WHAT SKILLS PAY MORE? THE CHANGING DEMAND AND RETURN TO SKILLS FOR PROFESSIONAL WORKERS

Cecily Josten, Corresponding author. London School of Economics, Houghton Street, WC2A 2AE, United Kingdom. Phone: +49-1624332004. Email: c.c.josten@lse.ac.uk

Helen Krause, Citi, Citigroup Centre, 33 Canada Square, London E14 5LB, United Kingdom. Email: helen.krause@citi.com

Grace Lordan, London School of Economics, Houghton Street, WC2A 2AE, United Kingdom. Phone: +49-1624332004. Email: g.lordan@lse.ac.uk

Brian Yeung, Citi, Citigroup Centre, 33 Canada Square, London E14 5LB, United Kingdom. Email: brian.yeung@citi.com

Abstract

Technology is disrupting labor markets. We analyze the demand and reward for skills at occupation and state level across two time periods using job postings. First, we use principal components analysis to derive nine skills groups: 'collaborative leader', 'interpersonal & organized', 'big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical'. Second, we comment on changes in the price and demand for skills over time. Third, we analyze non-linear returns to all skills groups and their interactions. We find that 'collaborative leader' skills become significant over time and that legacy data skills are replaced over time by innovative ones.

1. Introduction

The skills demanded by the labor market are currently being shaped by the Fourth Industrial Revolution and the pace of this change has been accelerated by the Covid-19 pandemic (Frey and Osborne 2017; Dingel and Neiman 2020; Campello, Kankanhalli, and Muthukrishnan 2021). In the past, technological advances have led to a hollowing out of the middle of the income distribution as jobs that require routinised tasks are codified and jobs that require more complex non-routine tasks gain in efficiency from new technologies coming on stream (Frey and Osborne 2017; Josten and Lordan 2022). We are currently experiencing the Fourth Industrial Revolution that started around 2015 (Schwab 2015), which is bringing with it artificial intelligence, robotics, quantum computing, genetic engineering and the Internet of Things, all of which are disrupting the nature of work. Overall, these labor markets developments are changing the tasks at the occupation level, and the corresponding skills required to perform specific occupations (Josten and Lordan 2021). In particular, there is evidence that employers are increasingly demanding and rewarding social skills (e.g. leadership and communication (Josten and Lordan 2021)), while continuing to reward cognitive skills. Examples of cognitive skills include decision-making (Deming 2021), critical thinking (Deming and Kahn 2018) in addition to emerging cognitive skills such artificial intelligence skills (Alekseeva et al. 2021; Deming 2021).

Overall, the demand for specific skills from humans are changing, as firms adopt the technologies available that complement and substitute for tasks previously done by their workforce. In the face of the Fourth Industrial Revolution and a rapidly changing market for skills, this study analyzes how the price of skills (measured at the occupation and state level) changes across two time periods, namely 2014-2015 and 2018-2020 Q1 using job flow data of professionals in the United States. We choose these times periods as they frame the outbreak of the Fourth Industrial Revolution 2015 very well. This is our first time period 2014-2015 marks the arrival of the technology later defining the Fourth Industrial Revolution and our second time period 2018-2020 summarizes its progression.

We obtain data on the demand for skills from a large platform of online job advertisements. We link each job advert to wage data based on the state and occupation a job was posted for. While job advert data in and of itself is a proxy for the demand for skills in the labor market (Carnevale, Jayasundera, and Repnikov 2014), linking it to actual wage outcomes informs on whether the demand for skills is changing the price of skills, in our case at the geographic region and occupation level.

The approach taken in this study builds on the research on the changing nature of work and the changing demand for skills. Our study is most closely related to Deming and Kahn (2018) who also analyze job advertisement data to measure the variation in skills demand for professionals between 2010 and 2015. They reduce the skills keywords mentioned in job postings from Burning Glass Technologies (BGT) down to 10 broad job skills following the task literature based on their assessment of how best to divide skills. The authors link cognitive and social skills[1] to wages and firm performance and find a positive correlation between both social and cognitive skills. They also find a strong complementarity between social and cognitive skills with the interaction of both skills positively and significantly correlating with wage and firm outcomes. Overall, they find that the demand for cognitive and social skills accounts for around 5% of the variation in wages and firm performance when controlling for occupation, industry, education and experience requirements and eight other skills requirements. They highlight that more research is needed on alternative skills such as interpersonal skills. Our work, extends and goes beyond Deming and Kahn (2018) in the following three ways:

1. We consider a more detailed list of skills groups that is statistically determined based on the skill requirements in job advertisement data rather than chosen by the authors. We take an inductive approach for the selection of skills groups. That is we derive skills groups using a principal components analysis (PCA). PCA allows us to group keywords that appear together in the skills requirement section of an advert in a meaningful way. We follow the academic literature in the choice of the keywords (Deming and Kahn 2018) and the professional literature as defined in a report by the management consulting company McKinsey (Dondi et al. 2021), in addition to the authors' expertise. We refine our keyword choices based on an analysis of co-occurrences. The PCA reduces to nine latent factors of which the following intuitive labels emerge: 'collaborative leader', 'interpersonal & organized', 'big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical'.

[1] Social skills are a subset of non-cognitive skills and defined as skills centered around human interaction in particular, including collaboration and communication skills, amongst others (Josten and Lordan, 2021).

2. We focus on two time frames for our analysis of 2014-2015 and 2018-2020 Q1[2] and can thereby comment on changes in the returns to skills over time. These two time periods are particularly interesting as they capture the start (2014-2015) of the Fourth Industrial Revolution, with the second period allowing a sufficient lag for the new technologies to have diffused and influenced the labor market. Studying the returns to skills over time at the occupation/state level for these two time periods is interesting as it informs on how labor market developments, such as technological innovation, are changing the value of skills both at the occupation level, but also across geography.
3. We analyze returns to our nine skills groups, allowing for intuitive complementarities across the nine skill groupings. For example, we expect that certain cognitive skills will be more valuable if a person has high levels on leadership skills also. This aligns with Weinberger (2014) who finds an increasing complementarity of social and cognitive skills, and Deming (2021) who finds that decision-making and cognitive ability are complementary, and their rewards are increasing over time. Our analysis extends these analyzes by looking at the interactions of a set of nine skills groups and contributes to the existing literature by shedding light on which specific skills and combinations thereof are rewarded in the labor market.

Drawing on job flow data we relate our nine skill groups ‘collaborative leader’, ‘interpersonal & organized’, ‘big data’, ‘cloud computing’, ‘programming’, ‘machine learning’, ‘research’, ‘math’ and ‘analytical’ to the logarithm of hourly wage in a linear regression including a set of demographic and industry controls, in addition to state and occupation fixed effects. Overall, we find changing prices of cognitive and non-cognitive skills over our two time periods that align with shifts in the labor market. The most interesting, stylized facts are as follows:

First, we find that the two non-cognitive skills groups ‘collaborative leader’ and ‘interpersonal & organized’ are differently rewarded. For the ‘collaborative leader’ skills group, we find that a 10 percentage point increase in this skill garners an increase in wages of 0.3% in 2018-2020 Q1 (this effect is not statistically significant in 2014-2015). This estimate implies a 0.15$

[2] The two specific time frames 2014-2015 and 2018-2020 Q1 frame the start of the Fourth Industrial Revolution. They are further chosen for three additional reasons: First, LinkUp only becomes sufficiently large in 2014 and we restrict the data to before the outbreak of the Covid-19 pandemic that changed labor market demand substantially from the end of March 2020 onwards. Second, we use two time frames that are two years apart due to the rolling averages of wages nature of the OEWS. Third, we pool the years within each time frame (e.g., 2014 and 2015) to account for changing trends in skills requirements we cannot control for such as labor market shocks (Deming and Kahn, 2018).

increase in hourly wages for which the mean is 49.49$ per hour in 2018-2020 Q1. For 'interpersonal & organized', a 10 percentage increase predicts a reduction in wages of -0.36% in 2014-2015 and of -0.73% in 2018-2020 Q1. This corresponds to a reduction of the mean hourly wage of -0.16$ in 2014-2015 (with the mean hourly wage being $44.79 in 2014-2015) and of -$0.36 in 2018-2020 Q1. So, to summarize for non-cognitive skills, we show that the skills group 'collaborative leader' exhibits positive and increasing returns while that of 'interpersonal & organized' exhibits negative returns. Both skills groups are increasing over time in terms of demand (i.e., exhibit increasing shares). This differentiation is in line with the literature (Deming and Kahn 2018; Calanca et al. 2019). Edin et al. (2022) analyze non-cognitive skills as defined by a psychologist-assessed measure of teamwork and leadership that relates closely to our 'collaborative leader' skills group. They find a strong increase in the return to non-cognitive skills among men in the private sector using Swedish military enlistment data from 1992 to 2013 combined with administrative wage data. The return to a one standard deviation increase in non-cognitive skills increased from 7 to 14 percent with this effect being even larger at the top end of the wage distribution. Similarly, Deming (2017) analyzes the returns to social skills in particular and finds that they are increasingly valued in the labor market in terms of wages and employment when analyzing US surveys from 1979 versus from 1997. Social skills refer to the ability to work with others and in particular skills related to coordination, negotiation or persuasion and are hence again most closely related to our 'collaborative leader' group that also captures overlapping keywords. Their finding is complemented by a later paper by Deming (2021) that focuses on decision-making skills only. The author uses online job advertisement data from Burning Glass Technologies (BGT) alongside newspaper advertisement data from Atalay et al. (2020) and finds that decision-making skills have increased in importance and gain larger wage premia. This finding points at the importance of skills that help in dealing with increasing complexity and open-endedness of job tasks. This again is similar to the 'collaborative leader' skills group that entails keywords such as negotiation or strategic. Our finding can also be explained with the automatability of occupations requiring particular non-cognitive skills, i.e. time management skills that are part of the 'interpersonal & organized' have been shown to be automatable (Josten and Lordan 2022), as compared to 'collaborative leader' skills, which have been shown to be, given current technology, automation proof (Atalay et al. 2020; Deming 2021).

Second, we find that the reward to data science skills is constantly evolving with the newest data science skills being rewarded and legacy data science skills being punished. This is a symptom of technology being currently in evolution, and with it demanding an evolving skill set. Concretely a 10 percentage point in the share of 'big data' is associated with an increase in wages of 1.85% in 2014-2015 (i.e. an extra 0.83$ above the mean hourly wage of 44.79$), turning negative in 2018-2020 Q1 with a 10 percentage point increase in the respective skills group leading to -1.21% lower wages in 'big data' corresponding to -0.6$. A similar trend of positive return in 2014-2015 turning negative in 2018-2020 Q1 is found for 'cloud computing' with a 10 percentage point increase in 'cloud computing' shares increasing wages by 0.57% (i.e. an increase of 1.47$ above mean hourly wages) in 2014-2015 and decreasing wages by -0.37%. (i.e., a decrease of -0.18$ below mean hourly wages) in 2018-2020 Q1. In contrast, by 2018-2020 Q1 the skills group that has increased the most relatively in share demanded and wage premium is 'machine learning'. This is a skill grouping that did not appear in job adverts in 2014-2015 and emerged in between these periods. In 2018-2020 Q1, 'machine learning' gained a wage premium with a 10 percentage point increase in this skills group increasing wages by 5.83%. At professionals' mean hourly wages of 49.49$ per hour in 2018-2020 Q1, this corresponds to an increase of 2.89$.

This shift across data science skills reflects a market in data science that is constantly evolving, with those that upskill in line with the sector trends being in shortage in the labor market, and as a result enjoying high wage premiums. The finding of 'machine learning' only appearing in the job advertisement data in the later time frame and exhibiting highly positive returns is also a reflection of the adoption of new technologies by companies and is reflected in the literature on AI, which corresponds closely to our 'machine learning' classification. Alekseeva et al. (2021) study skills requirements in online job advertisements between 2010 and 2019 using data from BGT with a specific focus on the demand for artificial intelligence (AI) skills in the labor market. AI skills are identified with keywords that are directly related to AI such as 'artificial intelligence' or 'keras' and the share of advertisements including at least one of these keywords is linked to shares and wages. They find an increased demand in AI skills across occupations, sectors and firms and a premium to those skills of 11% for job postings in the same firm and of 5% within the same job title. Their finding highlights developments in AI adoption in companies and shows substantial and increasing returns to AI. Our study also looks at machine learning more specifically. Similarly, Squicciarini and Nachtigall (2021) study occupations requiring AI using online job postings in Canada, Singapore, the United Kingdom and the United States using BGT job advertisements data. They also find that an increasing number of occupations require AI skills across all four countries. They find that over time skills related to legacy computing skills such as software engineering and development decreased in importance as compared to AI-specific skills like natural language processing. Deming and Noray (2020a) look at how the returns to specific skills acquired at university change over time. In their model, they show that individuals who study applied subjects such as computer science or engineering or business (as compared to economics or biology) are required to change their skill set more often throughout their career, which leads to lower returns in the long run. Their finding again highlights that rapidly changing applied skills are rewarded initially like 'big data' in our analysis but turn into legacy skills over time that are not rewarded anymore as compared to more stable skills.

Third, 'programming' has a substantive negative wage premium across both time frames (i.e., a 10 percentage point increase in 'programming' predicts a decrease of wages of -0.95% in 2014-2015 and of -1.09% in 2018-2020 Q1). A possible explanation is that programming skills such as java or SQL are pre-requisites in top programming occupations and are only explicitly mentioned in occupations that search for medium-skill workers familiar with low-level coding.

Fourth, the premium to 'research' skills increases over time with a 10 percentage point increase in the share of research skills increasing wages by 0.44% in 2014-2015 and by 0.59% in 2018-2020 Q1. Keywords that are part of the 'research' category overlap with the broader category of cognitive skills as for example 'research' or 'statistics' defined by Deming and Kahn 2018. Given their finding of a positive correlation of cognitive skills on wages, it is hence not surprising that 'research' correlates positively with the logarithm of wages in our study.

Our second set of models consider the returns to the skills interactions we identify using a Lasso (least absolute shrinkage and selection operator) regression approach. Overall, we can confirm the complementarity between soft skills and cognitive skills:

Concretely, 'collaborative leader' interacted with 'research' has a positive wage premium across both time frames (i.e., a 10 percentage point increase in the share of the 'collaborative leader' interaction with 'research' corresponds to an increase in wages of 0.01% in 2014-2015 that increases to 1.78% in 2018-2020 Q1). The later effect is substantially larger and has a dollar effect of 0.88$ above mean hourly wages of 2018-2020 Q1. That is occupations that require both 'collaborative leader' skills and 'research' experience a wage premium. As highlighted above, past research mainly focused on the interaction of social skills and cognitive skills. Our 'collaborative leader' skills group is, however, defined at a more detailed level. Specifically, this finding is in line with the findings by Deming and Kahn (2018) and Weinberger (2014) who also find a complementary effect of social and cognitive skills. Given that our skills groups are more narrowly defined, our finding indicates that 'collaborative leader' skills are particularly valuable social skills when combined with cognitive skills of 'research'. Such skills are centered in high-skill and high-paid occupations.

Second, there is a positive complementarity of 'big data' with 'cloud computing' in 2014-2015 but that becomes negative in 2018-2020 Q1. This finding points to the changing nature of cognitive skills that are rewarded highly but whose skills requirements are also subject to greater and more frequent changes.

Our work provides insights that are useful in a number of contexts. First, it provides information to firms and individuals on the skills that are becoming increasingly valuable in the advent of shifts caused by the Fourth Industrial Revolution. For firms this is useful in terms of hiring, planning, training and upskilling their workers for the daily tasks that they do, but equally providing training in emerging skills is useful as an amenity in the employee value proposition to attract and retain talent. For individuals, this is useful in terms of making choices regarding educating and upskilling themselves. Second, it provides information to firms and individuals on the volatility of prices for specific skills over time. Third, it provides a new lens through which investors can view firms. That is, they can analyze the skills being demanded by a company they are contemplating investing in and determine if this company is seeking the skills that are most relevant in today's economy for a specific occupation as a pulse point for their innovation and as a response to the macro-economic changes of the Fourth Industrial revolution.

This paper proceeds as follows: Section 2 describes the data used in this study. Section 3 describes the methods used ranging from i. principal component analysis to ii. linear regression to iii. Lasso regression. Section 4 summarizes and discusses the results and section 5 concludes.

2. Data

2.1. Overview

We draw on LinkUp Raw (LinkUp) (https://www.linkup.com/) provided by Citi. LinkUp is a large global job listing index of job openings with 165 million job postings listed since 2007 and sourced from employer websites worldwide (LinkUp 2022). LinkUp contains job advertisements from websites of publicly traded companies to be used as input for labor market analytics.[3] The data is continuously updated through crawling of public websites. The data contains detailed information on each advertisement including the state it was posted in, its Occupational Information Network (O*NET) occupation code, the Global Industry Classification Standard (GICS) codes at the 2-digit level[4], job and company attributes and raw job descriptions and job records. LinkUp is unique in retaining the full job description that allows us to focus on the section of skills requirements in particular. LinkUp has also been previously used to study the impact of the Covid-19 pandemic on hiring (Campello, Kankanhalli, and Muthukrishnan 2021) and the demand for software testing skills (Cerioli, Leotta, and Ricca 2020). The data has also been validated and shown to be representative (Campello, Kankanhalli, and Muthukrishnan 2021).[5]

Job advertisement data more generally is a useful data source to study labor market dynamics for its large sample size and the breadth of information it contains, including the detailed description of a job's requirements (Faberman and Kudlyak 2016). It further is a valuable addition to using, for example, self-reported survey-based labor market data (Carnevale, Jayasundera, and Repnikov 2014). This is because as compared to survey data that provides a snapshot view of the labor market at the point of collection and is costly to administer, job advert data represents readily available job flow data that reflects actual employment dynamics at the point at which they occur (Faberman and Kudlyak 2016). It hence also serves as a means to make predictions into the future as it shows which skills are in demand and what the employee of the future is like (Carnevale, Jayasundera, and Repnikov 2014). Job advertisement data has also been used frequently in past research using it to analyze the development of skills

[3] LinkUp scrapes 100% of publicly traded company websites but state that 15% of those companies do not post jobs on their website currently (LinkUp 2021).
[4] GICS is an industry classification developed by MSCI and S&P Dow Jones Indices that contains 11 sectors (i.e., 2-digit classification).
[5] Campbello et al. (2021) show that job postings in LinkUp predict firm job gains in the U.S. Census Bureau's Quarterly Workforce Indicators (QWI) data and in the BLS's Job Openings and Labor Turnover Survey (JOLTS) in subsequent time periods. This is true for small and large firms and high and low skill job postings.

requirements in occupations. For example, Modestino, Shoag and Balance (2020) use online job advertisement data from BGT to analyze skills requirements after the Great Recession and find that education and experience requirements increased; an effect that can be attributed to the increased supply of workers following layoffs during and after the recession. Blair and Deming (2020) also study BGT job advertisements after the Great Recession and find that skills demand has increased substantially following the recession.

We use LinkUp as data source that focuses on company websites only, as compared to the frequently studied Burning Glass Technologies data set (Deming and Kahn 2018; Hershbein and Kahn 2018; Forsythe et al. 2020; Samek, Squicciarini, and Cammeraat 2021) that also sources from job boards. Company websites are updated frequently, there is no risk of duplicate postings across different job boards (as is the case for Burning Glass data) (Campello, Kankanhalli, and Muthukrishnan 2021). We choose to restrict the LinkUp data to occupations of professionals that account for 52.5% of all job advertisements in the period studied.[6] Research on the use of job advert data has highlighted that there is a bias of jobs posted online towards white-collar industries and occupations that seek highly skilled individuals (Carnevale, Jayasundera, and Repnikov 2014). Job advertisements from company websites are hence inappropriate to study non-professional blue-collar labor markets (Deming and Kahn 2018). In addition, professional jobs are suited to the analysis of job advertisements as they have the largest variability with respect to skills requirements (Deming and Kahn 2018). We further restrict the LinkUp data to job postings in the US by companies that are listed in the MSCI World Index. The MSCI World Index is a stock market index that includes large and mid-cap companies that operate globally (MSCI 2022).[7] Job advert data has been shown to be less volatile and more consistent when focusing on a fixed set of job advertisement platforms (Carnevale, Jayasundera, and Repnikov 2014). To ensure stability we hence focus on the

[6] Professional occupations are restricted to the major SOC categories 11-29.

[7] The MSCI world index includes companies from 23 countries (i.e. Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the UK and the US) (https://www.msci.com/documents/10199/178e6643-6ae6-47b9-82be-e1fc565ededb). We are filtering for companies based on the MSCI World Index 2021. There are 1586 companies in total in the index; of which 853 are successfully mapped to/covered in the LinkUp job postings data feed. The MSCI World Index is updated annually (both the underlying constituents and the total number of companies in the index), however, for ease of standardization in our analysis we are keeping a constant panel of companies. The average annual panel change of the MSCI World Index is about 5-6% between 2014-2021. As compared to comparable major stock market indices, the MSCI World Index covers 94% of the companies of the NASDAQ 100, 94% of the S&P 500, 89% of the FTSE 100 and 70% of the STI 30.

websites of sufficiently large global companies listed in the MSCI world index even if this is traded off with coverage of a wider range of companies.

We restrict our analysis to data to 2014 to 2020 Q1. This time frame is chosen for three reasons. First, it marks the onset of the Fourth Industrial Revolution with 2014-2015 being the period shortly before the term Fourth Industrial Revolution is first coined in December 2015 and the second half until 2020 marking the rapid technological advances following its onset (Schwab 2015). Second, given that more companies posted job advertisements online over time, 2014 is the period when we assess we have comparable coverage with later years as per Table A1 in the appendix. Also, by 2014, between 60%-70% of all job openings are posted online (Carnevale, Jayasundera, and Repnikov 2014). Third, we also restrict our analysis to before the outbreak of the Covid pandemic as we see a large drop in advertisements from April 2020 onwards. In the US and also worldwide, labor markets contracted towards the end of March 2020 and saw a sharp decline in advertised jobs as well as a change in the distribution of skills demanded (OECD 2021).

2.2. Constructing our skill groups

Keywords are selected based on three different criteria: First, we take an inductive approach and filter for keywords in the context of skills requirements and focus on those that appear frequently in adverts. Second, we include a list of keywords related to skills used in Table 1 in Deming and Kahn (2018), who base their keyword selection for social and cognitive skills on the literature on nonroutine tasks. Third, we additionally include keywords as defined in a recent study by the McKinsey Global Institute that consider cognitive, interpersonal, self-leadership and digital distinct elements of talent (DELTAs) (Dondi et al. 2021).[8] With this keyword selection process, we try to be as inclusive of potentially relevant skills as possible by focusing on both the demand for skills as revealed by the data, in addition to the academic and professional literature. We narrow down keywords where they are too broad or ambiguous.[9] Building on the initial list of skills identified, we then expand the list to include relevant, associated skills that are similar in nature by identifying keywords that most

[8] McKinsey combines academic literature and their experience in adult training to define 56 skills and attitudes (e.g., adaptability or coping with uncertainty) that they then link to adult outcomes in a survey with 18.000 individuals. They find that individuals who score high on those skills have on average higher incomes, higher job satisfaction and are more likely to be employed.
[9] For example, we remove the word 'management' as it appears in about 50% of the job adverts and is an ambiguous term that can be used to describe a skill but also a person or company attributes.

frequently co-occur with the words in the list. The fundamental statistical method used for this exercise is Pointwise Mutual Information (PMI), a measure of association between two words.[10] This selection process yields 236 underlying keywords. We then further cluster keywords that are synonyms into skills categories.[11] The final set of keywords in skills categories includes 166 keywords. A list of all keywords used for the PCA, the synonym grouping, and their source can be found in table B1 in the appendix.

We extract the keywords from our job advert data using Bidirectional Encoder Representations from Transformer (BERT). This is a machine learning method for natural language processing (Devlin et al. 2019). Using BERT, we analyze 1.3 million job advertisements between 2014 and up to 2020 Q1. And classify job description sentences into five categories: responsibilities, skills, education, (legal) requirement, others. This model has a 80-90% accuracy in correctly classifying sentences into each of these categories. However, it takes a long time (i.e., a couple of weeks) to run the classification predictions. We hence derive a random stratified sample of 25% of the entire available data set of job advertisements that is stratified keeping the same distribution of jobs per company, state, O*NET occupation code and year combination. The keyword search is restricted to the section of the job advert where candidate skills requirements are listed and derived through the BERT natural language processing technique that is explained in more detail in appendix C. A keyword is a dummy variable that equals one if the keyword appears at least once in the skills requirement section of a given job advertisement. We further provide three exemplary job advertisements in section 4 of Appendix C.

2.3. Principal component analysis: Skills groups

Data that denotes the occurrences of the 166 keyword categories within our 1.3 million job advert dataset are inputs for a principal component analysis (PCA). While we could run wage regressions wages including all 166 skills keywords, this would likely lead to an overfitting of the regression and would certainly impede a straightforward interpretation of the estimates (Abdi and Williams 2010; Lordan and Pischke 2022). Further, by clustering skills we follow

[10] Pointwise Mutual Information (PMI) is defined as the ratio of the joint distribution (coincidence) relative to thee individual distributions (independence) of two words. For each word in the initial skills list, PMI is calculated against every English word that has appeared at least once across all job posting descriptions. The top 50 words with the highest PMI scores for each seed skill are manually reviewed and added to the list.

[11] Synonyms are grouped together as they likely appear in different advertisements despite describing the same skills group. The PCA would falsely classify such synonyms as being in different skills groups. An example for synonyms is strategy and strategist, which are grouped together.

the literature on tasks and skills that focuses on tasks/skills groups rather than a battery of individual tasks/skills items (Weinberger 2014; Deming and Kahn 2018; Atalay et al. 2020). Overall, our approach allows us to comment on which keywords cluster together in the underlying data and should be combined into broader skills groups based on principal components.

For the PCA we draw on the entire 1.3 million of LinkUp job advertisements for professionals for the years 2014 to 2020 Q1. We broadly follow the approach recommended by Heckman et al. (2012) and succeed in reducing the 166 variables to 9 skill groupings (see appendix D for more details). Specifically, we remove items that load on more than one component (cross-loadings) and items that have a loading of smaller than 0.32 (weak loadings).[12] The final components have no items that are weakly loading nor cross loading and they correlate freely. We use orthogonal rotations, that allow the components to be correlated, to find the optimal number of principal components subject to the following rules for the cut-off for the components: a cumulative variance explained of the components of at least 60%, examining a jump in the scree plot (i.e. a point at which the eigenvalue of a given component falls substantially) and choosing component cut-offs that are sensible and intuitive (Bartholomew et al. 2011). Each step of the PCA is explained in more detail in appendix D.

The overall PCA analysis results in nine latent factors being extracted. These represent skill groupings that we intuitively[13], based on the variables that loaded on each factor, labelled as follows: ‘collaborative leader‘, ‘interpersonal & organized‘, ‘big data‘, ‘cloud computing‘, ‘programming‘, ‘machine learning‘, ‘research‘, ‘math‘, ‘analytical‘.

Table 1 documents the nine skills components together with their underlying keywords and the corresponding loadings. We choose labels for the nine skills groups that best summarize the underlying keywords. In the case of ‘collaborative leader’, the group label captures keywords related to leadership (i.e., ‘strategic’, ‘leadership’, ‘influence’, ‘negotiation’) and those related to the collaborate nature of leadership (i.e. ‘collaborate’ ‘creativity’, ‘coaching’). In the case of

[12] The cut-off of 0.32 has been recommended in the literature (Tabachnick and Fidell, 2018) and our large sample size allows us choosing a relatively low loadings cut-off.
[13] A good example of an intuitive grouping is the skills group ‘machine learning’ where few keywords (i.e. ‘tensorflow‘, ‘pytorch‘ and ‘keras‘) that are clearly all machine learning programmes load very highly onto the component. The same is true of, for example, ‘big data’ and ‘cloud computing’.

the cognitive skills groups, the labels are even more descriptive where, for example, 'big data' only captures big data applications such as 'hadoop' or 'hive'.

Table 1 also illustrates the share of each keyword and the share of the overall component across advertisements. Overall, two of nine skills groupings describe non-cognitive skills and seven describe cognitive skills. The two non-cognitive skills groups, 'collaborative leader' and 'interpersonal & organized' follow closely what Deming and Kahn (2018) describe either as social or as character skills.[14] Similarly, the skills groups 'research' and 'analytical' resemble Deming and Kahn's (2018) cognitive skills.[15] The fact that the PCA yields slightly diverging results from, for example, Deming and Kahn (2018), who choose their groups based on the task literature and categorise the keywords manually, stems from the fact that the PCA results of this study are based on skill groupings as they appear in job advertisements. We note that non-cognitive and social skills more specifically have been found to be used very frequently in job advertisements (Calanca et al. 2019), which further explains the large shares of about 59% for 'collaborative leader' and of 30% for 'interpersonal & organized' but can also explain that there is overall a larger number of items that load moderately. In contrast, the cognitive skills groups have quite low shares as they describe more niche skills that either appear less frequently in job advertisement or are increasing over time in the case of 'machine learning'. Also, our list of related words is not exhaustive (e.g., we do not include all different programming languages as keywords or words like 'scraping', 'mining' etc.). Table E1 in the appendix further shows the shares of each of the nine skills groups for the two time frames 2014-2015 and 2018-2020 Q1. For example, 'collaborative leader' appears in 50.14% of job adverts in 2014-2015 and 61.07% in 2018-2020 Q1. In comparison, 'machine learning' does not appear in job advertisements in LinkUp in the earlier time frame and only appears in 0.19% of job advertisements in 2018-2020 Q1.

Table E2 in the appendix also shows the shares of the interactions of skills groups with each other. Overall, the share of all skills interactions have been increasing over the two time frames. Some interactions centre around zero in terms of shares (e.g. 'big data' interacted with math). The interaction of 'collaborative leader' and 'big data', for example, increased from 0.8% to

[14] Deming and Kahn (2018) classify social skills with the keywords: 'communication', 'teamwork', 'collaboration', 'negotiation', 'presentation'. And character skills are 'organized', 'detail oriented', 'multitasking', 'time management', 'meeting deadlines', 'energetic'.
[15] Deming and Kahn (2018) classify cognitive skills with the keywords: 'problem solving', 'research', 'analytical', 'critical thinking', 'math', 'statistics'.

1.6%. The interaction of ‘collaborative leader’ and ‘research’ increases by 3.8 percentage points from 9.5% to 14.2%, which points at the fact that with increasing automation, the complementarity between social skills (i.e., ‘collaborative leader’) and cognitive skills (i.e. ‘research’) increases. For example, doctors increasingly use technology such as Clinical Decision Support Software , but still need to understand statistics, which is a facet of ‘research’ skills alongside making final decisions drawing on their ‘collaborative leader’ skills.

The shares of each skills group below vary also significantly across occupations. For example, about 90% of all advertisements for marketing managers require ‘collaborative leader’ skills but only 5% of all advertisements for pharmacy technicians require the same skill. Financial examiners are among the top five highest shares in the skills group ‘interpersonal & organized’ and ‘analytical’ with 53% and 73% off all ads requiring these skills respectively. Logically, software developers for applications are required to have cognitive skills and are among the occupations with the largest shares in ‘big data’ (14%), ‘cloud computing’ (22%), ‘programming’ (63%) and ‘machine learning’ (0.4%). The skills group ‘research’ captures occupations such as statisticians or research scientists and the skills group ‘math’ is focused on occupations such as civil engineers or actuaries. A list of the top five and bottom five occupations according to their shares for each of the nine skills groups is shown in table E3 in the appendix.

Table 1: Cognitive skills and non-cognitive components resulting from principal components analysis

Non-cognitive skills components			Cognitive skills components					
Collaborative leader			Big data			Programming		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
strategic	0.59	24.14%	hadoop	0.75	1.19%	xml	0.64	1.26%
leadership	0.58	26.17%	spark	0.75	0.72%	json	0.6	0.67%
influence	0.51	12.75%	hive	0.73	0.55%	javascript	0.59	2.51%
collaborate	0.39	24.52%	hdfs	0.53	0.14%	java	0.56	6.08%
creativity	0.34	13.57%	scala	0.47	0.47%	sql	0.39	7.12%
negotiation	0.33	6.57%	nosql	0.34	0.81%	git	0.38	1.16%
coaching	0.32	5.14%				api	0.37	1.51%
Overall		58.89%	Overall		2.28%	Overall		13.45%

Interpersonal & organized			Cloud computing			Machine Learning		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
time management	0.4	4.65%	docker	0.74	0.64%	tensorflow	0.84	0.11%
competing priorities	0.39	12.30%	kubernetes	0.71	0.41%	pytorch	0.76	0.04%
interpersonal	0.38	17.20%	amazon web services	0.48	2.32%	keras	0.73	0.03%
organized	0.36	3.38%	terraform	0.45	0.15%			
			azure	0.41	1.04%			
			jenkins	0.41	0.95%			
			openshift	0.35	0.06%			
			containerization	0.35	0.12%			
			openstack	0.32	0.22%			
Overall		29.86%	Overall		3.89%	Overall		0.11%

Research			Analytical			Math		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
quantitative	0.58	3.45%	accounting	0.65	5.58%	calculus	0.73	0.05%
statistics	0.54	5.48%	finance	0.63	7.63%	algebra	0.63	0.12%
qualitative	0.43	1.06%	common software e.g. excel	0.41	16.66%	trigonometry	0.56	0.05%
research	0.37	12.86%	analytical	0.33	20.78%	stochastic	0.47	0.05%
Overall		18.58%	Overall		35.70%	Overall		0.21%

Notes Loadings show the loading of each skill grouping's keywords that are larger than or equal to 0.32 and form the respective principal component. Share shows the share of each keyword and component across advertisements.

2.4. Matching skills group data to wage data

Job advertisements in LinkUp do not state the wages paid to a given advertised role. Our wage data is therefore at the six-digit occupation and state level. Concretely, we match the LinkUp data with wage data from the Occupational Employment and Wage Statistics (OEWS) from the US Bureau of Labor Statistics (BLS) based on six-digit Standard Occupational Classification (SOC) codes and US states. The OEWS wage data is well-suited as it is provided at the state and detailed occupation level annually (U.S. Department of Labor 2022), which allows us to match it to the LinkUp data on state and detailed occupation code. We follow Deming and Kahn (2018) who also match wage data based on six-digit occupation code and geography. Their level of geography is, however, more detailed at the Metropolitan State Area that is not available in the LinkUp data set. Annual estimates of wages are adjusted to inflation using BLS consumer price index data. The OEWS wage estimates are based on rolling averages collected over three years, which it is why it is recommended to do comparisons three years apart. We hence pool the data for the regression analysis so that we have cross-sectional data but take two timeframes that are three years apart from each other with the first being 2014 to 2015 and the second being 2018 to March 2020 when considering changes to the returns to skills over our time window. Pooling the data within each time frame helps smooth changing trends in skills requirements over time that we cannot control for such labor market shocks (Deming and Kahn 2018).[16]

We follow Deming and Kahn (2018) exactly with respect to the control variables in our regressions. We obtain control variables from the American Community Survey (ACS) at the State level. Specifically, we control for state-level share of female, Black, Hispanic, Asian, married, moved in the last year, education (high school dropouts, exactly high school, some college, exactly BA) and age (less than 18, 19–29, 30–39, 40–49, 50–64) distributions. We further obtain data on the education and experience requirements by six-digit occupation from O*NET through the variable 'job zone' that captures how much preparation (i.e. education and experience) is needed for a given occupation.[17]

[16] We use a crosswalk provided by the BLS to standardise the coding structure to SOC2010 as there was a change in the occupational coding structure in 2019 and 2020 in the OEWS.

[17] O*Net's job zone variable captures how much education people need to do the work, how much experience people need to do the work and how much on-the-job training people need to do the work in a respective occupation. The variable is coded from 1-5 with 1 describing occupations that need little or no preparation and 5 describing occupations that need extensive preparation. See more information here: https://www.onetonline.org/help/online/zones

3. Methodology

3.1. Wages and skills groups

Our aim is to relate the share of skills demanded for nine skills groupings at an occupation/state level to wage data at the occupation/state level for two time periods 2014-2015 (shortly before start of the Fourth Industrial Revolution) and 2018-2020 Q1 (initial period of the Fourth Industrial Revolution). Our nine skill groupings are: 'collaborative leader', 'interpersonal & organized', 'big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical'. In a first instance, we run a regression of the logarithm of hourly wages on the nine skills groups as follows:

$$\log(wage)_{or} = \alpha + \beta skills_{ior} + Controls + \varepsilon_{or} \qquad (1)$$

where $\log(wage)_{or}$ is the inflation-adjusted logarithm of mean hourly wages in occupation o in state r. The main independent variable $skills_{ior}$ is a vector of the share of the nine skills groups where i denotes the share of skills group i in occupation o in state r.[18] All regression specifications are weighted by the number of observations in each state and occupation cell.

We are interested in quantifying whether occupations pay more for some or all of the nine skills groups. Wages are not, however, determined by skills requirements alone, which is why we run five specifications including increasingly detailed control variables. Specification (1) is only weighted by the number of observations in each state and occupation cell. Specification (2) additionally controls for O*Net job zone codes accounting for education and experience requirements in an occupation, basic state-level demographic controls from the American Community Service and the share of ads in each two-digit North American Industry Classification (NAICS) industry. It further includes SOC major occupation controls. Demographic controls at the state-level and education and experience requirements at the occupation-level help to account for factors that drive both skills requirements and wages. Occupations with higher education requirements, for example, also likely require more skills and also pay higher wages. Major occupation controls and industry shares account for occupation- and industry-specific differences in skills requirements. For example, the language

[18] The skills groups are as per Table 1 above derived from the PCA results. A skills group is equal to one if an advert contains at least one of the corresponding keywords that load onto the skills group component. The data is then matched to wage data and collapsed by state and occupation so that the skills group variable becomes a share.

used in job advertisements is likely different for the major occupation of 'Management' as compared to that of 'Computer and mathematics'. Specific skills requirements may have a different signaling effect in one industry as compared to another. Analytical thinking, for example, may be mentioned in 'Management' occupations but not in 'Computer and mathematical' occupations as it is simply assumed in the latter. In specification (3) we additionally control for minor SOC occupation fixed effects and in specification (4) for broad SOC occupation fixed effects. Specification (5) includes detailed SOC occupation fixed effects and state fixed effects. Controlling for state fixed effects in specification (5) controls for potentially higher skills requirements in states that are wealthier and have higher costs of living or pay higher wages because the workforce is more skilled overall.[19] An example is California that has the largest share of 'machine learning' skills and is a wealthy state. Controlling for increasingly detailed occupation codes from specification (2) to (5) accounts for within occupation differences. Even at the detailed occupation code, we can imagine an advert's phrasing for marketing managers to be different to those of sales managers who are all part of the major occupation management.

Specification (5) is the preferred specification of this study as we want to avoid capturing anything that has to do with state-specific or occupation-specific differences in, for example, culture. It also accounts for unobserved skills that are required together with our nine skills groups and affect wages. However, we cannot claim certainty over causality. First, the LinkUp data set does not provide reliable information on the Metropolitan State Area an advert was posted in, which is why we match the data to wages based on the broader state variable. There may be geographical differences in the use of skills words and in wage premia that we cannot account for. Second, even the most controlled environment may suffer from unobserved variable bias in the absence of external variation. That is, there may be unobserved variables that determine both the skills demand and the wage premium. An example could be that even within the same occupation in the same state there could be cultural differences in the way skills requirements are phrased and rewarded.

[19] Ideally, we would control for a more detailed level of region than state such as Deming and Kahn (2018) who control for Metropolitan State Area (MSA). However, the job advert data set LinkUp does not provide data on job adverts at the MSA-level.

We run two separate regressions for each time frame we consider. That is: 2014-2015 and 2018- 2020 Q1. We exclude the 'machine learning' skills group in 2014-2015 as there are too few job adverts for those years (i.e., less than 10).

3.2. Wages and skills groups: interactions and non-linearities

We also relate log wages to interactions between our nine skill groupings, in addition to allowing for non-linear returns. This is to account for potential complementarities across skills groups as well as non-linear effects. This amounts to the inclusion of 63 skills-related variables in addition to the most detailed controls in our regressions. To avoid overfitting and ease interpretation we estimate a Lasso (least absolute shrinkage and selection operator) model. Lasso is a shrinkage and variable selection method for linear regression models, that minimizes prediction error for a quantitative response variable. It causes some regression coefficients to shrink toward zero through an imposed constrain, and therefore selects the variables that are most relevant in predicting logarithm of wage. Specifically, we estimate the following regression:

$$\underset{\beta}{min}\left[\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p}\left|\beta_j x\right|\right] \tag{2}$$

Regression (2) shows the Lasso regression that minimizes the prediction error from running regression (1) with the most detailed set of controls as per specification (5) where all definitions are consistent with equation (1) and in addition it includes an interaction of the nine skills groups 'collaborative leader', 'interpersonal & organized', 'big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical' with each other and their second degree polynomial.

4. Results and Discussion

4.1. Wages and skills groups

Table 2 documents the results from equation (1), which relates log hourly wages at the occupation and state level to the share of nine skill groupings (i.e. 'collaborative leader', 'interpersonal & organized', 'big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical') demanded in job adverts in the time period 2014-2015 (specifications (1)-(5)) and then also 2018-2020 Q1 (specifications (6)-(10)).

Specifications (1) and (6) shows the raw unadjusted estimates for the respective time frames. As we move from specifications (2) to (5) and (7) to (10) the coefficients decrease as variation in the outcome variable of log hourly wages is picked up by our control variables. In specification (2)-(5) and (7) to (10) we include a full set of demographic controls as well as controls of job zone and industry shares and we further include occupation fixed effects at an increasing level of detail. Specifications (5) and (10) then further include detailed SOC occupation fixed effects and state fixed effects. The focus of this study is on specification (5) and (10) for 2014-2015 and 2018-2020 Q1 respectively that each include the most detailed occupation and state controls and hence account for any state- and occupation-specific differences in skills rewards. Specification (5) and (10) control for unobserved skills that correlate with the nine skills groups in the same occupation. They also control for occupation-specific differences in the use of skills keywords. Given that the skills groups are shares, we interpret the results from Table 2 as the impact of a 10 percentage point increase in skills share on the percent change in the logarithm of hourly wages.[20] To further illustrate our effect sizes we also look at the dollar value of such a 10 percentage point increase for the average hourly wage (i.e. the average wage across all occupations and states) and for exemplary occupations.

[20] To calculate the wage premium, we use the coefficient of the regression of the log of wage on the respective skills group from table 2. We multiply the coefficient with the share of the skill (i.e., 0.1 for a 10 percentage point increase) and exponentiate it. We then subtract 1 and multiply by 100. That is: wage premium = (exp(coefficient β * skills group share)-1)*100.

Table 2: Wage premium to nine skills groups in 2014-2015 and 2018-2020 Q1

	2014-2015					**2018-2020 Q1**				
	Dependent variable: Log of hourly wage in occupation state cells									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Collaborative leader	0.744**	0.032**	-0.017**	0.015**	-0.001	0.951**	0.107**	0.029**	0.042**	0.031**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Interpersonal & organized	-0.122**	-0.007*	-0.033**	-0.049**	-0.036**	0.105**	-0.064**	-0.057**	-0.074**	-0.073**
	(0.006)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Big data	1.488**	0.342**	0.314**	0.316**	0.183**	0.050**	-0.440**	-0.225**	0.126**	-0.122**
	(0.033)	(0.020)	(0.016)	(0.015)	(0.019)	(0.016)	(0.010)	(0.008)	(0.007)	(0.007)
Cloud computing	1.316**	0.633**	0.501**	0.369**	0.323**	0.858**	0.245**	0.197**	0.047**	-0.037**
	(0.030)	(0.019)	(0.016)	(0.013)	(0.018)	(0.008)	(0.005)	(0.005)	(0.004)	(0.004)
Programming	-0.099**	0.149**	0.173**	-0.161**	-0.095**	-0.194**	0.174**	0.218**	-0.249**	-0.110**
	(0.006)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
Machine learning	N/A	N/A	N/A	N/A	N/A	3.553**	1.946**	1.281**	1.081**	0.567**
						(0.051)	(0.030)	(0.026)	(0.030)	(0.025)
Research	0.024**	0.061**	0.085**	0.063**	0.044**	-0.045**	0.088**	0.095**	0.044**	0.059**
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Math	0.478**	0.309**	0.353**	0.124**	0.089**	0.907**	0.073**	0.231**	0.079**	0.250**
	(0.048)	(0.027)	(0.022)	(0.017)	(0.026)	(0.033)	(0.018)	(0.016)	(0.012)	(0.010)
Analytical	0.135**	-0.074**	-0.093**	-0.025**	-0.014**	-0.119**	-0.193**	-0.161**	-0.066**	0.002
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.033)	(0.018)	(0.016)	(0.012)	(0.010)
Weights	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
O*net job zone	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES
Demographics	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES
State FE	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES
Occupation FE	NO	MAJOR	Minor	Broad	Detailed	NO	MAJOR	Minor	Broad	Detailed
Industry shares	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES
Constant	3.383**	1.190**	2.132**	2.693**	3.666**	3.383**	1.190**	2.132**	2.693**	3.666**
Observations	142,284	142,284	142,284	142,284	142,284	142,284	142,284	142,284	142,284	142,284
R-squared	0.350	0.791	0.867	0.927	0.954	0.350	0.791	0.867	0.927	0.954

Notes: The table displays the output from running a regression of the inflation-adjusted logarithm of hourly wages on nine skills groups (the share of each skills group across states and occupations) in the years 2014-2015 and then 2018-2020 Q1. All regressions are weighted by the share of ads within each state and occupation cell. Specifications (2)-(5) further control for O*Net job zone codes accounting for education and experience requirements in an occupation, basic demographic controls from ACS and the share of ads in each two-digit North American Industry Classification (NAICS) industries. And they additionally include SOC occupation fixed effects at different levels of detail: major occupation codes (2), minor occupation codes (3), broad occupation codes (4) and detailed occupation codes (5). Specification (5) is most detailed and further includes state fixed effects. Standard errors in parentheses, ** $p<0.01$, * $p<0.05$

Non-cognitive skills results

We first look at the two non-cognitive skills ‘collaborative leader’ and ‘interpersonal and organized’. The skills group ‘collaborative leader’ is not significant and centered around zero for the year 2014-2015. For the time frame of 2018-2020 Q1 the effect of ‘collaborative leader’ skills becomes significant in the most detailed specification (10) where a 10 percentage point increase in the share of ‘collaborative leader’ leads to an increase in wages of 0.3%. This finding is similar to Squicciarini and Nachtigall (2021) who find that non-cognitive skills such as creativity, which is also a keyword of the ‘collaborative leader’ skills group, gain in importance over time. For the mean hourly wage of 49.49$ per hour of professionals in 2018-2020 Q1, a 10 percentage point increase in ‘collaborative leader’ share corresponds to a 0.15$ increase in hourly wages. If we look at occupations that require high levels of collaborative leadership such as marketing managers (i.e., above 90% of marketing manager job adverts require collaborative leadership) they enjoy a wage premium of $2.23 (see Table E3 in the appendix). Individuals are working more collaboratively than ever with collaboration becoming essential in today’s workplace. The increasing importance of collaborative leadership skills in terms of is intuitive. Facets of collaborative leadership have been previously highlighted as valuable such as creativity (Squicciarini and Nachtigall 2021), collaboration and negotiation (Deming and Kahn 2018), coaching (Dondi et al. 2021) and strategic and leadership (Josten and Lordan 2020).

The other non-cognitive skills group ‘interpersonal & organized’ has a negative correlation with hourly wages with the estimates implying that a 10 percentage point increase in the skills share demanded decreases hourly wages by 0.36%. This corresponds to a decrease of 0.16$ from the mean hourly wage of 48.95$. The ‘interpersonal & organized’ skills group becomes even more negative in 2018-2020 Q1 despite the mean share of it increasing over the two time frames from 25% to 31%. A 10 percentage point increase in the ‘interpersonal & organized’ share in 2018-2020 Q1 predicts a -0.73% decrease in hourly wages (i.e. this corresponds to a lower hourly wage of -0.36$ as compared to the mean wage for this time frame). Looking at the occupation of financial examiners with a skills share of 49% on average over the two fime frames, they would earn -$0.67 less in 2014-2015 and -$1.66 less in 2018-2020 Q1 (see Table E3 in the appendix).

Our finding of differential rewards to different non-cognitive skills (i.e., rewards to the skills group ‘collaborative leader’ and punishment to that of ‘interpersonal & organized’) is in line with Calanca et al. (2019) who also find that skills related to leadership are rewarded while skills related to the 'interpersonal & organized' skills groups are penalized. It is also in line with Deming and Kahn (2018) who find social skills to have a positive effect on wages at the occupation level while character skills that are similar to our 'interpersonal & organized' skills may be a signal of occupations that pay little and require obedience and may hence be rewarded less. Overall, there are two main aspects that help explain the divide in the non-cognitive skills premium:

1. **Occupations that require collaborative leadership skills are less likely to be automated than those requiring individuals to be interpersonal and organized** as explained by Josten and Lordan (2022). The occupations with the highest share in leadership skills requirements all belong to management occupations (e.g., marketing managers) for which many job tasks are open-ended making them less likely to be automated (Atalay et al. 2020; Deming 2021), and most likely to evolve in response to technology. We find the other non-cognitive skills group 'interpersonal & organized' to have negative wage returns (i.e., 'interpersonal & organized' skills have a negative wage premium that increases over time from -0.16$ in 2014-2015 to -0.36$ in 2018-2020 Q1). Facets of the 'interpersonal & organized' skills group like, for example, 'time management' have been shown to have a positive effect on automation (Josten and Lordan 2022). This is because such skills center around setting rules and gathering information, which are tasks that are likely to be automated as they are easily codified. The 'interpersonal & organized' skills group also appears frequently as a requirement for occupations that have been previously highlighted to be at least partly automatable such as financial managers or lawyers (Lordan 2018; Josten and Lordan 2020). The finding of differential rewards to different non-cognitive skills, i.e. 'collaborative leadership' versus 'interpersonal & organized', is in line with Calanca et al. (2019) who find that skills related to leadership such as strategic planning are rewarded while skills related to the 'interpersonal & organized' skills groups such as 'time management' are punished.
2. **Collaborative leadership fosters individual and company performance both directly and indirectly through fostering inclusion.** Collaboration is crucial for innovation. For innovation and idea creation, working collaboratively has been shown to be crucial in combination with working independently (Girotra, Terwiesch, and Ulrich 2010). It might be true that to solve a mathematical formula, it is enough to have one individual who is a math genius but to come up with innovative ideas, having multiple individuals working

well together enhances the output (Girotra, Terwiesch, and Ulrich 2010). But the positive effect of collaboration on innovation and performance depends crucially on the quality of the collaboration. Being a 'connector' is a key trait of a collaborative leader; someone who brings people together in a way that fosters success (Ibarra and Hansen 2011). Hence, a collaborative leader can determine the quality of collaboration by fostering creativity, diversity of thought, open discussions, debates, conflict, making decisions, amongst others. A collaborative leader can also create a safe space and inclusive environments where individuals feel safe to speak up about new ideas. Inclusion prevents groupthink and confirmation bias, which have been shown to hinder performance and innovation both in a team and also at the company level (Shore and Chung 2021).

Cognitive skills results

Of the set of cognitive skills, 'machine learning' has the largest positive and significant coefficient with a 10 percentage point increase in this skills share predicting an increase in wages of 5.83% in 2018-2020 Q1.[21] This wage increase of 10 percentage points corresponds to an increase of 2.89$ above mean hourly wages. To give some context, the estimates imply that 'Computer and Information Research Scientists' that have a relatively large demand for machine learning skills (i.e., 11% of adverts in this occupation require 'machine learning' skills) would gain a wage return of $3.93 above their mean hourly wage of $62.98 (see table E3 in the appendix). These findings for 'machine learning' are comparable to Alekseeva et al. (2021) who also study the demand for AI skills as defined by overlapping keywords (i.e., keras) and using job postings data. They find a strong positive effect of AI skills on wages with an AI advert increasing wages by 5% when including firm and occupation fixed effects.

In 2014-2015 the share of 'big data' demanded, was positively related to wages. Specifically, a 10 percentage point increase leads to an increase in wages of 1.8%. This wage increase of 10 percentage points corresponds to an increase of 0.88$ above mean hourly wages. Interestingly, this estimate is negative in 2018-2020 Q1 with a 10 percentage point increase in the 'big data' share leading to a decrease in wage by -1.21%. This corresponds to -0.6$ below the mean hourly wage. In the context of the occupation 'Computer and Information Research Scientists' that is among the occupations that require 'big data' skills the most with 29% of adverts over

[21] 'Machine learning' is not mentioned yet in the earlier time frame.

the two time frames require ‘big data’ skills, this corresponds to a $2.90 wage premium in 2014-2015 and a -$2.46 wage decrease in 2018-2020 Q1 (see table E3 in the appendix).

Similarly, ‘cloud computing’ has a positive coefficient in specification (5) in 2014-2015 suggesting an increase in mean hourly wages of 0.57% when the share demanded increases by 10 percentage points. However, in specification (10) in 2018-2020 Q1 this effect turns negative to -0.37%.

Decreasing returns to ‘big data’ over time and the appearance of ‘machine learning’ with a very large wage premium in the latest period is in line with the finding of Deming and Noray (2020) who find that applied computing skills are rewarded initially but turn into legacy skills more rapidly than stable skills as supply outstrips demand. In addition, trends in computing change rapidly. This is also in line with Squicciarini and Nachtigall (2021) who find that over time legacy computing skills, for example software engineering, decrease in importance as compared to newer AI skills.

‘Programming’ already has a wage penalty that amounts to a 10 percentage point increase in the share leading to -0.94% lower wages of occupations in 2014-2015 and -1.09% in 2018-20. This could be because the programming skills such as java or SQL are pre-requisites in top programming occupations and only explicitly mentioned in occupations such as web developers[22] that search for medium-skill workers familiar with low-level coding (Manyika et al. 2017).

The demand for ‘research’ skills increases as the share rises from 16% in 2014-2015 to 19% in 2018-2020 Q1. The wage premium of a 10 percentage point increase in ‘research’ share also increases from 0.44% higher wages in 2014-2015 to 0.59% in 2018-2020 Q1. The occupation ‘Statisticians’ is among those that require ‘research’ skills the most with a share of 73% of job adverts in this occupation requiring ‘research’ skills across the two time frames. The wage premium in this occupation increases from $1.45 in 2014-2015 to $1.95 in 2018-2020 Q1.

[22] The share of ‘programming’ skills is highest in the occupation Web Developer with 68% over the two time frames as per table E3 in the Appendix.

The effect of a 10 percentage point increase in 'math' remains small and positive increasing hourly wages from 0.89% to 2.53% between 2014-2015 and 2018-2020 Q1. The demand also increases slightly but remains overall very small (i.e., the share increases from 0.19% to 0.21% of all job adverts as per table E1 in the appendix). When looking at occupations that require relatively high levels of 'math' skills such as 'Chemical Technicians' with 7% in 2014-2015 and 4% in 2018-2020 Q1, the wage premium increases from \$0.14 to \$0.26 over the two time frames.

'Analytical' has a small negative coefficient in 2014-2015 with a 10 percentage point in demand implying a reduction of hourly wages by -0.13%. The coefficient is 'not significant and centered around zero in the most detailed specification (10) but negative across all other specifications. The keyword 'analytical' is part of the cognitive skills in Deming and Kahn (2018) that is positively related to wages. It is also generally highlighted as positive contributor to wages (Calanca et al. 2019; Ziegler 2021). Our definition of 'analytical' as defined by the PCA, however, also contains keywords such as 'accounting', 'finance', and common software (e.g. Excel) that may be explicitly stated in lower paid occupations and are simply assumed in higher paid occupations.

4.2. Wages and skill interactions and non-linearities

Table 5 below shows the results from running the Lasso regression equation (2) above separately for the time frames 2014-2015 and 2018-2020 Q1. The Lasso regression is run with the share of all nine skills groups, their second-degree polynomial and a double interaction of all nine skills groups. It further includes basic demographic controls from ACS, industry shares, job zone codes from O*NET and state and detailed occupation code fixed effects. In Table 5, we document only show non-zero coefficients that are selected through the Lasso shrinkage process and that are most relevant and significant in explaining the logarithm of hourly wage.

From Table 5 the 'collaborative leader' category significantly explains variation in occupational wage in the two time periods. 'Collaborative leader' interacted with 'research' has a positive wage premium in 2018-2020 Q1. As per Table E2 in the appendix, the share of this interaction also increases over time from 9.5% in 2014-2015 to 13.2% in 2018-2020 Q1. This suggests that those skilled in both research and collaborative leadership skills are becoming more valuable to employers as the Fourth Industrial Revolution progresses. This is intuitive because as automation increases and advanced technologies replace human cognitive abilities, it will ultimately still be important to understand the implications of such technological tools with the help of broad cognitive research skills and convey them to others with the help of social skills.

Specifically, a 10 percentage point increase in the share of 'collaborative leadership' together with 'research' predicts an increase in occupational wages by 0.01% in 2014-2015 increasing to an estimate of 1.78% in 2018-2020 Q1. The latter effect has a dollar value of an additional 0.88$ per hour when evaluated at mean hourly wages of 49.49$ in 2018-2020 Q1. An example from our data as per table E3 in the appendix is the occupation 'Sales Managers' for which on average 85% of all job adverts require 'collaborative leader' skills and 11% require 'research' skills that pays above average hourly wages of $63.68 in 2014-2015 (as compared to mean hourly wages of $44.79) and of $70.25 in 2018-2020 Q1 (as compared to mean hourly wages of $49.49.

The finding of an increasing wage premium and increasing share demanded of research interacted with collaborative leadership skills is in line with the findings by Deming and Kahn (2018) and Weinberger (2014) who also find a complementary effect of social and cognitive skills. Our 'collaborative leadership' skills group is, however, defined at a more detailed level. While it contains aspects of social skills such as negotiation or collaboration, it also captures creativity or strategic skills, which have also been highlighted as crucial non-linear thinking skills (Lordan and Pischke 2022). Similarly, the 'research' skills group resembles definitions as classified by the cognitive skills definition but focuses in more detail on broad cognitive skills such as the facets 'quantitative' or 'qualitative' rather than more niche cognitive skills related to, for example, data science. This finding is in line with the automation literature that highlights that automation increases the need for social skills alongside advanced cognitive skills like logical reasoning (Manyika et al. 2017). Logical reasoning is needed for our 'research' skills group as it forms part of each facet of it like 'statistics'. Professionals require soft skills but also need to understand the implications of numerical calculations (Manyika et al. 2017). So even if very advanced technologies come on stream and improve and replace human cognitive abilities, it will still be important to understand the implications of such technological tools. This can be seen in LinkedIn data also where, for example, more automatable cognitive skills like accounting have been decreasing over time while less automatable skills like management have been increasing in terms of employer demand

(Manyika et al. 2017). Our finding is also in line with Josten and Lordan (2022) who find that jobs that require ‘people’ skills together with ‘brain’ skills are less likely to be automated. Such skills are likely centered in high-skill and high-paid occupations.

Further, both ‘interpersonal & organized’ and ‘analytical’ show diminishing returns in 2014-2015 (i.e. a negative coefficient on the second-degree polynomial) but increasing returns in 2018-2020 Q1. That is despite them both having a negative wage premium in the linear regression as per tables 3 and 4 above, high shares of each of those two skills groups have a positive wage premium in 2018-2020 Q1. This finding suggests that high shares of both skills group are an indicator for higher paid occupations over time. This finding is in line with authors who highlight that non-cognitive skills are often non-linearly related to wage outcomes (Heineck and Anger 2010; Collischon 2020). Also, the share of ‘analytical’ and ‘interpersonal & organized’ skills is particularly high in business occupations such as Financial Examiners as per Table E3 in the appendix and has been increasing substantially over the two time frames. ‘Analytical’ shares increased for Financial Examiners from 63% to 74% and those of ‘interpersonal & organized’ for the same occupation group increased even more from 42% to 55%. This indiciates that higher shares are both demanded and increasingly rewarded.

The Lasso output further confirms the shift in returns to applied computing skills: the interaction of ‘big data’ with ‘cloud computing’ is positive in 2014-2015 but negative in 2018-2020 Q1. As explained above, the skills required for cloud computing and big data analysis are constantly evolving, so it is important for individuals to stay up to date with the latest technologies and techniques in order to remain competitive in the job market explaining the changing returns over time (Deming and Noray 2020). ‘Cloud computing’ interacted with ‘analytical’ has negative returns across both years. The increase of cloud computing technologies over time has come with increased automation of related tasks and outsourcing of labor that reduces wages, which is likely also true of analytical skills such as Excel skills (Berger and Frey 2016).

‘Big data’ interacted with ‘programming’ language skills has, however, a positive return in the later time frame in 2018-2020 Q1. Machine learning occupations that are rapidly increasing in terms of demand frequently require the combination of programming together with big data skills, which may explain the positive return of the interaction of these two skills groups (Verma, Lamsal, and Verma 2022).

Table 3: Lasso regression output 2014-2015 and 2018-2020 Q1

Dependent variable: Log of hourly wage in occupation state cells			
2014-2015		2018-2020	
Collaborative leader	0.02	Collaborative leader	0.01
		Interpersonal & organized	-0.08
Big data	-0.04		
		Cloud computing	0.20
Research	-0.02		
		Math	0.01
		Machine Learning	0.37
Interpersonal & organized squared	-0.04	Interpersonal & organized squared	0.05
		Programming squared	-0.16
		Research squared	-0.11
Analytical squared	-0.03	Analytical squared	0.03
Collaborative leader x Research	0.001	Collaborative leader x Research	0.18
		Collaborative leader x Math	0.20
Interpersonal x Programming	-0.02		
		Interpersonal x Analytical	-0.06
Big data x Cloud computing	2.66	Big data x Cloud computing	-1.02
		Big data x Programming	0.11
Cloud computing x Programming	0.92		
		Cloud computing x Research	0.99
Cloud computing x Analytical	-1.71	Cloud computing x Analytical	-1.46
Programming x Analytical	-0.11		
Research x Analytical	0.16		
Constant	2.89	Constant	3.54
R squared	0.94	R squared	0.95

Notes: The table displays the output from running the Lasso equation (2) above. It includes a full set of controls and state and detailed occupation fixed effects.

5. Conclusion

Amid a Fourth Industrial Revolution that is disrupting labor markets and the way in which we work, it becomes increasingly important to understand which skills will be in demand and rewarded. In this study I analyzed the skills demanded and rewarded in the labor market over time. Concretely, I look at skills as posted in job advertisements in the US shortly before the start of the Fourth Industrial Revolution 2014-2015 and compare their demand and reward to a later time period when new technologies have already settled more in 2018-2020 Q1. To do so I focus on professional occupations for the two time frames and run a regression of the logarithm of hourly wages on nine skills groups. Later, I also run a Lasso regression to test for non-linearities and interactions in the skills rewards.

This study considered nine skills groups: two non-cognitive skills ('collaborative leadership' and 'interpersonal & organized') and seven cognitive skills ('big data', 'cloud computing', 'programming', 'machine learning', 'research', 'math' and 'analytical'). Overall, I find that 'collaborative leadership' increased in importance over time in terms of predicting a positive wage premium and rising shares demanded. In contrast, the other non-cognitive skill 'interpersonal & organized' has a negative wage premium in both years, despite rising shares demanded. This difference in wages received is possibly explained by the higher automatability of occupations that require 'interpersonal & organized' skills (Josten and Lordan 2022). It can also be explained by the fact that in professional workplaces, collaboration has increased in importance requiring leaders to master this important non-cognitive skill (Ibarra and Hansen 2011; Allen, Belfi, and Borghans 2020). Collaborative leadership has the potential to foster individual and company performance both directly and indirectly through fostering inclusion. That is a leader that has collaborative leadership skills such as creativity also closely resembles an inclusive leader that determines the quality of collaboration by fostering these skills in others and creating an inclusive environment where individuals feel safe to speak up about new ideas (Nembhard and Edmondson 2006; Carmeli, Reiter-Palmon, and Ziv 2010).

Our findings demonstrate that data science is constantly evolving, causing data science skills that are in high demand and attracting a wage premium in one period, to lose their premium in the next period as individuals are required to upskill on new technologies. That is, given that technology is constantly evolving, so too are the skills demanded by those who work in the area. This underlines the importance of continuous learning for professional data scientists, in addition to wage premiums to encourage them to focus on learning the latest data science skills. Concretely, in this study, we find that 'big data' shifts from having positive returns in the earlier time frame of 2014-2015 to having negative returns in 2018-2020 Q1 while more recent technologies such as machine learning gain a wage premium in 2018-2020 Q1.

We find a complementarity between 'collaborative leadership' and 'research' skills. This finding is in line with past research that focused on the interaction of social skills and cognitive skills (Weinberger 2014; Deming 2017) and the fact that non-linear thinking becomes key for the future of work (Lordan and Pischke 2022). Our finding is also in line with Josten and Lordan (2022) who find that jobs that require 'people' skills together with 'brain' skills are less likely to be automated. Professionals in particular require non-cognitive skills but also need to understand the implications of numerical calculations (Manyika et al. 2017). Mastering even complex technologies requires a broad understanding of the underlying mechanisms and the ability to bring those across with the help of social skills. With increasing complexity in labor markets due to increased technology, it becomes crucial for workers to coordinate across production processes, to be interdisciplinary skilled and to keep an overview over the large number of machine-driven processes (Goos et al. 2019). Collaborative leaders with access to new technology can foster a culture of innovation and creative problem-solving (Goos et al. 2019).

The insights from our study are useful for companies and individuals in a number of ways. Understanding the demand and reward for skills is crucial for firms and individuals in the face of rapidly changing labor markets due to the Fourth Industrial Revolution. They provide information on the skills that are valuable in today's labor market such as 'collaborative leadership' in and of itself or in combination with broad cognitive research skills or such as up-to-date data science skills. This is useful in terms of successfully matching individuals with companies, but equally it is useful as an amenity in the employee value proposition to attract and retain talent.

In the advent of the Fourth Industrial Revolution, human capital and the attraction thereof has come to the top of the agenda of many companies. Hiring has evolved away from very specific education and experience criteria towards detailed skills requirements (Fuller et al. 2022). With a larger focus on skills-based hiring which reduces the focus on degrees, companies might hire more diversely and inclusively by broadening their talent pool to skilled non-degree holder (Fuller et al. 2022). Companies can invest in task-based assessments that focus on the skills they need and those that have been highlighted as relevant in this study. As company knowing which specific skills to hire and to invest in is hence key as demonstrated in this study.

Further, companies can invest in upskilling their workforce. As regards to non-cognitive skills, Josten and Lordan (2021) highlight that they are more malleable than cognitive skills throughout an individual's lifespan. They also highlight, however, that there is mixed evidence on the impact of non-cognitive skills training for knowledge workers. They argue that it is crucial for teaching programs to carefully design courses that are evidenced-based where possible. They also call for programs to be vigorously evaluated for their effectiveness. This can be achieved when courses are rolled out in a manner that mimics randomized control trials, to allow for clear evidence on the causal effect of such courses on the desired outcomes.

This study further highlights that this is particularly crucial in the field of data science that evolves rapidly. Examples for upskilling in data science are data coding bootcamps or short courses in data science (Deming and Noray 2020). Such upskilling tools have been on the rise, which shows that 'lifelong learning' is already at the top of the agenda for companies and individuals alike. Upskilling also helps to tackle skills shortages (Deming and Noray 2020). From this study, employers also learn that labor shortages are not necessarily about a shortage in workers but about a shortage in job-relevant skills. That is while employers who are skilled in legacy data science skills might lack job-relevant skills, they may still be able to upgrade their skill set to job-relevant skills that are paid well. Our work provides information on the volatility of prices for specific skills.

LONDON SCHOOL OF ECONOMICS
CITI

Appendix Skills paper

Appendix A

1. Frequency of job advertisements over time

Table A1: Frequency of job advertisements over time

	2013	2014	2015	2016	2017	2018	2019	2020	2021
					Frequency				
January	13	1,077	4,764	13,644	15,909	18,150	25,350	25,185	20,410
February	23	1,501	6,795	11,139	16,211	18,687	23,438	27,048	22,807
March	20	3,640	5,713	10,336	19,024	20,642	25,559	20,524	21,727
April	30	4,229	6,184	16,864	20,271	19,711	23,387	10,009	2,960
May	42	3,817	5,838	11,338	18,846	20,749	24,838	12,780	2,144
June	50	4,933	4,559	14,393	19,953	26,414	23,000	10,875	
July	57	5,214	6,102	14,877	19,746	21,125	28,885	13,712	
August	145	5,522	4,676	15,121	20,208	24,049	26,197	15,643	
September	193	4,212	12,310	15,724	17,877	22,680	11,926	16,559	
October	241	6,665	14,220	15,305	18,463	27,074	26,460	19,549	
November	750	4,196	11,191	13,718	15,574	20,221	25,394	17,788	
December	543	5,081	10,598	12,795	13,312	17,817	17,996	16,904	

Notes: This table shows the number of job advertisements over time. We restrict our data to the time frame 2014-March 2020. Before 2014, the data is small and after March 2020 there is a drop in advertisements due to the outbreak of the Covid-19 pandemic.

Appendix B: Keyword selection

1. Keyword selection

The 236 underlying keywords are chosen based on how they appear in the data but we also follow the academic literature in the choice of the keywords (Deming and Kahn 2018) and the professional literature as defined in a report by the management consulting company McKinsey (Dondi et al. 2021). Synonymous keywords are grouped into 166 broader skills categories. See Table B1 below.

Table B1: Keyword list and source

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
1	creativity	creative	
		innovative	
		innovation	McKinsey
		ingenuity	
		creativity	McKinsey
2	thought leader	thought leader	
		thought leadership	
3	visionary	visionary	
4	disruptor	disruptor	
5	entrepreneurial	entrepreneurial	
		entrepreneurship	McKinsey
6	conscientious	conscientious	
		meticulous	
		attention to detail	
		diligent	
		rigorous	
7	reliable	reliable	
8	competent	competent	
		competency	
9	self discipline	self discipline	
		disciplined	
10	organised	organised	
		methodical	
		organized	Deming and Kahn (2018), Table 1
11	detail oriented	detail oriented	Deming and Kahn (2018), Table 1
12	attentive	attentive	
13	dependable	dependable	
14	verbal skills	verbal skills	
15	stakeholder management	stakeholder management	
16	build rapport	build rapport	
		building rapport	
17	articulate	articulate	
18	Presentation Skills	Presentation Skill	
		presentation	Deming and Kahn (2018), Table 1
19	interpersonal	interpersonal	
20	collaborate	collaborate	
		collaborative	

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
		work closely	
		collaboration	Deming and Kahn (2018), Table 1, McKinsey
21	supportive	supportive	
22	inclusive	inclusive	McKinsey
23	strategic	strategic	
		strategy	
		strategize	
		strategist	
24	influence	influence	
		influential	
		influencing	
25	negotiation	negotiation	Deming and Kahn (2018), Table 1
		negotiator	
		negotiate	
26	gravitas	gravitas	
27	networking	networking	
		Developing relationship	
28	charismatic	charismatic	
29	persuasiveness	persuasiveness	
		persuasive	
		persuade	
		persuasion	
30	confident	confident	McKinsey
31	personal brand	personal brand	
32	self starter	self starter	
33	goal orientated	goal orientated	McKinsey
34	motivated	highly motivated	
		motivated	
35	autonomous	autonomous	
36	hardworking	hardworking	
		hard working	
37	multitask	multi tasker	
		multi task	
		multi tasking	Deming and Kahn (2018), Table 1
		multitasker	
		multitask	
38	competing priorities	competing priorities	

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
		prioritise	
		prioritize	
		prioritisation	McKinsey
		prioritization	
39	juggle	juggle	
		juggling	
40	time management	time management	Deming and Kahn (2018), Table 1, McKinsey
41	curiosity	curiosity	
		curious	
42	openness	openness	
43	tenacity	tenacity	
		tenacious	
44	imaginative	imaginative	
45	inquisitive	inquisitive	
		inquisitiveness	
46	persistence	persistence	McKinsey
		persistent	
47	empathy	empathy	McKinsey
		empathetic	
48	humble	humble	
		humility	McKinsey
49	tolerant	tolerant	
50	thoughtful	thoughtful	
51	mindful	mindful	
52	accommodating	accommodating	
53	empower	empower	McKinsey
54	emotional intelligence	emotional intelligence	
		EQ	
55	leadership	leadership	Deming and Kahn (2018), Table 1
		leader	
56	critical thinking	critical thinking	Deming and Kahn (2018), Table 1
57	critical decision making	critical decision making	
58	decisive	decisive	
		decisiveness	
59	analytical	analytical	Deming and Kahn (2018), Table 1
60	astute	astute	

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
61	logical	logical	
62	judgement	judgement	
63	observant	observant	
64	research	research	Deming and Kahn (2018), Table 1
65	scientific	scientific	
66	qualitative	qualitative	
67	quantitative	quantitative	
68	experimental	experimental	
69	math	maths	
		mathematics	
		mathematical	
		math	Deming and Kahn (2018), Table 1
70	algebra	algebra	
71	calculus	calculus	
72	calculation	calculation	
73	trigonometry	trigonometry	
74	numerate	numerate	
		numerical	
		numeracy	
75	discipline	discipline	
76	statistics	statistics	Deming and Kahn (2018), Table 1
		statistical	
77	econometric	econometric	
78	multivariate	multivariate	
79	anova	anova	
80	linear models	linear models	
81	biostatistics	biostatistics	
82	bayesian	bayesian	
83	stochastic	stochastic	
84	r studio	r studio	
85	spss	spss	
86	data-driven	data driven	
87	informatics	informatics	
88	actuarial	actuarial	
89	bioinformatics	bioinformatics	
90	Python	Python	Deming and Kahn (2018), Table 1
91	Amazon Web Services	Amazon Web Services	

Count	**Skills category**	**Keywords of skills requirements (236 in total)**	**Additional source**
		AWS	
92	apache	apache	
93	hadoop	hadoop	
94	azure	azure	
95	bigquery	bigquery	
96	containerization	containerization	
97	docker	docker	
98	GCP	GCP	
99	Google Cloud Platform	Google Cloud Platform	
100	dynamodb	dynamodb	
101	elasticsearch	elasticsearch	
102	kubernetes	kubernetes	
103	slack	slack	
104	terraform	terraform	
105	nlp	nlp	
106	hdfs	hdfs	
107	hive	hive	
108	jupyter	jupyter	
109	keras	keras	
110	machine learning	machine learning	
111	pytorch	pytorch	
112	scala	scala	
113	spark	spark	
114	scikit learn	scikit learn	
115	tensorflow	tensorflow	
116	Java	Java	Deming and Kahn (2018), Table 1
117	SQL	SQL	Deming and Kahn (2018), Table 1
118	mongodb	mongodb	
119	nosql	nosql	
120	jenkins	jenkins	
121	git	git	
122	openshift	openshift	
123	openstack	openstack	
124	api	api	
125	udeploy	udeploy	
126	vmware	vmware	
127	javascript	javascript	
128	json	json	
129	nginx	nginx	

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
130	xml	xml	
131	teamwork	teamwork	Deming and Kahn (2018), Table 1
		team work	
		team player	
		teamplayer	
132	objectivity	objectivity	
		objective	
133	Problem solving	Problem solving	Deming and Kahn (2018), Table 1, McKinsey
134	meeting deadlines	meeting deadlines	Deming and Kahn (2018), Table 1
135	energetic	energetic	Deming and Kahn (2018), Table 1
136	writing	writing	Deming and Kahn (2018), Table 1
137	Customer	Customer	Deming and Kahn (2018), Table 1
138	sales	sales	Deming and Kahn (2018), Table 1
139	client	client	Deming and Kahn (2018), Table 1
		client relationship	
140	project management	project management	Deming and Kahn (2018), Table 1
141	Supervisory	Supervisory	Deming and Kahn (2018), Table 1
142	mentoring	mentoring	Deming and Kahn (2018), Table 1
143	budgeting	budgeting	Deming and Kahn (2018), Table 1
144	accounting	accounting	Deming and Kahn (2018), Table 1
145	finance	finance	Deming and Kahn (2018), Table 1
146	cost	cost	Deming and Kahn (2018), Table 1
147	computer	computer	Deming and Kahn (2018), Table 1
148	common software (e.g. Excel, PowerPoint)	Excel	Deming and Kahn (2018), Table 1
		PowerPoint	Deming and Kahn (2018), Table 1
		spreadsheets	Deming and Kahn (2018), Table 1

Count	Skills category	Keywords of skills requirements (236 in total)	Additional source
149	mental flexibility	mental flexibility	
150	goal achievement	goal achievement	
151	self-awareness and self-management	self awareness	McKinsey
		self management	McKinsey
		self aware	
		self manage	
152	Active listening	Active listening	McKinsey
153	Public speaking	Public speaking	McKinsey
154	Synthesizing	Synthesizing	McKinsey
155	Consensus	Consensus	
156	Logical	Logical	McKinsey
157	Adaptability	Adaptability	McKinsey
158	Agile thinking	Agile thinking	McKinsey
159	trust	trust	McKinsey
		trustworthy	
160	Sociability	Sociability	McKinsey
		sociable	
161	Role model	Role model	
162	Coaching	Coaching	McKinsey
163	risk-taking	risk taking	McKinsey
164	Conflict	Conflict	McKinsey
165	Grit	Grit	McKinsey
166	Integrity	Integrity	McKinsey

Notes: This table shows the 236 keywords used in the third column 'keywords of skills requirements'. The keywords describe skills requirements derived from the skills requirements section in a respective job advertisement. For the principal component analysis, keywords that are synonyms or very similar are grouped into broader skills categories highlighted in the second column 'Skill category'. We group an overall of 166 skills categories. An example is the skills category 'trust' that is coded as a binary variable that is equal to one if either the word trust or trustworthy appears in an advert. The final column 'Additional source' flags if a keywords has also been mentioned by Deming and Kahn (2018) or published by Dondi et al. (2021). We exclude ambigious words; i.e. 'patient' and 'staff' that have been used by Deming and Kahn (2018) or 'senior' and 'planning and ways of working'. We further remove the keywords 'communication' and 'management' as they are frequently used also in non-skills contexts (i.e. they appear in 59% of all job advertisements).

Appendix C

1. **Description of the BERT model**

The Bidirectional Encoder Representations from Transformers (BERT) model (https://arxiv.org/pdf/1810.04805.pdf) was developed by Google AI Language, published in 2018 – one of the biggest recent breakthroughs in Natural Language Processing (NLP).

Transformers are models that convert text into vector embeddings via encoding, and back via decoding. BERT is a partial example of such, as the model only generates embeddings from text (i.e. the encoding process). Fundamentally, BERT is a pre-trained model based on two tasks – Masked Language Model (Masked LM) and Next Sentence Prediction (NSP). In the first task, a proportion of word tokens in a sentence are "masked" at random – either removed, replaced with another word token, or unchanged – and the model is trained to predict the masked token. This allows the model to learn input texts in a multi-layered context through a bidirectional approach, which is more powerful than traditional unidirectional, left-to-right or right-to-left approaches. In the NSP task, BERT received pairs of sentences as input, where 50% of inputs are pairs of consecutive sentences, and the other 50% being random pairs. BERT is trained to predict if the second sentence in the pair is the subsequent sentence in the original document, and hence learns the context and association of sentences. The goal in BERT's pre-training is to minimise the overall loss function from these two tasks.

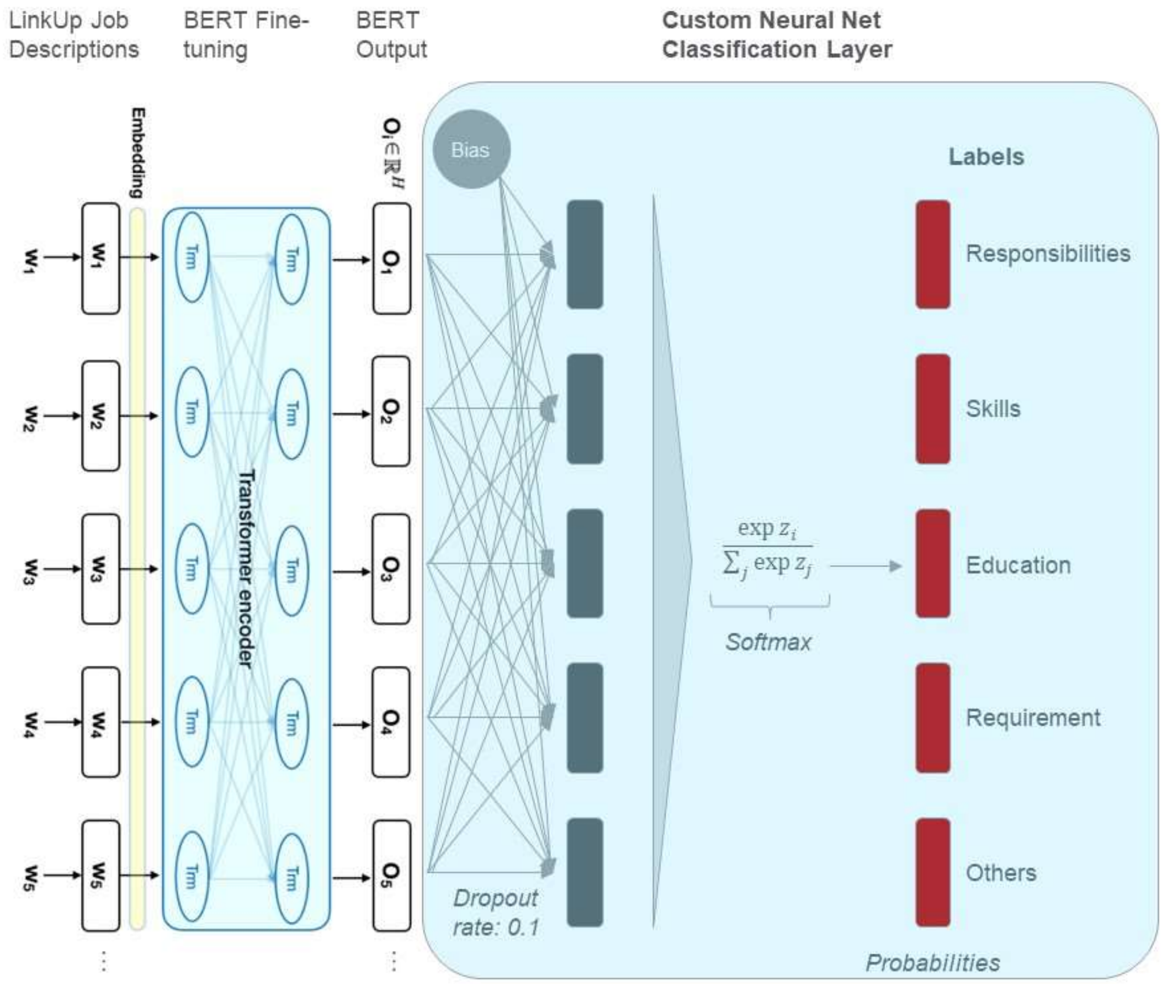
Following the pre-training, the BERT model can be adapted for downstream NLP tasks through fine-tuning, which is computationally inexpensive and straightforward. The model has been evaluated against 11 common NLP tasks – such as General Language Understanding Evaluation (GLUE) https://gluebenchmark.com/leaderboard, the Stanford Question Answering Dataset (SQuAD) v1.1 and v2.0 https://arxiv.org/pdf/1606.05250.pdf, and the Situations with Adversarial Generations (SWAG) https://arxiv.org/abs/1808.05326 – and has been shown to achieve state-of-the-art results.

2. **Model implementation**

DistilBERT (https://huggingface.co/docs/transformers/model_doc/distilbert and https://arxiv.org/abs/1910.01108), one of the many variations of BERT which features a reduced model size (40% fewer parameters) and improved speed (60% faster) while preserving 95% of BERT's performance, has been implemented in this paper. The model is used for a multi-class classification task – given any input sentence, the prediction output is to classify the sentence in one of the following five categories: responsibilities, skills, education, requirement and others. This mimics the structure and order of how a job description is typically written. In section 4 of Appendix C below we provide three exemplary job advertisements that are flagged for different overall keyword categories (i.e. machine leaning, collaborative leadership and cloud computing).

In the classification procedure, the pre-trained DistilBERT model is first used to generate the sentence embeddings on a small sample of manually pre-labelled dataset of 832 sentences from job descriptions, divided with a 80-20 train-test split. These embeddings and labels are then subsequently fed into a neural network architecture for the classification task training. The neural net architecture is of a standard construct with the inclusion of bias, a dropout rate of 0.1 and a softmax activation function for classification. After completion of the training phase and test data evaluation, DistilBERT is used to generate sentence embeddings at scale on the rest of the unseen data – sentences from millions of job postings – and then classified using the trained neural net model into one of the five sentence categories. This process is depicted in Figure C1 below.

Figure C1: Illustration of the DistilBERT model adapted for sentence classification



Notes: The figure shows an overview (from left to right) of how the DistilBERT model is used for classifying sentences from job postings descriptions into five categories: responsibilities, skills, education, requirement and others. Words in each sentence are tokenised and fed as input to the BERT embedding layer. The output is passed into a simple neural net classifier with dropout rate 0.1 and a softmax activation function.

3. Model performance and accuracy

In table C1 below we present the classification prediction results based on the 20% test set data from the train-test split above on the format of a confusion matrix.

The precision and recall metrics across each five categories are 80%+, with the education category achieving the best prediction results due to the distinct wordings typically used in education requirements, such as BSc, MSc and PhD. There are several misclassifications between the responsibilities and others categories, which is expected as BERT create individual embeddings per sentence and is prone to losing a contextual understanding on a paragraph level (e.g. an entire paragraph on responsibilities that goes before a skills paragraph in a typical job description which could otherwise be easily identified by a human) – especially in case of generic, short sentences. Nonetheless, the overall model accuracy for the multiclass classification is 90.4%[23] on the test set, which is sufficiently accurate to be used for narrowing down job descriptions data into skill-related sections only.

Table C4: Classification prediction results

		Predicted					Precision	Recall
		responsibilities	skills	education	requirement	others		
Actual	responsibilities	41	0	0	2	5	85.4%	85.4%
	skills	3	27	0	0	1	96.4%	87.1%
	education	0	0	5	0	0	100%	100%
	requirement	0	1	0	9	0	81.8%	90.0%
	others	4	0	0	0	69	92.0%	94.5%

The limiting factor in this model is the runtime of DistilBERT in creating sentence embeddings and classification through neural net. A 25% stratified sample of the raw job descriptions data was extracted, and it took four weeks in total to complete the classification predictions on all sentences in the sample. Although not implemented in the paper, the model runtime can be improved through incorporating the CUDA parallel computing model where GPUs are available.

[23] If one sums up the five values along diagonal of the table (41+27+5+9+69), i.e. the correct predictions, and divides it by the total table sum (correct + incorrect predictions) this gives the overall accuracy which is 90.4%.

4. Skills data processing from model output

After training the BERT model and reviewing the prediction accuracy, we utilise the model to classify job posting descriptions at scale and extract skills-related sentences from the descriptions. This is a crucial process that allows the skills analysis to focus on relevant sections of the job posting description text, and discard other irrelevant parts (i.e. education, legal requirement and others) which often contain generic descriptions of the company and team. For each job posting, we create a dummy variable for each skill keyword in Table B1, representing a Boolean flag of whether that skill keyword is present for at least once in the job description (0=not represent, 1=present). Filtering out parts of the job descriptions other than skills and responsibilities prevents false positive flags that are otherwise identified from the generic company and team descriptions that bear little relevance to the actual requirements of the role.

1. Three examples of job advertisements by flagged keyword category

a. Machine Learning

Machine Learning Engineer - Safety Product

Amazon.com, Inc.

About US:

Launched in 2011, Twitch is a global community that comes together each day to create multiplayer entertainment: unique, live, unpredictable ex periences created by the interactions of millions. We bring the joy of co-op to everything, from casual gaming to outstanding esports to anime marathons, music, and art streams. Twitch also hosts TwitchCon, where w e bring everyone together to celebrate and grow their personal interest s and passions. We're always live at Twitch. Continually learn about al l things Twitch on Linkedin, Twitter and on our Blog.

About the Role:

Are you passionate about making Twitch safer, more inclusive, and a nic er place to enjoy? This position lets you do exactly that! You will be part of a rapidly growing Machine Learning team which develops and depl oys algorithms that are the first line of defense of users' safety at T witch. You will work with passionate co-workers who live Twitch's missi on and put their hearts into their work. If this sounds like an environ ment where you will thrive, come and join our team!

You Will:

* Build machine learning products in the safety world to protect Twitch from bad behavior such as followbotting, spam, phishing, and violent or illegal content
* Design and build scalable infrastructure that enables deploying machi ne learning models on petabytes of data
* Develop data pipelines and other modern big data processing systems
* Build distributed services to power machine learning solutions
* Design databases and make storage choices for efficient ML data manag ement
* Bring operational excellence to MLOps/DevOps
* Work on event-driven data flows to evolve machine learning applicatio ns
* Partner with fellow engineering and science teams to accomplish compl ex projects together

You Have:

* Bachelors in Computer Engineering/Science or equivalent
* Outstanding programming skills
* Demonstrated ability to understand and contribute to large software s ystems
* Experience building distributed services or backend services and unde rstand scaling computation to thousands of machines
* Passion for machine learning

Bonus Points:

* 2+ years of industry experience or equivalent internship experience
* Experience working with Amazon Web Services or other cloud solutions
* Experience with ML libraries/frameworks such as Keras, Tensorflow, an
d AWS Sagemaker
* Understanding of MLOps or DevOps concepts
* Experience working with large-scale data and orchestration tools such
as Airflow, AWS Stepfunctions and Kubeflow
* Experience with streaming data and event-driven systems, and knowledg
e tools like Kinesis, Kafka, Flink, Spark, RabbitMQ and SQS
* You are a Twitch user who cares about safety

Perks:

* Medical, Dental, Vision & Disability Insurance
* 401(k), Maternity & Parental Leave
* Flexible PTO
* Commuter Benefits
* Amazon Employee Discount
* Monthly Contribution & Discounts for Wellness Related Activities
& Programs (e.g., gym memberships, off-site massages, etc.),
* Breakfast, Lunch & Dinner Served Daily
* Free Snacks & Beverages

Pursuant to the San Francisco Fair Chance Ordinance, we will consider f
or employment qualified applicants with arrest and conviction records.

We are an equal opportunity employer and value diversity at Twitch. We
do not discriminate on the basis of race, religion, color, national ori
gin, gender, sexual orientation, age, marital status, veteran status, o
r disability status.

b. Collaborative Leadership

Senior Business Development Mgr, Business Electronics

Amazon.com, Inc.

The Consumer Electronics (CE) team at Amazon is looking for a Senior Bu
siness Development Manager responsible for expanding the Business Elect
ronics (BE) categories on Amazon Business. Amazon Business is dedicated
to offering a broad selection of products and supplies to business, ind
ustrial, education, government and commercial customers at competitive
prices. The Business Electronics categories includes PC and office prod
ucts, networking equipment, professional video / audio equipment, secur
ity, camera and imaging equipment.

This Senior Business Development Manager creates new partnerships (both
internally and externally), grows existing relationships with Fortune 5
00 companies, and licenses assets to drive product / service improvemen
ts and innovation while reducing costs without sacrificing the customer
experience. This role is an ideal next step for a leader who is looking
to develop into a next career stage and to gain exposure to senior lead
ership both internally and externally.

This position has responsibilities that can create step-level changes t
o the business through adding strategic selection and introducing new v
endor programs and initiatives like pricing and service expansion. This
role requires an individual who can work autonomously in a highly deman
ding environment, with strong attention to detail and exceptional organ
izational skills.

The ideal candidate will have experience in negotiations, strategic pla
nning, forecasting, and a background in B2B, B2C or e-commerce business
es. The candidate must be able to work in an ambiguous but collegial en
vironment where teamwork is a priority to deliver results. The right ca
ndidate will be flexible, action and results oriented, self-starting an
d demonstrate a willingness to learn and react quickly. The candidate m
ust also be decisive and able to move with speed to implement their own
ideas. The candidate should be strong analytically and be comfortable g
enerating and evaluating forecasts and metrics to come up with recommen
dations and guidance to present to leadership. Strong communication ski
lls (both oral and written) are critical.

The Senior Business Development Manager will be responsible for the fol
lowing:

* Lead the signing, and on-boarding of new business and professional ve
ndors and expanding business and professional selection from existing C
E vendors
* Own high-level negotiations of agreements/deals with leading brands t
o drive business inputs
* Act as a leader and ambassador of Amazon and B2B across CE categories
, developing deep knowledge of supply/demand trends and success drivers
* Lead day-to-day operational aspects of the business, including gather
ing and addressing customer and vendor feedback, price management, and
business improvement initiatives
* Ability to see around corners and pioneer new initiatives with stakeh
olders across the company
* Work with a team charged with building, owning, and sharing financial
goals and deliverables for select group of vendors
* Develop and grow strong collaborative relationships internally and ex
ternally
* Bachelor's degree required
* 5+ years of relevant experience in sales, buying, account management,
consulting and/or marketing preferably in eCommerce or B2B industries
* Exceptional interpersonal and communication skills; strong writing an
d speaking skills
* Demonstrated ability to manage multiple projects - prioritization, pl
anning and time management
* Proactive attitude, detail oriented, fast learner and team player
* Strong influencing and negotiation skills
* Proven analytical skills \u2013 ability to analyze large data sets to
make strategic decisions
* Demonstrated success in situations with a high level of ambiguity
* Proven track record of delivering results in B2B or relevant category
* MBA
* Experience across categories and markets
* Business Development / Vendor Management experience

Amazon is an Equal Opportunity-Affirmative Action Employer - Minority /
Female / Disability / Veteran / Gender Identity / Sexual Orientation.

c. Cloud Computing

Application Innovation Specialist

Microsoft Corporation

We are currently looking for Application Innovation Specialists to join our teams across our various business groups: Enterprise, Small Medium & Corporate, as well as Regulated Industries. By applying to this role, you will be considered for multiple opportunities within Microsoft across the United States.

Microsoft is on a mission to empower every person and every organization on the planet to achieve more. Our culture is centered on embracing a growth mindset, a theme of inspiring excellence, and encouraging teams and leaders to bring their best each day. Growth mindset encourages each of us to lean in and learn what matters most to our customers, to create the foundational knowledge that enables us to make customer-first decisions in everything we do. In doing so, we create life-changing innovations that impact billions of lives around the world. You can help us achieve our mission.

Are you insatiably curious? Do you embrace uncertainty, take risks, and learn quickly from your mistakes? Do you collaborate well with others, knowing that better solutions come from working together? Do you stand in awe of what humans dare to achieve, and are you motivated every day to empower others to achieve more through technology and innovation? Are you ready to join the team that is at the leading edge of Innovation at Microsoft?

To learn more about Microsoft's mission, please visit: https://careers.microsoft.com/mission-culture

Check out all our products at: http://www.microsoft.com/en-us

We are currently hiring across a variety of teams with various levels of skills and experiences required. Below maps the minimum required qualifications to be considered for these positions.

Experiences Required: Education, Key Experiences, Skills and Knowledge:

Professional

* 5+ years of technology-related sales or account management experience OR a Bachelor's Degree in Computer Science, Information Technology, Business Administration, or related field AND 4+ years of technology-related sales or account management experience required
* Experienced. Relevant experience selling cloud services or application development services to medium and large enterprise customers with a focus on cloud application development required
* Account Management. Effective territory/account management: planning, opportunity qualification and creation, stakeholder and executive communication, needs analysis, services/partner engagement, opportunity management and pipeline management required
* Executive Presence. Experience and expertise selling to LOB decision makers, technical decision makers & enterprise solution architects by aligning & reinforcing the value of the solution to the customer

's overall business pain and/or strategic opportunities and decision cr
iteria preferred
* Problem Solver. Ability to solve customer problems through cloud tech
nologies, specifically solutions related to cloud native apps - contain
ers & serverless, microservices, developers tools and DevOps, low c
ode, migration to cloud required
* Collaborative. Orchestrate and influence virtual teams to pursue sale
s opportunities and lead v-teams through influence required

Technical

* Enterprise-scale technical experience with cloud and hybrid infrastru
ctures, architecture designs, migrations, and technology management. Su
bject matter expertise in one or more of the following: required
* Application development platforms on public clouds and/or Azure in de
velopment languages such as Java, JavaScript, Python, PHP, C#, Node.JS
targeting Android, iOS, Linux, Windows, public clouds or Azure.
* Scalable architectures using Azure App Service, API management, serve
rless technologies, container orchestration (e.g. AKS, Kubernetes, Red
Hat OpenShift etc.), microservice frameworks etc.

* Software development practices like DevOps and CI/CD tool chains (i.e
. Jenkins, Spinnaker, Azure DevOps, GitHub). required
* Understanding of Data & AI technologies in context of app develop
ment (e.g. SQL and NoSQL Databases, Big Data, Cognitive Service, Machin
e Learning etc.). preferred
* Understanding of Low code platform and technologies such as Power Pla
tform. preferred
* Competitive Landscape. Knowledge of cloud development platforms requi
red
* Partners. Understanding of partner ecosystems and the ability to leve
rage partner solutions to solve customer needs required

Education

* Bachelor's degree or equivalent work experience required
* Certification in the following technologies preferred: Cloud, mobile,
web application development, cloud-native application architecture (i.e
. containers, microservices, API management), modern software developme
nt techniques like DevOps and CI/CD tool chains (i.e. Jenkins, Spinnake
r, Azure developer services, GitHub) and container orchestration system
s (i.e. Docker, Kubernetes, Red Hat OpenShift, Cloud Foundry, Azure Kub
ernetes Service, GitHub), Low Code (Power Platform). Required
* Certification in sales, sales management, complex sales training, sal
es methodologies, broad evangelism through events (presentation skills)
, and consultative selling preferre

Microsoft is an equal opportunity employer. All qualified applicants wi
ll receive consideration for employment without regard to age, ancestry
, color, family or medical care leave, gender identity or expression, g
enetic information, marital status, medical condition, national origin,
physical or mental disability, political affiliation, protected veteran
status, race, religion, sex (including pregnancy), sexual orientation,
or any other characteristic protected by applicable laws, regulations a
nd ordinances. We also consider qualified applicants regardless of crim
inal histories, consistent with legal requirements. If you need assista
nce and/or a reasonable accommodation due to a disability during the ap

plication or the recruiting process, please send a request via the Acco
mmodation request form.

Benefits/perks listed below may vary depending on the nature of your em
ployment with Microsoft and the country where you work.

Microsoft is uniquely positioned to win "App Innovation" workloads to h
elp with customer's Digital Transformation journey. Microsoft apps port
folio spans Azure App Platform, PowerApps, GitHub/DevOps and Visual Stu
dio. Azure App Platform is one of the fastest growing businesses inside
the Azure platform and with tighter integrations with developer tools.
Microsoft is hiring Specialist sellers for Application Innovation to de
liver on Microsoft's aspirations and sales goals in this dynamic and fa
st-growing enterprise market.

As an App Innovation Specialist, you will be a senior solution sales le
ader within our enterprise sales organization working with our most imp
ortant customers selling entire Microsoft Apps portfolio. You will lead
orchestrating a virtual team of Cloud Solution Architects, partners and
other resources to advance the sales process and achieve/exceed sales a
nd usage/consumption targets for Application Innovation related workloa
ds in your assigned accounts. You will be a trusted advisor and a cloud
application development subject matter expert.

Primary accountabilities for this role include:

* Create "buy-in" vision with the Apps Decision Makers
* Take active role in defining and influencing the customer's business
challenges and opportunities.
* Understand the financial, qualitative and competitive drivers impacti
ng customer business.
* Create "buy-in" vision and gain sponsorship by generating excitement
around Microsoft solutions value.

* Map out customer's current and desired state and expectations
* Identify customer's digital transformation needs, business drivers, t
heir perspective, and concerns.
* Collaborate effectively with the customers to outline their business
problems, opportunities.
* Establish and understand the buying decision criteria and timeline to
make decision regarding the solution.
* Establish credibility and trust by demonstrating that Microsoft solut
ions not only solve customer business problems but also the value they
realized.

* Lead the solution design by assembling and orchestrating Sales &
technical resources.
* Proficient in delivering Microsoft Apps vision, strategy, value to C-
level and Apps Decision Makers
* Orchestrate the technical experts to create the solution, validate it
via Proof of Concept, and showcase how it meets customer business and t
echnology requirements.
* Able to handle customer objections and competitive differentiation.

* Substantiate the value of the solution (commercial discussion)

* Building business cases with TCO analysis and negotiate a deal that's based on value by leveraging various MS offers and programs.
* Secure customer commitment for the business proposal.

* Drive Sales Excellence
* Drive the App Innovation business to overachieve revenue, consumption and scorecard targets.
* Maintain excellence in pipeline management, accuracy of sales forecast, and deal close plans.

* Lead with subject matter expertise and be the Voice of the customer.
* Influence the Microsoft Application Innovation go to market strategies by providing feedback to sales, marketing, and engineering on product requirements and sales blockers.
* Stay sharp, attaining and maintaining required certifications. We encourage all our employees to continuously maintain and enhance their technical, sales, professional skills and competitive readiness. You will be recognized for sharing, learning and driving individual work that all result in business impact for customers, partners and within Microsoft. We encourage thought leadership and leadership from every employee.

Appendix D

1. **Principal Component Analysis**

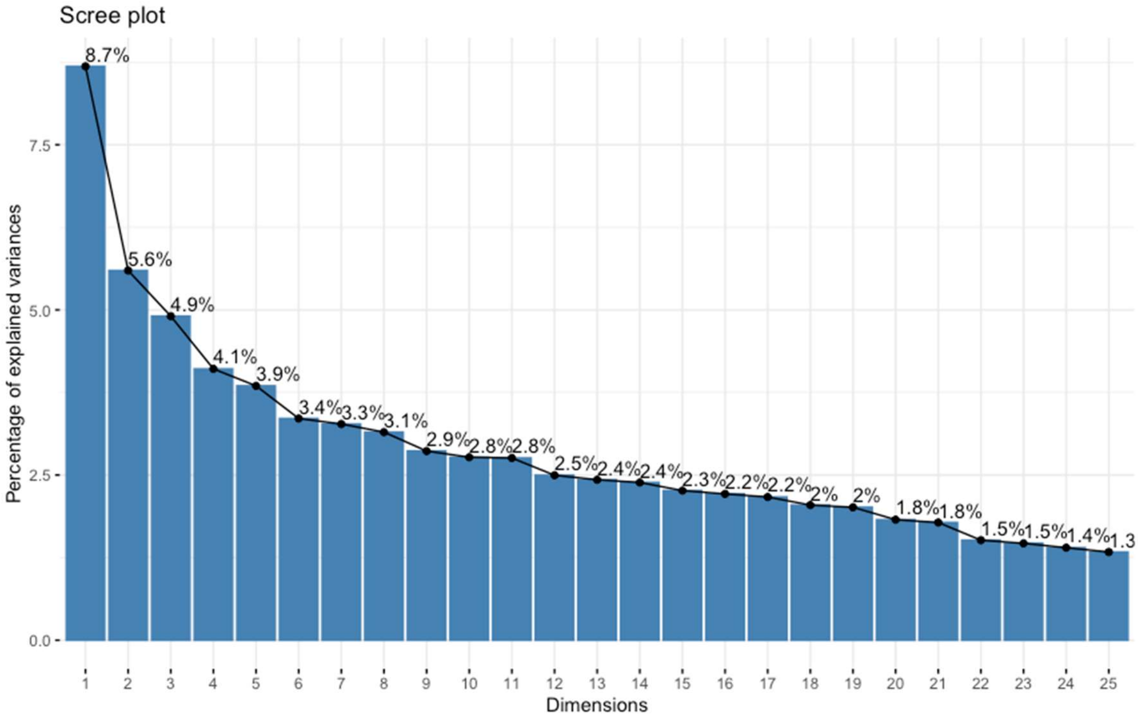
The skills requirement keywords from LinkUp are grouped using principal component analysis (PCA). PCA is a psychometric method used to reduce the dimensionality of variables (Jollife and Cadima 2016). The aim of PCA is to define variables that are linear combinations of the keywords of skills in the data and that are uncorrelated with each other (Jollife and Cadima 2016). It aims at simplifying the interpretation of the variables of interest; i.e. rather than analyzing the impact of a large number of individual keywords we are interested in the impact of components that are representative of the original sample. The sample for the PCA includes all 1.3 million of LinkUp job advertisements for professionals for the years 2014-2020 Q1. It further includes 166 skills category variables that are used as input for the PCA (see table B1 above). The analysis is run at the job advertisement level and the 166 skills category variables are coded as binary variables that are equal to one if the underlying keywords appear in the respective job advert and zero otherwise.

We run the principal component analysis initially using all 166 skill category variables. We then run oblique rotations that allows for correlation across components. If the first principal components account for a large part of the total variance, the remaining components can be dropped in a confirmatory factor analysis through rotations (Bartholomew et al. 2011; Heckman et al. 2012). The PCA is hence run in rotations of components to be retained to find the optimal number of principal components subject to the following rules for the cut-off for the components: a cumulative variance explained of the components of at least 60%, examining a jump in the scree plot (i.e. a point at which the eigenvalue of a given component falls substantially) and choosing component cut-offs that are sensible and intuitive (Bartholomew et al. 2011). In the final step, we drop keywords that are weakly associated with the components (i.e. have a loading of less than 0.32) and those that cross-load onto multiple components. The cut-off of 0.32 has been recommended in the literature (Tabachnick and Fidell 2018) and our large sample size allows us choosing a relatively low loadings cut-off. The overall PCA analysis results in a total number of nine skills components: ‘collaborative leadership‘, ‘interpersonal & organized‘, ‘big data‘, ‘cloud computing‘, ‘programming‘, ‘machine learning‘, ‘research‘, ‘math‘, ‘analytical‘.

Step 1: Principal Component Analysis using all 166 skills

Figure D1 below and table D1 together show the output of running the first principal component analysis. In step 2 below we choose to rotate with 19 components given a cumulative variance explained of greater than 60% in table D1 as well as a small jump in the scree plot in figure F1.

Figure D1: Scree plot principal component analysis with all skills keywords



Notes: The figure shows the scree plot from the principal component analysis and the percentage of the explained variance. A small jump in the scree plot can be detected after 19 components where the variance levels of. The z-axis shows the number of principal components and the y-axis the percentage of explained variances.

Table D2: Cumulative variance explained

Importance of components

	Comp.1	**Comp.2**	**Comp.3**	**Comp.4**	**Comp.5**	**Comp.6**	**Comp.7**
Standard deviation	0.60	0.48	0.45	0.41	0.40	0.37	0.37
Proportion of Variance	0.09	0.06	0.05	0.04	0.04	0.03	0.03
Cumulative Proportion	0.09	0.14	0.19	0.23	0.27	0.31	0.34
	Comp.8	**Comp.9**	**Comp.10**	**Comp.11**	**Comp.12**	**Comp.13**	**Comp.14**
Standard deviation	0.36	0.34	0.34	0.34	0.32	0.32	0.32
Proportion of Variance	0.03	0.03	0.03	0.03	0.02	0.02	0.02
Cumulative Proportion	0.37	0.40	0.43	0.45	0.48	0.50	0.53
	Comp.15	**Comp.16**	**Comp.17**	**Comp.18**	**Comp.19**	**Comp.20**	**Comp.21**
Standard deviation	0.31	0.30	0.30	0.29	0.29	0.28	0.27
Proportion of Variance	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Cumulative Proportion	0.55	0.57	0.59	0.61	0.63	0.65	0.67

Step 2: Orthogonal rotations

We first run an orthogonal rotation reduced to 19 components, which yields a matrix of standardised factor loadings. Of those 19 components, 8 components have less than three loadings of more than 0.32 and are hence weakly loading. We therefore rotate with 11 components in the next rotation. Again two components load weakly so we rotate again with 9 principal components. Table D2 below shows the output from the rotation of 9 principal components. Any cross-loadings (loadings of greater than 0.32 on more than one component) and weak loadings (loadings of smaller than 0.32) have been removed.

Table D3: Cognitive skills and non-cognitive components resulting from PCA

Non-cognitive skills components			Cognitive skills components					
Collaborative leader			Big data			Programming		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
strategic	0.59	24.14%	hadoop	0.75	1.19%	xml	0.64	1.26%
leadership	0.58	26.17%	spark	0.75	0.72%	json	0.6	0.67%
influence	0.51	12.75%	hive	0.73	0.55%	javascript	0.59	2.51%
collaborate	0.39	24.52%	hdfs	0.53	0.14%	java	0.56	6.08%
creativity	0.34	13.57%	scala	0.47	0.47%	sql	0.39	7.12%
negotiation	0.33	6.57%	nosql	0.34	0.81%	git	0.38	1.16%
coaching	0.32	5.14%				api	0.37	1.51%
Overall		58.89%	Overall		2.28%	Overall		13.45%

Interpersonal & organized			Cloud computing			Machine Learning		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
time management	0.4	4.65%	docker	0.74	0.64%	tensorflow	0.84	0.11%
competing priorities	0.39	12.30%	kubernetes	0.71	0.41%	pytorch	0.76	0.04%
interpersonal	0.38	17.20%	amazon web services	0.48	2.32%	keras	0.73	0.03%
organized	0.36	3.38%	terraform	0.45	0.15%			
			azure	0.41	1.04%			
			jenkins	0.41	0.95%			
			openshift	0.35	0.06%			
			containerization	0.35	0.12%			
			openstack	0.32	0.22%			
Overall		29.86%	Overall		3.89%	Overall		0.11%

Research			Analytical			Math		
Keywords	Loading	Share	Keywords	Loading	Share	Keywords	Loading	Share
quantitative	0.58	3.45%	accounting	0.65	5.58%	calculus	0.73	0.05%
statistics	0.54	5.48%	finance	0.63	7.63%	algebra	0.63	0.12%
qualitative	0.43	1.06%	common software e.g. excel	0.41	16.66%	trigonometry	0.56	0.05%
research	0.37	12.86%	analytical	0.33	20.78%	stochastic	0.47	0.05%
Overall		18.58%	Overall		35.70%	Overall		0.21%

Appendix E

1. **Skills shares over time**

There is a large variation in how often each skill grouping appears in the job adverts as seen in Table 2 below. For example 'Collaborative leadership' appears in 50.14% of job adverts in 2014-2015 and 61.07% in 2018-2020 Q1. In comparison, machine learning does not appear in job advertisements in LinkUp in the earlier time frame and only appears in 0.19% of job advertisements in 2018-2020 Q1. Overall, soft skills are overused in job advertisements and also across disciplines (Calanca et al. 2019), while cognitive skills are more specific.

Table E1: Share and absolute number of observations of job adverts that request each of the nine skills groups across two time periods (2014-2015 and 2018-2020 Q1)

		Collaborative leadership	Interpersonal & organized	Big data	Programming	Machine Learning	Cloud computing	Research	Math	Analytical
2014-2015	%	50.14%	24.70%	1.63%	13.60%	N/A	1.76%	16.19%	0.19%	31.19%
	#	71,718	35,335	2,325	19,448	N/A	2,524	23,152	273	44,607
2018-2020 Q1	%	61.07%	30.85%	2.49%	13.03%	0.19%	4.88%	18.94%	0.21%	35.83%
	#	374,087	188,981	15,246	79,827	1,155	29,911	115,996	1,288	219,461

Notes: This table shows the share of the nine skills groups across two time frames of 2014-2015 and 2018-2020 Q1. It further shows the absolute number of observations by skills group.

2. **Skills shares of interactions over time**

Table E2 below documents the shares of all skills group. Overall, the share of all skills interactions have been increasing over the two time frames. The interactions that centre around zero in terms of shares are highlighted in yellow and will not be considered for the regression analysis. The interaction of 'collaborative leadership' and 'big data' for example increased from 0.8% to 1.6%. The interaction of 'collaborative leadership' and 'research' increases by 3.8 percentage points from 9.5% to 14.2%, which points at the fact that with increasing automation, the complementarity between social skills (i.e. collaborative leadership) and cognitive skills (i.e. research) increases. For example, doctors increasingly use technology such as Clinical Decision Support Software , but still need to understand statistics, which is a facet of 'research' skills alongside making final decisions drawing on their 'collaborative leadership' skills.

Table E2: Share of job adverts that request selected interactions between the nine skills groups across two time periods (2014-2015 and 2018-2020 Q1)

Skills group interactions	**2014-2015**	**2018-2020 Q1**
Collaborative leadership x Interpersonal & organized	15.7%	22.3%
Collaborative leadership x Big data	0.8%	1.6%
Collaborative leadership x Cloud computing	1.0%	3.2%
Collaborative leadership x Programming	6.6%	8.0%
Collaborative leadership x Research	9.5%	13.2%
Collaborative leadership x Math	0.1%	0.1%
Collaborative leadership x Analytical	18.8%	24.9%
Collaborative leadership x Machine Learning	n/a	0.1%
Interpersonal & organized x Big data	0.2%	0.5%
Interpersonal & organized x Cloud computing	0.3%	0.9%
Interpersonal & organized x Programming	2.8%	3.4%
Interpersonal & organized x Research	5.1%	7.2%
Interpersonal & organized x Math	0.0%	0.1%
Interpersonal & organized x Analytical	11.9%	16.0%
Interpersonal & organized x Machine Learning	n/a	0%
Big data x Cloud computing	0.3%	0.9%
Big data x Programming	1.3%	2.0%
Big data x Research	0.4%	0.8%
Big data x Math	0.0%	0.0%
Big data x Analytical	0.4%	0.7%
Big data x Machine Learning	n/a	0.1%
Cloud computing x Programming	1.1%	2.8%
Cloud computing x Research	0.2%	0.6%
Cloud computing x Math	0.0%	0.0%
Cloud computing x Analytical	0.3%	0.9%
Cloud computing x Machine Learning	n/a	0%
Programming x Research	2.6%	3.2%
Programming x Math	0.0%	0.0%
Programming x Analytical	4.3%	4.7%
Programming x Machine Learning	n/a	0.1%
Research x Math	0.1%	0.1%
Research x Analytical	8.1%	9.8%
Research x Machine Learning	n/a	0.11%
Maths x Analytical	0.10%	0.08%
Maths x Machine Learning	n/a	0.01%
Analytical x Machine Learning	n/a	2%

3. Skills shares and wage premium by occupations

To give a clearer sense of the regression estimates we consider what these mean for a subset of occupations. We show occupations that require high versus low shares of each of the respective nine skills group and their respective skills premium in table below. The wage premium is calculated by multiplying the occupation share with the coefficient. So logically, the premium is larger for occupations with high shares of the respective skills group and lower for those with low shares. For the category collaborative leader, for example, job postings for managing occupations require large shares of collaborative leaders ranging from 83% to 93%. For 2014-15 the wage premium is insignificant but turns positive for 2018-2020 and makes up around 5% of the hourly wages in the top five occupations. The required shares for 'interpersonal & organised' are smaller but still between 42%-53% in the top five occupations that show increasing wage penalties over time of up to 11% of the hourly wage. An example is the occupation loan officers for which 56% of job adverts require 'interpersonal & organised' skills and the wage penalty is around 4$ of an hourly wage of 35$. An interesting case is big data: the occupations 'Computer and information research scientists' or 'Software Developer, Applications' that are large occupations in terms of absolute count show that the highest share required for 'big data' skills turn from a wage premium in 2014-2015 to a penalty in 2018-2020. The same is true for cloud computing occupations. Programming is another interesting example where the top occupations require large shares of programming skills of up to 67% of web developers but those skills are actually punished across both time frames. Wage premia to machine learning in 2018-2020 are quite large while the share required is still very low as in the 'math' skills group. Research occupations experience wage premia. Analytical occupations have negative returns in 2014-2015 but turn insignificant in 2018-2020.

Table E3: Wage premium by top and bottom (in terms of share) occupations for all nine skills groups, and hourly wages and overall count by occupation

Collaborative leader	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Marketing Managers	6,042	29,844	91%	93%	$70.33	$76.15	-0.09%	2.93%	-$0.06	$2.23
Computer and Information Systems Managers	2,336	10,986	78%	88%	$69.63	$77.27	-0.08%	2.76%	-$0.05	$2.13
Human Resources Managers	1,267	5,771	80%	88%	$56.74	$63.15	-0.08%	2.76%	-$0.05	$1.74
Sales Managers	3,444	15,811	83%	87%	$63.68	$70.25	-0.08%	2.74%	-$0.05	$1.93
Graphic Designers	327	1,297	83%	85%	$26.22	$28.83	-0.08%	2.68%	-$0.02	$0.77
Bottom 5										
Medical and Clinical Laboratory Technicians	421	2,815	8%	4%	$19.30	$31.47	-0.01%	0.13%	$0.00	$0.04
Pharmacy Technicians	323	1,936	11%	9%	$14.77	$16.95	-0.01%	0.29%	$0.00	$0.05
Surgical Technologists	577	1,187	6%	7%	$20.97	$23.64	-0.01%	0.21%	$0.00	$0.05
Radiologic Technologists	356	1,142	8%	10%	$26.46	$28.94	-0.01%	0.32%	$0.00	$0.09
Medical Records and Health Information Technicians	924	3,118	16%	27%	$18.44	$21.96	-0.02%	0.85%	$0.00	$0.19

Interpersonal & organized	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Financial Examiners	387	2,272	42%	55%	$44.63	$41.82	-1.50%	-3.97%	-$0.67	-$1.66
Loan Officers	1,363	7,135	54%	55%	$37.87	$37.33	-1.93%	-3.94%	-$0.73	-$1.47
Financial Managers	5,268	22,811	35%	53%	$69.34	$75.02	-1.26%	-3.76%	-$0.88	-$2.82
Compliance Officers	417	2,349	39%	43%	$34.25	$35.39	-1.40%	-3.08%	-$0.48	-$1.09
Lawyers	477	2,636	37%	43%	$67.76	$72.73	-1.32%	-3.10%	-$0.89	-$2.25
Bottom 5										
Health Technologists and Technicians, All Other	701	6,130	4%	1%	$21.51	$22.71	-0.13%	-0.05%	-$0.03	-$0.01
Interior Designers	213	485	3%	4%	$24.32	$28.71	-0.10%	-0.30%	-$0.02	-$0.09
Nurse Practitioners	65	970	18%	3%	$48.17	$54.20	-0.66%	-0.20%	-$0.32	-$0.11
Pharmacists	265	3,424	30%	7%	$56.86	$59.98	-1.07%	-0.50%	-$0.61	-$0.30
Healthcare Social Workers	169	1,060	12%	13%	$26.03	$28.46	-0.45%	-0.97%	-$0.12	-$0.28

Big data	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Computer and Information Research Scientists	566	4,498	26%	33%	$59.67	$62.98	4.87%	-3.90%	$2.90	-$2.46
Software Developers, Applications	13,114	56,083	11%	15%	$51.93	$55.94	1.96%	-1.87%	$1.02	-$1.05
Web Developers	1,671	4,296	4%	4%	$34.80	$39.15	0.65%	-0.54%	$0.23	-$0.21
Training and Development Specialists	970	4,753	0%	0%	$30.29	$32.13	0.00%	0.00%	$0.00	$0.00
Statisticians	187	1,096	3%	3%	$41.06	$47.31	0.49%	-0.34%	$0.20	-$0.16
Bottom 5										
Interior Designers	213	485	0%	0%	$24.32	$28.71	0%	0%	$0.00	$0.00
Paralegals and Legal Assistants	167	769	0%	0%	$26.69	$27.61	0%	0%	$0.00	$0.00
Transportation, Storage, and Distribution Managers	297	1,643	0%	0%	$45.77	$50.46	0%	0%	$0.00	$0.00
Purchasing Agents, Except Wholesale, Retail, and Farm Products	1,025	4,542	0%	0%	$32.08	$34.37	0%	0%	$0.00	$0.00
Industrial Engineering Technicians	591	3,565	0%	0%	$26.85	$28.63	0%	0%	$0.00	$0.00

Cloud Computing	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Software Developers, Applications	13,114	56,083	10%	28%	$51.93	$55.94	3.24%	-1.02%	$1.68	-$0.57
Network and Computer Systems Administrators	4,998	16,116	5%	15%	$40.56	$43.32	1.76%	-0.55%	$0.71	-$0.24
Web Developers	1,671	4,296	4%	14%	$34.80	$39.15	1.44%	-0.53%	$0.50	-$0.21
Database Administrators	610	1,511	1%	12%	$41.17	$46.77	0.42%	-0.43%	$0.17	-$0.20
Computer and Information Systems Managers	2,336	10,986	4%	8%	$69.63	$77.27	1.14%	-0.29%	$0.79	-$0.23
Bottom 5										
Paralegals and Legal Assistants	167	769	0%	0%	$26.69	$27.61	0%	0%	$0.00	$0.00
Occupational Health and Safety Specialists	965	938	0%	0%	$34.87	$37.21	0%	0%	$0.00	$0.00
Chemists	373	1,855	0%	0%	$38.01	$41.00	0%	0%	$0.00	$0.00
Medical and Clinical Laboratory Technicians	421	2,815	0%	0%	$19.30	$31.47	0%	0%	$0.00	$0.00
Registered Nurses	11,101	26,025	0%	0%	$32.45	$36.44	0%	0%	$0.00	$0.00

Programming	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Web Developers	1,671	4,296	72%	64%	$34.80	$39.15	-6.57%	-6.82%	-$2.29	-$2.67
Software Developers, Applications	13,114	56,083	65%	61%	$51.93	$55.94	-5.97%	-6.46%	-$3.10	-$3.61
Computer Programmers	425	1,674	53%	47%	$40.67	$44.72	-4.91%	-5.08%	-$2.00	-$2.27
Computer Systems Analysts	4,722	13,650	32%	34%	$43.03	$46.17	-2.99%	-3.67%	-$1.29	-$1.69
Management Analysts	5,717	25,543	15%	21%	$44.67	$45.36	-1.44%	-2.27%	-$0.64	-$1.03
Bottom 5										
Registered Nurses	11,101	26,025	0%	0%	$32.45	$36.44	0%	0%	$0.00	$0.00
Health Technologists and Technicians, All Other	701	6,130	0%	0%	$21.51	$22.71	0%	0%	$0.00	$0.00
Property, Real Estate, and Community Association Managers	188	2,065	0%	0%	$35.87	$37.05	0%	0%	$0.00	$0.00
Medical and Clinical Laboratory Technicians	421	2,815	0%	0%	$19.30	$31.47	0%	0%	$0.00	$0.00
Personal Financial Advisors	948	8,018	0%	0%	$53.95	$57.26	0%	0%	$0.00	-$0.01

Machine Learning	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Computer and Information Research Scientists	566	4,498	n/a	11%	$59.67	$62.98	n/a	6.24%	n/a	$3.93
Software Developers, Applications	13,114	56,083	n/a	1%	$51.93	$55.94	n/a	0.39%	n/a	$0.22
Software Developers, Systems Software	2,303	9,870	n/a	1%	$53.29	$56.04	n/a	0.38%	n/a	$0.21
Statisticians	187	1,096	n/a	0.2%	$41.06	$47.31	n/a	0.10%	n/a	$0.05
Computer Programmers	425	1,674	n/a	0.3%	$40.67	$44.72	n/a	0.17%	n/a	$0.08
Bottom 5										
Mental Health and Substance Abuse Social Workers	594	2,124	n/a	0%	$21.19	$23.34	n/a	0.00%	n/a	$0.00
Registered Nurses	11,101	26,025	n/a	0%	$32.45	$36.44	n/a	0.00%	n/a	$0.00
Health Technologists and Technicians, All Other	701	6,130	n/a	0%	$21.51	$22.71	n/a	0.00%	n/a	$0.00
Property, Real Estate, and Community Association Managers	188	2,065	n/a	0%	$35.87	$37.05	n/a	0.00%	n/a	$0.00
Human Resources Managers	1,267	5,771	n/a	0%	$56.74	$63.15	n/a	0.00%	n/a	$0.00

Research	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Statisticians	187	1,096	79%	68%	$41.06	$47.31	3.52%	4.11%	$1.45	$1.95
Computer and Information Research Scientists	566	4,498	69%	76%	$59.67	$62.98	3.10%	4.59%	$1.85	$2.89
Natural Sciences Managers	656	3,889	75%	68%	$69.32	$76.56	3.37%	4.08%	$2.34	$3.12
Medical Scientists, Except Epidemiologists	1,016	5,944	54%	59%	$46.52	$50.48	2.42%	3.56%	$1.13	$1.80
Sales Managers	3,444	15,811	12%	9%	$63.68	$70.25	0.54%	0.52%	$0.35	$0.37
Bottom 5										
Health Technologists and Technicians, All Other	701	6,130	0%	0%	$21.51	$22.71	0.00%	0.01%	$0.00	$0.00
Registered Nurses	11,101	26,025	2%	2%	$32.45	$36.44	0.11%	0.14%	$0.04	$0.05
Pharmacists	265	3,424	6%	3%	$56.86	$59.98	0.27%	0.19%	$0.15	$0.12
Construction Managers	254	1,719	6%	4%	$46.80	$51.13	0.24%	0.21%	$0.11	$0.11
Personal Financial Advisors	948	8,018	5%	5%	$53.95	$57.26	0.23%	0.28%	$0.12	$0.16

Maths	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Chemical Technicians	129	622	7%	4%	$22.97	$25.28	0.62%	1.01%	$0.14	$0.26
Civil Engineers	108	1,224	6%	1%	$42.62	$45.44	0.58%	0.25%	$0.25	$0.11
Computer and Information Research Scientists	566	4,498	3%	2%	$59.67	$62.98	0.25%	0.52%	$0.15	$0.33
Actuaries	215	1,147	4%	2%	$53.52	$56.63	0.33%	0.48%	$0.18	$0.27
Electrical and Electronics Engineering Technicians	724	2,771	1%	2%	$29.53	$32.05	0.07%	0.56%	$0.02	$0.18
Bottom 5										
Health Technologists and Technicians, All Other	701	6,130	0%	0%	$21.51	$22.71	0.00%	0.01%	$0.00	$0.00
Public Relations and Fundraising Managers	314	1,995	0%	0%	$59.53	$67.40	0.00%	0.00%	$0.00	$0.00
Licensed Practical and Licensed Vocational Nurses	302	1,260	0%	0%	$20.29	$23.09	0.00%	0.00%	$0.00	$0.00
Lawyers	477	2,636	0%	0%	$67.76	$72.73	0.00%	0.00%	$0.00	$0.00
Personal Financial Advisors	948	8,018	0%	0%	$53.95	$57.26	0.03%	0.02%	$0.02	$0.01

Analytical	Overall count by occupation		Skills share by occupation		Hourly wage		Wage premium (%)		Wage premium ($)	
Top 5	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1	2014-2015	2018-2020 Q1
Accountants and Auditors	3,434	13,285	88%	89%	$36.20	$38.13	-1.22%	0.18%	-$0.44	$0.07
Financial Analysts	3,075	11,400	81%	85%	$46.16	$47.52	-1.13%	0.17%	-$0.52	$0.08
Financial Examiners	387	2,272	63%	74%	$44.63	$41.82	-0.87%	0.15%	-$0.39	$0.06
Actuaries	215	1,147	67%	76%	$53.52	$56.63	-0.94%	0.15%	-$0.50	$0.09
Management Analysts	5,717	25,543	62%	69%	$44.67	$45.36	-0.86%	0.14%	-$0.38	$0.06
Bottom 5										
Health Technologists and Technicians, All Other	701	6,130	0%	0%	$21.51	$22.71	0.00%	0.00%	$0.00	$0.00
Surgical Technologists	577	1,187	1%	1%	$20.97	$23.64	-0.01%	0.00%	$0.00	$0.00
Pharmacists	265	3,424	8%	6%	$56.86	$59.98	-0.11%	0.01%	-$0.06	$0.01
Registered Nurses	11,101	26,025	4%	5%	$32.45	$36.44	-0.05%	0.01%	-$0.02	$0.00
Medical and Clinical Laboratory Technicians	421	2,815	4%	5%	$19.30	$31.47	-0.05%	0.01%	-$0.01	$0.00

Notes: The table displays the estimated wage premium by occupation. For each skills group we select the top 5 and bottom 5 occupations as measured by the share within occupation. We only select occupations that have a sizeable number of total observations across the years we study and whose occupation titles are familiar. The count shows the number of observations in each occupation group. The share of skills per occupation is measured for the years 2014 to 2020 Q1. The wage premium is measured as the multiplication of the coefficient on the skills group times the skills share in the occupation.

References

Abdi, Hervé, and Williams, Lynne J., "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2 (2010), 433–459.

Alekseeva, Liudmila, Azar, José, Giné, Mireia, Samila, Sampsa, and Taska, Bledi, "The demand for AI skills in the labor market," *Labour Economics*, 71 (2021), 102002 (Elsevier B.V.).

Allen, Jim, Belfi, Barbara, and Borghans, Lex, "Is There a Rise in the Importance of Socioemotional Skills in the Labor Market? Evidence From a Trend Study Among College Graduates," *Frontiers in Psychology*, 11 (2020).

Atalay, Enghin, Phongthiengtham, Phai, Sotelo, Sebastian, and Tannenbaum, Daniel, "The Evolution of Work in the United States," *American Economic Journal: Applied Economics*, 12 (2020), 1–34.

Bartholomew, David J, Steele, Fiona, Moustaki, Irini, and Galbraith, Jane, "Principal Component Analysis," in *Analysis of Multivariate Social Science Data* (2011).

Berger, Thor, and Frey, Benedikt, "Digitalization, jobs, and convergence in Europe: strategies for closing the skills gap," *European Commission - DG Internal Market, Industry, Entrepreneurship and SMEs* (2016).

Blair, Peter Q.., and Deming, David, "Structural Increases in Skill Demand after the Great Recession," *AEA Papers and Proceedings*, 110 (2020), 362–365.

Calanca, Federica, Sayfullina, Luiza, Minkus, Lara, Wagner, Claudia, and Malmi, Eric, "Responsible team players wanted: an analysis of soft skill requirements in job advertisements," *EPJ Data Science*, 8 (2019).

Campello, Murillo, Kankanhalli, Gaurav, and Muthukrishnan, Pradeep, "Corporate Hiring under COVID-19: Financial Constraints and the Nature of New Jobs," *Journal of Financial and Quantitative Analysis (JFQA), Forthcoming*, (2021).

Carmeli, Abraham, Reiter-Palmon, Roni, and Ziv, Enbal, "Inclusive Leadership and Employee Involvement in Creative Tasks in the Workplace: The Mediating Role of Psychological Safety," *Psychology Faculty Publications*, 30 (2010).

Carnevale, Anthony, Jayasundera, Tamara, and Repnikov, Dmitri, "Understanding Online Job Ads Data," *Georgetown University Centre on Education and the Workforce* (2014).

Cerioli, Maura, Leotta, Maurizio, and Ricca, Filippo, "What 5 Million Job Advertisements Tell Us about Testing: a Preliminary Empirical Investigation," *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, (2020), 1586–1594 (ACM).

Collischon, Matthias, "The Returns to Personality Traits Across the Wage Distribution," *LABOUR*, 34 (2020), 48–79 (Blackwell Publishing Ltd).

Deming, David, "The Growing Importance of Decision-Making on the Job," *NBER Working paper series*, w28733 (2021).

Deming, David J., "The Growing Importance of Social Skills in the Workplace," *The Quarterly Journal of Economics*, 132 (2017), 1593–1640.

Deming, David J, and Noray, Kadeem, "STEM careers and the changing skills requirements of work," *The Quarterly Journal of Economics*, 135 (2020), 1965–2005.

Deming, David, and Kahn, Lisa B, "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, 36 (2018), S337–S369.

Devlin, Jacob, Chang, Ming Wei, Lee, Kenton, and Toutanova, Kristina, "BERT: Pre-training of deep bidirectional transformers for language understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1 (2019), 4171–4186.

Dingel, Jonathan I., and Neiman, Brent, "How many jobs can be done at home?," *Journal of Public Economics*, 189 (2020), 104235 (Elsevier B.V.).

Dondi, Marco, Klier, Julia, Panier, Frederic, and Schubert, Jörg, "Defining the skills citizens will need in the future world of work," *McKinsey*, <https://mck.co/3b1wULK> (2021).

Edin, Per Anders, Fredriksson, Peter, Nybom, Martin, and Öckert, Björn, "The Rising Return to Noncognitive Skill," *American Economic Journal: Applied Economics*, 14 (2022), 78–100.

Faberman, Jason, and Kudlyak, Marianna, "What does online job search tell us about the labor market?," *Economic Perspectives*, 40 (2016), 1–15.

Forsythe, Eliza, Kahn, Lisa, Lange, Fabian, and Wiczer, David, "Labor Demand in the time of COVID-19: Evidence from vacancy postings and UI claims," *Nber Working Paper Series*, w27061 (2020) (Cambridge, MA).

Frey, Carl Benedikt, and Osborne, Michael A., "The future of employment: How susceptible are jobs to computerisation?," *Technological Forecasting and Social Change*, 114 (2017), 254–280 (Elsevier Inc.).

Fuller, Joseph, Langer, Christina, Nitschke, Julia, O'Kane, Layla, Sigelman, Matt, and Taska, Bledi, "The Emerging Degree Reset," *White Paper, Burning Glass Institute* (2022).

Girotra, Karan, Terwiesch, Christian, and Ulrich, Karl T, "Idea Generation and the Quality of the Best Idea," *Management Science*, 56 (2010), 591–605.

Goos, Maarten, Arntz, Melanie, Zierahn, Ulrich, Gregory, Terry, Carretero Gómez, Stephanie, González Vázquez, Ignacio, and Jonkers, Koen, "The impact of Technological innovation on the Future of Work," *JRC Working Papers Series on Labour, Education and Technology*, No. 2019/03 (2019).

Heckman, James J, Pinto, Rodrigo, Savelyev, Peter A, Blair, Clancy, Benjamin, Dan, Browning, Martin, Cattan, Sarah, Dodge, Kenneth, Duckworth, Angela, Finklestein, Amy, Gensowski, Miriam, Gentzkow, Matt, Grogger, Jeff, Kamenica, Emir, Meghir, Costas, Pischke, Jörn-Steffen, Raval, Devesh, Roberts, Brent, Sanandaji, Tino, Schweinhart, Larry, Waxman, Sandra, Williams, Ben, Yi, Junjian, Hansman, Christopher, Tan Teng Kok, Kegon, Ju Lee, Min, Lin, Xiliang, Pei, Yun, and Stosic, Ivana, "Web Appendix to Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, (2012).

Heineck, Guido, and Anger, Silke, "The returns to cognitive abilities and personality traits in Germany," *Labour Economics*, 17 (2010), 535–546 (Elsevier B.V.).

Hershbein, Brad, and Kahn, Lisa B, "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings," *American Economic Review*, 108 (2018), 1737–1772.

Ibarra, Herminia, and Hansen, Morten T., "Are you a collaborative leader?," *Harvard Business Review*, 3 (2011).

Jollife, Ian T., and Cadima, Jorge, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374 (2016) (The Royal Society Publishing).

Josten, Cecily, and Lordan, Grace, "Robots at Work: Automatable and non-automatable jobs," in *Handbook of Labor, Human Resources and Population Economics* (Springer International Publishing, 2020).

——, "The Accelerated Value of Social Skills in Knowledge Work and the COVID-19 Pandemic," *LSE Public Policy Review*, 1 (2021), 1–10.

——, "Automation and the changing nature of work," *PLOS ONE*, 17 (2022), e0266326 (Public Library of Science).

LinkUp, "Job Market Data," (2022).

Lordan, Grace, "Robots at work - A report on automatable and non-automatable employment shares in Europe," *European Commission - Directorate-General for Employment, Social Affairs and Inclusion* (2018).
Lordan, Grace, and Pischke, Jörn Steffen, "Does Rosie Like Riveting? Male and Female Occupational Choices - Appendix," *Economica*, 89 (2022), 110–130.
Manyika, James, Lund, Susan, Chui, Michael, Bughin, James, Woetzel, Jonathan, Batra, Parul, Ko, Ryan, and Sanghvi, Saurabh, "Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation," *McKinsey Global Institute* (2017).
Modestino, Alicia Sasser, Shoag, Daniel, and Balance, Joshua, "Upskilling: Do employers demand greater skill when workers are plentiful?," *Review of Economics and Statistics*, 102 (2020), 793–805.
MSCI, "MSCI World Index (USD)," <https://www.msci.com/documents/10199/178e6643-6ae6-47b9-82be-e1fc565ededb> (2022) (Jan. 12, 2022).
Nembhard, Ingrid M, and Edmondson, Amy C, "Making It Safe: The Effects of Leader Inclusiveness and Professional Status on Psychological Safety and Improvement Efforts in Health Care Teams," *Journal of Organizational Behavior*, 27 (2006), 941–966.
OECD, "An Assessment of the Impact of COVID-19 on Job and Skills Demand Using Online Job Vacancy Data," *OECD*, (2021), 1–19.
Samek, Lea, Squicciarini, Mariagrazia, and Cammeraat, Emile, "The human capital behind AI," *OECD Science, Technology and Innovation Policy Papers* (2021).
Schwab, Klaus, "The Fourth Industrial Revolution- What It Means and How to Respond," *Foreign Affairs*, (2015) (2015).
Shore, Lynn M., and Chung, Beth G., "Inclusive Leadership: How Leaders Sustain or Discourage Work Group Inclusion," *Group and Organization Management*, 47 (2021).
Squicciarini, Mariagrazia, and Nachtigall, Heike, "Demand for AI skills in jobs: Evidence from online job postings," *OECD Science, Technology and Innovation Policy Papers*, 03 (2021).
Tabachnick, Barbara, and Fidell, Linda, "Using Multivariate Statistics," (Pearson Education, 2018).
U.S. Department of Labor, "Bureau of Labor Statistics," *Occupational Employment Statistics* (2022).
Verma, Amit, Lamsal, Kamal, and Verma, Payal, "An investigation of skill requirements in artificial intelligence and machine learning job advertisements," *Industry and Higher Education*, 36 (2022), 63–73.
Weinberger, Catherine, "The Increasing Complementarity between Cognitive and Social Skills," *The Review of Economics and Statistics*, 96 (2014), 849–861.
Ziegler, Lennart, "Skill Demand and Wages. Evidence from Linked Vacancy Data," *IZA Discussion Paper Series*, No. 14511 (2021).