR_s$ and $SCRO_s$ separately for ESCO

25 and ESCO 35 occupations with the indication of the reusability level and type of the skill/knowledge.

Table 7: Skills/knowledge with the highest values for *$SCRR_s$* and *$SCRO_s$*. ESCO 25 occupations

Across regions				
$SCRR_s$	Digital	Description	Reuse level	Type
0.662	NO	*project management*	sector-specific	knowledge
0.632	YES	*use software design patterns*	sector-specific	skill
0.613	NO	*adapt to change*	transversal	skill
0.605	YES	*analyse software specifications*	sector-specific	skill
0.574	YES	*administer ICT system*	sector-specific	skill
0.539	YES	*business ICT systems*	sector-specific	knowledge
0.534	NO	*work in teams*	transversal	skill
0.531	YES	*computer programming*	transversal	knowledge
0.526	YES	*use a computer*	cross-sectoral	skill
0.514	NO	*teamwork principles*	cross-sectoral	knowledge
Across 4-digit occupations				
$SCRO_s$	Digital	Description	Reuse level	Type
0.536	YES	*analyse software specifications*	sector-specific	skill
0.501	YES	*unified modeling language*	sector-specific	knowledge
0.489	YES	*SQL*	sector-specific	knowledge
0.473	YES	*JavaScript*	sector-specific	knowledge
0.423	YES	*Java*	sector-specific	knowledge
0.386	YES	*process data*	cross-sectoral	skill
0.370	YES	*use spreadsheets software*	transversal	skill
0.362	YES	*implement front-end website design*	sector-specific	skill
0.349	YES	*SQL Server*	sector-specific	knowledge
0.345	YES	*use software design patterns*	sector-specific	skill

Table 8: Skills/knowledge with the highest values for *$SCRR_s$* and *$SCRO_s$*. ESCO 35 occupations

Across regions				
$SCRR_s$	Digital	Description	Reuse level	Type
0.879	NO	*assist customers*	cross-sectoral	skill
0.873	NO	*customer service*	cross-sectoral	knowledge
0.672	NO	*adapt to change*	transversal	skill

0.581	NO	*communication*	cross-sectoral	knowledge
0.576	NO	*work in teams*	transversal	skill
0.545	YES	*use MS office*	cross-sectoral	skill
0.507	NO	*team building*	occupation-	knowledge
0.500	YES	*use a computer*	cross-sectoral	skill
0.488	NO	*provide leadership*	cross-sectoral	skill
0.457	YES	*administer ICT system*	sector-specific	skill
		Across 4-digit occupations		
$SCRO_s$	Digital	Description	Reuse level	Type
0.538	YES	*CSS*	sector-specific	knowledge
0.524	YES	*use a computer*	cross-sectoral	skill
0.429	YES	*implement front-end website design*	sector-specific	skill
0.421	YES	*use markup languages*	sector-specific	skill
0.412	NO	*English*	transversal	knowledge
0.361	YES	*Java*	sector-specific	knowledge
0.346	YES	*business ICT systems*	sector-specific	knowledge
0.326	NO	*solve problems*	transversal	skill
0.323	NO	*create solutions to problems*	cross-sectoral	skill
0.322	YES	*use MS Office*	cross-sectoral	skill

6. CONCLUSION

The statistical analysis conducted in this work examines a measure of change in the demand for skills/knowledge for the ESCO 25 and 35 occupations. Although both refer to ICT, they are of different complexity: ESCO 25 (ICT professionals) is a group of occupations with a high density of skills/knowledge compared to ESCO 35 (ICT technicians) and includes a large part of the Lightcast knowledge concepts. A composition effect influences the skill change indexes SCI_o and SCI_r. SCI_o is affected by the shift in the distribution across regions of job ads demanding occupation o. Similarly, SCI_r is affected by the change in the distribution of job ads across the occupations within region r. In any case, the skill change is still evidence of a movement in the labour market at the level of the required skills.

The study at the 4-digit occupation level shows that the occupational profiles examined do not have a stable configuration regarding the changes in skill/knowledge demand, as the value of the skill change index is composed of an almost equivalent number of positive and negative variations, with a few exceptions.

A similar pattern occurs even at the regional level, with a heterogeneous picture of Italy. Still, no clear north-south divide emerges, and the results are difficult to interpret, given the numerous underlying factors that can influence the comparisons, as anticipated in Section 2. The study of this kind of data at the territorial level proves to be a complex task, as already pointed out in other empirical analyses (ILO, 2021; Kahlawi et al., 2022; Kahlawi et al., 2023, among others).

In any case, the analysis of ICT occupations seems to be a good training ground for assessing the information content of job ads. In particular, the skill perspective, derivable from OJAs, could be helpful to provide a more focused, comprehensive and detailed understanding of a context characterised by the rapid obsolescence of skills/knowledge and the emergence of new ones. In fact, OJAs specify occupations in terms of the required skills (verbs) and knowledge (nouns) according to the principles of ESCO taxonomy. However, the fact that most non-ESCO skills turn out to be knowledge concepts would suggest that the text of the job ads does not contain action verbs or, better, that the text mining algorithms simplify the statements, extracting only the keywords and decontextualising them. Still, the emphasis on knowledge contents is relevant in ICT occupations, where the tools to complete tasks continuously evolve. Furthermore, the reference to knowledge can be seen as a sound strategy to implement effective planning of education and training policies to limit the risk of skill mismatch.

Paper Funding: This work was supported by the Italian Ministry of University and Research (MUR), Department of Excellence project 2018-2022 - Department of Statistics, Computer Science, Applications - University of Florence.

REFERENCES

Aica, Anitec-Assinform, Assintel, Assinter Italia (2019). *Osservatorio delle competenze digitali*. https://www.assintel.it/assinteldownloads/osservatorio-delle-competenze-digitali-2019/. Last access: 19/01/2024.

Arregui Pabollet, E., Bacigalupo, M., Biagi, F., Cabrera Giraldez, M., Caena, F., Castano Munoz, J., Centeno Mediavilla, C., Edwards, J., Fernandez Macias, E., Gomez Gutierrez, E., Gomez Herrera, E., Inamorato Dos Santos, A., Kampylis, P., Klenert, D., López Cobo, M., Marschinski, R., Pesole, A., Punie, Y., Tolan, S., Torrejon Perez, S., Urzi Brancati, C. and Vuorikari, R. (2019). *The changing nature of work and skills in the digital age*. EUR 29823 EN. Publications Office of the European Union, Luxembourg.

Beręsewicz M. and Pater R. (2021). *Inferring job vacancies from online job advertisements*. Publications Office of the European Union, Luxembourg.

Cammeraat, E. and Squicciarini, M. (2021). Burning Glass Technologies' data use in policy-relevant analysis: an occupation-level assessment. *OECD Science, Technology and Industry Working Papers.* 2021/05.

Carnevale, A.P., Jayasundera, T. and Repnikov, D. (2014). *Understanding online job ads data: a technical report*, Georgetown University, Technical Report.

Cedefop (2019). *Online job vacancies and skills analysis. A Cedefop pan-European approach*, The European Centre for the Development of Vocational Training, Thessaloniki.

Cedefop (2023). *Skills in transition. The way to 2035*. Publications Office of the European Union, Luxembourg.

Council of the European Union (2017). *Council Recommendation on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning*. 2017/C 189/03.

de Pedraza P., Visintin, S., Tijdens, K. and Kismihók, G. (2017). Survey vs scraped data: comparing time series properties of web and survey vacancy data. *AIAS Working Paper*, 175, Universiteit van Amsterdam.

Deming, D.J. and Noray, K. (2020). Earnings dynamics, changing job skills, and STEM careers. *Quarterly Journal of Economics*. 135(4): 1965–2005.

European Commission (2019). *ESCO handbook*, Directorate-General for Employment, Social Affairs and Inclusion.

Eurostat (2023a). ICT specialists in employment. *Statistics Explained.* https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/47162.pdf. Last access: 07/12/2023.

Eurostat (2023b). ICT specialists – statistics on hard-to-fill vacancies in enterprises. *Statistics Explained*. https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/40327.pdf. Last access: 07/12/2023.

Fabo, B. and Mýtna Kureková, L. (2022). Methodological issues related to the use of online labour market data. *ILO Working Papers* 68. International Labour Office, Geneva.

Giambona, F., Khalawi, A., Buzzigoli, L., Grassini, L. Martelli, C. (2021). Big data analysis of Italian online job vacancies data. In: B. Bertaccini, L. Fabbris, and A. Petrucci, Editors, *Statistics and Information Systems for Policy Evaluation*. FUP, Firenze: 117-120.

ILO (2001). *The public employment service in a changing labour market*, International Labour Organization, Geneva.

ILO (2012). *International Standard Classification of Occupations*. International Labour Office, Geneva.

ILO (2020). *The feasibility of using big data in anticipating and matching skills needs*. International Labour Organization, Geneva.

ILO (2021). *Public Employment Services pressing ahead with digitalisation should be aware of the digital divide*. ILO Policy Note.

Kahlawi A., Buzzigoli L., Grassini L., Martelli C. (2022). Skill similarities and dissimilarities in online job vacancy data across Italian regions. In: A. Balzanella, M. Bini, C. Cavicchia, and R. Verde, Editors, *51st Scientific Meeting of the Italian Statistical Society-Book of Short Papers*, Pearson Italia: 284-291.

Kahlawi, A., Buzzigoli, L., Giambona, F., Grassini, L., Martelli, C. (2023). Online job ads in Italy: a regional analysis of ICT professionals, *Statistical Methods & Applications*, DOI: 10.1007/s10260-023-00735-9.

Lightcast (2022). *European Data Methodology*. Lightcast.

López Cobo, M., Rohman, I.K., De Prato, G., Cardona, M., Righi, R., Samoili, S., Vázquez-Prada Baillet, M. (2020). *ICT specialists in employment. Methodological note*, Seville: European Commission, JRC119846.

Lovaglio, P.G., Cesarini, M., Mercorio, F., Mezzanzanica, M. (2018). Skills in demand for ICT and statistical occupations: Evidence from web-based job vacancies, *Statistical Analysis and Data Mining*. 11:78–91.

Lovaglio P.G., Mezzanzanica, M., Colombo, E. (2020). Comparing time series characteristics of official and web job vacancy data. *Quality & Quantity*. 54: 85–98.

Lucarelli, A., Righi, A. (2023). Enriching job vacancy official information with online job advertisements: chances and limits. In A. Bucci, A. Cartone, A.Evangelista, and A. Marletta, Editors, *IES 2023 - Statistical Methods for Evaluation and Quality: Techniques, Technologies and Trend*s (T^3), Ed. Il Viandante, Pescara: 119-125.

Magrini, E., Pelucchi, M. and Brown, D. (2023). How is the digital transformation changing demand for skills in apprenticeships-typical occupations? Cedefop / OECD Symposium on apprenticeships and the digital transition, 15 - 16 June 2023.

Mezzanzanica, M., Mercorio, F. (2019). *Big data for labour market intelligence: An introductory guide.* European Training Foundation, Turin.

Napierala, J.; Kvetan, V. and Branka, J. (2022). Assessing the representativeness of online job advertisements. *Cedefop Working Paper*, No 17. Publications Office of the European Union, Luxembourg.

O'Kane, L., Narasimhan, R., Nania, J., Burning, Taska, B. (2020). *Digitalisation in the German labor market*, Bertelsmann Stiftung, Gütersloh.

OECD (2020). *What skills do employers seek in graduates? Using online job posting data to support policy and practice in higher education*. OECD Education Working Papers, No. 231.

Schmidt, J., Pilgrim, G. and Mourougane, A. (2023). What is the role of data in jobs in the United Kingdom, Canada, and the United States? A natural language processing approach. *SDD Working Papers.* No. 119, OECD.

Unioncamere and ANPAL (2022). *Le competenze digitali. Analisi della domanda di competenze digitali nelle imprese. Indagine 2022*. https://excelsior.unioncamere.net/sites/default/files/pubblicazioni/2022/B5-Competenze_Digitali_2022_DEF.pdf. Last access: 18/01/2024.

Vannini, I., Rotolone, D. and Di Stefano, C. (2019). Online job vacancies in the Italian labour market. *Statistica Applicata - Italian Journal of Applied Statistics*. 31(1): 97–118.