// NO.20-090 | 12/2020

DISCUSSION PAPER

// MAREIKE SEIFRIED, ROMAN JUROWETZKI, AND TOBIAS KRETSCHMER

Career Paths in Online Labor Markets: Same, Same but Different?

Career Paths in Online Labor Markets:

Same, same but different?

Seifried, Mareike (ZEW Mannheim)

Jurowetzki, Roman (Aalborg University)

Kretschmer, Tobias (LMU Munich)

March 2020

Abstract

The emergence of online labor markets calls the validity of traditional career models into question. Given the volatility and digital nature of this environment, short-term employment relationships and heterogeneity of workers, employers and tasks in these markets, it is unclear how careers might unfold – whether they are largely random and accidental or whether there are distinct trajectories and patterns in online careers. We document dominant career paths and develop a taxonomy of novel career patterns in OLMs. This addresses recent calls for research to update and refine our theories and understanding of careers (Rahman et al., 2016). We adopt a quantitative-inductive approach to describe workers' careers in terms of their task and skill specialization. This helps us understand the different types of careers on a continuum between more stable and random, unsystematic careers. Our results provide an innovative way of thinking about career development in OLMs and open up several fruitful areas of future research.

Keywords: *online labor markets, platforms, gig economy, career theory, freelancing, labor specialization, human capital.*

JEL class: J24, J44, O30

Acknowledgements: We are grateful for helpful comments from participants of the CESifo and LINER-AUEB Workshop "The Effects of the Digital Transformation on the Workplace and the Labor Market", Reshaping Work Conference (Platform Economy), Strategy, Entrepreneurship & Innovation Doctoral Consortium at KU Leuven, Academy of Management Annual Meeting Paper Development Workshop "AI and Machine Learning as a Tool and Topic", DRUID Conference, London Business School Trans-Atlantic Doctoral Conference, and DRUID PhD Academy.

1 Introduction

The concept of a career has usually been studied in organizational settings structured by internal labor markets. However, over the past decade, we have observed a steady growth of contingent work (Kässi & Lehdonvirta, 2018; Manyika *et al.*, 2016) and an increasing number of workers pursues a career outside traditional employment settings. This trend was enabled and further fueled by the emergence and rapid growth of online labor market (OLM) platforms such as Upwork, Fiverr and Freelancer.com, allowing a diverse set of workers to enter external labor markets. OLMs refer to markets "where labor is exchanged for money, the product of that labor is delivered over a wire and the allocation of labor and money is determined by a collection of buyers and sellers operating within a price system" (Horton, 2010: 516). Given their distinct features, it is unclear how careers predominately unfold: Digital workers have more flexibility due to the heterogeneity of jobs and short-term nature of work. Further, online workers self-assemble their work portfolios instead of working on firm- or occupation-specific task bundles. As such, they face an almost infinite pool of job opportunities, opening up many different pathways to structure their careers. Since the digital workforce is very diverse in terms of backgrounds and motivations to enter platforms (Manyika *et al.*, 2016), we might observe a plethora of differential career paths. Despite recent calls to study how technology redefines work patterns and whether we have to update career theories in a changing work environment (Barley, Bechky, & Milliken, 2017; Rahman, Valentine, Bechky, Leung, & Levina, 2016), career development paths in OLMs are not yet fully understood. Thus, we ask: *How do online workers structure their careers on OLM platforms? Put differently, which career paths emerge in this novel type of labor market?*

To answer our research question, we study specialization patterns of online workers on a task and skill level. Although there exist many ways to study the structure of careers, the coherence of job-related experiences can be seen as the primary and the most obvious indicator of career paths and whether a work history follows an overall logic that links jobs over time (Rosenfeld, 1992). Documenting specialization patterns further helps us understanding whether some workers follow fairly stable career paths, despite the volatility of these markets, or whether all careers are necessarily multidirectional, consisting of disorderly sequences of jobs. In our exploratory, purely descriptive study, we apply a broad range of innovative methods to examine online workers' career trajectories. We use an extensive dataset from Upwork.com, the world's largest online freelancing website. Our sample includes the work histories of over 74,000 freelancers and more than 2.6 million jobs accomplished on the platform.

We find evidence for both specialized and stable as well as diversified and multidirectional career paths. The dominant pattern lies somewhere in between: There is a tendency towards frequent movements across job boundaries but rather between related task and skill categories. We find that the peculiarities of OLM platforms have enabled novel career paths: Both increasingly fragmented and narrow careers due to the skill-based nature of work but also multidirectional and opportunistic trajectories that lack an overall logic and structure. Thus, we develop a taxonomy of four career types: a task specialist offering a narrow set of skills (*Micro career*), a task specialist offering a broad set of skills (*Task expert*), a category generalist who carries a specific skill across task categories (*Skill expert*), and a category generalist with a very broad skillset (*Opportunist*). Whereas the former three can be seen as fairly stable careers in terms of the jobs workers hold, the latter is multidirectional and characterized by frequent movements between diverse tasks and skillsets. We argue that these four paths have emerged due to the diverse motivations and shorter time horizons of online workers.

We contribute to a number of research streams: First, we add to career theory by showing that conventional career models can be applied to OLMs but that there are nuances. Some online workers pursue relatively stable and fairly "traditional" careers, characterized by linear progression and coherent work histories (Rosenfeld, 1992). Others cross boundaries often and span less structured sequences of jobs, as reflected in the existing boundaryless career model (Arthur & Rousseau, 1996; DeFillippi & Arthur, 1994). Our taxonomy of OLM-specific career paths can be seen as extreme versions of them and can serve as a starting point for analyzing careers in an online, platform-enabled environment and their respective motivations and outcomes. Second, we add to the growing literature on OLMs and the debate on the role these platforms may have in the future (Rahman *et al.*, 2016). Our study suggests that OLMs can be used in several different ways: From building a fairly stable career to developing specific skills while earning some money as well as experimenting with different jobs to find the best match. Third, we extend the emerging literature on careers in external labor markets by being among the first to study them quantitatively and on a large scale. Likewise, very few studies exist that explore freelancing careers of more high-skill (knowledge) workers. Further, whereas existing work usually assumes that most freelancing careers are necessarily multidirectional and unstructured sequences of jobs (O'Mahony & Bechky, 2006; Reilly, 2017), we find evidence for fairly stable careers that may enable linear career progression. Fourth, we contribute to the growing literature that studies the task and skill content of jobs in the labor market (Acemoglu & Autor, 2011; Autor, 2001; Gathmann & Schönberg, 2010). By clearly distinguishing between tasks and skills, we are able to portrait the coherence of work histories in a much nuanced way. In sum, our results provide an innovative way of thinking about career development in OLMs. With our descriptive approach to detect work patterns, we open up several fruitful areas of future research to study the causal mechanisms and dynamics behind them.

2 Career Theory

2.1 Traditional and Boundaryless Career Model

Career theory is not a unified theory but all attempts to explain the unfolding of human work experience over time can be considered career theories (Arthur et al., 1989; O'Mahony & Bechky, 2006). Career research is generally concerned with identifying and explaining recurring patterns of work histories (Rosenfeld, 1992; Schein, 1978; Spilerman, 1977). A career is defined as a sequence of jobs or work-related experiences within an individual's work history

(Spilerman, 1977). Career paths then refer to prototypical career sequences of individuals with similar career patterns (Joseph, Fong Boh, Ang, & Slaughter, 2012).

The notion of stable and multidirectional career paths originate from two existing career concepts: The traditional and boundaryless career model. Both implicitly describe different degrees of specialization as described in the following.

In the traditional career model, careers unfold within certain job and organizational boundaries, providing careers an overall logic. An "organization man's" career (Wythe, 1956) involves climbing a clearly defined and stable path to the top. Careers are thus characterized by a linear direction of advancement in objective career outcomes such as the rate of upward mobility (promotion) and other external indicators of achievement such as salary increases (Rosenbaum, 1979; Rosenfeld, 1992; Wilensky, 1964). The traditional career model implicitly describes more specialized careers because the coherence of work histories is often necessary for (linear) progression (Rosenfeld, 1992). More experience in a particular field develops more skills, which in turn qualifies a worker for more challenging, more "valuable" roles in the future. Prior work has shown that specialization is related with income growth and productivity (Ferguson & Hasan, 2013; Rosen, 1983), implicitly caused by increases in workers' skill proficiency and returns to those skills (Shaw, 1987). Remaining within certain job boundaries (i.e. specializing) thus illustrate career stability because workers do not frequently switch to new paths that potentially require new skills, routines or work environments (Feldman & Ng, 2007).

The boundaryless career model suggests that workers may span less structured job sequences that cross organizational, industry, occupational and geographical boundaries (Arthur & Rousseau, 1996; DeFillippi & Arthur, 1994; Greenhaus, Callanan, & DiRenzo, 2008). The boundaryless career is thus associated with higher levels of career mobility, which can potentially have a negative effect on objective career outcomes but is often positively associated with subjective career outcomes such as job satisfaction (Hall & Chandler, 2005; King, Burke, & Pemberton, 2005). The boundaryless career model is related with a certain degree of diversification because workers may move to different types of jobs, the boundary we focus on in this paper. Diversification can generally be advantageous: Prior work highlights that a greater diversity of work settings enhances a person's ability to acquire work from a greater number of employers (O'Mahony & Bechky, 2006). The capability of adapting to new environments, changing tasks and responsibilities or switching to the production of completely different task outputs is an important component of workers' human capital (Sunde, 2008). By developing diverse skills through on-the-job learning and by building professional and social networks (Arthur & Rousseau, 1996; Higgins, 2001; Sullivan, 1999), workers strengthen their employability, the ability to market oneself on the labor market (Kanter, 1989). Thus, diversification helps workers to remain flexible. However, if workers cross job boundaries simultaneously or often, they frequently shift their careers and it is more difficult to interpret the overall career so that work histories might become "erratic" (Leung, 2014) or "random" (Rosenfeld, 1992).

2.2 Career Paths in External Labor Markets

Career development in external labor markets is generally less well understood than careers in organizational settings (O'Mahony & Bechky, 2006). Existing work stresses that freelancing careers often involve disorderly and loose sequences of jobs and are thus closer to the boundaryless career model. For example, O'Mahony and Bechky (2006) analyze qualitatively how independent workers manage the so-called "career progression paradox", i.e. the problem of finding a job to develop one's skills in new segments (crossing job boundaries) without prior related work experience. They find that workers overcome this paradox by crossing boundaries through "stretchwork" that overlaps to some degree with prior jobs. Put differently, they study how workers are able to cross boundaries in a structured way. Their implicit assumption is that workers are generally motivated to diversify and cross job boundaries. Work on freelancing careers also stresses the role of social networks for career progression. Faulkner and Anderson

(1987) find that career opportunities in the film industry differ for those in the core versus the periphery of the industry's external labor market network. Similarly, Reilly (2017) studies stand-up comedians and develops a model of a layered career where each layer consists of a durable social infrastructure. He finds that individuals move progressively through three layers but also often move backwards, illustrating the multidirectional nature of freelancing careers.

Although these studies advance our understanding of careers in external labor markets, the OLM context differs from offline freelancing in several ways. Thus, the implications of these studies may be only partly applicable to careers on OLM platforms, which might not necessarily develop in similar ways. We explain these differences and implications for career development in the following.

3 Career Paths in Online Labor Markets

3.1 Defining Online Labor Markets

Online labor markets (OLMs) have emerged as digital matching platforms that facilitate the allocation of labor across global economies (Agrawal *et al.*, 2015). Put differently, work is performed entirely online rather than by physically collocated workers (Chen & Horton, 2016). We focus on spot markets for tasks, a particularly powerful new way of accomplishing work online (Horton, 2010). In these markets, workers switch between different employers to work on a project basis.

To enter the platform, both freelancers and clients register and create a profile where they provide basic information on their profiles. For freelancers, profile pages further include a description of skills, education, work experience outside of the platform, skill test scores, certifications, agency affiliation, portfolio items, and platform work history and feedback scores. Employers can post any number of jobs and hire as many freelancers as they like, in general and for a single project. Postings include a task description, the employer's location, the type of contract offered (fixed price or hourly wage), and other job features. For fixed-price projects, employers must specify outcome, budget and deadline. For hourly-wage projects, employers give the expected number of weeks and hours per week to complete the project. Freelancers apply by submitting cover letters and bids to jobs. Employers can interview and negotiate over bids with applicants before hiring. When hiring multiple freelancers for single projects, employers send separate job offers and outline the terms of each contract, which can then deviate from the original job posting. Once hired, the freelancer completes tasks remotely. Submission of deliverables and payments are done via the platform, which then charges a service fee. After completion, employers evaluate the project performance with a feedback score from 1 to 5, on six criteria related to process and outcome quality.

3.2 Differences between Offline and Online Freelancing

Online work differs from offline freelancing in several ways, with implications for emerging career paths.

Flexibility. OLMs are characterized by a large degree of flexibility due to the job and employer heterogeneity and project-oriented nature of work. Given that workers self-assemble their work portfolios instead of working on pre-defined task bundles, they face a broader opportunity space than offline workers in terms of structuring their careers. On the one hand, the flexibility of OLMs enables more specialization because they remove some of the barriers towards a finer division of labor (Agrawal *et al.*, 2015): First, OLMs aggregate the global demand for certain skills and tasks. Second, platforms reduce the costs of finding and coordinating workers with different specialties by the improved matching and digital nature of work relationships. Finally, the feasibility of contracting out smaller jobs encourages specialization because workers can narrow down their work focus. As such, career stability might be possible due to the focus on certain types of jobs. On the other hand, OLMs are platforms that allow for easy switching across tasks requiring diverse skills. The task heterogeneity enables workers to cover diverse task categories with potentially very distinct skill requirements (Kokkodis & Ipeirotis, 2016).

This is enabled by low entry barriers to novel job families due to the short-term nature of jobs and varying skill levels. At the same time, OLMs are very volatile environments, in which the demand for certain skills can change frequently (Horton & Tambe, 2019). Given that workers have to frequently apply for new jobs, OLMs can also lead to precarious, multidirectional work patterns and higher levels of career mobility.

Motivations. The digital workforce is very diverse in motivations, career time horizons, demographics, cultural and educational background (Agrawal *et al.*, 2015; Mayinka *et al.*, 2016). Thus, OLM careers need not necessarily follow the same overall logic or path. For example, some workers may enter not primarily for financial reasons but for the flexibility in work schedules or to practice their skills. Prior work on careers in external labor markets implicitly assumes that workers in an occupation (e.g. comedians, actors or alike) structure their careers in the same way to progress in similar career outcomes. Relatedly, online freelancers do not necessarily work full-time (Manyika *et al.*, 2016) so that their career horizons and the way they navigate their careers may vary.

Nature of employment. Opposed to offline freelancing, online work is mediated by a platform. Although workers may change clients frequently, they may still work within the boundaries of a single platform, providing their careers some sort of "backbone". There is even an ongoing debate about the legal status of workers, whether they are independent workers or employees (*see Chapter 2*). If online labor primarily mimics traditional employment relationships, careers might also develop in different ways than "true" freelancing careers. Finally, online jobs are often comparable to "employee-like" jobs, e.g. customer service representative, personal assistant, architect or marketing specialist. As such, workers may already have an idea what it takes to progress in their respective field (Horton & Tambe, 2019).

Market transparency. Given that freelancers' profiles on the platform include detailed work histories, workers are able to collect data on the career paths of other freelancers. Offline freelancers can consult their local social networks but they will not be able to gather such detailed information on how other freelancers structure their careers and the respective outcomes. Likewise, job descriptions and price tags are observable so that workers can identify more complex tasks and the respective skill requirements easily. This will help freelancers in deciding which skills to develop and what to do next. Further, platforms use the data they gather from online transactions to identify in-demand skills and tasks, thereby assisting workers in identifying interesting career opportunities. For example, Upwork publishes quarterly "Skill Reports" with the most in-demand skills.

In sum, the work environment of online freelancers is fundamentally different from offline freelancing, calling for more specialized research on careers in OLMs. Although an emerging stream of research has begun to study mobility patterns of workers on OLM platforms, the question of how careers predominantly unfold and differ from offline careers remains unanswered. For example, Horton and Tambe (2019) find that web developers react to skill shocks (i.e. the sudden decline of Adobe Flash) by quickly adapting their skill portfolio through on-the-job learning, online tutorials and alike. Their results suggest that careers can develop in very nuanced ways. Other researchers study the task bundles of workers (Kokkodis & Ipeirotis, 2016; Leung, 2014). For example, the closest study to ours examines whether the chronological order of job categories a freelancer has worked in (the *career trajectory*) affects employers' hiring decisions (Leung, 2014). The author finds that employers choose applicants who move incrementally between similar jobs over those who do not move (specialize in one job category) or those with highly diverse job histories (move between highly dissimilar job categories). These results provide some initial evidence that workers construct their careers differently and that career histories are meaningful.

4 Data and Method

4.1 Data & Sample

We focus on Upwork.com, the world's largest freelancing website. Opposed to many other OLMs, Upwork encourages longer-term projects and high-value, ongoing work (Pofeldt, 2016). Examples for work on Upwork are developing an online marketing strategy (Sales and Marketing), porting an Android app from an iOS app and adding new features to an existing app designed with C++ (both Web, Mobile, and Software Development).

Upwork is an ideal empirical context for two reasons. First, Upwork brokers jobs that are diverse in terms of task types, length, skill level and work setting (e.g. contract type). This allows for differential career types to emerge and creates potential for some form of progress (e.g. through working on more complex tasks over time). Second, Upwork is the largest freelancing website worldwide. This means freelancers may be less likely to multi-home given the supply of work on the platform. Consequently, it is more likely that we indeed document freelancers' complete online career because they work only on a single platform.

The original dataset was gathered in 2017 and includes data on 255,393 freelancers with a minimum of one job. The sample was then restricted to 74,516 freelancers with a minimum of 10 jobs to ensure a certain career trajectory, resulting in 2,662,983 observations (transactions). The project starting dates are between May 2004 and October 2017 because some workers had already been active on prior versions of the platform.

4.2 Variables

In order to understand the differential ways of structuring work histories or put differently, the opportunity space workers face, we build on a growing body of research studying worker mobility at the unit of the task (e.g. Acemoglu & Autor, 2011; Autor, Murnane, Levy, 2003; Autor & Handel, 2013; Gathmann & Schönberg, 2010; Yamaguchi, 2012). Importantly, the task approach clearly distinguishes between the skills and tasks. Usually they have been treated interchangeably, often due to data availability issues (Autor, 2001). The distinction between tasks and skills is important for the study of careers in OLMs because both task and skill specialization provide careers an overall logic and illustrate career stability. A *task* refers to a unit of work activity that produces output. In contrast, a *skill* is a worker's stock of capabilities for performing tasks. This definition includes practical skills as well as theoretical and practical knowledge (Eggenberger, Rinawi, & Backes-Gellner, 2018). Workers apply skills to tasks in exchange for wages; skills applied to tasks produce output (Acemoglu & Autor, 2011). This distinction has important implications for career development. Specifically, it suggests that the human capital decision of workers on the platform is twofold: On the one hand, workers decide on the tasks they perform, i.e. where they apply their skills, ranging from a very narrow to a broad range of activities to produce differential outcomes. For example, a graphic designer may decide to apply her graphic design skills to design only children's books or produces a more range of outputs. On the other hand, workers can decide on the skills to perform a task, i.e. which skills they develop and apply. For example, using different types of technologies to produce a certain output. This suggests that workers performing similar tasks might differ in the actual skills they apply. Conversely, workers having similar skills might perform different tasks and thus produce different outcomes.

Tasks. In line with prior research (Leung, 2014), we use the level of the sub task category to proxy tasks, which is the most fine-grained level of task boundaries on the platform. On Upwork, there are 89 subcategories available, ranging from Transcription and Data Entry to Software Development and Logo Design & Branding. A full list of the sub categories and their frequencies is provided in the *Appendix*.

Skills. In our study, skills refer to the knowledge and competencies necessary to perform a task, i.e. the concrete skills required to perform a job as indicated by employers. Thus, we are able to capture the multidimensionality of skills (Speer, 2017). For example, this includes technical skills (Python, Java, AutoCAD) or knowledge in certain areas (chemistry, statistics, engineering). Specifically, we use the skill tags attached to job postings, which is in line with prior work (e.g. Anderson, 2017). Skills can be more or less specific, i.e. applied to or productive in different areas. For example, Microsoft Office applications are rather “transversal” (Djumalieva & Sleeman, 2018) in the sense that they can be transferred to a broad range of tasks. However, some skills such as voice talent are rather specific because they can only be applied to a narrow range of tasks on the platform.

However, we initially faced the problem of many missing values on the skill level. On average, at least one skill is specified in approx. 16 per cent of the cases. 11 per cent of the jobs specify 2 skills and 8 per cent 3. Missing values are to a large extent due to privacy setting of jobs (i.e. the data could not be gathered). However, another reason for that low number could be that an employer may think that specifying main and sub task category is enough input to solicit qualified workers and skill specification may appear like a too strong restriction.

To extend our data foundation, we interpolate these skill labels for all job listings, using the text data in titles and descriptions. We model this problem as a multi-label predictive exercise. We restrict the number of possible skills to 1000 (from originally 3,222), which cumulatively make up 97% of all skill tags. We use modern recurrent neural network architecture to predict a variable amount of skills associated with a job posting[1]. Given the vast number of classes to predict as well as the variable number of labels per observation it is hard to point out a numerical performance measure. Also, the association of descriptions with labels provided by employers cannot be regarded as a gold standard on micro-level. Thus, it is possible that the neural net performs better at finding fitting skill tags for individual listings than the employer.

[1] Bidirectional Long-Short-Term-Memory network with 256 cells, Embedding Layer initiated with custom-trained Fasttext vectors not frozen, ~18 million trainable parameters., trained over 7 epochs

Table 3.1 shows a prediction on data the neural network has not been exposed to (test set). While there are some differences, the overall skill requirements are well predicted and even niche skills like "Japanese" identified.

Table 4.1: Comparison between Predicted and Original Skill Tags to Job Postings

Prediction	Original
[architectural-design, autocad]	[3d-design, cad-design, drafting, autodesk-rev...]
[css, html, php, web-design, website-developme...]	[web-design, wordpress, website-development, …]
[php, web-design, website-development, wordpress]	[css, html, php, wordpress]
[japanese, translation]	[japanese]
[article-writing, blog-writing, content-writing...]	[content-writing, internet-research]

We further describe freelancers in our full sample and associated career clusters (result of our cluster analysis) in terms of the following characteristics.

Freelancer characteristics. *Advanced economy* refers to the economic situation of a freelancer's location based on data by the International Monetary Fund (World Economic Outlook, 2017)[2]. It is 1 for workers from advanced economies and 0 for freelancers from emerging and developing economies. *High education* is a dummy equal 1 if the freelancer reports having undergraduate, graduate, or PhD education.

Online activity and outcomes. We construct a set of variables to capture the time dimension of online careers. We are interested in the overall time spent on the platform and the intensity. First, we use the number of jobs on Upwork to measure *career length*. We also integrate a variable capturing the *share of hourly paid* compared to fixed-price jobs. The contract type has implications for autonomy and income stability and is thus an important characteristic of an online freelancer's career and work environment. We finally describe workers associated career

[2] World Economic Outlook, 2017: https://www.imf.org/external/pubs/ft/weo/2017/02/weodata/groups.htm

outcomes with career paths. We create the variable *average job size*, capturing the average size of projects in terms of total amount charged. We further measure the *average hourly pay* a worker earns on the job.

4.3 Method Pipeline

We apply several methods to explore workers' career patterns and identify differences in terms of their task and skill portfolios. As a first step and exploratory analysis, we cluster workers task portfolios to uncover similarities in task bundles. To do so, we apply techniques to reduce the dimensionality of our data to prepare for our cluster analysis. We then use a novel cluster algorithm to identify workers who bundle their tasks in similar ways. Next, we matched the emerging clusters with skill data to examine the skill content associated with them. In our second analysis, we explore career development paths over time by calculating average task and skill walks of workers. That is, we measure whether workers move, on average, frequently and between rather distant or related tasks and skillsets. This will give us an idea how specialization patterns look like. We explain our method pipeline in detail in the following.

(1) Cluster Analysis

In freelancer's task portfolios, each job is described by one of the 89 sub task categories used on the platform. Portfolios are of variable length and each task category may occur multiple times in one portfolio. We transformed these portfolios into a sparse matrix using a bag of words approach, commonly seen in natural language processing (NLP). Just as in the domain of NLP, we assume that co-occurrence generally indicates similarity. We reduce the dimensionality of the matrix using Non-negative Matrix Factorization (NMF), an algorithm that is comparable to more traditional approaches for dimensionality reduction such as Principle Component Analysis (PCA) or Singular Value Decomposition (SVD) but performs better on sparse and positive data such as in the present case (Lee & Seung, 1999)[3]. The dimensionality reduced dense matrix has 20 features, each of which is a composition of the initial 89 sub-categories[4].

We then apply UMAP (Uniform Manifold Approximation and Projection for Dimension Reduction) to prepare the data for visualization and clustering in a single step. UMAP is a novel and highly efficient dimension reduction technique proposed by McInnes, Healy, and Melville (2018) that let us use angular distance metrics and thus produces a projection that preserves the "quasi-semantic" features of our data. We used HDBSCAN (Campello, Moulavi, & Sander, 2013) as a clustering algorithm on top of the two-dimensional embeddings produced by UMAP. Combining dimensionality reduction with clustering in this order has been proposed by the authors of UMAP and has led to the identification of well interpretable clusters. HDBSCAN is a high-performance density based hierarchical clustering technique that is geared toward "bottom-up" identification of clusters. Compared to commonly used approaches such as K-means, HDBSCAN allows for "irregularly" shaped clusters, which are common when working with data describing social rather than natural phenomena.

To understand the composition of clusters we use an approach common in NLP, in which we apply Term Frequency Inverse Document Frequency (TF-IDF) scaling to identify the most prominent elements present in a cluster in comparison to all others (Salton & Buckley, 1988):

$$w_{i,j} = tf_{i,j} \cdot \log\left(\frac{N}{df_i}\right)$$

Here $w_{i,j}$ is the weight of an element *i* in some container *j* (in our case a task category or a skill in a freelancer's portfolio; alternatively aggregated up to cluster level), $tf_{i,j}$ is the number of occurrences of element *i* in the aggregation of interest *j*, *N* is the overall number of containers

[3] We obtained comparable results using truncated SVD, a variation of SVD for sparse matrices.

[4] Unfortunately, no straightforward methodology exists for choosing the dimensionality of the NMF as it would be with PCA, where one would select the number of components where the marginal increase of the explained variance in the data is low. We ran experiments with other values but the choice of 20 components seemed appropriate with regard to computational efficiency, stability of latent patterns and given the use of the dense matrix as input for further dimensionality reduction with UMAP.

(portfolios), df_i the total number of occurrences of element *i*. Suppose that element *i* occurs in 9 out of 10 containers – a rather general element. For container *j* with 2 occurrences of *i* TF-IDF would discount the value to 2 * log (10 / 9) = 0.09.

The HDBSCAN technique asks for several parameters that influence the number of identified clusters. This stands in contrast to K-means, which requires the specification of a number *k* of clusters to be identified. The tuning of these parameters, specifically *minimum cluster size* and *minimum samples* has to be considered. While the former parameter is rather straightforward, the latter defines how conservative the clustering should be. Higher values result in denser areas being considered for clustering, leaving more unclustered observations (i.e. noise). Optimal parameters had to be determined empirically. We aimed at reducing the share of noise, while avoiding excessive fragmentation but also (task and skill) overlap for job category clustering.

UMAP generates a 2-dimensional projection of the initially 20-dimensional matrix resulting from NMF. The two dimensions and axes in the visualizations (see *Figure 3.1*) can be interpreted as orthogonal compositions of various discrete features present in freelancers' career portfolios. The distance between points and clusters can be interpreted as distance or similarity on different and potentially latent levels.

For our second analysis, in which we study career paths over time, we calculate the average distance between two consecutive jobs both on a task and a skill level. We refer to these as "skill walks" and "task walks".

(2) Skill Walks

In this step, we find indicators for variation or lack hereof in the career paths of freelancers. We adopt a dynamic approach to measuring skill diversity on freelancer level, that tries to capture the behaviour of individuals over time and cannot be operationalised by a diversity measure such as Herfindahl-Hirschman Index that is inherently static. Workers that only take on jobs, which require very similar skillsets, are likely to be very specialized while freelancers, whose jobs exhibit very varying skill requirements can be considered diversified with multidirectional paths. As jobs in our data are ordered by time, we aim at calculating “skill distances” between all consecutive jobs in workers' job portfolios. For this, we represent each job as a “basket” of skills encoded as a vector of length 1000 (the skills to which we limited the skill predictions). We apply Non-negative Matrix Factorisation (30 dimensions) to reduce the dimensionality. Each dimension here represents a scaled thematic grouping of related skills (e.g. marketing, translation). From this, we obtain short, dense vectors that represent each job as a combination of skill themes with a positive variable value assigned to it. This approach allows us to rely on the predicted skills since the presence or absence of a similar skill in a basket (the kind of variability introduced by the neural network) does not really alter this reduced representation. To determine the distance between each pair of consecutive jobs in a freelancer’s work history, we calculate:

$$\textit{Skill Walk} = 1 - \cos(\theta) \text{ with } \cos(\theta) = \frac{\vec{J}_t \cdot \vec{J}_{t+1}}{\lVert J_t \rVert \lVert J_{t+1} \rVert}$$

where $\vec{J}_t$ is a job performed by a freelancer represented as a vector of skills, and $\vec{J}_{t+1}$ the job performed right after, similarly represented as a skill vector. This distance measure can be interpreted as the skill distance between two consecutive jobs performed by a freelancer, which we refer to as skill walks. The smaller the angle, the higher the cosine similarity, the lower the distance. Values are between 0 and 1, where 0 indicates a perfect overlap.

(3) Task Walks

We use a very similar approach to calculate corresponding measures in terms of task categories on the freelancer level. Since each job is associated with exactly one task sub category, we need to observe cases of co-occurrence, assuming that co-occurring task sub categories are similar. Given that Upwork supports more high-skill work, we assume that there is a “natural” limit to movements across boundaries. That is, we assume that cases of workers jumping from 3D modelling to Chinese translation to java programming, will be an exception. Related work

(Leung, 2014) also makes the assumption of co-occurrence to assess similarity of task categories.

Task categories are (just as previously for the clustering exercise) aggregated on the freelancer level and encoded as a sparse freelancer-task-category matrix with freelancers as rows and sub-tasks as columns. The dot-product between this matrix and its transpose returns a square matrix – similar to a correlation matrix. Each row can be interpreted as a vector that represents the task sub-category in this row as a linear combination of all other tasks. For example, the cosine similarity between the task category “Graphic design” and “Photography” is 0.9538 and they co-occur very frequently in workers’ histories. Conversely, the similarity between “Graphic design” and “Game Development” is lower, 0.5268, so that they co-occur less frequently but are still related. Having such representations of each task category allows us to calculate task category distances over time just as in the case of skills. With that, we have dynamic indicators for freelancers’ skill and task variation. Comparing these values, we can identify and categorise freelancers given their career trajectories. Again, values lie between 0 and 1, where 0 indicates a perfect overlap between all task categories.

5 Results

5.1 Descriptive Statistics

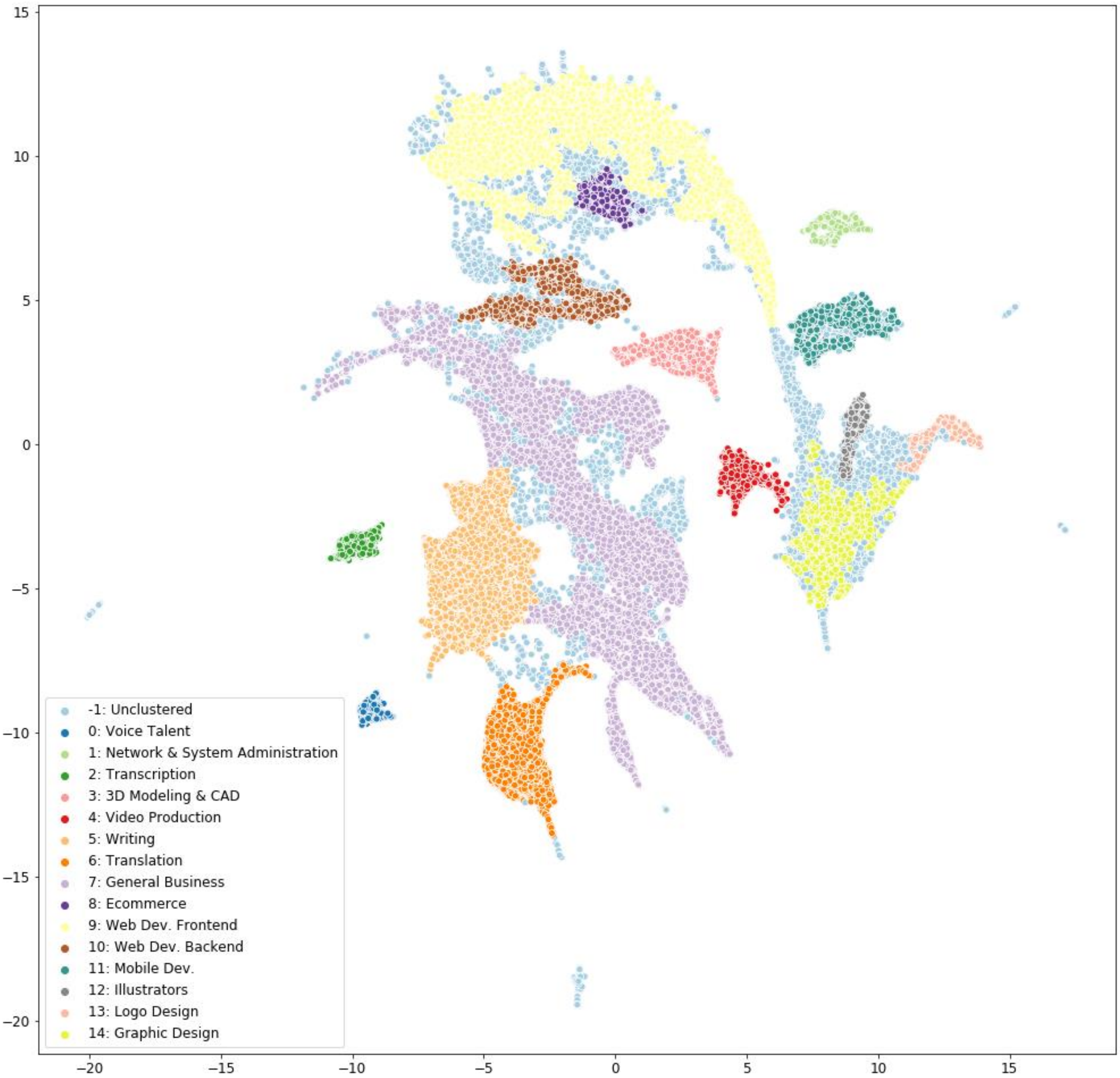
In our sample, workers have on average accomplished 89 jobs (median: 51) on the platform with an average total charge of 469 USD (median: 66.7) per job. This suggests that workers seem to stay on the platform for quite some time but also that some workers earn a relatively stable income. Further, the average hourly wage is 15 USD (median: 11.11 USD), suggesting that jobs on the platform are relatively high-skill. This is further supported by the average education of workers: 80% report having a postgraduate degree (i.e. Bachelor, Master, PhD). The top three task categories (assessed by number of jobs in respective task category in our sample) are Web Development (542,625 jobs: 20.38%), Graphic Design (341,809 jobs:

12.84%) and Article & Blog Writing (170,696 jobs: 6.41%). A large market size generally enables a finer division of labor (Rosen, 1983) so that some workers may indeed build relatively stable and specialized careers. In line with prior work, the majority of workers comes from less developed countries (72%). Half of the freelancers (51%) are located in Asia, followed by Europe (23%), North America (19%), Africa and South America (3%), and Oceania (1%). The top five freelancer countries are India (20%), USA (15%), Philippines (12%), Bangladesh (7%), and Ukraine (5%). In sum, the digital workforce is very diverse in terms of location but tends to be relatively high skill. Whereas some workers seem to have relatively stable careers in terms of average job size, some workers have to hop more frequently to new jobs to have a stable income.

5.2 Career Clusters by Task Categories

The result of our clustering of freelancer's work histories by their task subcategories is shown in *Figure 5.1*. The clustering identifies task combinations that co-occur frequently and lets us compare clusters by their characteristics. The names of the cluster are identified based on the task category with the highest tf-idf score, i.e. the most prominent element in worker's portfolios. *Table A.1.A* in our *Appendix* gives some detailed descriptives of our career clusters. Although the clustering does not take into account the timely order of tasks or explicitly the skills attached to tasks, the resulting task map and our exploration of these clusters already provides some interesting insight on workers' careers as well as the task and skill space.

Figure 5.1: Cluster Analysis Task Categories



The emerging career landscape (see *Figure 5.1*) can be described as a task space where one point refers to one freelancer's task portfolio. The numbers on the axes cannot be interpreted directly but rather describe the position of a freelancer's task portfolio in the projected space in relation to all other freelancers. If two clusters are adjacent to each other, they share many of the tasks and are only differentiated by some tasks almost exclusive to one career cluster (suggested by high tf-idf scores) and/or the relative frequency of particular tasks in a job portfolio. A career is a position on the task category map, and career development comprises movements through the task space (which we study in the second part of our paper). *Figure 5.1* and *Table A.1.A (Appendix)* suggest that career clusters differ in terms of their *distance*, *size*, and *specialization* with implications for career development.

Distance. Some clusters are closer than others, which can be interpreted as relatedness. The emerging career clusters can be divided broadly into language-oriented (Transcription, Writing, Translation, Voice Talent), business-oriented (General Business, Ecommerce), IT-related (Web Backend, Web Frontend, Network & System Administration, Mobile Development, 3D Modeling & CAD), and creative task areas (Logo Design, Graphic Design, Video Production, Illustration). For example, the cluster "Transcription" is close to both "General Business" and more language-oriented areas such as "Writing" or "Translation". Taking a closer look at the attached skills to jobs in these clusters reveals why: Transcription tasks require knowledge in transcription, and skills such as typing, English (or another language), and Microsoft Word. These skills are also demanded for tasks in writing and translation but also for admin support tasks such as data entry, web research or writing business plans. This is further reflected in the low tf-idf scores for these skills (see *Table A.1.A*), i.e. they do not uniquely identify transcription tasks but are rather transversal skills. Consequently, workers focusing on these skills can move between different tasks while sustaining a coherent skill portfolio. Since clusters are based on the co-occurrence of the associated task categories in workers' task portfolios, the intuitively meaningful distance between career clusters may also suggest that workers do not predominantly combine very distinct task categories. Nevertheless, the low tf-idf scores of some task categories within some clusters, e.g. Web Development, also imply that these task categories occur in many of the freelancers' task portfolios that are in potentially distant clusters. For example, even a freelancer in the cluster "Translation" may work on Web Development tasks. We take a closer look at these within-cluster differences when describing differences in specialization and in the second part of our analysis.

Size. Most freelancers are in the career clusters General Business (18,458) and Web Frontend (13,162), followed by Writing (9,023), Graphic Design (5,343), and Translation (4,925). Interestingly, whereas Writing, Graphic Design and Translation reflect rather traditional freelance occupations readily found in the offline world, for some clusters there is no obvious offline equivalent. For example, General Business comprises various support tasks such as data entry, web research, and personal/virtual assistance. Offline, these tasks were difficult to contract out on a flexible basis and large scale. If anything, business services agencies may have provided these services offline, but it was rare for individuals offering, e.g. accounting services, personal assistance and web research services simultaneously. There are also smaller clusters based on "hot skills" (e.g. 3D Modeling) and niche skills (e.g. Voice Talent). Due to the global aggregated demand, workers can decide to build an entire career on these tasks and respective skills. In sum, in addition to simple transfers from offline to online freelancing/employee jobs, new career clusters emerge in OLMs.

Specialization. Career clusters seem to vary in their degree of specialization. This is based on the fact that task categories generally differ in their broadness. For example, Transcription is a rather narrowly specified category, whereas Web Development is broader and more complex and thus subsumes many different subtasks and skills. The fact, that we still observe workers that focus entirely on narrow task categories such as Transcription suggests that some workers indeed become very specialized. However, since the names of our clusters are based on tf-idf scores, i.e. most prominent but not necessarily most frequent elements in each worker's portfolio, and to understand within-cluster specialization, we descriptively explored the cluster "Transcription" to understand the differential ways of constructing careers before we conducted our second analysis (task and skill walks). We observe that 22 percent of freelancers in this cluster have a share of 75% or higher of Transcription tasks in their work portfolio; 70 percent of freelancers have an intermediate share (25-75%) and 8% less than 25% Transcription tasks. To understand the types of tasks that freelancers combine with Transcription tasks, we list the top ten task categories present in workers' task bundles (cluster "Transcription") in *Table 5.1.*

Table 5.1: Top 10 Task Categories in Workers' Portfolios in Cluster "Transcription"

Task Category	Freq.	Share	Cumulative
Transcription	21,815	60.16	60.16
Other - Writing	3,855	10.63	70.79
Web Development	1,633	4.50	75.29
Article & Blog Writing	1,233	3.40	78.69
Data Entry	914	2.52	81.21
Editing & Proofreading	604	1.67	82.88
General Translation	596	1.64	84.52
Personal / Virtual Assistant	573	1.58	86.10
Web Research	441	1.22	87.32
Audio Production	417	1.15	88.47

As evident from *Table 5.1*, workers combine Transcription with a broad range of tasks but predominantly with those that overlap in their skill requirements, either more language-oriented (Other-Writing, Article & Blog Writing, Editing & Proofreading, General Translation, Audio Production) or more administrative (Data Entry, Personal / Virtual Assistant, Web Research), suggesting two different development paths that would both create coherent work histories. Nevertheless, some workers also combine Transcription with rather unrelated tasks (Web Development). This suggests that at least some workers may move in structured ways and have more specialized and coherent work histories but that some are multidirectional and "truly boundaryless".

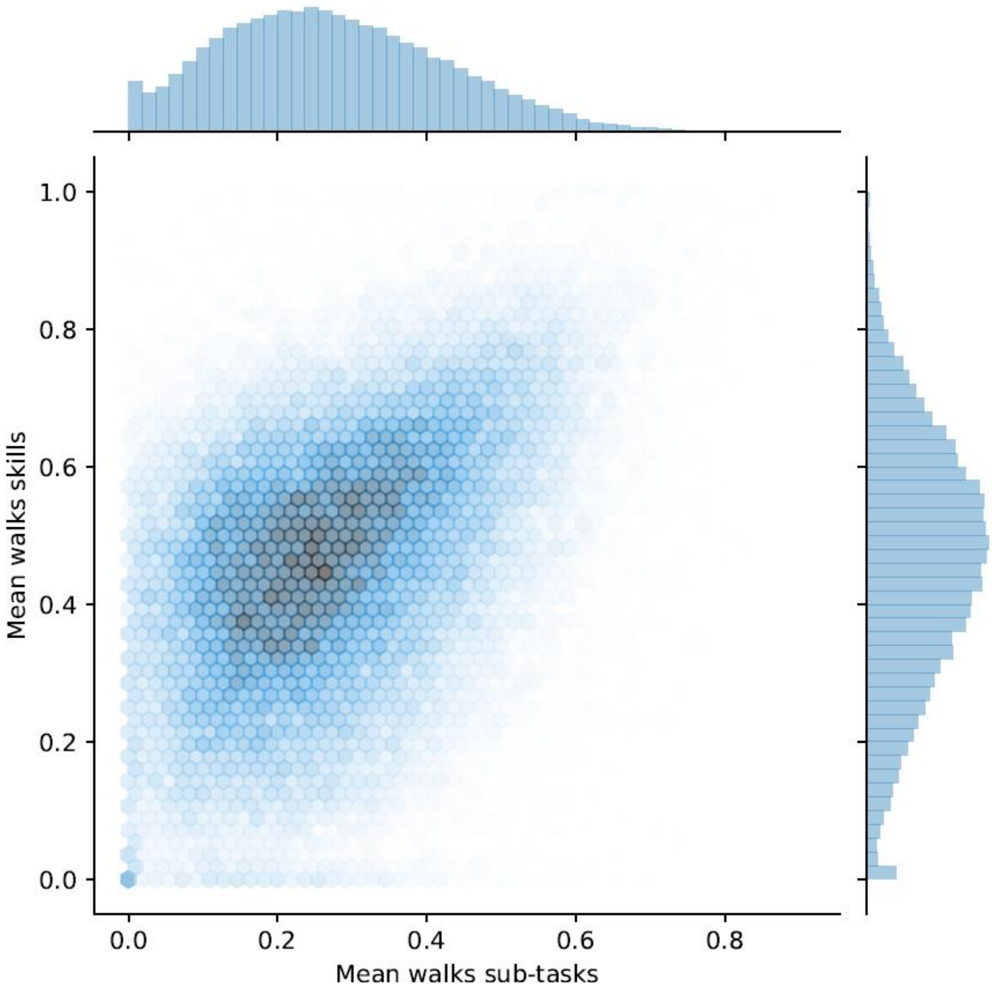
Note that a fraction of freelancers (12%) could not be categorized (-1). As evident from the task portfolios of these unclustered freelancers (*Table A.1.A, Appendix*), they seem to move around or across high-demand categories such as web development, general business jobs, and graphic design. Since these task categories have low tf-idf scores, workers do not have highly distinct or prominent elements in their task portfolio and could thus not be categorized.

In sum, our exploratory analysis suggests that there is variation in how workers construct their careers, ranging from more specialized to diversified work histories. We further find initial evidence that the distinction between tasks and skills is meaningful to understand career development paths in OLMs. Since our cluster analysis excludes the timely order of jobs and only implicitly studies skills, we further cluster workers based on the aforementioned movements through the task and skill space to identify differential career paths. This will help us understand the stability versus multi-directionality of careers.

5.3 Task and Skill Walks: A Taxonomy of Career Paths

Figure 5.2 illustrates the mean task and skill walks workers accomplish throughout their platform careers. The density bars indicate where most freelancers lie on a continuum between not moving at all (close to 0 on both dimensions) or moving opportunistically between tasks and skills (close to 1 on both dimensions).

Figure 5.2: Mean Task and Skill Walks of Workers



As evident from *Figure 5.2*, most freelancers show an intermediate degree of diversification on both dimensions. Thus, freelancers move between different types of tasks and skills from prior jobs but not at an extreme level. Remember that a jump from Graphic Design to Game Development corresponds to a cosine similarity score of 0.5268 (i.e. a task walk of 1-0.5268=0.47). Intuitively, these tasks are not completely unrelated. Since most workers show an average task walk score of 0.2 to 0.25, they move between rather related task categories. We further observe that a fraction of freelancers specializes very narrowly on almost single tasks and skills (left corner). At the same time, we find evidence for frequent movements between unrelated tasks and skillsets (right corner). Finally, there are workers diversifying only on one of the two dimensions. They have high values for task (skill) movements but not skills (tasks). Given that offline workers neither work only on one task or skill in a repetitive way nor are able to jump between extremely different types of tasks and skills within a single job, we argue that these career paths represent novel and OLM-specific careers. Even though some manufacturing workers may perform certain activities repetitively, this may be a novel phenomenon in the context of knowledge work. Based on these observations, we developed a taxonomy of career paths (*Table 5.2*).

The four prototypical career types are: a task specialist offering a narrow set of skills (*Micro career*), a task specialist offering a broad set of skills within a task category (*Task expert*), a task generalist who carries a specific skill across task categories (*Skill expert*), and a task generalist who possesses a broad set of skills (*Opportunist*). Since these four types have their roots in conventional career models but represent new, OLM-specific types of careers paths, they can be considered “same same but different”.

Table 5.2: Taxonomy of Career Paths

	Skill Specialist	Skill Generalist
Task Specialist	*Micro career* Skill focus within one category **Example:** focus on task category based on narrow skillsets: e.g. transcription, logo design; focus on narrow skillset in broader category (rare programming language within web development)	*Task expert* Working on different skills in one category **Example:** graphic designer with broad skillset, from logo design to print and web design; multilingual translator in Translation
Task Generalist	*Skill expert* Skill focus but across categories **Example:** Python or Microsoft Excel expert working in various categories	*Opportunist* Broad collection of tasks and skills **Example:** Freelancer working on data entry, translation, web development, data science & analytics

Each of these types has a distinct development path. A freelancer on a *micro career* path is likely to develop expertise in the narrow skillset she possesses, thus drawing on economies of scale by repetitively working on the same task. Stability in terms of the types of jobs taken is likely to be high, and pay is likely to increase due to more experience and skill proficiency. Following this path full-time and with a long-term orientation is, however, only possible if there are many tasks requiring the respective skill. Nevertheless, due to the repetitiveness and the risk of boredom and alienation, this career path might serve a different purpose than developing a long-term career on the platform (at least if the workers continues this path). So although this path is very coherent and stable, it implies the risk of being a “dead end”. At the same time, this hyper-specialization entails some risk if demand declines. An example is a transcriptionist working repetitively on the same types of transcriptions without variation in skills (e.g. different types of transcriptions that additionally require knowledge in medicine, music or legal aspects). Micro careers are obviously a special case of skill and task specialization.

A *task expert* can move within the task category and fill out different and likely broader roles within the task category. This reduces the risk of her skills becoming obsolete as a task expert covers a broad range of skills needed in a task category. Of course, this also depends on the breadth or rather skill variety of the task category. An example is a multilingual translator who can switch between different directions and combinations of translations so that the skill distance between jobs is relatively high although the freelancer focuses on one task category. Although the freelancer could also apply her language skills to different task categories such as "Proofreading", she rather remains within single task boundaries. Another example is a web developer specializing in "Web Development" who can either focus on one specific skill and a subtask of web development or develop are broader range skillset for various stages of the web development process. A potential threat to this path is the automation of tasks by algorithms, e.g. translation through deep learning translators such as www.DeepL.com/Translator. This also applies to micro careers.

A *skill specialist* is an expert in an in-demand and broadly applicable skill. For example, we observe skill tracks for MicrosoftExcel experts. They are able to apply their skills to a broad range of tasks to produce various outputs: Input for business plans, financial data analysis, data visualization, fixing bugs in macros etc. By applying the skill to a broad range of problems, the freelancer becomes very proficient in the skill. Building careers entirely on single skills or a narrow skill set is a fairly novel phenomenon. This is enabled by OLMs aggregating global skill demand, offering skill-based matching through the outsourcing of small tasks on easy-to-access/-use platforms. Traditionally, occupations and jobs in organizations have a broader task bundle (e.g. an employee responsible for the IT infrastructure of a company). Similarly, offline freelancers typically need a broader skill set to be successful and employable in the long run (O'Mahony & Bechky, 2006). However, the path carries the risk of the skill losing importance. Freelancers building their careers on certain skills may only be successful in the long run if they can predict demand of a skill accurately (Horton & Tambe, 2019). They have to observe the market to eventually becoming a first mover if a new skill emerges and develop strategies more similar to firms rather than employees. In sum, even though workers may pursue “fairly traditional” careers within certain boundaries, career paths seem to have a shorter time horizon and limited long-term career prospects.

Finally, an *opportunist* takes jobs as they present themselves. This career type covers a large range of task categories and skills so that careers can be seen as extreme case of boundaryless careers: Freelancers cross job boundaries frequently (even between unrelated jobs) and in a seemingly unstructured way, illustrating the multi-directionality of their careers. For example, a freelancer in our sample works on tasks in web development, scripts & utilities, data entry, other admin support, graphic design, web research, QA testing; data extraction & scraping and frequently hops between them. The required skills range from Adobe Photoshop to Python and Excel-VBA so that the work history is highly erratic. In offline settings, it would be very difficult if not even impossible to combine those tasks in one job so that this career type also represents a new form of career. Despite the lack of an overall logic, this path can still serve a certain purpose, e.g. access to employers from certain countries (remaining within geographical boundaries) or increased job satisfaction through flexible work schedules and job variety. At the same time, this path enables picking up general skills (e.g. time management, self-marketing, negotiation, communication) on the job. Then, the types of jobs become less important but another less visible factor that guides their career choices.

6 Discussion and Conclusion

6.1 Discussion

Building on a novel and extensive dataset from one of the world’s leading freelancing websites, we use a quantitative-exploratory approach to identify differential career paths in OLMs. Specifically, we apply a novel and advanced cluster algorithm as well as an innovative approach to show variations in workers’ careers.

Coming back to our question of whether we observe primarily multidirectional careers just like in offline freelancing settings, our results suggest that the careers of some freelancers appear fairly stable based on the types of jobs they work. The dominant pattern seems to be that workers apply their skills across task categories to ensure future employability without moving too far. Whereas some seem to do so in a more structured way (i.e. low average walks) others move more frequently and/or between more distant jobs. The former might illustrate what prior work refers to as “stretchwork” (O’Mahony & Bechky, 2006) or “incremental moves” (Leung, 2014). However, careers are also different from offline settings because we observe extreme degrees of specialization and diversification. Building on this insight, we developed a taxonomy of four particularly interesting and novel types of freelancer careers: a task specialist offering a narrow set of skills (*Micro career*), a task specialist offering a broad set of skills (*Task expert*), a task generalist who carries a specific skill across task categories (*Skill specialist*), and a task generalist with a broad skillset (*Opportunist*). On the one hand, careers become very fragmented and narrow due to the skill-based and more short-term oriented nature of work. Similar to the traditional and more stable career, workers focus on a narrower set of jobs and remain within task and skill boundaries. On the other hand, the boundaryless career becomes so diversified and multidirectional up to a point where it becomes difficult to objectively assess the overall logic and direction. Nevertheless, these careers can serve a specific purpose that might be more subjective and personal.

The increasing fragmentation and specialization as well as the diversification of work histories is a novel phenomenon. Because firms can outsource discrete, small chunks of work online, freelancers can build careers on a narrow skillset and repetitive tasks. While the existence of specialists and generalists is not new, the distinction between skill specialists and generalists is specific to and enabled by digital labor markets. In addition, OLMs allow for the emergence of novel occupations, such as 3D designer and ecommerce developers. Freelancing careers in OLMs are further different because career development online looks similar to firms’ strategic behavior in competitive markets. That is, freelancers can rely on economies of scale and scope in skill provision and development, develop differentiation vs. price leadership strategies to capture demand from different types of employers (quality vs. price sensitive), analyze skill demand and supply like traditional resources and production input factors and react to competitive forces such as the emergence of new entrants due to the introduction of novel task categories or other design features.

Our results resonate with previous research on the effects of skill shocks on freelancer behavior in OLMs, which finds evidence for skill-based careers of freelancers in Web Development (Horton & Tambe, 2019). Further, whereas research in organizational settings only observes tasks to proxy for skills (e.g. Gathmann & Schönberg, 2010), we identify the exact skills needed to carry out a task. In doing so, we break down job specialization further than prior work. Our taxonomy further complements existing research on external labor markets (O'Mahony & Bechky, 2006) by identifying career paths that are not designed to develop one's future employability by diversifying into different areas. We argue that this may stem from the fact that workers do not necessarily work full-time on the platform or use OLMs between "regular" (offline) jobs (e.g. a college graduate working as a translator after graduation to earn supplemental income while traveling).

The identified career types further suggest that platforms may become the "organizational backbone" of freelancers' careers: By defining task boundaries to structure the market for tasks, platforms shape workers' careers and help them structuring them. This further implies that the way platforms define these categories can have important consequences for career development. While narrowly defined task boundaries may help workers and employers to quickly find a match, those workers focusing on narrowly defined categories could also signal some sort of career stagnancy to employers (Leung, 2014). Platforms further have the potential to provide workers with more active career support. For example, by using big data from transactions, platforms could develop new tools and features (e.g. algorithmic career support) or identify and constantly update prototypical career patterns to guide workers in what to do next. Platforms could further help developing careers by integrating learning tools and online courses similar to Coursera, Datacamp or EdX.

6.2 Limitations and Future Research

Although we provide novel insights into OLM careers, a detailed assessment of each career paths and the factors that driver their emergence as well as associated outcomes is beyond the exploratory scope of our paper. The focus of our descriptive paper is rather to detect certain patterns to guide future research on the causal mechanisms and dynamics behind them.

Activity Patterns. In this study, we only explore the coherence of work portfolios to illustrate stable versus multidirectional careers but do not explicitly study the time dimension of careers. Thus, we do not know about their career stability in terms of time between jobs to assess continuous employment on the platform or the degree of engagement they have (full-time, part-time or as needed). However, our primary interest was to describe the structure of jobs in terms of their content. But future research could explore the engagement level and (in-)activity patterns of freelancers in combination with specialization patterns. This may also help us understanding how workers integrate sequences of online jobs into their regular, offline career, e.g. whether they only work online during times of unemployment (Borchert et al., 2018).

Drivers of Career Paths. In our study, we explore how career paths look like but not explicitly the motivations to follow these paths. As such, future research should explore this question in more detail. For example, narrowly defined careers may have their roots in the diverse motivations of workers to enter OLMs. If a worker uses OLMs only for a limited amount of time or intermittently to earn supplemental income, future employability and flexibility are not as important. Rather, these short-term oriented workers “exploit” a narrow skillset (e.g. transcription) with limited motivation to develop new ones, which then results in efficiency gains and economies of scale. However, freelancers focusing on niche skills may also realize higher prices due to high demand and low (global) supply. Then the worker possessing a rare skill has also no incentive to develop new skills and diversify into another area if she only has a short-term perspective. For freelancers aiming at developing a long-term career, this implies that starting their career in a broadly defined market with more diverse skill requirements will let them gradually move into different directions if demand changes. Conversely, there is no logical path or journey for a voice talent as they reside in isolated clusters and will experience a micro career within a small skill spectrum.

Career Direction and Outcomes. We study career progression only implicitly through the coherence of portfolios rather than explicitly analyzing growth in objective career measures and how they develop from job to job. We describe careers in terms of their task and skill content, assuming that the way of bundling and ordering them will affect career progression. How exactly different career patterns relate to financial career outcomes is beyond the scope of our paper. Future research could assess whether freelancers pursuing micro careers have a more stable income than field experts due to the similarity of jobs. Task experts, on the other hand, may realize higher hourly wages because they work on complex projects requiring unique skill portfolios. Studying wage mobility of developing country workers would also be a promising line of research. More generally, studying the career paths of top earners and whether there are multiple ways of reaching the top of the pay distribution would be promising for future work. Finally, online workers may have a variety of different measures, both objective and subjective, to assess their career progression. Likewise, how workers recognize career progression or more specifically, their visible career steps and environmental markers, e.g. moving to a co-working space instead of working from home, could represent a research opportunity. In sum, the study of career outcomes is important because it helps us to understand whether specialization is in fact related with career stability and progression or rather a proxy for career stagnation.

Coherence Patterns. We study the coherence of careers based on the jobs a person holds. However, coherence can also stem from working for certain types of employers (e.g. from specific countries). For example, anecdotal evidence suggests that workers use online jobs to transition into local labor markets by working only for employers from the respective geographical area. This helps gaining work experience and subsequent job transitions in an offline setting. Workers pursuing an opportunist path may indeed have a coherent set of employers and remain within geographical boundaries while crossing job boundaries. This insight would also help us understand the motivations to follow these seemingly erratic career paths. Relatedly, we study careers within platform boundaries. Although we do not believe that this will affect our results on the coherence of task portfolios, future research may dig deeper into freelancers' multiple online careers and identities. For example, future research could explore the role of platforms in shaping work patterns by comparing career paths of workers on multiple platforms. Studying mobility between platforms and whether workers build different types of careers on different platforms also represents a promising line of research.

In sum, our exploratory study complements and extends existing research in several ways by taking a novel perspective on career development in (external) online labor markets. We hope that our work will provide a springboard for interesting future research opportunities.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043-1171): Elsevier.

Anderson, K. A. (2017). Skill networks and measures of complex human capital. *Proceedings of the National Academy of Sciences*, 201706597.

Arthur, M. B., & Rousseau, D. M. (1996). A career lexicon for the 21st century. *Academy of Management Perspectives, 10*(4), 28-39.

Autor, D. H. (2001). Wiring the Labor Market. *Journal of Economic Perspectives, 15*(1), 25-40. doi:10.1257/jep.15.1.25

Autor, D. H., & Handel, M. J. (2013). Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, *31*(2), S59-S96. doi:10.1086/669332

Barley, S. R., Bechky, B. A., & Milliken, F. J. (2017). The Changing Nature of Work: Careers, Identities, and Work Lives in the 21st Century. *Academy of Management Discoveries, 3*(2), 111-115. doi:10.5465/amd.2017.0034

Borchert, K., Hirth, M., Kummer, M. E., Laitenberger, U., Slivko, O., & Viete, S. (2018). Unemployment and online labor. *ZEW-Centre for European Economic Research Discussion Paper*(18-023).

Campello, R. J., Moulavi, D., & Sander, J. (2013). *Density-based clustering based on hierarchical density estimates*. Paper presented at the Pacific-Asia conference on knowledge discovery and data mining.

DeFillippi, R. J., & Arthur, M. B. (1994). The boundaryless career: A competency-based perspective. *Journal of organizational behavior, 15*(4), 307-324.

Djumalieva, J., & Sleeman, C. (2018). *An Open and Data-driven Taxonomy of Skills Extracted from Online Job Adverts*. Retrieved from

Eggenberger, C., Rinawi, M., & Backes-Gellner, U. (2018). Occupational specificity: A new measurement based on training curricula and its effect on labor market outcomes. *Labour Economics*, *51*, 97-107. doi:https://doi.org/10.1016/j.labeco.2017.11.010

Faulkner, R. R., & Anderson, A. B. (1987). Short-Term Projects and Emergent Careers: Evidence from Hollywood. *American Journal of Sociology, 92*(4), 879-909. Retrieved from http://www.jstor.org/stable/2780042

Ferguson, J.-P., & Hasan, S. (2013). Specialization and Career Dynamics: Evidence from the Indian Administrative Service. *Administrative Science Quarterly, 58*(2), 233-256. doi:10.1177/0001839213486759

Gathmann, C., & Schönberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics, 28*(1), 1-49. doi:10.1086/649786

Greenhaus, J. H., Callanan, G. A., & DiRenzo, M. (2008). A boundaryless perspective on careers. *Handbook of organizational behavior*, *1*, 277-299.

Hall, D. T., & Chandler, D. E. (2005). Psychological success: When the career is a calling. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *26*(2), 155-176.

Higgins, M. C. (2001). Changing careers: the effects of social context. *Journal of organizational behavior, 22*(6), 595-618. doi:10.1002/job.104

Joseph, D., Fong Boh, W., Ang, S., & Slaughter, S. A. (2012). The career paths less (or more) traveled: A sequence analysis of IT career histories, mobility patterns, and career success *MIS Quarterly, 36*(2), 427-A424. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=74717015&site=ehost-live

Kanter, R. M. (1989). The new managerial work.

King, Z., Burke, S., & Pemberton, J. (2005). The 'bounded'career: An empirical study of human capital, career mobility and employment outcomes in a mediated labour market. *Human Relations, 58*(8), 981-1007.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788-791.

Leung, M. D. (2014). Dilettante or Renaissance Person? How the Order of Job Experiences Affects Hiring in an External Labor Market. *American Sociological Review, 79*(1), 136-158. doi:10.1177/0003122413518638

McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

O'Mahony, S., & Bechky, B. A. (2006). Stretchwork: Managing the career progression paradox in external labor markets. *Academy of Management Journal, 49*(5), 918-941. doi:10.5465/AMJ.2006.22798174

Rahman, H. A., Valentine, M., Bechky, B., Leung, M. D., & Levina, N. (2016). Online Labor Markets: Fad, Fringe, or Future of Work? *Academy of Management Annual Meeting Proceedings*, 1-1. doi:10.5465/AMBPP.2016.14694symposium

Reilly, P. (2017). The Layers of a Clown: Career Development in Cultural Production Industries. *Academy of Management Discoveries, 3*(2), 145-164.

Rosen, S. (1983). Specialization and Human Capital. *Journal of Labor Economics, 1*(1), 43. doi:10.1086/298003

Rosenbaum, J. E. (1979). Tournament Mobility: Career Patterns in a Corporation. *Administrative Science Quarterly, 24*(2), 220-241. doi:10.2307/2392495

Rosenfeld, R. A. (1992). Job Mobility and Career Processes. *Annual Review of Sociology, 18*(1), 39-61. doi:10.1146/annurev.so.18.080192.000351

Schein, E. H. (1978). *Career dynamics: Matching individual and organizational needs* (Vol. 6834): Addison Wesley Publishing Company.

Shaw, K. L. (1987). Occupational Change, Employer Change, and the Transferability of Skills. *Southern Economic Journal, 53*(3), 702-719. doi:10.2307/1058765

Speer, J. D. (2017). Pre-Market Skills, Occupational Choice, and Career Progression. *Journal of Human Resources, 52*(1), 187-246. doi:10.3368/jhr.52.1.0215-6940R

Spilerman, S. (1977). Careers, labor market structure, and socioeconomic achievement. *American Journal of Sociology, 83*(3), 551-593.

Sullivan, S. E. (1999). The Changing Nature of Careers: A Review and Research Agenda. *Journal of Management, 25*(3), 457-484. doi:10.1177/014920639902500308

Sunde, U. (2008). Human Capital Formation, Education and Earnings Inequality. *Applied Economics Quarterly, 54*(1), 7-26.

Wilensky, H. L. (1964). The Professionalization of Everyone? *American Journal of Sociology, 70*(2), 137-158. Retrieved from www.jstor.org/stable/2775206

Wythe, W. H. (1956). *The Organization Man*: Simon & Schuster.

Yamaguchi, S. (2012). Tasks and Heterogeneous Human Capital. *Journal of Labor Economics, 30*(1), 1-53. doi:10.1086/662066

Appendix

A.1 Appendix

A. 1.A: Descriptive Statistics Career Clusters

No.	***Career Cluster***	***Task Portfolio (Top3)***	***Skill Portfolio (Top5)***	***Size (Count)***	***Skill Space***	***Cluster Density***	***Demographics***	***Activity***	***Outcomes***
0	Voice Talent	Voice Talent, 0.99, Audio Production, 0.15, Video Production, 0.04,	voice-talent 0.69 voice-over 0.65 audio-production 0.19 audio-editing 0.17 audio-mixing 0.07	729	Language-oriented	High (0.995)	Freelancers from advanced economies: 81.46% Education high: 64.62%	Share hourly: 14,36%	Avg. Hourly Rate (overall): 32.06 USD Avg. job size: 154 USD
1	Network & System Administration	Network & System Admin, 0.99, Other - IT & Networking, 0.08, Web Development, 0.03	linux-system-administration 0.56 network-administration 0.27 asterisk 0.22 centos 0.20 amazon-ec2 0.18	907	Technology-oriented	High (0.973)	Freelancers from advanced economies: 28.45% Education high: 73.54%	Share hourly: 57,86%	Avg. Hourly Rate (overall): 24.89 USD Avg. job size: 629 USD
2	Transcription	Transcription, 0.99, Other - Writing, 0.091, Audio Production, 0.026	transcription 0.91 typing 0.18 English 0.15 microsoft-word 0.15 proofreading 0.15	939	Language-oriented	High (0.965)	Freelancers from advanced economies: 42.17% Education high: 72.79%	Share hourly: 31,36%	Avg. Hourly Rate (overall): 10.28 USD Avg. job size: 208 USD
3	3D Modeling & CAD	3D Modeling & CAD, 0.98, Architecture, 0.096, Animation, 0.095	3d-modeling 0.51 3d-design 0.39 3d-rendering 0.36 Autocad 0.29 cad-design 0.25	1,810	Technology-oriented / Creative	Rather high (0.936)	Freelancers from advanced economies: 29.45% Educ high: 76.65%	Share hourly: 35,60%	Avg. Hourly Rate (overall): 17.08 USD Avg. job size: 418 USD
4	Video Production	Video Production, 0.91, Animation, 0.37, Audio Production, 0.15	video-editing 0.64 video-production 0.43 adobe-after-effects 0.34 animation 0.29 video-postediting 0.19	1,725	Technology-oriented / Creative	Rather high (.89)	Freelancers from advanced economies: 37,97% Education high: 76.65%	Share hourly: 28,30%	Avg. Hourly Rate (overall): 16.35 USD Avg. job size: 298 USD

5	Writing	Article & Blog Writing, 0.80, Creative Writing, 0.33, Editing & Proofreading, 0.30	content-writing 0.44 article-writing 0.41 blog-writing 0.37 creative-writing 0.36 writing 0.24	9,023	Language-oriented	Rather high (.94)	Freelancers from advanced economies: 65.32% Education high: 81.88%	Share hourly: 24,23%	Avg. Hourly Rate (overall): 19.39 USD Avg. job size: 279 USD
6	Translation	General Translation, 0.99, Article & Blog Writing, 0.05, Other - Writing, 0.04	Translation 0.64 translation-english-spanish 0.25 translation-english-french 0.22 translation-english-german 0.21 French 0.21	4,925	Language-oriented	Rather high (.93)	Freelancers from advanced economies: 57.56% Education high: 76.16%	Share hourly: 22,61%	Avg. Hourly Rate (overall): 13.68 USD Avg. job size: 234 USD
7	General Business	Data Entry, 0.52, Web Research, 0.41, SEO - Search Engine Optimization, 0.34	data-entry 0.46 microsoft-excel 0.36 internet-research 0.33 lead-generation 0.21 virtual-assistant 0.18	18,458	Business-oriented	High (.98)	Freelancers from advanced economies: 20.32% Education high: 81.60%	Share hourly: 54,53%	Avg. Hourly Rate (overall): 10.47 USD Avg. job size: 561 USD
8	Ecommerce	Ecommerce Development, 0.99, Web Development, 0.12, Web & Mobile Design, 0.08	Magento 0.70 shopify 0.37 php 0.29 css 0.20 website-development 0.20	739	Business-/ Technology-oriented	High (.999)	Freelancers from advanced economies: 14.75% Education high: 83.33	Share hourly: 46,84%	Avg. Hourly Rate (overall): 17.98 USD Avg. job size: 559 USD
9	Web Frontend	Web Development, 0.68, Web & Mobile Design, 0.66, Ecommerce Development, 0.25	Wordpress 0.45 Php 0.39 Css 0.33 Html 0.32 web-design 0.27	13,162	Technology-oriented / Creative	Rather high (.93)	Freelancers from advanced economies: 14.81%	Share hourly: 46,67%	Avg. Hourly Rate (overall): 15.54 USD Avg. job size: 565 USD
10	Web Backend	Web Development, 0.45, Network & System Administration, 0.44, Mobile Development, 0.42	Php 0.33 javascript.33 c# 0.26 mysql 0.21 python 0.20	2,724	Technology-oriented	Medium (.84)	Freelancers from advanced economies: 18.39% Education high: 82.85	Share hourly: 49,54%	Avg. Hourly Rate (overall): 22.36 USD Avg. job size: 1297 USD
11	Mobile Dev	Mobile Development, 0.98, Game Development, 0.13, QA & Testing, 0.09	android-app-development 0.49 iphone-app-development 0.45 ios-development 0.44	2,206	Technology-oriented	Rather high (.94)	Freelancers from advanced economies: 17.09% Education high: 86.16%	Share hourly: 41,20%	Avg. Hourly Rate (overall): 19.52 USD

			mobile-app-development 0.27 objective-c 0.27						Avg. job size: 1213 USD
12	Illustrators	Illustration, 0.98, Graphic Design, 0.14, Animation, 0.08	Illustration 0.69 Cartooning 0.33 adobe-illustrator 0.29 drawing 0.27 character-design 0.22	1,084	Creative	Rather high (.94)	Freelancers from advanced economies: 35.70% Education high: 69.04%	Share hourly: 21,51%	Avg. Hourly Rate (overall): 15.84 USD Avg. job size: 252 USD
13	Logo Design	Logo Design & Branding, 0.99, Graphic Design, 0.15, Illustration, 0.03	logo-design 0.77 graphic-design 0.40 adobe-illustrator 0.28 adobe-photoshop 0.25 illustration 0.15	1,408	Creative	Rather high (.92)	Freelancers from advanced economies: 30.35% Education high: 71.28%	Share hourly: 17,73%	Avg. Hourly Rate (overall): 15.74 USD Avg. job size: 104 USD
14	Graphic Design	Graphic Design, 0.93, Logo Design & Branding, 0.28, Photography, 0.1	adobe-photoshop 0.49 graphic-design 0.47 adobe-illustrator 0.38 photo-editing 0.26 print-design 0.25	5,343	Creative	Medium (0.78)	Freelancers from advanced economies: 29.98% Education high: 71.19	Share hourly: 30,09%	Avg. Hourly Rate (overall): 14.67 USD Avg. job size: 204 USD
-1	Unclustered	Graphic Design, 0.54, Logo Design & Branding, 0.40, Presentations, 0.40,	graphic-design 0.37 adobe-photoshop 0.33 adobe-illustrator 0.25 php 0.24 logo-design 0.22	9,334		0	Freelancers from advanced economies: 25.85% Education high: 77.90%	Share hourly: 0.41%	Avg. Hourly Rate (overall): 17.46 USD Avg. job size: 563 USD

A. 1.B: Frequency Table Task Categories

No.	Sub task category	Frequency
1	3D Modeling & CAD	39,601
2	A/B Testing	76
3	Academic Writing & Research	14,329
4	Accounting	19,456
5	Animation	23,641
6	Architecture	2,326
7	Article & Blog Writing	170,696
8	Audio Production	10,612
9	Chemical Engineering	88
10	Civil & Structural Engineering	1,240
11	Contract Law	3,954
12	Contract Manufacturing	24
13	Copywriting	18,420
14	Corporate Law	219
15	Creative Writing	37,205
16	Criminal Law	23
17	Customer Service	7,835
18	Data Entry	133,128
19	Data Extraction / ETL	1,866
20	Data Mining & Management	2,206
21	Data Visualization	2,748
22	Database Administration	3,463
23	Desktop Software Development	14,769
24	Display Advertising	3,497
25	ERP / CRM Software	1,136
26	Ecommerce Development	53,100
27	Editing & Proofreading	30,158
28	Electrical Engineering	2,757
29	Email & Marketing Automation	9,530
30	Family Law	43
31	Financial Planning	4,040
32	Game Development	4,927
33	General Translation	135,726
34	Grant Writing	826
35	Graphic Design	341,809
36	Human Resources	2,348
37	Illustration	41,865
38	Information Security	630
39	Intellectual Property Law	1,055
40	Interior Design	1,176
41	Lead Generation	25,805
42	Legal Translation	1,568
43	Logo Design & Branding	116,469
44	Machine Learning	255
45	Management Consulting	1,464
46	Market & Customer Research	7,600
47	Marketing Strategy	4,255
48	Mechanical Engineering	1,108

49	Medical Translation	437
50	Mobile Development	45,154
51	Network & System Administration	35,943
52	Other - Accounting & Consulting	4,055
53	Other - Admin Support	17,051
54	Other - Customer Service	1,432
55	Other - Data Science & Analytics	2,440
56	Other - Design & Creative	16,047
57	Other - Engineering & Architecture	631
58	Other - IT & Networking	3,355
59	Other - Legal	5,289
60	Other - Sales & Marketing	5,059
61	Other - Software Development	26,227
62	Other - Web & Mobile Development	1
63	Other - Writing	55,654
64	Paralegal Services	582
65	Personal / Virtual Assistant	47,311
66	Photography	8,834
67	Presentations	18,225
68	Product Design	3,166
69	Product Management	1,863
70	Project Management	2,956
71	Public Relations	1,469
72	QA & Testing	11,509
73	Quantitative Analysis	6,218
74	Resumes & Cover Letters	5,269
75	SEM - Search Engine Marketing	11,583
76	SEO - Search Engine Optimization	44,272
77	SMM - Social Media Marketing	16,737
78	Scripts & Utilities	25,871
79	Technical Support	1,157
80	Technical Translation	1,986
81	Technical Writing	8,029
82	Telemarketing & Telesales	14,791
83	Transcription	33,724
84	Video Production	42,486
85	Voice Talent	23,920
86	Web & Mobile Design	137,395
87	Web Content	37,132
88	Web Development	542,626
89	Web Research	78,642
	Total	**2,662,983**

Download ZEW Discussion Papers from our ftp server:

http://ftp.zew.de/pub/zew-docs/dp/

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de