PSIHOLOGIJA, 2023, Online First, UDC
©2023 by authors DOI: https://doi.org/10.2298/PSI210817001I

How do we SEE Digital Platform Workers' Skill Patterns? Evidence from South-Eastern Europe*

Vladan Ivanović[1,2,3], Branka Anđelković[1], Tatjana Jakobi[1], Zoran Kalinić[1,2], Ljubivoje Radonjić[1,2], & Petar Čolović[4]

[1]Center for Public Policy Research, Belgrade

[2]Faculty of Economics, University of Kragujevac, Kragujevac

[3]CERGE-EI fellow

[4]Department of Psychology, Faculty of Philosophy, University of Novi Sad

The study aims to identify and describe the most relevant professional skill patterns among the digital platform workers from the selected Southeastern European (SEE) countries. Such orientation is based on the relatively modest presence of SEE countries in large pan-European studies, and on the lack of information regarding the applicability of existing online job taxonomies in observed countries. Applying a topic approach as a natural language processing technique, we analyzed the sets of self-reported skills provided by digital platform workers registered at the Upwork platform. Seven distinctive skill profiles were extracted, which only partly overlapped with the standard Oxford's Online Labour Index of digital job taxonomy.

Corresponding author: petar.colovic@uns.ac.rs

* This is an early electronic version of the manuscript that has been accepted for publication in Psihologija journal but has not yet been technically prepared for publication. Please note that this is not the final version of the paper as it has yet to be technically prepared for publication and minor changes to the text are possible before the final print. The final version of the article can be subjected to minor changes after proof reading and before final print. Please cite as: Čolović, P., Ivanović, V., Anđelković, B., Jakobi, T., Kalinić, Z., & Radonjić, Lj. (2023). How do we SEE Digital Platform Workers' Skill Patterns? Evidence from South Eastern Europe. *Psihologija*. Advance online publication. https://doi.org/10.2298/PSI210817001I

Results are indicating clear distinctions between highly specialized and general job categories, and between creative and technical professions. Mapping of the skills and national affiliations reveal differences between EU and non-EU countries in the region regarding professional inclinations. Partly in line with the findings of previous studies, the results pave the way for future research on this topic.
Keywords: online work, digital labour platforms, skills taxonomy, career boundaries

Highlights

- Professional skills of digital platform workers from nine South East European countries were examined.
- Skill patterns were derived from workers' public profiles on the Upwork digital platform.
- Seven skill clusters were identified using topic modeling.
- Skill distribution differs across SEE countries.

Pretext: The Rising Impact of Digital Labour Platforms

As the 4th Industrial revolution continues, disruptive technological change alternates the way people and societies organize themselves in order to meet personal and societal goals. While the impact of technology is visible in every aspect of life, society and economy, one of the most visible breakthrough transformations in the last few years is happening in the exorbitrary growing digital labour market. Around 163 million people have registered profiles on digital labour platforms, such as Upwork, Freelancer or Mechanical Turk, while 19 million have obtained work at least once via digital platforms (Kässi, et al., 2021). Although there are many professional, regulatory, security and societal challenges that accompany digital work

(Ozkan-Ozen & Kazancoglu, 2021; Englert et al., 2019; Zadik, et al., 2019; Donini et al., 2017), the pandemic of Covid-19 just sped up the change toward online work, making platform work an evermore attractive choice both to employers and employees.

Beyond turbulent structural change in the economy and in particular the labour market, digital platforms mean a completely new environment in which individuals are seeking and obtaining jobs. A global, transparent and very competitive environment made workers' set of skills crucial for professional success. Not only that the rise of digital platform work threatens traditional, long term, employment and work based on strict contractual regulation of labour relations, but also in new (digital) work arrangements an emphasis is more on the right set of skills, their interdisciplinary nature and altering professional profiles of workers. New competencies, mirroring new knowledge, skills and abilities are the prerequisites for a successful career development in a platform work (Mesquita et al., 2019). The changes digital workers are confronted with are ubiquitous, the emphasis is on a combination of general and technical skills, with a growing importance of higher-order skills. It is dramatical change with the dominant practice in the past where narrower technical skills and high specialization were dominantly demanded. Moreover, rapidly changing technology changes the skills being rewarded on the labour market (Ranđelović & Jandrić, 2018), which creates strong incentives for digital workers to constantly improve and upgrade the set of skills they possess - making them accurate, technically sophisticated and transdisciplinary. In order to exploit opportunities created by digital transformation and to find technology augmented "jobs of tomorrow" (WEF, 2021), digital platform workers are continuously under the pressure to reskill and upskill, which makes the boundaries between different professions less clear and more fluid.

The Relevance of Digital Skills

During its almost three-decades-long presence in literature, the concept of boundaryless career has been steadily evolving (Sullivan & Arthur, 2006). While its meaning is multifold, its defining feature is flexibility in vertical and horizontal professional mobility (Johns & Gratton, 2013; Sullivan & Arthur, 2006). Hence, the tendency is to see the career change as a permanent process inherent to professional life rather than a sequence of events (Arthur & Rousseau, 2001, 2001; Koch et al., 2021). A reconceptualization (Sullivan & Arthur, 2006) postulated that highly sustainable career trajectories are built around two continua: physical and psychological mobility. -They interact with three "meta-skills," named "knowing-why," "knowing-how," and "knowing-whom." To clarify their model's structure, Sullivan & Arthur (2006) describe the prototypical "profiles," a.k.a. the four quadrants defined by basic dimensions.

Such a conceptualization highlights the crucial role of skill development in a knowledge-based economy (Anđelković et al., 2019), the relevance of skill development for successful career management, lack of information about the structure and frequency of skill patterns in observed SEE countries, and, not unimportantly, possible effects of the COVID-19 pandemic, appear to constitute a sufficiently convincing rationale for the exploration of the current state of affairs regarding skills of digital platform workers' skills in the SEE countries. In this study, we consider it from the viewpoint of the "boundaryless career" concept.

Hence, it seems that "know-how" competencies (a.k.a. specialized professional skills) are the key drivers of career success in the digital realm. This is reflected in several recent studies that have attempted to devise taxonomies of online platform work skills based on specific methodological procedures (e.g., Djumalieva & Sleeman (2018); Bourgeon et al., (2019), Giabelli et al., (2020)). Djumalieva and Sleeman (2018) suggest a tentative taxonomy of six hierarchically organized clusters, whereby the broadest are named "Education, sales and marketing", "Information technology", "Science and research", "Engineering, construction and transport", "Health and social care", and "Business administration." Bourgeon et al. (2019) study is not based exclusively on digital skills, instead proposing three broad skill categories (soft, job-specific, and digital), divisible into thirty-five specific categories, and conceptually rooted in the previous literature (the authors cite e.g. Cedepof, 2017, and Burning Glass Technologies, 2018). Giabelli et al. (2020), demonstrating their newly proposed NEO methodology, point to forty-three emerging occupations in the digital domain. While the three studies only partially agree regarding their findings, one could outline non-negligible methodological similarities. Such points of agreement include relying on online data sources, and using natural language processing techniques to mine textual data. This may point to text-oriented analytic strategies as the way to explore a highly diverse and rapidly developing field such as digital skills. Additionally, the findings mentioned above suggest that conceptual consensus on the number and content of digital platform workers' skills has not been reached yet.

One of the current skill classification systems, namely The Online Labour Index (OLI) taxonomy (Kassi & Lehdonvirta, 2016, 2018) is essential for this study, as it has been used in the studies that involved online platform workers in countries we observe in our study. OLI's system of digital work skills comprises six classes, which can be regarded as professions: software development, creative services and multimedia, writing and translation, clerical work and data entry, sales and marketing, and professional services. A somewhat more general taxonomy contains the classes of professional (e.g., creative, translation, and software jobs) and non-professional services (e.g., clerical work and sales) (Pesole, et al., 2018). This approach has been used in high-profile applied econometric studies (e.g., Pouliakas & Branka (2020)), while the OLI taxonomy (if somewhat expanded) was successfully applied in seminal European studies on digital and platform work (e.g., Urzì Brancati et al., (2020).

Current Study

While geographical boundaries do not seem to be an issue in digital work (Koch et al., 2021), they seem to matter in the SEE region, where online platform work has been thriving in recent years (Anđelković et al., 2019; Radonjić, 2020.). Only recently, the development of specialized online resources has enabled monitoring of digital work on a systematic basis in the selected SEE countries: Serbia, Croatia, Bosnia and Herzegovina, Northern Macedonia, Hungary, Montenegro, Albania, Bulgaria, and Romania. The "Gigmetar" project results

(Gigmetar Team, 2021), within which standardized and special technical reports are produced on a regular basis, suggest that, by OLI taxonomy, creative and multimedia services feature most prominently in all countries, followed by software development and sales and marketing support. The distribution of skill patterns, as conceptualized by OLI, appears to be strikingly similar across the region.

Gigmeter monitoring results (concerning the OLI skill categories) called for further research. Namely, discordances were noticed among within-country results and figures in the overall sample. One example is the second most frequent "profession", software development, which has a 26.4% share overall, while its within-country distributions span from approximately 21% to approximately 32%. To our understanding, such results may suggest that national specificities may have more impact than it has been assumed so far. At the same time, one -may see this as a challenge for OLI taxonomy, but also as an important step in getting to know the region - specific features of skill structures. In addition to that, the results were so far focused either on the global picture, or on the national-specific skill distributions, between-country comparisons are still to be made. The lack of information about the structure and frequency of skill patterns in observed SEE countries, and, not unimportantly, possible effects of the COVID-19 pandemic, appear to constitute a sufficiently convincing rationale for the exploration of the current state of affairs regarding the skills of digital platform workers in SEE.

Since skill patterns (their content, structure, and mutual relations) are essential for the career-oriented approach to work, and particularly for the flexibility in career management, the ambiguous results call for clarification. There are additional pieces of "anecdotal" evidence that point to the importance of knowing how the skill patterns work in observed countries. A recent Serbian study, published as a technical report (Gigmetar Team, 2021) pointed to the aversive effects of the COVID-19 pandemic on the Serbian online labour market. Staying within a chosen profession proved to be vulnerable to pandemic-related influences, with the risk of abandoning one's profession rising approximately three months after the official proclamation of the pandemic in Serbia. This result speaks little about the changes in job content but cautiously poses questions about the possible courses that online work in Serbia may take during and in a post-pandemic environment. They also are in line with the conclusions of Kost et al. (2019), who are pointing out the sensitivity of digital platform workers' careers to external influences.

Therefore, our study is focused on the two crucial unknowns. To the best of our knowledge, there are no conclusive results regarding the digital platform workers' skill profiles in observed countries. Even country-specific studies are missing. Thus, we attempt to answer the provisional questions: what *is the structure of skill patterns reported by the digital platform workers in SEE region? Do such skill patterns comply with the OLI taxonomy, as the most transparently present (and testable) taxonomy at this moment? What is their distribution across the region?* From the perspective of the "boundaryless career" concept, one may argue that we examine first the boundaries among the professional clusters consisting of specific skills, then the potential boundaries between empirically derived skill structures and the classes of an existing skill taxonomy, and finally the boundaries on a map where skills and countries are projected into the same space. Thus, we examine the flexibility of the founding elements of career management cross-conceptually and cross-nationally. We aim to map the most relevant skill patterns, examine their interrelations, and determine their position within and across national borders. The intended conceptual contribution of this study concerns deriving an empirically based skill taxonomy, based on the self-reported skills extracted from the public profiles of digital platform workers at the major digital labour platform. One possible methodological contribution of this study may be the application of a natural language processing technique to identify empirical skill clusters, which is in line with the approaches employed by contemporary studies in the field.

In light of all conceptual issues and recent global developments, we expect to derive the structures similar to OLI categories, whose distribution across the region will be in line with the results of the Gigmeter monitoring studies so far.

Method

Participants and Procedure

Worker Characteristics

Digital platform workers observed in the study are registered at the Upwork platform (Global Inc. 2015-2021). Upwork, as any digital labour platform, may be defined as "*a set of digital resources–including services and content–that bring together buyers and sellers of intangible knowledge and service work*" (Malik, Heeks, Masiero and Nicholson, 2021, p. 1820). It is a virtual, online working space where the online workers create a profile defining what they can do, i.e., what skills they possess (up to 14 skills are listed on their webpage), and other data regarding the history of their work on the platform, while the employers post job offerings with detailed description of tasks to be executed. When the job is posted on the platform by an employer, digital workers compete for the job and a worker who has particular (set of) skills needed for the job and gives the most favorable offer (regarding the price) gets the job. We use web-scraped data from (public) web pages of digital platform workers registered on Upwork in May 2020 as a part of regular monitoring within the Gigmetar project (Public Policy Research Center, 2020). Out of the 12146 digital platform workers, 6.95% were from Albania, 15.07% from Bulgaria, 6.93% from Bosnia and Herzegovina, 6.42% from Croatia, 6.94% from Hungary, 1.88% from Montenegro, 8.05% from North Macedonia, 10.77% from Romania, and 36.99% from Serbia. For text analysis, we excluded 417 (3.43%) participants who did not provide sufficient textual data, naming a single skill or no skills. Figure 8 contains country and topic statistics for the reduced sample.

The sample comprised 33.49% women, and approximately 0.16% gender-neutral participants. The data on age was not available, as it was not publicly accessible.

Measures

The participants provided data on Job Description, Country, City, Hourly Rate, Total Hours Billed, Total Portfolio Items, Skills, Top Rated Status, Combined Total Earning, Gender, and Occupation. While all the data are taken over directly from the platform (personal web pages of digital workers), only the occupation is derived. Namely, after web-scraping, we have obtained, among others, the data regarding the skills a digital worker has listed on his web page. Starting from the OLI taxonomy of digital professions (Kassi & Lehdonvirta, 2018), we used those skills to assign a profession to every worker. According to data, 4.51% participants are working in professional services domain, 11.05% are clerical workers, 38.41% work in creative services and multimedia, 8.27% in sales and marketing, 23.76% software development, 13.97% writing and translation, while there was not enough data to determine their occupation for the remaining 0.03% participants.

Data Analysis

We used self-reported skills contained in digital workers' profiles at Upwork. Given that skill-related information is reported as an open response, and thus is coded as text, we used topic modeling, a natural language processing technique that uncovers the semantically coherent word patterns (Jockers & Thalken, 2020). As input, topic modeling utilizes a text-document matrix, where rows are source documents (in this study, participants) and columns are words extracted from texts (words and phrases representing participants' self-reported skills). Matrix cells contain frequencies of each word's occurrence in each participant's statement. Topics identification draws on the regularities of word distributions across documents, not dissimilar to latent semantic analysis or factor analysis. Similar to the mentioned methods, topic modeling assumes a latent construct underlying each topic. However, the probabilities of topic occurrences within documents enable document classification, more precisely, assigning each document to its most likely topic in a categorical manner.

Latent Dirichlet allocation with Gibbs sampler was applied (Blei et al., 2003; Pavlinek & Podgorelec, 2017). We used raw instead of transformed frequencies since the number of occurrences directly reflects a skill's importance. According to the recommendations, we used four commonly applied coefficients to decide on the number of topics. Two of the coefficients (Deveaud et al., 2014; Griffiths et al., 2004) yield maximum values for the best model, and two (Arun et al., 2010; Cao et al., 2009) yield minimum.

To help decide on the best topical model, we attempted to facilitate the interpretation of the four- coefficients graph. We a) summed up the standardized values of the two "maximum" and the two "minimum" coefficients, respectively; b) we subtracted the standardized "minimum" sum from the standardized "maximum" sum and standardized the result. This way, we pointed to the topic models that imply the most pronounced difference between the desired maxima and minima, thus pointing to the model with the optimal "convergence" of coefficient values. Additionally, we computed a provisional measure of topic overlap by calculating the squared average Fisher-transformed correlations among the topics (as continua, containing the topic probabilities for each text) within each model. Although the more extensive topic overlap does not generally suggest a less valid model, it may point to potential challenges in topic distinction in our study. We used multidimensional scaling to visually highlight the topic position within a shared "conceptual" space.

We examined the relations between skill patterns and geographical locations (countries of residence) using correspondence analysis (Garson, 2013; Roux & Rouanet, 2009). Both variables had the "principal" status in the analysis; hence there were no supplementary variables. Such an analytic strategy is rooted in the previous studies. Namely, there is no straightforward suggestion of asymmetry between geographical and vocational aspects.

We performed the text analyses using the packages "quanteda" (Benoit et al., 2018) and "seededlda" (Watanabe & Xuan-Hieu, 2021) in R (R Core Team, 2021).

Results

Descriptive Statistics and Word Frequencies

The participants' reports contained 4379 unique skill descriptions (words and phrases), which varied considerably in frequency. To exclude the infrequent and thus comparably less important terms, we inspected the frequencies plot (Figure 1), determining the frequency of 50 as the optimal demarcation line between substantially and "unsubstantially" frequent skill descriptions. The most frequent ten were *adobe photoshop, adobe illustrator, data entry, graphic design, javascript, wordpress, logo design, 11icrosoft excel, translation, html5*. The Figure 3 shows substantially (over $f = 50$) frequent terms.

Topics: Number and Contents

We tested the 2- to 15-topic solutions, the range suggested by Benoit et al. (2021) as a default in the "quanteda textmodels" R package (Benoit et al., 2021). Validity measures (as shown in Figure 1) highlighted twelve- to fifteen-, as well as a seven-topic solution, as acceptable. The eleven, three, and ten-topic models had also had preferable coefficient values. The five best solutions showed comparably small topic overlap, whereby Figure 4 shows that from the seven-topic model onwards the overlap decreases by a monotonic trend. Given that the seven-topic model is the comparably simplest ("most parsimonious") model with good coefficient values and small overlap, we decided to retain it as the optimal solution.

Table 1 shows the most representative terms, or skills, for each topic. Topic 1 comprises the terms related to programming (knowledge of c, javascript, react), and more specifically to web programming (such as html, php, css, css3), suggesting that coding skills in this topic are those applicable in web settings. WordPress also appears within this topic, adding to the relevance of web context. Hence the first topic can be best described as programming and web-administration skills. Writing and translation define Topic 2. This topic comprises "generic" writing skills, creative writing applied to different types of textual output (articles, blogs), but also writing applied to marketing and branding (content writing, copywriting). Proofreading, translation, and proficiency in English appear in this topic as well. Topics 3 and 7 are similar content-wise, but while the former contains primarily the terms typical for graphic design, the latter points to modeling methods used in architectural and engineering design. Thus, the skills that feature most prominently in Topic 3 imply proficiency in software tools used for photo editing and design, “generic” photography skills, and expertise in photo and video editing. On the other hand, Topic 7 is saturated by a range of design-related skills, that do not include photo or video post-production. The Topic 5 emphasizes highly specialized skills in architectural and industrial design, as well as in architectural and 3D modeling. One additional connection with the broader field of architectural design is interior design, also included in this topic. The Topic 6 refers to proficiency in the use of office software packages, combined with administrative support and clerical work. Topic 4 comprises the skills required in digital marketing, such as (but not limited to) social media marketing and management, wordpress skills, and proficiency in Google analytics.

Empirically Derived Topics and OLI Categories

Topics in a Common Space

Multidimensional scaling (Figure 5) points to the apparent closeness between specific pairs of topics and highlights the relatively isolated positions of the remaining ones. The first two dimensions account for 29.82% and 18.76% inertia respectively, favoring the two-dimensional solution as the optimal representation of the topical structure. The three topics comprising design and visual skills are positioned high on dimension 1 (x-axis), and around the middle of dimension 2 (y-axis). Design and photo editing are highly close (despite the comparably low average between-topic overlap in this model) and slightly but noticeably distant from industrial modeling and architecture. Opposite from this group of skills (that we may tentatively name the “creative skills cluster” are clerical work, software development, and web programming, as well as writing and digital marketing. Thus, dimension one presents a distinction between visual creative skills and professional patterns that require verbal skills, formal logic, and dexterity in administrative work. The second dimension seems to highlight the differences between clerical work and administration (high coordinates) and programming (low coordinates), with other topics in the middle. Possibly the most striking difference among the polarized topics is in the level of specialized skills required to successfully practice the chosen professions.

Topics and Countries in a Common Space

The results of correspondence analysis (Figure 7) suggest that the first dimension is marked by the polarization between translation/writing and engineering design/architecture. Less formally, one could describe this distinction as a continuum between specialized "soft" and specialized "hard" skills. Hungary and Bulgaria seem to be the closest to the "soft skills" pattern, while Serbia is the closest to the topic representing highly specialized "hard" skills, but also fairly close to other skill sets from the "creative" cluster. Programming is positioned close to the axes' intersection, indicating the "universality" of this professional choice. What may be a slightly more important result is that there is no straightforward overlap between a skill and a country. Somewhat this is no surprising result: the skills of workers mean nothing if they are not demanded by employers. So, digital platform workers in every country we have observed have an incentive to develop the skill set(s) globally demanded. In that sense, it comes over time to the convergence of skill sets between the countries.

Table 2 points to substantial overlap between the extracted topics and OLI categories assigned to the online platform workers. The largest congruences are between the translation/writing, software development, and digital marketing categories, while the largest discrepancies are in the categories of creative and professional services.

Discussion

This study aimed to outline the most relevant skill patterns of online platform workers in observed SEE countries and the dispersion of such concepts across national borders. At the same time, it is an attempt to contribute to the emerging efforts to develop reliable and comprehensive taxonomies of digital platform workers' skills. This study does not aim to come up with a single conclusive taxonomy of digital skills. However, following both conceptual and methodological principles of contemporary taxonomic studies in the field, it tends to contribute to an emerging research program by exploring skill patterns using the material that has seldom been utilized (self-report data from Upwork) and conducting a study in the SEE region where such research effort has not been undertaken.

One could argue that, by using an exploratory strategy to extract broader clusters from the digital platform workers' self-reported skill profiles, we effectively performed a tentative validation of the OLI taxonomy in a specific context (the increasing impact of digital work in SEE countries). Our results point to seven plausible skill clusters, which showed both substantial overlaps and non-negligible differences to the taxonomies proposed so far, whereby the comparison with the OLI categories (Kassi & Lehdonvirta, 2016) is particularly important due to their impact both conceptually and "in the field."

Seemingly, the more technically saturated professional patterns are more "conceptually" robust (a.k.a. overlap more across OLI and exploratory-driven skill clusters) than creative skills. This could mean that creative professions are more flexible content-wise, but the question how "permeable" their boundaries are in light of market and educational requirements remains open. Namely, it is still to be clarified whether one can shift between creative professions without extra effort (it seems more likely when photo editing and visual design are concerned) or acquiring additional skills would be necessary. Thus, a slight polarity appears to be present between the creative and "technical" domains. On the other hand, it does not seem to be only an issue of soft versus hard skills, since the extracted creative skills require expertise. One could argue that the discrepancies among skill domains and between OLI and "empirical" professional categories may be specific for the region. This question remains open and requires further examination since, at present, we have insufficient empirical evidence to address the issue with proper the methodological rigor.

In light of the OLI taxonomy, the results call for finer profiling in the creative department and a more exact definition of the vague professional categories. Thus, if one considered altering the existing taxonomy to better reflect the state of affairs in the region, it would not simply be enough to involve the inclusion or exclusion of professional classes. Instead, it would probably encompass a slight re-conceptualization and reorganization to better adjust it to the situation in the field. A similar polarity is transparent on a more general conceptual level when the skill patterns are placed into a common two-dimensional space. One dimension discriminates between creative and non-creative skills, regardless of the required level of expertise. At the same time, the other separates the "professions" that require a high degree of specialization from those that demand general skills. Creative skills are in between these polarities. By the OLI and COLLEEM standards (Pesole et al., 2018), this could represent a distinction between professional and non-professional services, but in this context, the problem appears to be somewhat more complex. First, it may point to the interconnectedness between the type of profession and the level of specialization. Subsequently, it may potentially underline two founding elements of the skill-based professional taxonomy – the required level of specialization/expertise and the continuum stretched between liberal arts and technical disciplines. Not only the present but the missing elements of this map are worth pointing out. As large portions of this "map" are unpopulated, one may wonder about the tentative position of highly specialized but highly creative skill patterns (perhaps related to music production and composition?) or pose a question about the creative jobs that do not require particular specialization. They may be derived from the self-reported skills that were not sufficiently frequent to make it to the map. Thus, a more liberal criterion for term choice may be employed in the forthcoming studies. Nevertheless, this "map" is a device enabling us not only to recognize and organize the existing concepts, but also to make assumptions about the plausibility of new ones. Finally, the constellation of the space that countries and skills share, adds to the assumptions derived from previous results and possibly clarifies certain issues. Here we have the combinations of specialized "soft" and "hard" skills, but also the hints of which professional clusters are more "typical" for a country. Generally, a polarization is evident between four EU countries (Bulgaria, Croatia, Hungary, Romania) and the five non-EU countries (Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, Serbia). Online platform workers from the former group appear to be more inclined to translation/writing. In contrast, those from the latter group strive towards clerical work (North Macedonia) and applied industrial design (Serbia). While this is by no means a definitive account of the current state of affairs in the region (if it is an account at all), it certainly shows the importance of research of online labour market requirements and the professional profiles of digital platform workers. It is also slightly out of line with some previous results which suggested that Romania was a leading provider of non-professional services (Pesole et al. (2018); the study did not include the other countries in our sample). Most importantly, these results apparently show that the country of residence and professional profile are not independent, at least in the region we have observed. How strong those boundaries are, remains to be examined in future studies.

While bringing novel insights into the realm of skills within digital platform workers in Southeastern Europe, the study has several limitations as well. First, attempting to provide the results that could be comparable across countries, we opted for between-group comparisons, which required skill patterns to be extracted from a common set of terms, and then compared in a "breakdown" analysis. This excluded the valuable within-country data but could be a topic for future studies. Thereby, a possible hierarchical or multilevel structure of the data could be considered and examined. Secondly, the study utilized a transversal design, thus not incorporating a time-related element. The design is unbalanced across countries and, although the respective numbers of participants are sufficient for the analyses we applied, a more balanced design would be preferable. The study encompassed the participants from only one digital labour platform (Upwork). Although it is undoubtedly the single most popular digital platform in the region and most accessible for researchers, more variation in this segment would be welcome. On the other hand, there are no research or database which are covering all the existing digital labour platforms. even the previously mentioned Oxford's Online Labour Index.Paradoxally, Oxford's Online Labour Index examines the profiles of digital platform workers from numerous of digital labour platforms, but not Upwork. An additional conceptual issue is whether it is justified to refer to the skill patterns as professions. While that is the case within the OLI taxonomy, our principal reason for not https://doi.org/ng so is the need to further examine the patterns' additional features that would assert them as professions. By that, we refer primarily to their temporal stability and standing in the digital labour market. To resolve those issues, new research efforts are required.

Conclusion

Conceived as an exploratory insight into the structure of online job-related skills in the SEE countries, this study managed to point to several important issues that should be explored further. Such issues concern: online professional profiles and their slight, but important differences from the most popular taxonomy; the distinction between groups of skill patterns, based on specialization and creative orientation; and finally, the current national skill profiles. From the perspective of boundaryless careers, we could argue that the boundaries present in our results are evident, but figuratively may be termed as semi-permeable due to their positions around the higher-order continua. Thus, one may conclude that the sound foundations for transgressing national boundaries exist, but that further studies are necessary to assess their actual strength.

References

Anđelković, A., B., Šapić, Š., J., & Skočajić, S., M. (2019, February). *Gig ekonomija u Srbiji: Ko su digitalni radnici i radnice iz srbije i zašto rade na globalnim platformama*. Centar Za Istraživanje Javnih Politika. Retrieved October 20, 2022, from https://publicpolicy.rs/publikacije/b7772ff6bcdb10f723f1b94dfceff2757bdb19b6.pdf

Arthur, M. B., & Rousseau, D. M. (2001). *The boundaryless career: A new employment principle for a new organizational era*. Oxford University Press on Demand.

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In M. J. Zaki, J. X. Yu, B. Ravindran, & V. Pudi (Eds.), *Advances in knowledge discovery and data mining* (pp. 391–402). Springer. https://doi.org/10.1007/978-3-642-13657-3_43

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). Quanteda: An r package for the quantitative analysis of textual data. *Journal of Open Source Software*, *3*(30), 774. https://doi.org/10.21105/joss.00774

Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A (2018). “quanteda: An R package for the quantitative analysis of textual data.” *Journal of Open Source Software*, 3 (30), 774. <https://https://doi.org/.org/10.21105/joss.00774>, <https://quanteda.io>.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3(null)*, 993–1022.

Bourgeon, P., Champarou, C., Filipek, K., & Martin, L. (2019). Submission prepared for the 2nd IZA/CREA workshop on “exploring the future of work” 5-6 december 2019, IZA Bonn, Germany.

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, *72*(7), 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numerique*, *17*(1), 61–84. https://doi.org/.org/10.3166/dn.17.1.61-84

Djumalieva, J., & Sleeman, C. (2018). An open and data-driven taxonomy of skills extracted from online job adverts. In *Developing skills in a changing world of work* (pp. 425–454). Rainer Hampp Verlag. https://doi.org/.org/10.5771/9783957103154-425

Garson, G. D. (2013). *Correspondence analysis* (2012th edition.). Statistical Associates Publishers. ASIN number : B00B0P0WB6

Giabelli, A., Malandri, L., Mercorio, F., Mezzanzanica, M., & Seveso, A. (2020). NEO: A Tool for Taxonomy Enrichment with New Emerging Occupations. The Semantic Web – ISWC 2020, 568–584. https://doi.org/10.1007/978-3-030-62466-8_35

Gigmetar Team (2021). Gig workers in the region - February 2021. *Public Policy Research Center*. Retrieved from http://gigmetar.publicpolicy.rs/en/region4-2/

Global Inc.", "Upwork®. (2015-20212015-2021). Upwork - the world‘s work marketplace for freelancing. *Upwork*. Retrieved from https://www.upwork.com/

Griffiths, T. L., Steyvers, M., Blei, D. M., & Tenenbaum, J. B. (2004). Integrating topics and syntax. Retrieved from https://openreview.net/forum?id=BkW-S_ZdWH

International Labour Office. (2021). World Employment and Social Outlook: The role of digital labour platforms in transforming the world of work. Geneva: International Labour Organization.

Jockers, M. L., & Thalken, R. (2020). *Text analysis with r: For students of literature* (2nd edition.). Springer. https://doi.org/.org/10.1007/978-3-030-39643-5

Johns, T., & Gratton, L. (2013). The third wave of virtual work. *Harvard Business Review*, *91*(1), 66–73.

Kassi, O., & Lehdonvirta, V. (2016). Building the online labour index: a tool for policy and research. [Paper presentation]. CSCW and the Sharing Economy: The Future of Platforms as Sites of Work Collaboration and Trust. 2016-02-27, available online: https://futureofplatforms.wordpress.com/position-papers/ Kässi, O., & Lehdonvirta, V. (2018). Online labour index: Measuring the online gig economy for policy and research. *Technological Forecasting and Social Change*, 137, 241–248. https://doi.org/10.1016/j.techfore.2018.07.056

Kässi, O., Lehdonvirta, V., & Stephany, F. (2021). How many online workers are there in the world? A data-driven assessment. *arXiv preprint arXiv:2103.12648*. https://doi.org/10.48550/ARXIV.2103.12648

Koch, M., Park, S., & Zahra, S. A. (2021). Career patterns in self-employment and career success. *Journal of Business Venturing*, *36*(1), 105998. https://doi.org/.org/10.1016/j.jbusvent.2019.105998

Kost, D., Fieseler, C., & Wong, S. I. (2019). Boundaryless careers in the gig economy: An oxymoron?. *Human Resource Management Journal*, *30*(1), 100–113. https://doi.org/10.1111/1748-8583.12265

Pavlinek, M., & Podgorelec, V. (2017). Text classification method based on self-training and LDA topic models. *Expert Systems with Applications*, *80*, 83–93. https://doi.org/10.1016/j.eswa.2017.03.020

Pesole, A., Brancati, U., Fernández-Macías, E., Biagi, F., & Gonzalez Vazquez, I. (2018). Platform workers in europe. *Luxembourg: Publications Office of the European Union*.

Public Policy Research Center. (2020). *Gig workers in the region*. Centar za istraživanje javnih politika. Retrieved from https://gigmetar.publicpolicy.rs/region2/

Pouliakas, K., & Branka, J. (2020). EU Jobs at Highest Risk of Covid-19 Social Distancing: Will the Pandemic Exacerbate Labour Market Divide? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3608530

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Radonjić, L. (2020). *Od mita do istine.* Centar za istraživanje javnih politika. Retrieved from https://www.publicpolicy.rs/publikacije/777122d08c497901c47e1fa37e5ce97832d98a66.pdf

Roux, B. L., & Rouanet, H. (2009). *Multiple correspondence analysis* (1st edition.). SAGE Publications, Inc. https://doi.org/10.4135/9781412993906

Sullivan, S. E., & Arthur, M. B. (2006). The evolution of the boundaryless career concept: Examining physical and psychological mobility. Journal of Vocational Behavior, 69(1), 19–29. https://doi.org/10.1016/j.jvb.2005.09.001

Urzì Brancati, M. C., Pesole, A.others. (2020). *New evidence on platform workers in europe: Results from the second COLLEEM survey*. Joint Research Centre (Seville site).

Watanabe, K., & Xuan-Hieu, P. (2021). *Seededlda: Seeded-LDA for topic modeling*. Retrieved from https://CRAN.R-project.org/package=seededlda

Kako vidimo obrasce veština radnika na digitalnim platformama? Rezultati iz jugoistočne Evrope

Vladan Ivanović[1,2,3], Branka Anđelković[1], Tatjana Jakobi[1], Zoran Kalinić[1,2], & Ljubivoje Radonjić[1,2], Petar Čolović[4]

[1]Centar za istraživanje javnih politika, Beograd

[2]Ekonomski fakultet, Univerzitet u Kragujevcu, Kragujevac

[3]CERGE-EI fellow, Prag

[4]Odsek za psihologiju, Filozofski fakultet, Univerzitet u Novom Sadu

Ova studija je imala za cilj da identifikuje i opiše najrelevantnije obrasce profesionalnih veština među radnicima na digitalnim platformama iz selektovanih zemalja jugoistočne Evrope (eng. Southeastern European, SEE). Ova orijentacija je bazirana na dosta slaboj zastupljenosti SEE zemalja u velikim panevropskim studijama, kao i na nedostatku informacija o primenljivosti postojećih online poslovnih taksonomija u posmatranim zemljama. Pristupajući ovoj temi preko tehnike obrade prirodnog jezika (eng. natural language processing technique), analizirali smo skupove izveštaja o veštinama koje su dali sami radnici na digitalnim platformama registrovani na platformi Upwork. Dobijeno je sedam različitih profila veština, koji se samo delimično preklapaju sa standardnim Oksfordskim online poslovnim indeksom taksonomije onlajn poslova. Rezultati ukazuju na jasne razlike između visoko specijalizovanih i opštih poslovnih kategorija, kao i između kreativnih i tehničkih profesija. Mapiranje veština i nacionalnih pripadnosti je pokazalo razliku između EU zemalja i zemalja koje nisu članice EU u pogledu profesionalnih inklinacija. Delimično u skladu sa nalazima prethodnih studija, rezultati otvaraju put za buduća istraživanja ove teme.

Ključne reči: rad onlajn, digitalne platforme za rad, taksonomija veština, karijerne granice

RECEIVED: 17.08.2021.

REVISION RECEIVED: 21.04.2022.

ACCEPTED: 05.10.2022.

Figure 1

Word frequencies

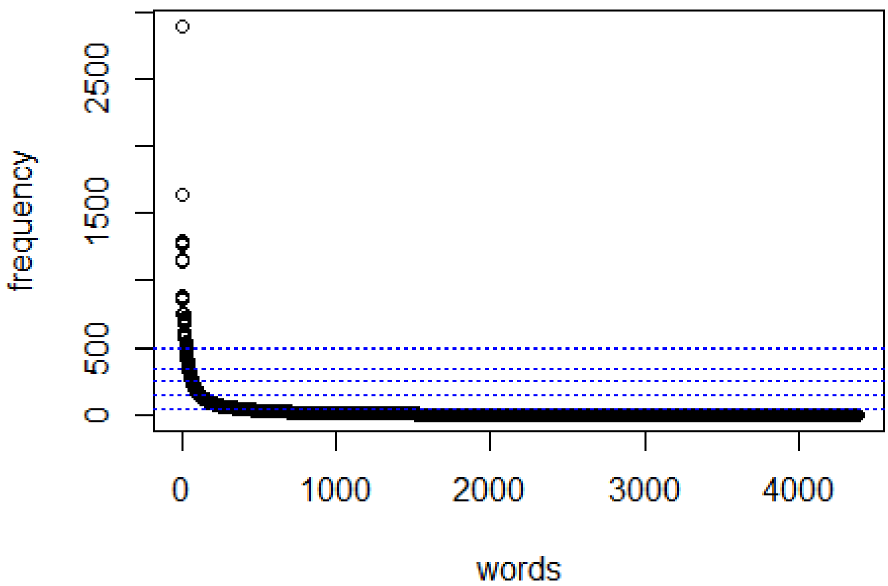


Figure 2

Topic overlap

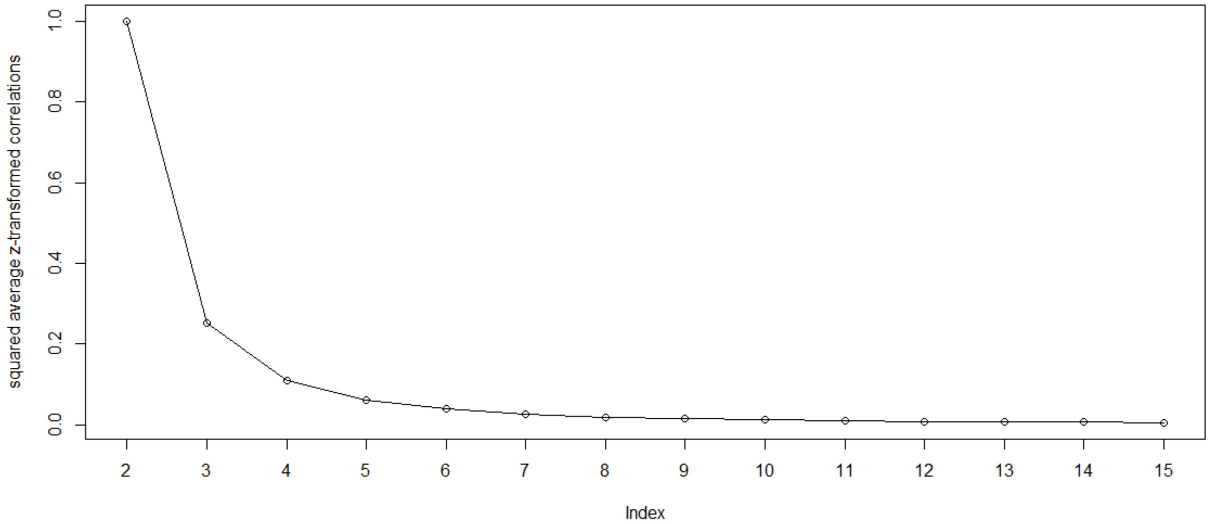
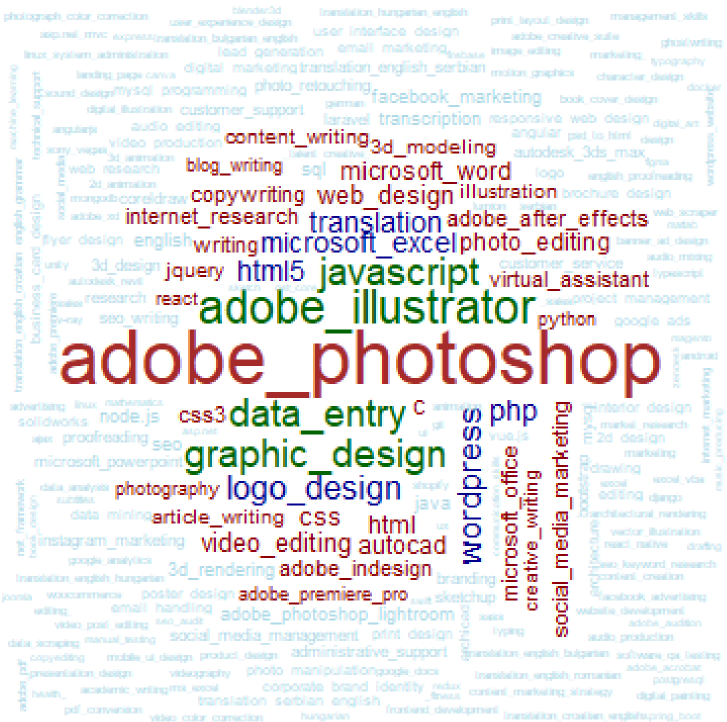


Figure 3

Word frequency: Wordcloud



Determining the number of topics

Figure 4

Topic number: Coefficients

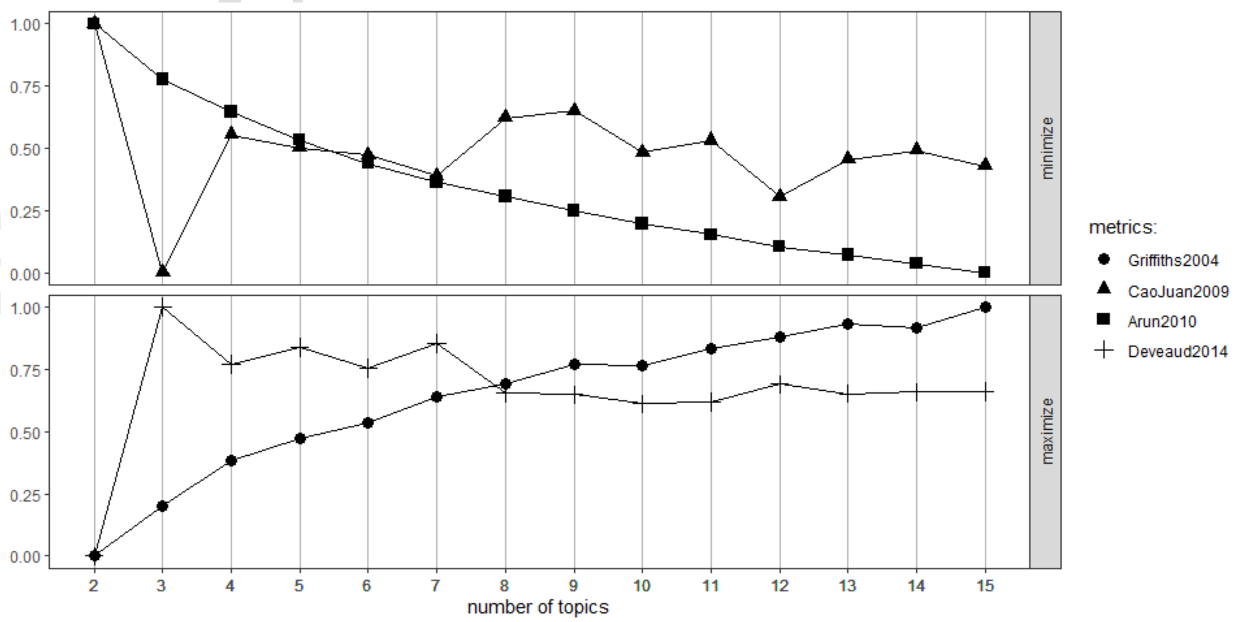


Table 1

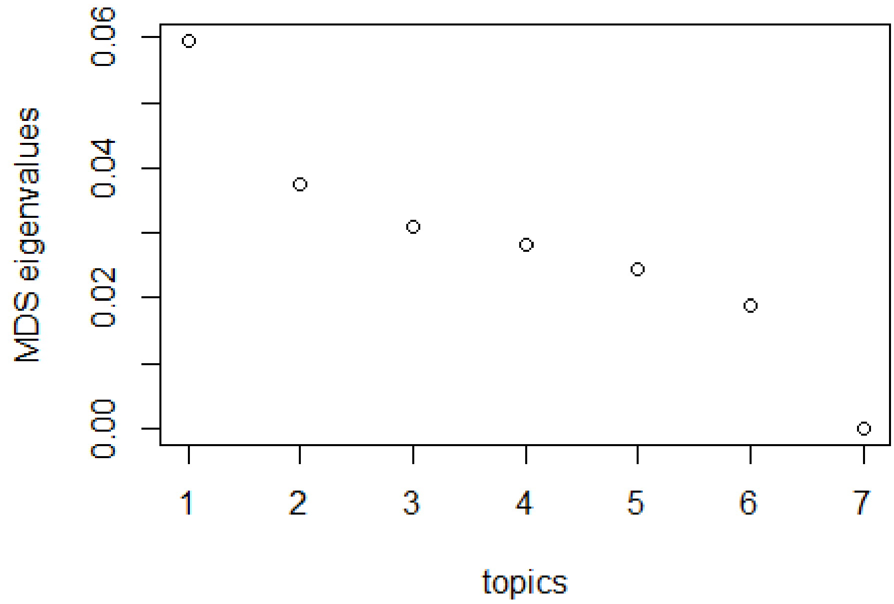
Topic contents

topic1[1]	topic2[2]	topic3[3]	topic4[4]	topic5[5]	topic6[6]	topic7[7]
Javascript	translation	adobe photoshop	social media marketing	adobe photoshop	data entry	adobe illustrator
html5	writing	video editing	facebook marketing	autocad	microsoft excel	adobe photoshop
php	article writing	photo editing	seo	3d modeling	microsoft word	logo design
css	creative writing	adobe after effects	wordpress	3d rendering	virtual assistant	graphic design
wordpress	copywriting	adobe premiere pro	social media management	3d design	internet research	web design
c	blog writing	photography	instagram marketing	sketchup	microsoft office	adobe indesign
html	content writing	adobe illustrator	google ads	autodesk 3ds max	microsoft powerpoint	illustration
css3	translation english serbian	adobe photoshop lightroom	email marketing	interior design	customer support	branding
jquery	proofreading	photo retouching	digital marketing	architecture	customer service	print design
react	english	graphic design	google analytics	2d design	administrative support	logo

[1]programming/software development; [2]writing and translation; [3]photo and video editing; [4]digital marketing; [5]industrial design and modeling; [6]clerical work; [7]graphic design

Figure 5

Multidimensional scaling: Eigenvalues



OLI and empirically derived topics: overlap

Table 2

Overlap between OLI dimensions and topic patterns

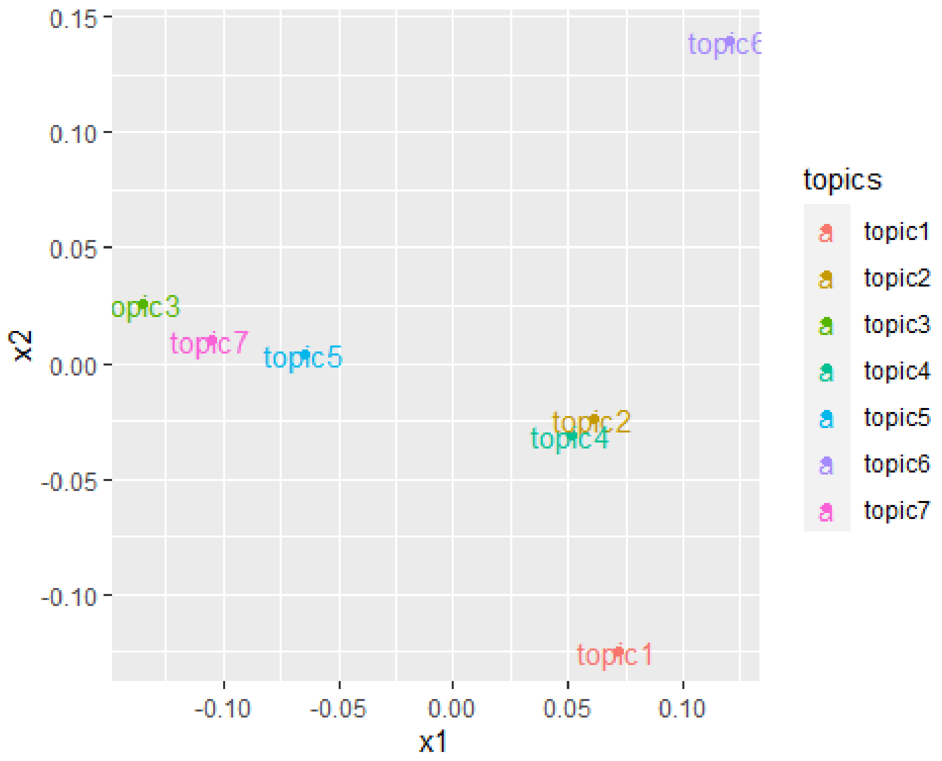
PRO	CLE	CRM	SMA	SDT	WTR
23	39	150	22	2,151	16
105	172	104	94	52	1,329
10	17	1,202	15	31	32
99	84	245	711	148	90
17	8	1,074	4	59	8
224	976	154	98	246	157

PRO	CLE	CRM	SMA	SDT	WTR
7	10	1,638	23	65	19

Note. Columns: PRO - Professional services, CLE - Clerical and data entry, CRM - Creative and multimedia, SMA - Sales and marketing support, SDT - Software development and technology, WTR - Writing and translation; Rows: Topic 1: programming/software development, Topic 2: writing and translation, Topic 3: photo and video editing, Topic 4: digital marketing, Topic 5: industrial design and modeling, Topic 6: clerical work, Topic 7: graphic design.

Figure 6

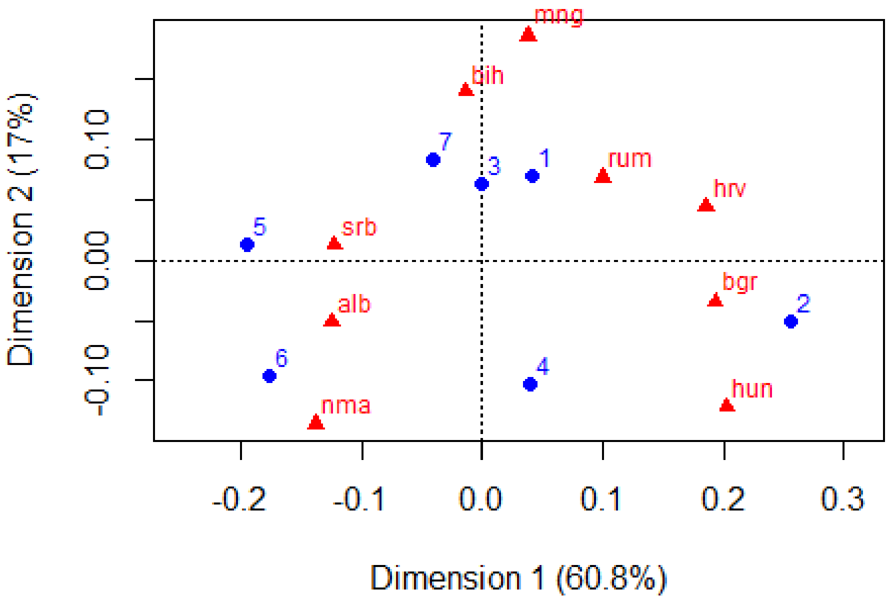
Seven topics in two-dimensional space



Note. Topic 1: programming/software development; Topic 2: writing and translation, Topic 3: photo and video editing; Topic 4: digital marketing", Topic 5: industrial design and modeling, Topic 6: clerical work, Topic 7: graphic design.

Figure 7

Topics and countries in a common space: correspondence analysis

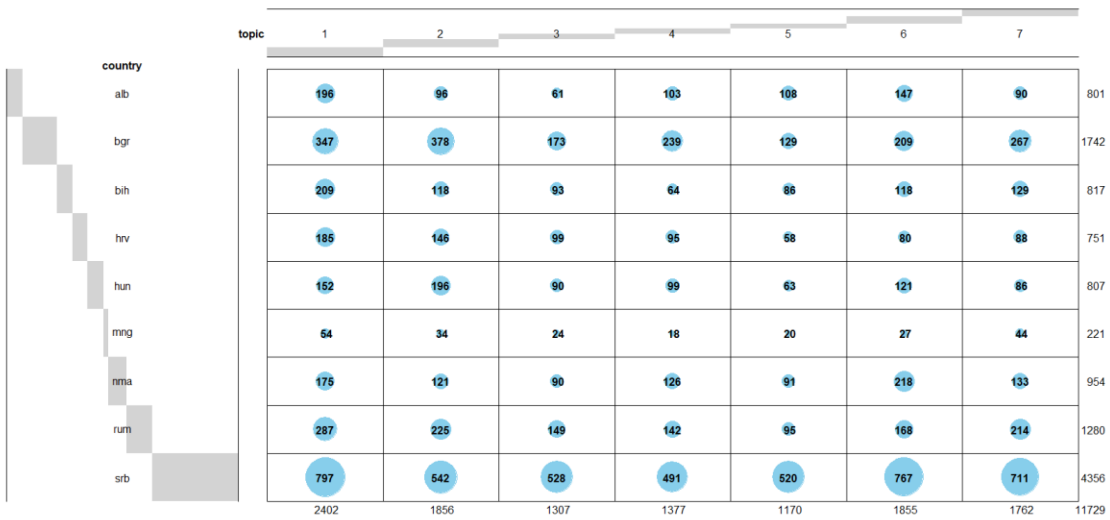


Note. Topic 1: programming/software development; Topic 2: writing and translation, Topic 3: photo and video editing; Topic 4: digital marketing", Topic 5: industrial design and modeling, Topic 6: clerical work, Topic 7: graphic design. ALB - Albania; BIH - Bosnia and Herzegovina; BGR - Bulgaria; HRV - Croatia (Hrvatska); HUN - Hungary; NMA - North Macedonia; RUM - Romania; SRB - Serbia.

Appendix

Figure 8

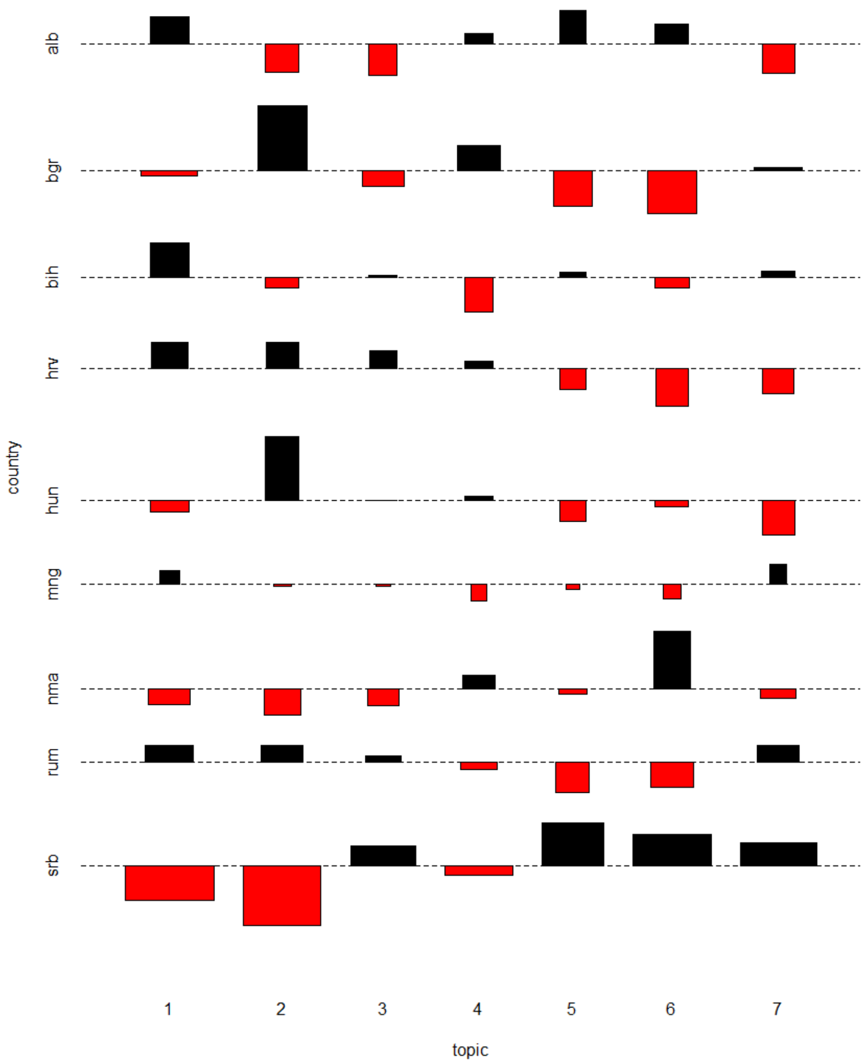
Frequencies: topic by country, raw data



Note. Topic 1: programming/software development; Topic 2: writing and translation, Topic 3: photo and video editing; Topic 4: digital marketing", Topic 5: industrial design and modeling, Topic 6: clerical work, Topic 7: graphic design. ALB - Albania; BIH - Bosnia and Herzegovina; BGR - Bulgaria; HRV - Croatia (Hrvatska); HUN - Hungary; NMA - North Macedonia; RUM - Romania; SRB - Serbia.

Figure 9

Standardized residuals - topic by country



Note. Topic 1: programming/software development; Topic 2: writing and translation, Topic 3: photo and video editing; Topic 4: digital marketing", Topic 5: industrial design and modeling, Topic 6: clerical work, Topic 7: graphic design. ALB - Albania; BIH - Bosnia and Herzegovina; BGR - Bulgaria; HRV - Croatia (Hrvatska); HUN - Hungary; NMA - North Macedonia; RUM - Romania; SRB - Serbia.