

RIVISTA DI
**POLITICA
ECONOMICA**
**L'IMPRESA DELLE COMPETENZE
I NUOVI SAPERI E IL LAVORO**

INTRODUZIONE STEFANO MANZOCCHI, ANTONIO RANIERI

Alfonso Balsamo
Rossella Cappetta
Serenella Caravella
Rosario Cerra
Emilio Colombo
Valeria Costantini
Francesco Crespi
Maurizio Del Conte
Cristina Grieco

Dario Guarascio
Giovanni Marin
Mirko Menghini
Joanna Napierala
Elena Paglialunga
Manuela Samek Lodovici
Renata Semenza
Francesco Trentini
Antonella Zuccaro

N. 1-2023


CONFINDUSTRIA

Rivista di
Politica Economica

Direttore: Stefano Manzocchi

Advisory Board

Cinzia Alcidi

Barbara Annicchiarico

Mario Baldassarri

Riccardo Barbieri

Leonardo Becchetti

Andrea Boitani

Massimo Bordignon

Marina Brogi

Elena Carletti

Alessandra Casarico

Stefano Caselli

Lorenzo Codogno

Luisa Corrado

Carlo Cottarelli

Sergio Fabbrini

Alessandro Fontana

Giampaolo Galli

Nicola Giammarioli

Gabriele Giudice

Luigi Guiso

Francesco Lippi

Francesca Mariotti

Marcello Messori

Salvatore Nisticò

Gianmarco Ottaviano

Ugo Panizza

Andrea Prencipe

Andrea Filippo Presbitero

Riccardo Puglisi

Pietro Reichlin

Francesco Saraceno

Fabiano Schivardi

Lucia Tajoli

Maria Rita Testa

Fabrizio Traù

Gilberto Turati

RIVISTA DI

POLITICA ECONOMICA

L'IMPRESA DELLE COMPETENZE
I NUOVI SAPERI E IL LAVORO

Introduzione pag. 5
Stefano Manzocchi, Antonio Ranieri

PARTE PRIMA

I NUOVI SAPERI

**Il cambiamento delle professioni come cambiamento
delle competenze: evidenze dagli annunci di lavoro
online in cinque paesi europei** » 17
Emilio Colombo, Francesco Trentini

**La diffusione delle competenze digitali in Italia e in
Europa: divari territoriali e sociali** » 39
Serenella Caravella, Rosario Cerra, Francesco Crespi, Dario Guarascio, Mirko Menghini

Domanda di competenze e transizione ecologica » 63
Valeria Costantini, Giovanni Marin, Joanna Napierala, Elena Paglialunga

PARTE SECONDA

UN SISTEMA DA RIDEFINIRE

**Formarsi tutti, formarsi sempre. Un sistema di formazione
di massa per la qualità e la produttività dei lavori** » 89
Rossella Cappetta, Maurizio Del Conte

**La responsabilità educativa nella *vision* e progettualità
delle imprese: focus su orientamento e ITS** » 115
Alfonso Balsamo

ITS Academy: uno strumento per costruire il futuro » 145
Cristina Grieco, Antonella Zuccaro

**Formazione, competenze e lavoro: il grande divario
di genere** » 167
Manuela Samek Lodovici, Renata Semenza

Il cambiamento delle professioni come cambiamento delle competenze: evidenze dagli annunci di lavoro online in cinque paesi europei

Emilio Colombo, Francesco Trentini*

- *In questo lavoro presentiamo un'indagine sul cambiamento delle competenze richieste per le professioni in cinque grandi paesi europei nel periodo 2019-2021. Il lavoro è condotto su dati degli annunci di lavoro prodotti dalla piattaforma WIH-OJA di Eurostat e Cedefop, utilizzando sia il dato grezzo in linguaggio naturale sia il dato classificato secondo standard ESCO. Il testo degli annunci di lavoro è usato per generare un modello linguistico specifico per il contesto del mercato del lavoro e testato sul task di classificazione delle competenze. Il modello consente di misurare la similarità delle nuove competenze osservate nel 2021 e le competenze esistenti nel 2019.*
- *Il panorama che emerge è caratterizzato da una significativa differenziazione tra diverse lingue e mercati del lavoro, con specificità nazionali in merito ai tipi di competenze che stanno cambiando maggiormente e alle professioni interessate. Al netto delle differenze linguistiche, le nostre analisi danno evidenza del fatto che la domanda di competenze nei mercati nazionali tende a diversificarsi nel tempo e questo effetto è più forte per le professioni poco qualificate.*

JEL Classification: J24, C45, C55.

Keywords: skill requirements, skill change, online job advertisements, NLP, word embedding.

* emilio.colombo@unicatt.it, Università Cattolica del Sacro Cuore di Milano e CRISP; francesco.trentini@unimib.it, Università di Milano-Bicocca, CRISP e LABORatorio "R. Revelli".

1. Introduzione

Un termine sintetico per catturare i fenomeni che stanno investendo il mercato del lavoro dei principali paesi avanzati, e dunque anche dell'Italia, è indubbiamente “cambiamento strutturale”; “strutturale” in quanto riguarda la struttura stessa della produzione e in particolare il rapporto tra impiego dei fattori di produzione (capitale e lavoro) e output.

I fattori di cambiamento (detti comunemente *megatrend*) sono stati ben identificati dalla letteratura e sono ascrivibili alla digitalizzazione e progresso tecnologico, alla globalizzazione, all'invecchiamento della popolazione e alla transizione verde.

La rivoluzione tecnologica non è circoscritta alla cosiddetta industria 4.0 e all'impiego di robot in sostituzione del lavoro manuale, ma si riferisce anche allo sviluppo dell'intelligenza artificiale con le implicazioni per il mondo dei servizi, coinvolgendo professioni che sino a pochi anni fa sembravano immuni dalla minaccia tecnologica.

La globalizzazione, nell'ambito produttivo, ha accentuato i processi di localizzazione quali *outsourcing* e *offshoring* con la nascita della cosiddetta *global value chain*. Nel mercato del lavoro ne è conseguita una diversa domanda di competenze nelle varie fasi del processo di produzione, generalmente tradotta in maggior domanda di occupazioni elementari e a bassa qualifica nei paesi meno avanzati dove vengono localizzate le attività a minore valore aggiunto e una maggiore domanda di occupazioni altamente qualificate nei paesi più avanzati, dove tendono a concentrarsi le attività a maggiore valore aggiunto. A seguito degli shock del Covid-19 e dell'invasione russa dell'Ucraina molte filiere si stanno riorganizzando con processi di *reshoring* aggiungendo ulteriore variazione della domanda.

L'invecchiamento della popolazione, oltre a sollevare il problema dell'obsolescenza delle competenze legato alla maggiore longevità della forza lavoro¹, aumenta in modo rilevante la domanda di alcune specifiche competenze come quelle legate alle attività di cura, riabilitative, ma anche di uso del tempo libero ecc.

Infine, il potente impatto che il cambiamento climatico ha sulle società avanzate e la spinta dei poderosi programmi di investimento europei indirizzati verso la transizione verde favorirà da una parte la nascita di nuove mansioni, e dall'altra lo sviluppo di competenze *green* che saranno richieste in modo trasversale da molte occupazioni².

¹ De Grip A., Van Loo J., “The Economics of Skills Obsolescence: A Review”, *The Economics of Skills Obsolescence*, Research in Labor Economics, 2002, vol. 21, pp. 1-26; Freeman R., “Is A Great Labor Shortage Coming? Replacement Demand in a Global Economy”, in *Reshaping the American Workforce in a Changing Economy*, DC Urban Institute Press, 2007.

² Vona F., “Labour Markets and the Green Transition: A Practitioner Guide to the Task-Based Approach”, *Report for the JRC-EU Commission*, 2021.

Utilizzando una terminologia cara agli economisti possiamo pensare che l'impatto di questi *megatrend* sul mercato del lavoro agisca su due livelli.

Un primo livello è costituito dal cosiddetto margine estensivo e ha a che fare con la creazione e distruzione di posti di lavoro. Tutti i fenomeni sopra descritti hanno l'effetto di contribuire sia alla scomparsa di alcuni lavori sia alla creazione di nuove professioni. Così come i robot svolgono mansioni sempre più sofisticate, e dunque sostituiscono l'uomo, la tecnologia consente lo sviluppo di nuovi lavori che non esistevano sino a pochi anni fa: si pensi ad esempio ai lavori connessi al settore della sicurezza informatica, e in generale a tutto ciò che riguarda la tumultuosa evoluzione dei social media o dei videogiochi.

Un secondo livello è costituito dal margine intensivo, che riguarda la trasformazione dei lavori esistenti. La tecnologia cambierà le competenze richieste per svolgere non solo i nuovi lavori ma anche quelli che continueranno a esistere. La crescente disponibilità di dati e l'uso sempre più diffuso di macchine e di strumenti sofisticati ha modificato notevolmente anche professioni basate sui rapporti personali, come quella del medico e dell'insegnante.

L'attenzione dei media e dei ricercatori si è rivolta soprattutto verso il margine estensivo anche in virtù della grande preoccupazione che hanno generato alcuni studi come quello di Frey e Osborne³ che hanno stimato che negli USA circa il 47% delle occupazioni sia a "rischio automazione", ovvero sono composte da *task* che possono essere svolti da macchine o da computer. Frey e Osborne sottolineano che le occupazioni caratterizzate da attività semplici e ripetitive siano maggiormente a rischio; tali occupazioni possono riguardare sia l'ambito amministrativo, sia attività manuali semplici che possono efficacemente essere svolte da robot e macchine. Al contrario, professioni ad alta qualifica che richiedono attività complesse, sia di carattere intellettuale (analisi, discernimento, valutazione), sia di carattere relazionale (si pensi al rapporto con i pazienti di un medico o di un infermiere), non sono considerate a rischio in quanto caratterizzate da *task* difficilmente sostituibili dalla tecnologia.

Nonostante le stime iniziali di Osborne e Frey siano state ridimensionate da studi successivi⁴, la distinzione tra attività ripetitive "a rischio" e attività intellettuali e relazionali "non a rischio" è rimasta. Recentemente le innovazioni nell'ambito dell'intelligenza artificiale e del *machine learning* hanno notevolmente cambiato questo quadro. Queste innovazioni, infatti, consentono a macchine e ad algoritmi

³ Frey C.B., Osborne M.A., "The Future of Employment: How Susceptible Are Jobs to Computerisation?", *Technological Forecasting and Social Change*, 2017, n. 114, pp. 254-280.

⁴ Nedelkoska L., Quintini G., "Automation, Skill Use and Training", *OECD Social, Employment Effective and Well-Targeted Adult Learning and Migration Working Paper*, 2018.

di svolgere attività complesse, che sino a pochi anni fa si pensava potessero essere svolte solo dalle persone. Nell'ambito sanitario le tecniche di analisi delle immagini hanno raggiunto un livello di precisione e di efficienza talmente elevato che un algoritmo è in grado di leggere una radiografia e riconoscere un tumore con un grado di precisione superiore a quello di un buon medico. La stessa attività di guida, nonostante sia svolta da moltissime persone, è una attività estremamente complessa, in quanto implica una serie di valutazioni e discernimenti articolati e sofisticati. Tuttavia, sappiamo che in un futuro prossimo le automobili saranno a guida autonoma, il che metterà a rischio alcune occupazioni come, ad esempio, quella degli autisti.

A fronte della notevole attenzione rivolta al margine estensivo, molto inferiore è stata quella dedicata al margine intensivo. Tuttavia, esso avrà un impatto maggiore sul mercato del lavoro poiché interesserà l'intero *stock* di occupati. Questa minore attenzione è anche spiegata da una difficoltà intrinseca di questo tipo di analisi, la costruzione di misure di cambiamento delle competenze richieste dalle occupazioni.

L'obiettivo di questo studio è fornire un contributo in questo senso, analizzando la direzione del cambiamento che le occupazioni subiscono in termini di composizione delle competenze per alcuni paesi europei. Un aspetto innovativo di questa ricerca è l'applicazione di metodi di intelligenza artificiale, come il *word embedding*, per condurre un'analisi economica. Misuriamo i cambiamenti negli insiemi di competenze delle occupazioni sulla base delle parole utilizzate per pubblicizzare le offerte di lavoro negli annunci di lavoro online, per misurare i cambiamenti tra e all'interno dei paesi europei, considerando la specificità delle occupazioni.

2. Dati e metodi

In questo studio vengono utilizzati i dati provenienti dal *Web Intelligence Hub* (WIH), una piattaforma creata da Eurostat per uniformare metodi e strumenti per la raccolta di dati dal web al fine di produrre statistiche ufficiali. Il caso di utilizzo più avanzato è costituito dal WIH-OJA, sviluppato congiuntamente con il Centro europeo per lo sviluppo della formazione professionale (Cedefop)⁵, che raccoglie in modo automatico gli annunci di lavoro pubblicati sul web (OJA, Online Job Advertisements). L'obiettivo del progetto è duplice. Da una parte arrivare a portare le OJA a essere parte delle statistiche ufficiali di Eurostat (le cosiddette *trusted smart statistics*)⁶, dall'altra analizzare

⁵ L'attuale piattaforma eredita ed estende il lavoro di raccolta e studio iniziato dal Cedefop, i cui risultati sono pubblicati nell'ambito dell'osservatorio sulla domanda di lavoro Skills-OVATE.

⁶ https://ec.europa.eu/eurostat/cros/content/trusted-smart-statistics-nutshell_en.

i fabbisogni di competenze e *skill* all'interno di una occupazione al fine di supportare le politiche nel mercato del lavoro.

I dati vengono raccolti da circa 500 siti web in tutta Europa, che includono tutti i principali attori del mercato del lavoro, pubblici e privati: siti web specializzati (*job board*), siti delle agenzie per il lavoro pubbliche e private, sezioni dedicate alle offerte di lavoro nei siti dei principali quotidiani nazionali⁷. Le offerte di lavoro pubblicate su questi siti vengono estratte tramite software dedicati (*web scraper*) che identificano le pagine contenenti le OJA e ne scaricano il contenuto. Le pagine vengono analizzate sintatticamente (*parsing*) e viene estratto il loro contenuto al fine di ottenere:

- il titolo dell'annuncio,
- la descrizione,
- l'area geografica a cui si riferisce,
- il settore economico dell'azienda richiedente.

Tutte queste informazioni sono normalmente presenti negli annunci anche se generalmente sparse in differenti posizioni delle pagine web in funzione della struttura della pagina di ciascun sito. Per questo motivo i *software scraper* sono dedicati a ciascun sito in modo da rispettarne e comprenderne la struttura. Successivamente, si applicano tecniche di validazione dei dati e deduplica degli annunci, riconoscendo gli annunci pubblicati più volte su siti differenti o sullo stesso sito attraverso analisi di similarità dei testi e altri parametri come, ad esempio, la data di pubblicazione.

Il titolo e la descrizione dell'annuncio vengono utilizzati per classificare gli annunci. Al fine di rendere la banca dati leggibile secondo convenzioni e, quindi, interoperativa, il dato viene classificato in riferimento alla tassonomia ufficiale dell'Unione europea, la *European Skills, Competences, Qualifications and Occupations* (ESCO), sia per le competenze sia per le occupazioni⁸. Il processo avviene attraverso l'utilizzo di algoritmi di *machine learning* seguendo l'approccio sviluppato da Boselli e coautori⁹. Questa attività è preceduta da un processo che utilizza un approccio denominato *bag-of-words*, che consente di estrarre parole o insiemi di parole (n-grammi) rilevanti ai fini della classificazione di un testo scritto in linguaggio naturale. Il processo sopra descritto prevede diverse fasi di validazione delle

⁷ Cedefop, "The Online Job Vacancy Market in the EU", *Cedefop Research Paper*, n. 72, Luxembourg, Publications Office, 2019, <https://data.europa.eu/doi/10.2801/16675>.

⁸ L'attuale classificazione delle occupazioni è fatta al quarto livello ESCO, che combacia con il quarto livello della classificazione ISCO08.

⁹ Boselli R., Cesarini M., Mercorio F., Mezzanzanica M., "Classifying Online Job Advertisements through Machine Learning", *Future Generation Computer Systems*, 2018, 86, pp. 319-328, <https://doi.org/10.1016/j.future.2018.03.035>.

conoscenze derivate dagli annunci, effettuate da esperti di dominio. All'interno del progetto WIH-OJA di Eurostat e Cedefop sono impiegati 32 esperti nazionali, uno per ogni paese UE in aggiunta a un esperto per UK, Islanda, Norvegia, Liechtenstein e Svizzera.

L'analisi che presentiamo in questo lavoro è basata sull'utilizzo, oltre che del dato classificato, anche del dato grezzo dell'annuncio, costituito dal titolo e dalla descrizione in linguaggio naturale. Abbiamo utilizzato un campione creato da Eurostat¹⁰, stratificato per lingua e per sette variabili presenti nelle OJA (occupazione, tipo di contratto, salario, livello di istruzione, esperienza richiesta, settore di attività economica). Per ogni cella sono state selezionate 50 OJA creando un campione di 9.262.005 osservazioni per 25 lingue. L'analisi contenuta in questo articolo è ristretta a cinque paesi: Francia, UK, Olanda, Germania e Italia.

Per quel che concerne il dato classificato, le OJA seguono, come già esposto, la tassonomia ESCO sia con riferimento alla classificazione delle occupazioni che alla classificazione delle *skill*. Per queste ultime, ogni *skill* è stata assegnata a un gruppo appartenente alla gerarchia ESCO v1.0.8.

La tassonomia ESCO consente di raggruppare le competenze in categorie per una migliore rappresentazione e analisi. La prima distinzione importante è tra *hard* e *soft skill*. Le *hard skill* sono tipicamente abilità e competenze specifiche, necessarie per svolgere un lavoro o un compito specifico (ad esempio, la conoscenza di un software o di strumenti specifici, abilità manuali specifiche ecc.). Le *soft skill*, invece, sono trasversali e si riferiscono alla capacità degli individui di interagire con gli altri e con l'ambiente (esempi sono le capacità di comunicazione, di risoluzione dei problemi ecc.).

Nell'ambito delle *hard skill*, distinguiamo ulteriormente tra *hard skill* digitali (HD) e *hard skill* non digitali (HND). Le competenze digitali consentono a un individuo di utilizzare gli strumenti ITC a diversi livelli, dall'uso, la manipolazione e l'interazione con strumenti ITC standard alla progettazione, l'implementazione e la distribuzione di sistemi e servizi ITC complessi. Le competenze *hard non digital* sono competenze professionali specifiche che caratterizzano un'occupazione e sono sufficienti a differenziare un'occupazione da un'altra.

La nostra classificazione si basa sui gruppi ESCO, ovvero competenze essenziali/facoltative, competenze linguistiche, competenze trasversali e competenze digitali, con una classificazione personalizzata che estende la classificazione ESCO per le competenze digitali e trasversali. Le *skill* linguistiche sono state escluse e abbiamo ristretto l'analisi

¹⁰ Il dataset è stato pubblicato come audit dataset `wih_oja_sample_v1_2021_r20220215`.

delle *skill* professionali (le cosiddette *hard non digital skill*) a quelle definite come essenziali o opzionali da ESCO¹¹. Inoltre, sono state escluse le occupazioni del sesto grande gruppo ISCO08 (personale specializzato addetto all'agricoltura, alle foreste e alla pesca) per mancanza di osservazioni. Sono altresì state escluse anche le occupazioni al quarto livello ISCO08 dei gruppi "non classificati altrove". La composizione del nostro dataset classificato è riportata nella Tabella 1.

Tabella 1 - Numero di OJA utilizzate nell'analisi

Paese	2019	2021	Totale
DE	5.642.670	6.711.163	12.353.833
FR	2.264.716	7.178.613	9.443.329
IT	876.058	2.776.858	3.652.916
NL	462.405	2.408.782	2.871.187
UK	2.627.806	8.125.521	10.753.327
Totale	11.873.655	27.200.937	39.074.592

Fonte: elaborazioni degli autori su dati WIH-OJA.

3. Misurare la similarità attraverso i *word embedding*

Il contributo principale del nostro studio è quello di fornire una metodologia e delle stime di similarità tra le competenze esistenti e quelle nuove per valutare il grado di innovazione che caratterizza i diversi profili professionali, utilizzando i *word embedding*.

Questo metodo presenta diversi vantaggi nel considerare i numerosi problemi che emergono nel confrontare la composizione delle competenze e il cambiamento delle competenze tra i diversi paesi e nel tempo. Il problema principale è costituito dalla dimensione linguistica. Ogni lingua ha infatti strutture sintattiche e semantiche peculiari, che sono idiosincratice e, metaforicamente parlando, definiscono la densità della lingua. Le parole sono quindi inserite (*embedded*) in una rete di relazioni con altre parole. Tale rete di relazioni viene catturata dai modelli linguistici. La seconda motivazione per l'uso di modelli di *word embedding* addestrati sulle descrizioni della OJA è quella di rappresentare le specificità non solo di una lingua ma del dialetto

¹¹ Le *skill* ESCO sono agganciate alle occupazioni al livello ESCO 0 che corrisponde al 5° o 6° livello ISCO08. Poiché le occupazioni nel WIH-OJA sono classificate al 4° livello ISCO08 sono state considerate le *skill* opzionali o essenziali per tutte le occupazioni contenute nel 4° livello ISCO08 corrispondente.

degli annunci di lavoro. In questo modo possiamo catturare in modo ampio la similarità tra le parole nel testo degli annunci, tenendo conto quindi della loro co-occorrenza e producendo una rappresentazione multidimensionale di questa rete di relazioni.

In questo studio utilizziamo questa rappresentazione per calcolare una misura quantitativa della distanza tra due competenze. Utilizziamo i nomi (le etichette ESCO) usate per classificare le *skill* negli annunci come input, valutando la distanza tra queste all'interno del modello linguistico. Il modello linguistico, quindi, offre una misura sintetica della relazione di prossimità tra le competenze rilevate negli annunci di lavoro. Poiché il modello è stato addestrato sull'intero *corpus* di annunci di lavoro in una lingua, le relazioni tra le *skill* sono catturate indipendentemente dai gruppi professionali¹². La similarità viene misurata utilizzando la *cosine similarity* tra ogni coppia di competenze.

In questo lavoro, abbiamo addestrato 100 modelli per ogni lingua e abbiamo selezionato il miglior *embedding*, seguendo i recenti lavori di Giabelli, Malandri, Mercorio, Mezzanzanica e Nobani¹³. Questo metodo permette di eseguire una valutazione intrinseca dei modelli per diversi *task* su diversi *benchmark* e genera un punteggio di valutazione globale. Il punteggio riassume le performance dei modelli di *word embedding*, fornendo all'utente una valutazione completa per la scelta del modello da utilizzare a seconda dei *task*. Nel nostro caso, il *task* per il quale è stato selezionato il miglior modello di *embedding* è stato la classificazione delle competenze.

Le misure che emergono dai modelli di lingue diverse non sono direttamente confrontabili, perché la densità dello spazio vettoriale all'interno del quale ogni modello linguistico è rappresentato varia da lingua a lingua. Nel nostro caso, questo renderebbe i risultati non comparabili tra lingue. Questo problema è ben noto in letteratura con il nome di *embedding alignment* ma, ad oggi, non esiste un consenso sulla migliore strategia per risolverlo. Una delle direzioni di ricerca più promettenti è l'*anchors alignment*. Questo metodo richiede di trovare diverse parole con la traduzione più esatta possibile con una lingua di riferimento, in modo che la corrispondenza tra queste parole fornisca un ancoraggio tra i modelli linguistici. La distanza tra le ancore all'interno di ciascun modello linguistico viene usata per determinare

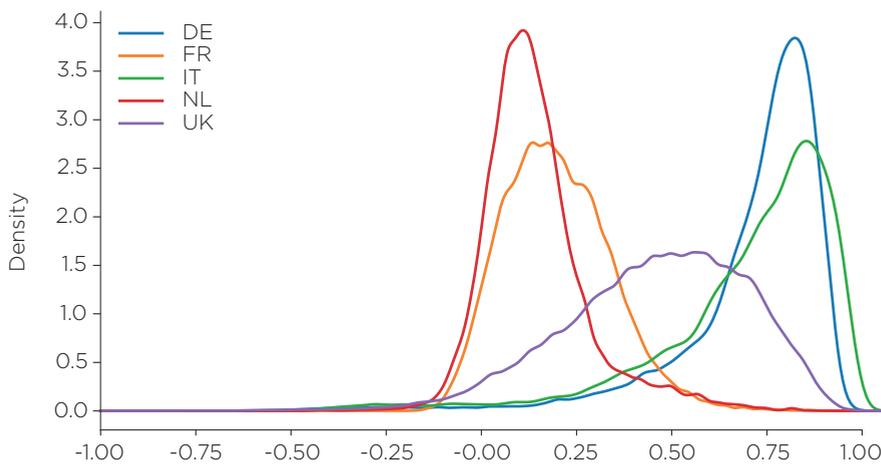
¹² Una interessante estensione sarebbe quella di addestrare i modelli linguistici per gruppi professionali per verificare le differenze nel grado di similarità di ciascuna competenza in contesti diversi. Questo processo richiederebbe una assunzione forte, non richiesta dal nostro metodo, cioè quella di prendere come data la classificazione delle professioni.

¹³ Giabelli A., Malandri L., Mercorio F., Mezzanzanica M., "WETA: Automatic Taxonomy Alignment via Word Embeddings", *Computer in Industry*, 2022, 138, 103626, <https://doi.org/10.1016/j.compind.2022.103626>; Giabelli A., Malandri L., Mercorio F., Mezzanzanica M., Nobani N., "Embeddings Evaluation Using a Novel Measure of Semantic Similarity", *Cognitive Computation*, 2022, 14 (2), pp. 749-763, <https://doi.org/10.1007/s12559-021-09987-7>.

la densità dello spazio vettoriale e viene utilizzata per standardizzarlo. Il risultato è costituito da modelli linguistici comparabili, assumendo che la corrispondenza tra le ancore sia esatta e corpi testuali sottostanti su cui vengono generate i *word embedding* siano coerenti. Purtroppo, questo approccio di ricerca è ancora nelle fasi preliminari e non consente di ottenere metodi utili ai nostri fini.

In questo articolo proponiamo un approccio statistico per superare questa limitazione, utilizzando un approccio in analogia con la standardizzazione fornita dal metodo di *anchors alignment*. Calcolando la *cosine similarity* tra le competenze nuove e quelle esistenti si ottengono distribuzioni di similarità diverse in ogni lingua, che possono essere utilizzate come *proxy* della densità dello spazio vettoriale specifico del *word embedding*. La Figura 1 presenta la distribuzione delle similarità tra le *skill* esistenti e quelle nuove osservate nelle OJA in tutti i modelli linguistici.

Figura 1 - Distribuzione della similarità tra *skill* esistenti e *skill* nuove per paese



Fonte: elaborazioni degli autori su dati WIH-OJA.

Nonostante si tratti di un test statistico formale, la figura mostra chiare differenze tra i paesi. La nostra proposta prevede quindi l'utilizzo di una misura di tendenza centrale, come la mediana, e successivamente il calcolo della distanza lineare di ogni osservazione da questo valore. La distanza relativa dal valore di riferimento fornisce, anche se in termini qualitativi, il grado di similarità tra *skill* esistenti e *skill* nuove (Tabella 2).

Tabella 2 - Similarità mediana tra *skill* nuove ed esistenti

Paese	<i>cosine similarity</i> mediana
DE	0.771
FR	0.187
IT	0.765
NL	0.122
UK	0.481

Fonte: elaborazioni degli autori su dati WIH-OJA.

4. Risultati

Le tabelle dalla 3 alla 7 mostrano i risultati per paese, occupazione¹⁴ e gruppi di competenze. I colori nella colonna “Similarità” sono utili per identificare i risultati qualitativi, dato che nei modelli linguistici non è possibile ottenere una corrispondenza esatta tra i valori cardinali dei vari paesi, nonostante la standardizzazione.

Considerando il conteggio delle competenze uniche (colonne “Conteggio competenze” e “Nuove competenze” di ciascuna tabella), osserviamo che esiste una elevata numerosità e variabilità di competenze uniche in ogni paese. Complessivamente, la maggiore varietà di competenze si osserva tra le occupazioni ad alta qualifica (grandi gruppi 1, 2, 3 ISCO08), seguite dalle occupazioni a media e bassa qualifica. Il Regno Unito è il paese con la maggiore varietà di competenze osservate. Tuttavia, il Regno Unito è anche il paese con il valore più basso in termini di richiesta di competenze nuove, in termini sia assoluti sia relativi per quasi tutte le occupazioni, ad eccezione del gruppo principale 9, che mostra un grado di cambiamento paragonabile a quello di altri paesi. Al contrario, i paesi con il minor numero di competenze uniche, ossia Italia e Paesi Bassi, si collocano ai primi posti in termini di variazione assoluta e relativa.

La composizione del set di competenze per tipo di competenze varia a seconda dell'occupazione a livello di paese. A questo proposito è pos-

¹⁴ Nel seguito verrà effettuata una distinzione tra i grandi gruppi professionali ISCO08: dirigenti (1), professioni intellettuali e scientifiche (2), professioni tecniche intermedie (3), impiegati di ufficio (4), professioni nelle attività commerciali e nei servizi (5), artigiani e operai specializzati (7), conduttori di impianti e macchinari, addetti al montaggio (8), professioni non qualificate (9). Dall'analisi sono stati esclusi le forze armate (gruppo 0) e personale specializzato nell'agricoltura, foreste e pesca (gruppo 6) a causa della bassa numerosità di osservazioni. Questi gruppi sono ulteriormente suddivisi in occupazioni ad alta qualifica (*high skill*, grandi gruppi ISCO08 da 1 a 3), occupazioni *medium skill* (grandi gruppi ISCO08 4 e 5) e occupazioni a bassa qualifica (*low skill*, grandi gruppi ISCO08 da 7 a 9).

sibile delineare alcune tendenze generali. Le professioni intellettuali e scientifiche (2) e le professioni tecniche intermedie (3) mostrano la più alta incidenza di competenze *hard digital* (HD) e *hard non digital* (HND), e il cambiamento nelle loro competenze è dovuto principalmente alle competenze HND. Al contrario, le occupazioni dei gruppi ISCO08 1, 4, 5, 7, 8 e 9 sono quelle che stanno cambiando di più in termini di competenze digitali. Alcune eccezioni degne di nota sono il Regno Unito, che in generale mostra un basso cambiamento nelle occupazioni ad alta qualifica, e Francia e Germania, che mostrano un alto grado di innovazione nelle occupazioni a media e bassa qualifica. È interessante notare che la Germania mostra anche un importante aumento della varietà di *soft skill* richieste per le occupazioni degli stessi gruppi.

Ci concentriamo ora sulla similarità tra le competenze esistenti e quelle nuove (colonna "Similarità"). L'analisi attraverso i *word embedding* consente di costruire delle misure basate sui rapporti tra le parole appartenenti al corpo testuale degli annunci. Pertanto, possiamo interpretare la similarità tra una competenza osservata nel 2019 e una competenza del 2021 mai osservata prima come il grado di novità di quest'ultima¹⁵. In termini di lettura di tali risultati, è importante considerare separatamente il confronto tra le competenze esistenti e quelle nuove all'interno dello stesso gruppo professionale, rappresentate dagli elementi diagonali di ciascuna sottomatrice 3x3.

Le tabelle mostrano che, nel complesso, la similarità tra le competenze esistenti e quelle nuove appartenenti allo stesso gruppo ha un valore superiore alla mediana; questo è conforme alle aspettative dato che ogni nuova competenza aggiuntiva è, nel complesso, più simile alle competenze dello stesso gruppo rispetto alle competenze di altri gruppi. Tuttavia, vi sono notevoli differenze tra professioni e paesi.

In Francia, ad esempio, si osserva la più alta somiglianza tra i gruppi. È opportuno sottolineare come le competenze HD per gli impiegati d'ufficio (gruppo 4 ISCO08), artigiani e operai specializzati (7) e i lavoratori delle professioni non qualificate (9) siano le più simili alle competenze HD già presenti, il che significa che il grado di novità di queste competenze è basso; al contrario, la maggiore novità si osserva nelle competenze S rispetto alle competenze HD e HND. Le professioni intellettuali e scientifiche (2) condividono con artigiani e operai

¹⁵ Presentiamo la deviazione relativa della similarità del gruppo rispetto alla mediana complessiva. Ne segue che i valori negativi rappresentano i casi in cui le nuove competenze hanno una minore similarità rispetto a quelle esistenti ossia sono caratterizzati da maggiore innovatività. Al contrario, i valori positivi, dove la gradazione del colore va dal giallo verso il rosso, indicano i casi di maggiore similarità semantica tra le nuove competenze rispetto alle competenze possedute nel precedente anno di osservazione. Inoltre, per mettere in risalto il risultato qualitativo abbiamo utilizzato gradi di colore: le combinazioni più innovative (con valori negativi) sono evidenziate in verde, mentre le combinazioni meno innovative (con valori positivi) sono evidenziate in rosso, passando per il colore giallo che evidenzia la combinazione a cambiamento in linea con l'andamento complessivo.

specializzati (7), conduttori di impianti macchinari e addetti al montaggio (8) un alto grado di novità delle competenze HND rispetto alle competenze di tutti i tipi. Le nuove competenze S tendono a essere altamente innovative rispetto alle competenze HD e HND esistenti.

In Germania, le nuove competenze S hanno un grado di innovazione inferiore rispetto alle competenze S esistenti. Al contrario, si osserva una tendenza delle nuove competenze HD e HND, che tendono a innovare rispetto alle competenze esistenti, soprattutto per le professioni ad alta qualifica.

In Italia, le competenze S presentano il grado più basso di novità rispetto a quelle esistenti per le professioni ad alta qualifica. Il più alto grado di innovazione si osserva per i gruppi a più bassa qualifica (gruppi 8 e 9), in particolare per le competenze HD con tutti i tipi di competenze, indicando una radicale innovazione per queste occupazioni. Una tendenza simile, anche se con un grado di intensità inferiore, si osserva per le professioni ad alta specializzazione e per gli artigiani e operai specializzati (7), a indicare che la digitalizzazione è una tendenza caratterizzante l'evoluzione dei profili professionali italiani e sta modificando in modo rilevante le competenze esistenti.

Nei Paesi Bassi, l'innovazione delle competenze nelle occupazioni a media e bassa qualifica è molto limitata, sia per le competenze HD che per quelle S, a differenza di quanto osservato in altri paesi.

Nel Regno Unito, nel complesso si registra un basso grado di innovazione tra le nuove competenze, che rafforza il risultato sul conteggio delle competenze uniche: il cambiamento non è solo limitato in termini di novità, ma anche il grado di connessione delle nuove competenze con quelle esistenti è alto. La maggiore novità è concentrata nelle competenze S. È interessante notare che per le professioni intellettuali e scientifiche (2) le competenze HND sono le più innovative, mentre, a differenza di altri paesi, le competenze HND presentano un basso grado di innovazione per gli impiegati di ufficio (4).

Tabella 3 - Francia: composizione delle *skill* e variazione 2019-2021

ISCO08-I	Gruppo	N. <i>skill</i>		Nuove <i>skill</i>		Similarità		
		2019	2021	N	%	HD	HND	S
1	HD	141	151	31	22%	16.2%	-17.7%	-14.7%
	HND	118	122	23	19%	-17.4%	5.0%	-2.5%
	S	120	102	13	11%	-4.1%	20.3%	15.5%
2	HD	227	216	10	4%	13.2%	-29.2%	-14.9%
	HND	253	273	43	17%	-32.2%	-7.4%	-23.0%
	S	129	119	20	16%	-8.1%	-17.8%	16.1%
3	HD	188	175	14	7%	15.2%	-15.2%	-6.8%
	HND	188	208	39	21%	-22.3%	8.5%	-10.8%
	S	116	113	12	10%	-9.9%	-6.0%	23.1%
4	HD	86	94	30	35%	43.5%	-0.3%	-12.5%
	HND	46	49	12	26%	-13.7%	11.4%	8.1%
	S	75	88	21	28%	-2.1%	27.1%	16.9%
5	HD	88	118	45	51%	10.4%	-18.4%	-17.2%
	HND	57	57	10	18%	-20.8%	22.7%	-5.1%
	S	79	94	21	27%	-8.7%	20.0%	12.2%
7	HD	77	89	29	38%	47.8%	-3.5%	-5.6%
	HND	65	83	21	32%	4.8%	23.1%	-18.2%
	S	63	68	17	27%	-2.3%	-13.1%	2.9%
8	HD	52	68	28	54%	6.2%	-25.4%	-30.6%
	HND	43	46	10	23%	-9.2%	16.7%	-8.4%
	S	71	83	21	30%	-4.5%	-33.3%	8.8%
9	HD	32	43	25	78%	32.3%	-21.5%	-19.7%
	HND	17	22	6	35%	-0.9%	48.0%	4.4%
	S	42	58	18	43%	0.5%	-32.0%	1.9%

Fonte: elaborazioni degli autori su dati WIH-OJA.

Tabella 4 - Germania: composizione delle skill e variazione 2019-2021

ISCO08-I	Gruppo	N. skill		Nuove skill		Similarità		
		2019	2021	N	%	HD	HND	S
1	HD	183	156	16	9%	-4.8%	-6.8%	-0.3%
	HND	168	140	13	8%	-7.3%	-3.1%	0.5%
	S	163	122	15	9%	-2.9%	0.3%	6.3%
2	HD	250	233	5	2%	-5.1%	-8.5%	-3.8%
	HND	356	322	28	8%	-8.8%	-7.2%	-6.6%
	S	186	133	9	5%	1.0%	-1.7%	5.8%
3	HD	211	188	7	3%	-4.5%	-8.2%	-4.8%
	HND	229	222	28	12%	-6.5%	-6.7%	-3.7%
	S	138	117	11	8%	0.7%	-1.4%	5.1%
4	HD	121	84	11	9%	2.4%	1.8%	0.2%
	HND	67	63	9	13%	-0.4%	3.2%	4.3%
	S	88	79	15	17%	3.8%	5.3%	7.1%
5	HD	139	105	17	12%	-1.5%	-2.1%	-0.9%
	HND	72	67	9	13%	-0.6%	0.5%	5.8%
	S	116	93	14	12%	1.1%	4.3%	6.3%
7	HD	131	116	19	15%	-1.8%	-4.3%	-3.4%
	HND	88	96	14	16%	-4.0%	-6.5%	-0.7%
	S	79	79	17	22%	2.1%	-0.6%	5.2%
8	HD	128	114	15	12%	-2.7%	-3.2%	-0.6%
	HND	64	67	11	17%	-7.3%	-2.5%	-3.4%
	S	102	90	13	13%	0.5%	2.3%	6.3%
9	HD	77	68	18	23%	-1.6%	-20.8%	-5.1%
	HND	19	22	5	26%	-0.3%	1.4%	3.6%
	S	46	55	17	37%	0.6%	-4.2%	5.0%

Fonte: elaborazioni degli autori su dati WIH-OJA.

Tabella 5 - Italia: composizione delle skill e variazione 2019-2021

ISCO08-I	Gruppo	N. skill		Nuove skill		Similarità		
		2019	2021	N	%	HD	HND	S
1	HD	103	91	13	13%	-5.0%	-3.9%	5.2%
	HND	112	105	15	13%	-0.2%	0.8%	4.3%
	S	101	88	12	12%	-2.8%	0.4%	4.0%
2	HD	217	210	16	7%	-7.6%	-4.4%	1.4%
	HND	252	252	32	13%	3.7%	1.2%	2.9%
	S	117	107	18	15%	-1.7%	-1.2%	2.5%
3	HD	163	162	18	11%	-8.1%	-5.3%	1.7%
	HND	163	182	33	20%	-2.4%	-2.0%	2.0%
	S	106	95	11	10%	-3.0%	-4.2%	2.3%
4	HD	53	66	30	57%	-1.3%	-4.2%	5.8%
	HND	51	49	5	10%	-2.8%	7.2%	2.1%
	S	67	75	14	21%	1.8%	-1.0%	1.1%
5	HD	40	49	27	68%	0.4%	-5.1%	4.0%
	HND	52	63	14	27%	2.5%	2.5%	4.2%
	S	66	73	17	26%	0.4%	-2.2%	-1.0%
7	HD	41	53	21	51%	-1.4%	1.7%	0.7%
	HND	64	76	18	28%	-2.7%	0.6%	3.7%
	S	44	57	19	43%	-2.2%	-1.7%	0.2%
8	HD	27	30	13	48%	-21.1%	-10.3%	-4.1%
	HND	38	55	20	53%	-7.3%	5.2%	3.9%
	S	54	61	18	33%	-6.2%	3.0%	0.2%
9	HD	9	11	6	67%	-16.9%	-33.9%	-16.4%
	HND	22	23	2	9%	2.6%	8.1%	6.5%
	S	30	40	15	50%	-8.4%	-15.6%	-1.0%

Fonte: elaborazioni degli autori su dati WIH-OJA.

Tabella 6 - Paesi Bassi: composizione delle skill e variazione 2019-2021

ISCO08-I	Gruppo	N. skill		Nuove skill		Similarità		
		2019	2021	N	%	HD	HND	S
1	HD	109	108	21	19%	4.3%	-22.1%	-16.0%
	HND	115	114	20	17%	-11.2%	10.0%	2.3%
	S	117	100	15	13%	-20.0%	16.5%	27.7%
2	HD	224	211	8	4%	6.4%	-28.6%	-22.2%
	HND	262	263	42	16%	-14.5%	6.6%	-21.6%
	S	143	107	15	10%	-17.0%	-15.5%	23.1%
3	HD	143	144	22	15%	25.2%	-17.9%	-17.4%
	HND	164	186	39	24%	-15.6%	16.1%	-11.7%
	S	109	102	14	13%	-22.0%	-10.3%	27.3%
4	HD	59	78	28	47%	24.1%	-15.4%	-7.5%
	HND	56	58	10	18%	-11.8%	13.5%	7.1%
	S	79	94	28	35%	-6.1%	7.1%	33.1%
5	HD	59	95	40	68%	11.5%	2.4%	-7.7%
	HND	45	55	12	27%	-22.1%	36.8%	10.9%
	S	80	97	30	38%	-18.4%	33.3%	32.2%
7	HD	67	74	23	34%	31.4%	-4.6%	-12.5%
	HND	63	74	15	24%	-5.5%	54.6%	-9.6%
	S	53	61	15	28%	-13.7%	-5.9%	35.6%
8	HD	63	81	34	54%	5.0%	-13.1%	-12.9%
	HND	44	48	8	18%	-8.1%	26.9%	0.8%
	S	96	92	21	22%	-23.5%	-14.2%	25.0%
9	HD	41	61	26	63%	33.9%	1.4%	-4.3%
	HND	12	16	5	42%	-6.4%	47.4%	17.3%
	S	50	63	20	40%	-23.9%	0.2%	34.4%

Fonte: elaborazioni degli autori su dati WIH-OJA.

Tabella 7 - Gran Bretagna: composizione delle skill e variazione 2019-2021

ISCO08-I	Gruppo	N. skill		Nuove skill		Similarità		
		2019	2021	N	%	HD	HND	S
1	HD	242	194	3	1%	-0.7%	1.2%	-5.6%
	HND	307	232	13	4%	-1.5%	14.2%	-0.6%
	S	384	268	7	2%	2.1%	7.8%	-1.6%
2	HD	277	259	1	0%	-0.2%	-6.0%	-5.4%
	HND	667	585	27	4%	-4.0%	-2.7%	-5.1%
	S	399	286	8	2%	4.3%	-2.0%	0.2%
3	HD	259	226	2	1%	1.0%	-5.8%	-5.0%
	HND	475	418	35	7%	-2.1%	2.9%	-3.4%
	S	394	284	6	2%	4.1%	-0.5%	0.2%
4	HD	154	143	20	13%	5.8%	17.5%	-2.4%
	HND	109	98	4	4%	4.7%	27.3%	-3.1%
	S	336	231	9	3%	4.8%	18.6%	-2.2%
5	HD	135	131	18	13%	1.1%	-1.9%	-4.3%
	HND	140	128	11	8%	-1.4%	-4.5%	-6.1%
	S	296	232	17	6%	5.8%	4.0%	-2.2%
7	HD	152	149	22	14%	-1.6%	-1.0%	-4.8%
	HND	163	163	23	14%	-0.6%	3.6%	-5.4%
	S	266	197	16	6%	1.5%	1.7%	-2.8%
8	HD	213	184	10	5%	-2.1%	-7.4%	-6.9%
	HND	114	118	16	14%	5.2%	9.1%	-3.1%
	S	325	220	11	3%	4.5%	1.9%	-2.2%
9	HD	70	87	29	41%	-4.5%	4.1%	-0.8%
	HND	39	52	17	44%	-4.6%	4.9%	-6.9%
	S	210	179	14	7%	0.9%	-0.7%	-1.4%

Fonte: elaborazioni degli autori su dati WIH-OJA.

Per approfondire meglio le analisi descritte nelle tabelle precedenti abbiamo implementato una semplice analisi multivariata in cui abbiamo regredito la misura di similarità tra le nuove competenze nel 2021 rispetto alle competenze del 2019 come funzione lineare di una serie di variabili esplicative. La distanza è calcolata con la *cosine similarity*: un maggior valore indica un minor grado di innovatività delle nuove competenze. Nel dettaglio, abbiamo stimato la seguente regressione OLS:

$$sim_{ij} = a + \beta_1 new + \beta_2 skill_group_i + \beta_3 skill_group_j + \beta_4 d_country + \beta_5 d_occupation + \varepsilon_{ij}$$

dove sim_{ij} identifica la similarità semantica (*cosine similarity*) tra una nuova competenza i (2021) e la competenza j (2019), new definisce la proporzione di nuove competenze nel 2021 rispetto alle competenze del 2019 all'interno dell'occupazione 4 livello ISCO; $skill_group$ definisce il raggruppamento di competenze (HD, HND, S) a cui appartengono le competenze i e j , e i termini $d_$ identificano delle *dummy* paese e occupazione (al 4° livello ISCO08).

La Tabella 8 ne riporta i risultati. Possiamo notare come in media la similarità sia inversamente correlata alla quota di nuove competenze (new): tanto maggiore è la quota di nuove competenze richieste nel 2021 rispetto al 2019 tanto minore è la loro similarità con le competenze del 2019. Questo suggerisce che le imprese tendono a espandere lo *skill set* richiesto verso competenze diverse rispetto a quelle precedentemente ricercate. L'interazione della nuove *skill* con il livello di occupazione (seconda colonna della regressione) rivela che questo effetto tende a essere concentrato nelle professioni meno qualificate rispetto a quelle qualificate. Una rappresentazione grafica di questo fenomeno è offerta dalla Figura 2 che mostra l'effetto marginale della variabile new per ogni livello del primo digit ISCO08. La spiegazione di questo fenomeno è da ricercarsi nel fatto che lo *skill set* delle professioni maggiormente qualificate è già ampio e diversificato e, dunque, le nuove *skill* tendono ad essere maggiormente simili a quelle esistenti. Viceversa, per le professioni meno qualificate, caratterizzate da una minore ampiezza dello *skill set*, le nuove competenze tendono a essere maggiormente eterogenee.

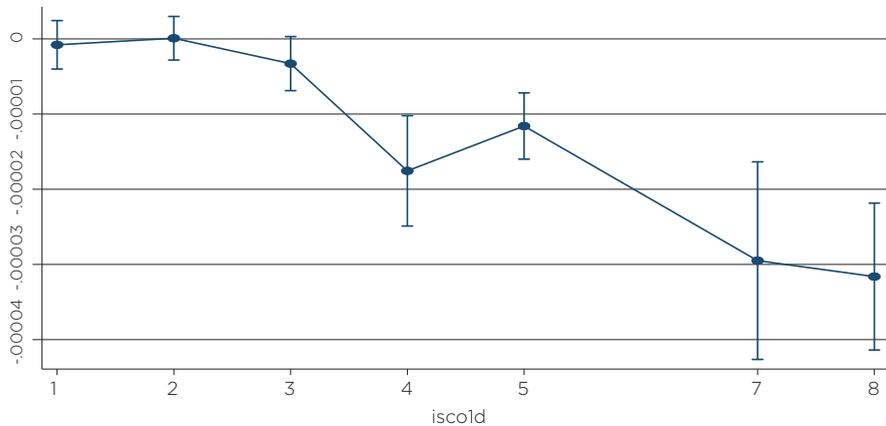
Le variabili relative ai raggruppamenti di competenze ($skill_2019$ e $skill_2021$) suggeriscono che rispetto al gruppo delle competenze digitali, che costituisce il gruppo di riferimento nella regressione, le nuove competenze tendono a essere più simili. Le *dummy* paese suggeriscono che la similarità delle nuove competenze, rispetto a UK, paese di riferimento, è maggiore in Germania e Italia, mentre è minore in Francia e Paesi Bassi. Infine, le *dummy* per livello di occupazione ($isco1d$) suggeriscono un chiaro schema, in quanto la similarità è più alta, con un più alto livello di qualifica. Per le occupazioni *high skill* le nuove *skill* sono più semanticamente simili a quelle esistenti rispetto alle occupazioni *low skill*. Questo risultato suggerisce quindi che le occupazioni a bassa qualifica siano soggette in media a un maggior cambiamento in termini di competenze rispetto alle occupazioni a più alta qualifica.

Tabella 8 - Regressione OLS del grado di innovatività delle nuove competenze su caratteristiche delle competenze, dell'occupazione e del paese*Anni 2019 e 2021*

Variabili esplicative	Regressione 1	Regressione 2
<i>new</i>	-0.0001***	0.0001***
<i>occ</i>	0.0001***	0.0001***
<i>new#occ</i>		-0.0001***
<i>skill_2019</i>		
HND	0.0134***	0.0134***
S	0.0352***	0.0352***
<i>skill_2021</i>		
HND	-0.00328***	-0.00332***
S	0.00705***	0.00705***
<i>country</i>		
DE	0.265***	0.265***
FR	-0.255***	-0.255***
IT	0.248***	0.248***
NL	-0.311***	-0.311***
<i>isco1d</i>		
1	0.140***	0.139***
2	0.113***	0.113***
3	0.106***	0.105***
4	0.105***	0.105***
5	0.0781***	0.0778***
7	0.0392***	0.0389***
8	0.0242***	0.0240***
Costante	0.273***	0.273***
R2	0.433	0.434
N	1.18m	1.18m

Fonte: elaborazioni degli autori su dati WIH-OJA.

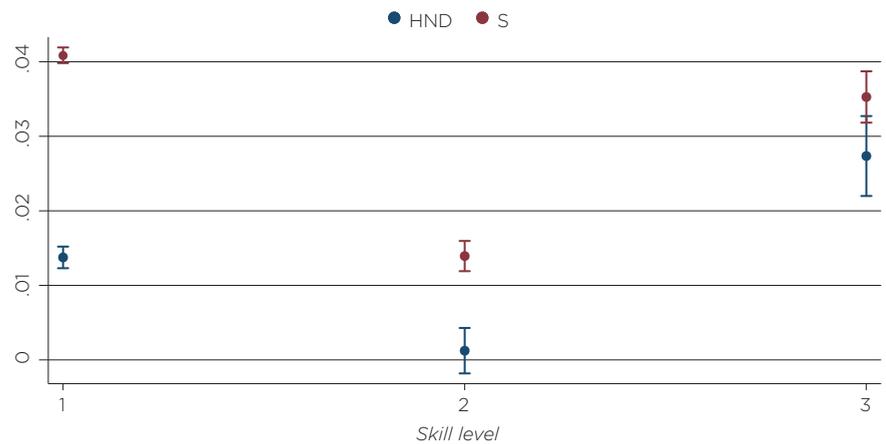
Figura 2 - Similarità marginale media delle nuove skill, per grande gruppo professionale, 1 digit ISCO08



Fonte: elaborazioni degli autori su dati WIH-OJA.

Infine, la Figura 3 mostra l'interazione tra il livello di qualifica delle occupazioni (1 *high skill*, 2 *medium skill*, 3 *low skill*) e il gruppo di competenze (HD, HND e S). Prendendo le competenze digitali come riferimento, la similarità semantica delle nuove competenze è più accentuata per le *soft skill* e *hard non digital*, in particolare per le professioni a elevata e bassa qualifica.

Figura 3 - Effetto marginale medio di livello di qualifica (skill level) e gruppo di competenze, in riferimento alle competenze digitali



Fonte: elaborazioni degli autori su dati WIH-OJA

5. Discussione e conclusioni

In questa sezione discutiamo in prima battuta le principali caratteristiche e limitazioni del dato degli annunci di lavoro online e i metodi di trattarlo, per poi esporre i risultati emergenti della nostra analisi del testo in linguaggio naturale.

Il dato raccolto dagli annunci di lavoro copre virtualmente l'universo della domanda di lavoro rivolta dai datori tramite canali formali ed espressa online. Esclude quindi dati sulle posizioni che il datore può coprire tramite canali informali o promozioni interne. Il dato raccolto sul web raggiunge un livello di dettaglio molto elevato, in quanto il datore di lavoro potenziale può esprimere tutti i dettagli qualitativi del lavoro offerto, includendo specifiche su competenze richieste e condizioni di lavoro. Rispetto ad altre fonti di dati, quindi, questa offre un vantaggio in termini di tempestività dei dati e di finezza dell'informazione. Allo stesso modo, il dato ha anche caratteristiche che ne restringono le possibilità di utilizzo. Gli annunci contengono sempre un elemento discrezionale, in quanto il potenziale datore di lavoro mette in evidenza i tratti del profilo occupazionale più caratterizzanti e utili al suo fine e non il profilo occupazionale completo, che viene spesso mantenuto implicito. Dunque, il dato mette in evidenza le competenze di più difficile reperimento o di maggiore rilevanza per la mansione da svolgere. Inoltre, il datore di lavoro promuove contestualmente l'immagine del proprio marchio, cercando di rendere appetibile il lavoro per ottenere un più ampio bacino di candidati. In aggiunta, il dato rappresenta unitamente la domanda di espansione (*expansion demand*) e domanda di sostituzione (*replacement demand*) della forza lavoro. Infine, negli annunci viene colta una varietà di tecniche produttive, con una variabilità interna alle occupazioni che dipende dalla dotazione di capitale e dall'organizzazione del lavoro da parte dello specifico datore di lavoro.

Questi elementi devono essere presi in considerazione nel trattamento e nell'interpretazione del dato perché ne circoscrivono le possibilità di utilizzo. Questo tipo di dato è molto adatto a catturare alcune dinamiche della direzione del cambiamento della domanda, ossia delle competenze che sono relativamente più richieste dal mercato. A livello aggregato tutto ciò si traduce in un generale orientamento della domanda di specifiche competenze. Dunque, si tratta di un dato diverso dalle realizzazioni di mercato, e che anzi ne anticipa la direzione di cambiamento con un elevato livello di dettaglio. In questo si rivela un dato utile sia per la ricerca, sia per applicazioni operative alle politiche attive del lavoro, per identificare le tipologie di formazione più interessanti per il mercato, sia per studiare nel dettaglio i profili di occupabilità delle persone in cerca di lavoro.

Questo contributo ha una duplice valenza. Da una parte illustra una metodologia basata sui *word embedding* per analizzare il testo de-

gli annunci e per costruire degli indici di similarità tra occupazioni e competenze. Dall'altra, l'applicazione di queste metodologie permette di analizzare alcune delle dinamiche tra paesi europei. In generale, osserviamo che, nel periodo a cavallo della pandemia da Covid-19, la domanda di competenze nei mercati nazionali si è diversificata. I profili che mostrano una maggiore innovazione sono quelli appartenenti alle professioni a più bassa qualifica, mentre il cambiamento nelle professioni a media e alta qualifica è maggiormente diversificato tra paesi e tipo di competenze in analisi. In Italia le nuove competenze digitali richieste sono molto diverse dalle competenze richieste precedentemente per tutte le occupazioni. In Francia e in Germania, diversamente, le competenze digitali cambiano soprattutto per le professioni ad alta qualifica e, in Germania, congiuntamente alle competenze caratterizzanti, le *hard non digital*.

Considerando infine le competenze *soft*, risulta interessante l'importanza che esse svolgono nel mercato inglese, considerato tradizionalmente innovatore in termini di processi produttivi. Il trend più importante in questo paese è proprio quello relativo a questo tipo di competenze, a fronte di una generale minore innovazione nelle competenze richieste.

Rivista di Politica Economica

La Rivista di Politica Economica è stata fondata nel 1911 come “Rivista delle società commerciali” ed ha assunto la sua attuale denominazione nel 1921. È una delle più antiche pubblicazioni economiche italiane ed ha sempre accolto analisi e ricerche di studiosi appartenenti a diverse scuole di pensiero. Nel 2019 la Rivista viene rilanciata, con periodicità semestrale, in un nuovo formato e con una nuova finalità: intende infatti svolgere una funzione diversa da quella delle numerose riviste accademiche a cui accedono molti ricercatori italiani, scritte prevalentemente in inglese, tornando alla sua funzione originaria che è quella di discutere di questioni di politica economica, sempre con rigore scientifico. Gli scritti sono infatti in italiano, più brevi di un paper accademico, e usano un linguaggio comprensibile anche ai non addetti ai lavori. Ogni numero è una monografia su un tema scelto grazie ad un continuo confronto fra l'editore e l'*Advisory Board*. La Rivista è accessibile online sul sito di Confindustria.

Redazione Rivista di Politica Economica

Viale Pasteur, 6 - 00144 Roma (Italia)

e-mail: rpe@confindustria.it

<https://www.confindustria.it/home/centro-studi/rivista-di-politica-economica>

Direttore responsabile

Silvia Tartamella

Coordinamento editoriale ed editing

Gianluca Gallo

Paola Centi

Adriana Leo

La responsabilità degli articoli e delle opinioni espresse è da attribuire esclusivamente agli Autori. I diritti relativi agli scritti contenuti nella Rivista di Politica Economica sono riservati e protetti a norma di legge. È vietata la riproduzione in qualsiasi lingua degli scritti, dei contributi pubblicati sulla Rivista di Politica Economica, salvo autorizzazione scritta della Direzione del periodico e con l'obbligo di citare la fonte.

Edito da:



Confindustria Servizi S.p.A.

Viale Pasteur, 6 - 00144 Roma