Digital skills in emerging economies: Evidence from Tunisian online job postings

Teo Firpo[ab*], Julia Baumann[a], Timm Teubner[c], Anastasia Danilov[ab]

[a] Humboldt-Universität zu Berlin, School of Business and Economics, Unter den Linden 6, 10099 Berlin, Germany

[b] Einstein Center Digital Future, Wilhelmstraße 67, 10117 Berlin, Germany

[c] Trust in Digital Services, Technische Universität Berlin, Einsteinufer 17, 10587, Berlin, Germany

[*] Corresponding author: teo.firpo@hu-berlin.de

Abstract

The ICT revolution has had wide ranging impacts across the world. A growing literature has focused on how these technological changes have affected labor markets, both in advanced and emerging economies. Our paper contributes to this research by studying the demand for digital skills in an emerging economy. We collect a novel dataset of online job postings from Tunisia and use their text content to identify digital skills. We find that 41.7% of all ads mention at least one digital skill, with the most common category of digital skills being "programming." Ads mentioning digital skills are also generally more likely to mention a series of non-digital skills, and to be located in the capital region. Our study provides the first benchmark measures of digital and non-digital skills based on online job postings from an emerging economy.

Keywords: digital skills; information and communication technologies; digitization; employment.

1 Introduction

The rise of Information and Communication Technologies (ICT) has radically altered how firms operate (Bresnahan et al., 2002) and perform (Ahmad et al., 2004; Atalay & Sarada, 2020). As a result, the skills required of workers have also changed (Brynjolfsson & McAfee, 2011; Michaels et al., 2014). Understanding these changes is crucial for policymakers tasked with preparing their populations to navigate modern labor markets.

A large literature has analyzed the connection between ICT and the demand for skilled work (for overviews see Spitz-Oener, 2006; Acemoglu and Autor, 2011). While much of it has focused on developed countries (e.g., Gallipoli and Makridis, 2018; Frank et al., 2019; Cirillo et al., 2021), increasing attention has been paid to the context of emerging and developing economies (see Vivarelli, 2012; Oberdabernig, 2016; Martins-Neto et al., 2021 for overviews).

However, researchers and policymakers interested in the latter face two challenges. First, much of this research agenda relies on country surveys, which are not available for many developing and emerging economies (Lewandowski et al., 2020). For instance, the STEP surveys, which aim to provide evidence on skills in low- and middle-income countries equivalent to the PIAAC surveys in OECD economies, only cover 17 countries. A common workaround is to impute measures of the task-content of occupations from surveys of developed economies; however, recent research has shown this approach is not necessarily representative for a range of skills (Lo Bello et al., 2019). Second, even when survey data is available, it is often not up-to-date or detailed enough to inform policies on specific skills – such as upskilling or reskilling programs. This has led to calls for the collection of more granular data (Caunedo et al., 2021).

In this paper we address both these concerns by providing empirical evidence on the demand for digital skills in an emerging economy. We create a new dataset of skills in Tunisia by crawling job postings and using text mining techniques to identify and study the skills embedded within them. Online job ads are a unique data source, as they allow researchers to directly observe the skill requirements posted by employers, often covering a large percentage of all new roles. As a result, they have been used study the changing composition of skills and tasks (Atalay et al., 2020; Deming & Kahn, 2018; Hershbein & Kahn, 2018), often with regards to particular industries or technologies (Papoutsoglou et al., 2022; Pejic-Bach et al., 2020), mostly focusing on advanced economies. We contribute to the literature on skills by extending their use to an emerging economy, directly addressing calls to do so in prior research (Martins-Neto et al., 2021).

To do so, we adapt an approach previously used to study skills in the US and UK (respectively, Atalay et al., 2018; Djumalieva & Sleeman, 2018b). Our approach relies on 'text mining', a technique to analyze large unstructured textual datasets (Gandomi & Haider, 2015) with a large number of practical applications (Kumar et al., 2021).

Our paper contributes key labor market insights on Tunisia in particular, with a special focus on a set of skills – digital competences– at the center of the country's policy discussions. Tunisia, like many emerging economies, currently lacks up-to-date data on the skills requested by employers from workers. The North African country is not covered by international skills' surveys such as the STEP program. The most recent nationally representative labor market survey dates back to 2014 and does not present disaggregated data at the skills level (OAMDI, 2016). At the same time, the country's policymakers have attempted to position Tunisia as a digital leader in the region

(MTCEN, 2016), and an IT offshoring destination.[1] A key goal of this strategy is to ensure Tunisians are equipped with the digital skills demanded by employers through reskilling programs (The World Bank & MTCEN, 2020). Our analysis provides direct, up-to-date evidence on what digital skills employers look for in Tunisia. We also study their complementarity with non-digital skills, and geographic distribution. In doing so, our research contributes key insights for policymakers, particularly those aiming to design reskilling and retraining programs.

The empirical evidence we present is especially relevant given Tunisia's current labor market conditions. The country faces high graduate unemployment (30%, National Institute of Statistics, 2021). While there are likely multiple causes behind this high unemployment rate, a potentially significant factor is the existence of a mismatch between the skills demanded by employers and those supplied by Tunisian job seekers (Assaad et al., 2018). This explanation is supported by evidence from the World Bank Enterprise Survey, which finds that 34.8% of Tunisian firms identify an inadequately educated workforce as a major constraint – a substantially higher rate than across all countries (20.5%; The World Bank, 2020). Digital skills in particular have been identified as an area that potentially suffers from this labor market mismatch throughout North Africa (AUC & OECD, 2021). Indeed, in a nation-wide survey, Tunisian firms named technical and digital skills as by far the most important competency when filling job posts (IACE, 2019). Our study provides direct evidence on the absolute number of roles requiring digital skills across a range of categories. This information can be used to guide labor market policies in this area.

In sum, our study presents results from a new dataset on digital skills requested in Tunisia. In doing so, we illustrate how an increasingly popular data source – job postings – can be used in the context of emerging economies, with a view to inform labor market policies. Finally, we assess the

[1] Through the 'Smart Tunisia' initiative (for more information, see http://www.smarttunisia.tn/).

advantages and limitations of this approach, providing insights for future studies. The remainder of this paper is structured as follows: in Section II, we describe the data collection and method used to extract skills from the text of job ads in detail. Section III presents the results of the data analysis. Section IV discusses the value and limitations of our approach. Section V concludes.

II. Data and method

In this section, we explain how our dataset was constructed from the raw text of Tunisian online job ads. We start by describing how the data was collected and cleaned. We then outline the procedure used to identify a list of digital skills from the data. We continue with a description of how we classify the digital skills into eight broad categories. Finally, we expand our analysis by classifying our data into nine additional categories of non-digital skills.

Collecting and cleaning the data

Our dataset consists of ads collected from the largest job portal website offering Tunisia-based roles.[2] Over the period from May 2020 to May 2021, we collected 280,865 raw ads using web crawling techniques. The data includes the title of the job role being advertised, as well as the content of the job posting. Many ads were active over a prolonged period and so were crawled more than once. Thus, we remove duplicates, i.e., ads with the same URL. We also deduplicate postings that contain identical titles and job descriptions.[3] This procedure leaves us with 63,719 unique job postings. Ads in our dataset may appear in different languages, primarily French,

[2] The job portal 'OptionCarriere.tn' was the largest in terms of available job ads, as of May 2020, when data collection began. It aggregates job postings from several smaller portals in Tunisia. We selected this website because of the number of job postings for Tunisia available on it, and to avoid duplication issues (as ads reposted on several portals can vary in non-systematic ways, which could lead to double counting).

[3] There are different reasons why these ads are duplicated, e.g., employers might decide to take down the job offer and then repost it. Our final sample is therefore a conservative estimate of the absolute number of vacancies.

Arabic and English. We restrict our attention to French-language ads which constitute the bulk of our data (92% of the total). This leaves us with 59,254 unique postings.

Next, we clean and process the text of the ads. We lowercase all text, replace all accented letters with equivalents (e.g., 'é' with 'e'), and remove punctuation. We also remove non-essential words (also known as 'stop words'), such as pronouns and prepositions, as they can interfere with the textual analysis.[4]

Identifying digital skills

Our second step is to identify and extract digital skills from our clean text data. There are many definitions of digital skills, depending on context and purpose (Djumalieva & Sleeman, 2018b; van Laar et al., 2017). For instance, the European Commission's ESCO (European Skills, Competences, Qualifications and Occupations) classification defines digital competencies as the "ability to use information and communication technologies effectively to achieve work objectives" (European Commission, 2017), while the OECD defines digital skills as "the capacity to use ICT devices and applications to access and manage information and solve problems" (OECD, 2015). Relatedly, for UNESCO digital skills are "a range of abilities to use digital devices, communication applications, and networks to access and manage information" (UNESCO, 2018). We employ an approach developed by Djumalieva and Sleeman (2018b), which allows us to identify digital skills in our data that match these definitions, rather than using a pre-existing list of skills that might be too restrictive. This process involves three steps.

[4] This is especially the case in French, where words will often be connected to their definite article, e.g., "l'application" ("the application"). Failing to remove the stop words (in this case, the article "la", shortened "l'"), can lead to terms being excluded from the textual analysis, as they are parsed as different words.

In a first step, we manually identify a small number of digital skills (either single words or bigrams). To do so, we select five ads from our dataset for a range of jobs roles explicitly related to digital tasks (web development, web marketing, software engineering) and extract 59 digital skills (the text of the ads and the skills extracted are presented in Figure A.1 in the Appendix). These digital skills range from software or programming languages (‘Python’, ‘PHP’, ‘HTML’), digital marketing skills (‘SEO’ (Search Engine Optimization), the process of optimizing traffic to websites) to broader digital skills (‘e-commerce’). These 59 skills form our starting set from which we identify a larger number of skills.

In a second step we employ a word embeddings model to identify synonymous skills. This is a machine learning model that turns words from a corpus of text into vectors; these vectors can be used to find other words that share a number of characteristics, such as appearing in similar contexts.[5] We train our model on the text of all of our ads and use it to identify skills similar to our starting set of 59 skills; in each case, the resulting skills are manually checked against the ESCO, UNESCO and OECD definitions of skills given above, retaining the skills that fall under at least one definition. Through this step, we identify an additional 544 keywords, for a total of 603 digital skills.

Grouping digital skills

We next group our digital skills into categories. We do this for three reasons. Firstly, classifying a large number of skills into categories can make the analysis more understandable (Atalay et al., 2020; Deming & Kahn, 2018; Deming & Noray, 2020). Secondly, we use our categorization to separate skills that are specific to certain job types (such as web development or graphic design)

[5] A similar approach is used by Atalay et al. (2018) to expand the set of tasks they extract from newspaper job ads.

and skills that are more transversal and appear in a wide variety of jobs (such as basic office software); this distinction allows us to classify ads that contained more than general digital skills. Finally, we use our categories of digital skills to explore their relation to non-digital skills, using a taxonomy derived from Deming and Noray (2020).

We start by categorizing our 603 digital skills into broad categories. We adapt a taxonomy developed by Djumalieva and Sleeman (2018a). The advantage of this taxonomy is that it is also derived from job postings data and thus the keywords in each category closely match our digital skills. The disadvantage is that the original taxonomy includes a large number of categories, which can reduce its analytical value; we therefore further consolidate the clusters that match our digital skills into eight broad categories.

Table 1 provides a breakdown of the eight categories. The first category ("programming") includes skills related to web, app, and software development; it comprises programming languages, frameworks, and software. "Technical support and IT" brings together a series of skills relating to system administration, IT support and security, and networks. The category "office software" covers typical IT tools, such as the Microsoft Office Suite, as well as references to common office IT skills ("emailing"). "Graphic design and digital content creation" combines graphic design-related tools (such as "Photoshop") with skills associated with the creation of content for the web (such as blogs or newsletters). Finally, we create four additional categories for "data science and data engineering," "digital marketing," "data and business analysis" and "machine learning and AI." A small number of skills that do not fall under these eight categories are left unclassified as "other."[6]

[6] These are relatively broad digital skills, such as "digital skills", or narrower, uncategorized skills, such as "IoT" (Internet of Things).

As in the original Djumalieva and Sleeman (2018a) taxonomy, the categories are not mutually exclusive. The reason is that there are skills that span a variety of uses. For instance, "SQL" – a computer language used to manage databases – is related to both "data engineering" and "programming," as well as "technical support and IT."

[Table 1 about here.]

Additionally, we identify nine non-digital skills categories in our data to examine their co-occurrence with digital skills. We adopt the taxonomy used by Deming and Noray (2020), focusing on their nine non-digital skills categories.[7] The first skills category is "social", which includes references to "communication," "teamwork" and "collaboration." The next category is "cognitive" skills, which is modelled after "nonroutine analytical" tasks described in Autor et al. (2003b), and includes keywords such as "thinking," "researching" and "analyzing." A third category, "character," includes a number of "soft skills" and personality traits, such as "time management" and "attention to details." The "management" category includes skills related to project and people management (e.g., "supervising," "mentoring," "staff development"). Further categories include skills related to "creativity," "writing" and "finance." The "business systems" category includes techniques used in business processes, such as "Six Sigma" or "KPIs." Finally, "customer service" includes keywords related to interacting with clients (such as "customer" or "sales"). (Details of all keywords used to identify each non-digital category are presented in Table A.1 in the Appendix). We translate the keywords in each category into French and identify alternative spellings in our data. Following Atalay et al. (2018), we also use our word embeddings model to

[7] We do not use their five remaining categories which cover digital skills, as they are already covered by our (more precise) taxonomy. In fact, these five categories map well onto our eight digital skills categories: "office software", "technical support", "data analysis", "specialized software" (which directly relates to our "programming" category, as well as software skills in our "data science and data engineering" and "graphic design and digital content creation" categories) and "machine learning and AI".

uncover synonymous skills in our data. In total, we identify 104 keywords and phrases across the nine non-digital skills categories.

In the next section, we turn to our analysis of the distribution of digital and non-digital skills in the job postings in our dataset.

III. Results

The aim of this paper is to understand the demand for digital skills in the labor market of an emerging economy. We start by exploring the number of ads that mention any digital skills. We then study the distribution of ads by digital skills category, and the relationship between digital and non-digital skills. Finally, we look at the geographic distribution of ads mentioning digital skills.

We begin by coding ads by the number of digital skills they mention. We find that 41.7% of all ads include at least one digital skill. This number is significantly lower than the fraction of digital skills in job postings in the UK (82%; Nania et al., 2019) and Germany (79%; O'Kane et al., 2020). Conditional on mentioning any digital skills, three quarters mention at least two skills with an average of 6.9 digital skills. Figure 1 shows the distribution of mentions. As discussed in Section II.C, two categories include more general or broad skills ("office software" – which combines a number of common office IT tools such as the Microsoft Suite, and "other" – which includes broad terms such as "digital skills"). If we exclude these two categories, we find that 27.1% of all ads mention at least one skill from the remaining categories.

[Figure 1 about here.]

We next discuss the demand for digital skills by category. We assign a category to an ad if it mentions at least one of the skills in that category. Therefore, an ad can be assigned multiple categories of digital and non-digital skills. Table 2 reports the count of ads in each category, as well as the rate as a percentage of all ads (59,254). The category "programming" has the highest share in our data, being mentioned in 16.1% of all ads. The "office software" category appears in 9.1% of all ads.[8] For comparison, in their job posting data from the US, Deming and Noray (2020) find that office software skills were mentioned in 8.8% of all ads in the year 2007; by 2019, that number had increased to 12%.[9] Moreover, the "machine learning and AI" category appears in 1.3% of all ads. Again, this is similar to the 2007 figure in Deming and Noray (2020) – 1.8%, rising to 7.4% in 2019. While the other categories in Deming and Noray (2020) do not map as closely onto ours, a similar pattern emerges. In their data, "technical support" (which comprises a subset of our "technical support and IT" keywords) is present in 17.7% of all ads in 2007 (13.9% in 2019); while "data analysis" (a subset of our "data and business analysis" category) appears in 6.7% of all ads in 2007 (11.6% in 2019). Both figures for 2007 are closer to those in our Tunisian data (9.2% and 4.5% respectively). Overall, our results show that the demand for digital skills categories is closer (and often lower) to the US figures for 2007, rather than 2019, as presented by Deming and Noray (2020), pointing to a potential lag in the adoption of digital skills in Tunisia.

Nevertheless, the relative shares of different categories seem to correspond to existing evidence from the US. For instance, investigating a cluster of IT tasks with US data, Das et al. (2020) show that "SQL Databases and Programming," "Java" and "JavaScript & jQuery" have the highest shares of mentions, especially among high- and mid-wage occupations. "AI" and "Big Data," conversely, have the lowest shares. This corresponds to our findings for Tunisia when examining

[8] While the latter number might appear low, this could indicate that this type of skills is sufficiently common in certain job categories as to be assumed by employers (and therefore not mentioned in job postings), as has been previously suggested (Lamb et al., 2019).
[9] Tables A1 and A2 in the Online Appendix (Deming & Noray, 2020).

individual skills within our categories (as shown by the top ten most common individual skills in our data, presented in Table A.2 in the Appendix).[10]

[Table 2 about here.]

We then turn to the interaction between digital and non-digital skills. We investigate what non-digital skills tend to appear together with digital skills, exploring differences by category. We rely on nine non-digital skills categories used by Deming and Noray (2020) described in Section II.C. As with digital skills, we assign a non-digital category to an ad if it mentions at least one of the keywords in that category. We run separate Probit regressions with each non-digital skills category as the dependent variable, and all our digital skills categories as independent variables.

Table 3 presents the resulting marginal effects for each digital skills category; these correspond to an ad's change in probability of mentioning the relevant non-digital skills category conditional on mentioning all other digital skills categories. For example, an ad has a 10.1% higher probability of mentioning "social" skills, when it mentions "programming" skills compared to an ad that does not mention programming.

There are two takeaways from this table. The first is a positive association between the "social," "character," "cognitive" and "writing" categories and all digital skills categories. The marginal effects on these four categories are all positive, relatively large, and all statistically significant at the 1% level.[11] These results are in line with previous research highlighting the complementarity of 'soft skills' and cognitive or technical skills (Deming, 2017; Deming & Kahn, 2018). The connection between these creative and digital skills – sometimes called "Createch" – has been identified as an increasingly important skillset for the so-called "Creative Industries" (Bakhshi et

[10] In particular, "Javascript" appears in 3.7% of all ads, "Java" in 2.7%, "SQL" in 2.5% (all among the top ten most common individual skills), whereas "AI" and "Big Data" each appear in 0.3% of all ads.
[11] The only exception is the marginal effect of "machine learning and AI" on "character," which is positive but not significant.

al., 2019). Alternatively, it could be due to shorter ads including fewer keywords overall (and thus any type of skill). Indeed, the average length of ads mentioning at least one digital skill is greater than those that do not (1053.8 and 592.4 characters, respectively[12]). However, we observe a clear jump in the number of non-digital skills mentioned in ads with no digital skills (2.0 on average) vs with at least one digital skill (5.2) – as can be seen in Figure A.2 in the Appendix.[13]

[Table 3 about here.]

The second takeaway is that while most marginal effects are positive, the strength of the association between digital and non-digital skills varies by category. For instance, several digital skills categories have negative marginal effects on "finance" skills, while the relationship is positive for "office software" and "data and business analysis." "Graphic design and digital content creation" and "digital marketing" share positive marginal effects on "creativity" but differ in their associations with "customer service" skills, the former being negative (-0.081) and the latter positive (0.182). This could reflect the more customer-facing aspects of digital marketing roles.

Finally, we study the locations for which the job roles in our data are advertised. We find a concentration of online job ads in the capital, Tunis (Figure 2 below). The share of ads from Tunis (42.8%) is significantly higher than what we would expect, even accounting for the capital's share of the population (14%, National Institute of Statistics, 2014). This difference is even larger when we focus on job ads mentioning at least one digital skill (52.0%). A potential explanation would be differential rates of job creation by region. However, if we compare job ads in our data to other sources on the distribution of job vacancies, we find that Tunis does not stand out in these other datasets. The rate of offers for Tunis received by the national employment agency (ANETI) in

[12] The difference – tested with a t-test – is also statistically significant at the 1% level.

[13] The difference in the number of non-digital ads mentioned, when tested with a t-test, is statistically significant at the 1% level.

2019 was only 11.4% of all offers (National Institute of Statistics, 2019b). A survey of medium and large enterprises in the same year put the figure of vacancies in the capital at 22.9% (IACE, 2019). Table A.3 in the Appendix presents the data on the number of job ads and vacancies by region. Our results are in line with previous research for the US, which has shown that technological requirements mentioned in job offers increase with city size (Atalay et al., 2022).

[Figure 2 about here.]

IV. Discussion

Thus far we have quantified the demand for digital skills in Tunisia based on our dataset of job postings. In this section, we discuss the implications for the Tunisian labor market. We also appraise the value of job postings as a data source to inform policy, in particular with respect to the representativeness of our data.

Findings and their implications

Our first contribution is to document the prevalence of digital skills in a significant portion of job ads (41.7%). In absolute terms, we find that 24,718 out of 59,254 ads mention at least one digital skill. If we exclude our two broader categories ("office software," which includes the "Microsoft Office" suite, and "other," which includes general skills such as "computer skills"), the figure is 16,049. With the number of unemployed Tunisians standing at 676,600 at the end of 2020 (National Institute of Statistics, 2021), there is a clear mismatch between the demand for jobs and the supply of jobs in Tunisia. There are simply more unemployed Tunisians than jobs currently in supply for them as we and other sources[14] find. Nevertheless, it is important that the open positions

[14] A nationally representative survey of medium and large enterprises conducted by IACE estimated that, in 2019, there were 47,026 vacancies in the country (IACE, 2019). In the same year, the employment bureaus run by the Tunisian National Employment Agency (ANETI) received 87,701 job offers requests from employers across all sectors (National Institute of Statistics, 2019b).

can be filled by suitable applicants. Therefore, if there is a mismatch induced by unsuitable skill sets of applicants, it may be mitigated by promoting digital skills that are frequently demanded. Our study benchmarks the absolute number of such roles, providing an evidence base for policymakers designing programs to create digital-related jobs.[15]

Our second set of findings – on the *type* of digital skills mentioned in job ads – provides further policy insights. We find a significant number of ads mentioning relatively advanced digital skills. For instance, if we combine all ads mentioning skills related to the "programming," "data science and data engineering," "technical support and IT" and "machine learning and AI" categories, we find 11,727 ads in our data, or 19.8% of all ads. This is a considerable figure when compared to the total number of (formal) jobs in the ICT sector, which in 2019 stood at 31,264 (National Institute of Statistics, 2019a). While the ICT sector has been growing steadily in Tunisia, averaging 8.4% employment growth over the 2010-2019 period (National Institute of Statistics, 2019a), the relatively large number of job ads mentioning programming and other ICT-related skills could indicate a wider spread of these skills across sectors. These results – together with the relative importance of different digital skills categories – can provide guidance for the development of so-called 'upskilling' or 'reskilling' programs.[16]

We also find an association between digital and non-digital skills. These findings are in line with previous research, which highlights the connection between social and cognitive skills (Deming, 2017; Deming & Kahn, 2018). It follows that the Tunisian labor market may benefit from

[15] This is especially relevant given past government plans have established the goal of creating thousands of 'digital jobs.' For instance, the national strategy 'Digital Tunisia 2020,' created in 2016, set out the aim of creating 80,000 jobs by 2020 and 25,000 a year from then on (MTCEN, 2016).

[16] A number of such programs, often publicly funded, have been developed in Tunisia. Two relevant examples are the Tunisian National Employment Agency's "ANETI Twaken" program, and the German-Tunisian Chamber of Commerce's reskilling program "CORP," both of which include a focus on digital skills. More information can be found, respectively, at https://www.emploi.nat.tn/formations/Formation_en_ligne.html and http://www.corp.tn/le-corp/.

promoting digital skills together with social skills, a point already raised in recent policy debates in Tunisia (PWC et al., 2021). Our contribution is to provide a detailed breakdown of which soft skills are most requested alongside each digital category.

Finally, our findings point to a concentration of jobs requiring digital skills in the capital. This highlights the need for policymakers to adapt their strategies to bridge the gap between Tunis and other areas of the country.

Representativeness of the dataset

To contextualize our findings in the Tunisian labor market, it is important to consider the extent to which our dataset covers job ads in Tunisia. Naturally, not all job openings are advertised online, a concern raised by prior literature (Davis et al., 2013; Hershbein & Kahn, 2018). A nation-wide survey of medium and large Tunisian businesses in 2019 found that only 31.5% of firms filled non-managerial roles through ads (online and/or offline), against 44.3% choosing local employment agencies (IACE, 2019). In light of these facts, the absolute figures presented in this paper likely understate the total number of job ads requiring digital skills, as we cannot exclude that job offers not included in our dataset also mention digital skills. Despite this, we believe that our dataset is likely to capture a sizable share of digital jobs in Tunisia. Previous research has shown that online job ads tend to overrepresent high-skilled and STEM occupations (Carnevale et al., 2014). These more high-skilled occupations are generally associated with a higher demand for digital skills, as suggested for example by industry reports (Nania et al., 2019; O'Kane et al., 2020). Moreover, surveys of employers show that firms in the ICT sector are more likely to fill positions through job ads, compared to the overall enterprise population (IACE, 2019). This implies that our online job ads dataset is more representative for the digital skills labor market than it is for the total labor market.

Our dataset can be used to establish estimates for the absolute number of Tunisian jobs requiring digital skills, as well as their share of all job offers in the country. Since digital jobs are more likely to be advertised online, we believe that the dataset gives us a good understanding of the overall demand for digital skills in Tunisia; even with the limitation that we are not able to get a complete picture of the Tunisian labor market. This yields an improved understanding of the geographical distribution of digital jobs, the demand for different types of digital skills and the non-digital skills appearing alongside digital skills. Moreover, we have shown that online job ads data is a valuable data source for the Tunisian market. This paper therefore serves as a starting point for further analyses utilizing such and similar data in other North African countries. As we have shown, this type of data is already in popular use in advanced economies. With the above-mentioned caveats regarding representativeness, online job ads can likewise generate valuable insights for developing and emerging economies.

V. Conclusion

In this paper, we use a novel dataset of online job postings from Tunisia to study the demand for digital skills in the Tunisian labor market. We collect data from the largest portal in the country over the period from May 2020 to May 2021, obtaining 59,254 unique jobs ads. Following an approach developed by Djumalieva and Sleeman (2018a), we identify 603 digital skills in our data. We classify them in eight digital skills categories, ranging from “office software” to “machine learning and AI.” We find that 41.7% of all ads mentioned at least one digital skill. The digital skills category appearing in the highest number of job postings is “programming,” which is mentioned in 16.1% of all ads, and includes skills related to web, software and app development. Moreover, we examine the co-occurrence between our digital skills and nine non-digital skills categories used by Deming and Noray (2020). We find that ads mentioning digital skills also tend to mention non-digital skills at a higher rate than the other job postings, and in particular the categories for "cognitive," "social," "writing" and "character" skills.

While job postings provide a useful novel approach to studying the demand for digital skills, there are a few limitations. Because they likely only represent a portion of all vacancies, it is important to consider whether job ads posted online differ systematically from other vacancies. We find that ads in our dataset are more concentrated in the capital area compared to job requests received by the National Employment Agency and constitute a higher share of the total even after accounting for the larger population of the capital. Nevertheless, our dataset is helpful in providing lower bounds for the total number of vacancies requiring digital skills. Moreover, if job offers containing digital skills are indeed more likely to be posted online, our analysis of the associations between digital and non-digital skills are also likely to be more representative of digital jobs.

Future research could expand this analysis by matching data from job postings to further data sources. Two promising avenues, following Deming and Kahn (2018), include: matching data to firm-level indicators, which would allow for an estimation of the connection between skill requirements and firm performance; and matching job postings to data on employment and wage growth by location and sector. A further avenue for research on the role of digital skills in the Tunisian labor market would involve expanding the analysis to include the supply of these skills.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, *4* (PART B), 1043–1171. https://doi.org/10.1016/S0169-7218(11)02410-5

Ahmad, N., Schreyer, P., & Wölfl, A. (2004). ICT investment in OECD countries and its economic impacts. In *The Economic Impact of ICT: Measurement, Evidence and Implications* (Vol. 9789264026). https://doi.org/10.1787/9789264026780-5-en

Assaad, R., Ghazouani, S., & Krafft, C. (2018). The Composition of Labor Supply and Unemployment in Tunisia. In R. Assaad & M. Boughzala (Eds.), *The Tunisian Labor Market in an Era of Transition*. Oxford University Press.

Atalay, E., Phongthiengtham, P., Sotelo, S., & Tannenbaum, D. (2018). New technologies and the labor market. *Journal of Monetary Economics*, *97*, 48–67. https://doi.org/10.1016/j.jmoneco.2018.05.008

Atalay, E., Phongthiengtham, P., Sotelo, S., & Tannenbaum, D. (2020). The evolution of work in the United States. *American Economic Journal: Applied Economics*, *12*(2), 1–34.

Atalay, E., & Sarada. (2020). Firm Technology Upgrading Through Emerging Work. Working Paper.

Atalay, E., Sotelo, S., & Tannenbaum, D. (2022). *The Geography of Job Tasks*. Working Paper.

AUC, & OECD. (2021). *Africa's Development Dynamics 2021: Digital Transformation for Quality Jobs*.

Autor, D., Levy, F., & Murnane, R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, *118*(4), 1279–1333.

Bakhshi, H., Djumalieva, J., & Easton, E. (2019). *The Creative Digital Skills Revolution*. Nesta Report.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, *117*(1), 339–376. https://doi.org/10.1162/003355302753399526

Brynjolfsson, E., & McAfee, A. (2011). *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. Digital Frontier Press.

Carnevale, A. P., Jayasunder, T., & Repnikov, D. (2014). *Understanding Online Jobs Data: A Technical Report*. Georgetown University, Centre on Education and the Workforce.

Caunedo, J., Keller, E., & Shin, Y. (2021). *Technology and the Task Content of Jobs across the Development Spectrum* (No. 28681; NBER Working Paper Series). http://www.nber.org/papers/w28681

Cirillo, V., Evangelista, R., Guarascio, D., & Sostero, M. (2021). Digitalization, routineness and employment: An exploration on Italian task-based data. *Research Policy*, *50*(7), 104079. https://doi.org/10.1016/j.respol.2020.104079

Das, S., Steffen, S., Clarke, W., Brynjolfsson, E., & Fleming, M. (2020). Learning Occupational Task-Shares Dynamics for the Future of Work. *AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 36–42.

Davis, S. J., Faberman, R. J., & Haltiwanger, J. C. (2013). The Establishment-Level Behavior of Vacancies and Hiring. *Quarterly Journal of Economics*, *128*(2), 581–622. https://doi.org/10.1093/qje/qjt002

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10.1093/qje/qjx022

Deming, D. J., & Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, *36*(S1), S337–S369. https://doi.org/10.1086/694106

Deming, D. J., & Noray, K. (2020). Earnings dynamics, changing job skills, and stem careers. *Quarterly Journal of Economics*, *135*(4), 1965–2005. https://doi.org/10.1093/qje/qjaa021

Deming, D., & Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, *36*(S1), S337–S369. https://doi.org/10.1086/694106 https://doi.org/10.1086/694106

Djumalieva, J., & Sleeman, C. (2018a). An Open and Data-driven Taxonomy of Skills Extracted from Online Job Adverts. In C. Larsen, S. Rand, A. Schmid, & A. Dean (Eds.), *Developing Skills in a Changing World of Work* (Issue August, pp. 425–454). Nomos. https://doi.org/10.5771/9783957103154-425

Djumalieva, J., & Sleeman, C. (2018b). Which digital skills do you really need? Exploring employer demand for digital skills and occupation growth prospects. Nesta report. https://media.nesta.org.uk/documents/Which_digital_skills_do_you_really_need.pdf

European Commission. (2017). *ESCO Handbook: European, Skills, Competences, Qualifications and Occupations* (Issue September). Directorate-General for Employment, Social Affairs and Inclusion. https://ec.europa.eu/esco/portal/document/en/0a89839c-098d-4e34-846c-54cbd5684d24

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(14), 6531–6539. https://doi.org/10.1073/pnas.1900949116

Gallipoli, G., & Makridis, C. A. (2018). Structural transformation and the rise of information technology. *Journal of Monetary Economics*, *97*, 91–110. https://doi.org/10.1016/j.jmoneco.2018.05.005

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Hershbein, B., & Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, *108*(7), 1737–1772. https://doi.org/10.1257/aer.20161570

IACE. (2019). *Rapport National de l'Emploi*. https://iace.tn/rapport-national-sur-lemploi-2019/#:~:text=71%2C7%25%20pr%C3%A9voient%20la%20stagnation,Industrie%20%C2%BB%20et%20%C2%AB%20Services%20%C2%BB.

Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights*, *1*(1), 100008. https://doi.org/https://doi.org/10.1016/j.jjimei.2021.100008

Lamb, C., Vu, V., & Willoughby, R. (2019). *Digital, Defined: Understanding the demand for digital skills in Canada*. Brookfield Institute Report.

Lewandowski, P., Park, A., & Schotte, S. (2020). The Global Distribution of Routine and Non-Routine Work The Global Distribution of Routine and Non-Routine Work. In *IZA Discussion Paper* (Issue 13384).

Lo Bello, S., Sanchez Puerta, M. L., & Winkler, H. (2019). From Ghana to America: The Skill Content of Jobs and Economic Development. *World Bank Policy Research Working Paper;No. 8758* https://doi.org/10.1596/1813-9450-8758

Martins-Neto, A., Mathew, N., Mohnen, P., & Treibich, T. (2021). Is There Job Polarization in Developing Economies? A Review and Outlook. In *CESifo Working Paper* (No. 9444; CESifo Working Paper Series, Issue November). https://doi.org/10.2139/ssrn.3979349

Michaels, G., Natraj, A., & Van Reenen, J. V. (2014). Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years. *Review of Economics and Statistics*, *96*(1), 60–77. https://doi.org/10.1162/REST_a_00366

MTCEN. (2016). *Plan National Stratégique Tunisie Digitale 2020*. https://www.mtcen.gov.tn/index.php?id=14&L=594

Nania, J., Bonella, H., Restuccia, D., & Taska, B. (2019). *No Longer Optional: Employer Demand for Digital Skills* (Issue June). https://www.gov.uk/government/publications/current-and-future-demand-for-digital-skills-in-the-workplace

National Institute of Statistics. (2014). *Census 2014*.

National Institute of Statistics. (2019a). *Employed population and distribution of employed population by industry 2010-2019*.

National Institute of Statistics. (2019b). *Employment agency statistics, job offers received*.

National Institute of Statistics. (2019c). *Employment agency statistics, job offers received 2019*.

National Institute of Statistics. (2021). *Unemployment rate and higher education graduates unemployment rate 2006-2021*.

O'Kane, L., Narasimhan, R., Nania, J., & Bledi Taska. (2020). *Digitalization in the German Labor Market: Analyzing Demand for Digital Skills in Job Vacancies*. https://www.bertelsmann-stiftung.de/en/publications/publication/did/digitalization-in-the-german-labor-market-en

OAMDI. (2016). *Labor Market Panel Survey (LMPS) Version 2.0 of Licensed Data Files; TLMPS 2014*. Economic Research Forum (ERF). http://erf.org.eg/data-portal/

Oberdabernig, D. A. (2016). *Employment effects of innovation in developing countries : A summary* (2016/2; R4D WORKING PAPER).

OECD. (2015). *Does having digital skills really pay off?* https://doi.org/https://doi.org/10.1787/24121401

Papoutsoglou, M., Rigas, E. S., Kapitsaki, G. M., Angelis, L., & Wachs, J. (2022). Online labour market analytics for the green economy: The case of electric vehicles. *Technological Forecasting and Social Change*, *177*, 121517. https://doi.org/https://doi.org/10.1016/j.techfore.2022.121517

Pejic-Bach, M., Bertoncel, T., Meško, M., & Krstić, Ž. (2020). Text mining of industry 4.0 job advertisements. *International Journal of Information Management*, *50*(August), 416–431. https://doi.org/10.1016/j.ijinfomgt.2019.07.014

PWC, Huawei, & UNESCO. (2021). *Talents TIC en Tunisie: l'adéquation entre l'off re et la demande*.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24*(2), 235–270. https://doi.org/10.1086/499972

The World Bank. (2020). *Enterprise Surveys*.

The World Bank, & MTCEN. (2020). *Diagnostic de l'Économie numérique de la Tunisie*.

UNESCO. (2018). *Digital skills critical for jobs and social inclusion*.

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588. https://doi.org/10.1016/j.chb.2017.03.010

Vivarelli, M. (2012). Innovation , Employment and Skills in Advanced and Developing Countries: A Survey of Economic Literature. *Journal of Economic Surveys*. https://doi.org/10.2753/JEI0021-3624480106

List of Tables

Table 1: Description of digital skills categories 22
Table 2: Distribution of job ads across digital skills categories 23
Table 3: Marginal impact of digital skills categories on the probability of mentioning non-digital skills categories 24
Table A.1: Description of non-digital categories 28
Table A.2: Top 10 most common individual skills, by number of mentions 29
Table A.3: Geographic distribution of online ads, ANETI vacancies, and population by governorate 30

Table 1: Description of digital skills categories

Category	*Description*
Programming	Web development (e.g., HTML, CSS, PHP); app development (e.g., Apache Cordova, Flutter, iOS development); software development (e.g., Java, software engineering, C++)
Data science and data engineering	Data science (e.g., Python, Big Data); data engineering (e.g., Azure, AWS, cloud computing)
Technical support and IT	IT support (e.g., VLAN, antivirus); IT security (e.g., firewalling, Fortinet); system administration (e.g., PowerShell, Linux); servers and middleware (e.g., Hibernate, SoapUI, server administration); networks (e.g., WLAN, MPLS)
Office software	Office software (e.g., the Microsoft suite); general office digital tools (e.g., emailing, computer tools)
Graphic design and digital content creation	Graphic and digital design (e.g., Adobe, Illustrator, Photoshop); digital content creation (e.g., blogs, optimized digital contents); web content management (e.g., newsletter, web banners)
Digital marketing	Digital marketing tools (e.g., Google Analytics, Google Ads); web content optimization (e.g., Search Engine Optimization/SEO); general digital marketing keywords (e.g., e-marketing, Netlinking campaigns)
Data and business analysis	Data analysis skills (e.g., Power BI, data visualization); business analysis (e.g., CRM, Microsoft Dynamics)
Machine learning and AI	Machine learning (e.g., Random Forests, PyTorch); Artificial Intelligence (e.g., Deep Learning, AI)

Note: This table presents the categorization of digital skills derived from the text of 59,254 ads.

Table 2: Distribution of job ads across digital skills categories

Category	*Count*	*Rate – as % of all ads*	*Rate – as % of digital ads*
Programming	9,526	16.1%	38.5%
Data science and data engineering	5,714	9.6%	23.1%
Technical support and IT	5,443	9.2%	22.0%
Office software	5,410	9.1%	21.9%
Graphic design and digital content creation	4,696	7.9%	19.0%
Digital marketing	3,795	6.4%	15.4%
Data and business analysis	2,681	4.5%	10.9%
Machine learning and AI	770	1.3%	3.1%

Note: The count column presents the absolute number of ads that mention at least one skill in the relevant category. The rate column corresponds to the percentage of ads mentioning at least one skill from the relevant category (i.e., the count divided by the total number of ads, N = 59,254).

Table 3: Marginal impact of digital skills categories on the probability of mentioning non-digital skills categories

Mentioned digital skills categories	*Change in the probability of mentioning the non-digital skills categories*								
	Social	Character	Customer Service	Cognitive	Finance	Writing	Creativity	Mgmt	Business systems
Programming	0.101*** (0.008)	0.182*** (0.007)	0.068*** (0.007)	0.152*** (0.007)	-0.027*** (0.004)	0.028*** (0.004)	0.002 (0.003)	0.044*** (0.004)	0.002 (0.002)
Data science & data engineering	0.049*** (0.012)	0.076*** (0.010)	0.005 (0.010)	0.104*** (0.010)	-0.013** (0.006)	0.043*** (0.006)	-0.005 (0.003)	-0.013*** (0.003)	0.013*** (0.003)
Tech. support & IT	0.098*** (0.010)	0.096*** (0.010)	0.088*** (0.009)	0.075*** (0.009)	-0.037*** (0.004)	0.030*** (0.005)	-0.001 (0.003)	0.005 (0.003)	0.012*** (0.003)
Office software	0.330*** (0.006)	0.231*** (0.008)	0.125*** (0.007)	0.147*** (0.007)	0.104*** (0.006)	0.079*** (0.005)	0.000 (0.003)	0.017*** (0.003)	0.004*** (0.001)
Graphic design & content creation	0.115*** (0.009)	0.072*** (0.009)	-0.081*** (0.007)	0.015* (0.008)	-0.061*** (0.003)	0.049*** (0.005)	0.191*** (0.008)	-0.028*** (0.002)	-0.008*** (0.001)
Digital marketing	0.264*** (0.009)	0.136*** (0.010)	0.182*** (0.010)	0.204*** (0.010)	-0.055*** (0.004)	0.143*** (0.008)	0.071*** (0.006)	0.023*** (0.005)	0.055*** (0.005)
Data & business analysis	0.172*** (0.012)	0.158*** (0.011)	0.131*** (0.011)	0.251*** (0.011)	0.117*** (0.009)	0.066*** (0.007)	0.015*** (0.004)	0.041*** (0.006)	0.021*** (0.003)
ML & AI	0.150*** (0.022)	0.011 (0.018)	-0.047*** (0.017)	0.119*** (0.020)	0.004 (0.013)	0.060*** (0.012)	0.085*** (0.012)	0.024*** (0.009)	0.012*** (0.004)
Observations	59,254	59,254	59,254	59,254	59,254	59,254	59,254	59,254	59,254

Note: Each column reports the average marginal effects from a Probit regression with the non-digital skills category as dependent variable, and all eight digital skills categories as independent variables. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Digital skills category names (column 1) are abbreviated but follow the same order as Table 1 above. Mgmt=Management.

List of Figures

Figure 1: Number of digital skills mentioned per job ad 26
Figure 2: Geographic locations of online digital ads. Ads (per 1,000 inhabitants) by governorate (left) and relationship between population and shares of online digital ads and ANETI vacancies by governate (right)........ 27
Figure A.1: Job ads used to identify a starting set of digital skills 31
Figure A.2: Relationship between the number of digital and non-digital skills mentioned........ 33

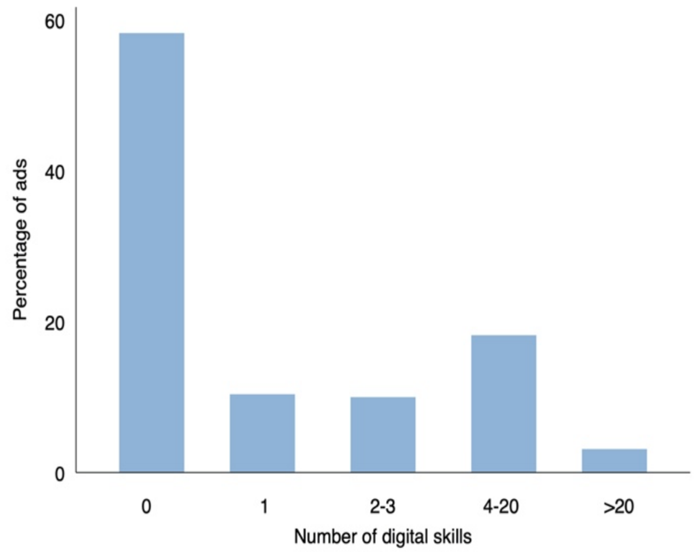


Figure 1: Number of digital skills mentioned per job ad

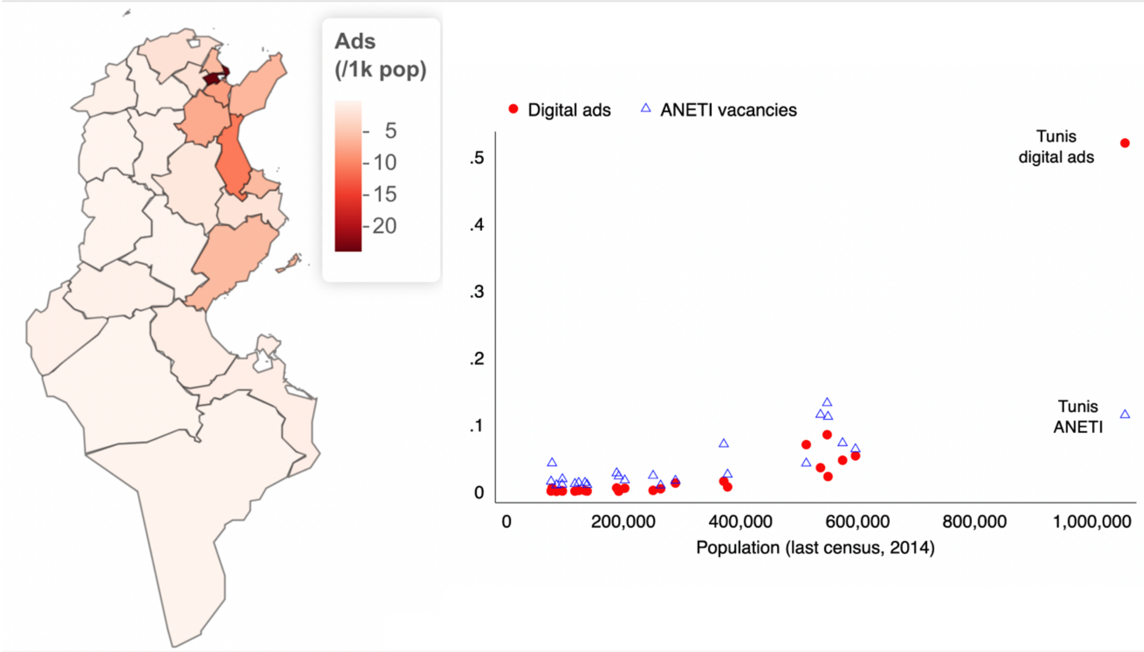


Figure 2: Geographic locations of online digital ads. Ads (per 1,000 inhabitants) by governorate (left) and relationship between population and shares of online digital ads and ANETI vacancies by governate (right).

Note: The figure on the left presents the distribution of ads (per 1,000 inhabitants) by Tunisian governorate (N=59,254). The figure on the right shows the relationship between population by governorate and the rate of digital online ads (i.e., online ads mentioning at least one digital skill) as a percentage of all ads (red dots). The blue triangles represent shares of job ads received by the Tunisian National Employment Agency – ANETI in 2019 (National Institute of Statistics, 2019b) also at the governorate level. All population figures from the last census (National Institute of Statistics, 2014).

Appendix

Table A.1: Description of non-digital categories

Category	*Description*
Social	Communication (“communication”, “communiquer”), collaboration (“collaboratif”, “collaborer”), negotiation (“negotier”), “Team” (“sens equipe”, “travailler equipe”, “esprit equipe”), persuasion (“persuader”), listening (“ecoute”), presentation (“presenter”), English* (“anglais”), French* (“francais”), sharing spirit* (“esprit partage”), interpersonal skills* (“relationnel”)
Cognitive	Solving (“resoudre”), research (“recherche”), “analy” (“analyser”, “analyse”), decision (“choisir”, “decider”, “decideurs”), thinking (“penser”, “pensez”, “reflechir”, “reflexion”), math (“mathematique”), “statistic” (“statistique”), calculation (“calcul”, “calculer”)
Character	Organization skills (“organisation”, “organise”), time management (“gestion temps”), detailed-oriented (“attention detail”, “attention details”, “sens detail”, “souci detail”), meeting deadlines (“honorer deadlines”, “respecter deadlines”, “respectez deadlines”), multi-tasking, energetic (“energetique”), self-starter, initiative (“esprit initiative”, “prendre initiative”, “prise initiative”, “prise initiatives”), self-motivation (“motivation personnelle”), rigorous* (“rigoureux”, “rigueur”), autonomy* (“autonomie”, “autonome”), self-management*
Creativity	“Creativ” (“creatif”, “creative”, “creative”)
Writing	Writing (“ecrire”, “redactionnelle”, “redactionelles”), editing (“editer”, “editeur”, “edition”), preparing reports (“elaborer rapport”, “redaction documentations”, “redaction rapports”, “rediger rapports”), preparing proposals (“redaction proposition”)
Management	Supervisory (“superviser”), leadership (“team leader”), mentoring (“mentorat”, “coaching”), staff supervision/development (“supervision personnel”, “developpement personnel”), performance/personnel management (“organisation personnel”)
Finance	“Financ” (“finance”, “financiere”), budgeting (“budgetisation”, “elaborer budget”, “gere budget”, “gestion budget”), accounting (“comptabilite”), cost (“cout”)
Business systems	Systems Developpement/Integration/Architecture (“architecture systeme”, “architecture systemes”, “developpement systeme”, “developpement systemes”, “integration systeme”, “integration systemes”), business intelligence/systems/planning/strategy, Six Sigma, KPIs
Customer service	Customer service (“ventes”, “vente”), patient (“patients”, “patience”), client

Note: This table presents the categorization of non-digital skills following Deming and Noray (2020). The first column presents the name of the non-digital skills category. The second column presents the keywords and phrases used to identify each category. The English-language keywords and phrases are taken from Deming and Noray (2020); in parentheses we present the French-language translations (where not identical to the original). * indicates the keywords are not present in the original category; rather, they were identified by the authors using a word embeddings model (see text for details).

Table A.2: Top 10 most common individual skills, by number of mentions

Skill	*Category*	*Count*	*Rate – as % of all ads*
Microsoft Excel	Office software	2,340	0.039
Javascript	Programming	2,169	0.037
Adobe Photoshop	Graphic design and digital content creation	1,840	0.031
PHP	Programming	1,759	0.030
Microsoft Word	Office software	1,613	0.027
Java	Programming	1,596	0.027
CSS	Programming	1,505	0.025
HTML	Programming	1,452	0.025
SQL	Programming, data science and data engineering, technical support and IT	1,448	0.024
Illustrator	Graphic design and digital content creation	1,426	0.024

Note: The table presents the ten most common skills in our eight digital categories, by the number of ads mentioning them. The category column describes the category (or categories) under which the skill falls. The count column presents the absolute number of ads that mention the skill. The rate column corresponds to the percentage of ads mentioning the skill (i.e., the count divided by the total number of ads, N = 59,254).

Table A.3: Geographic distribution of online ads, ANETI vacancies, and population by governorate

Governorate	*Online ads (count)*	*Online ads (rate)*	*Digital online ads (count)*	*Digital online ads (rate)*	*ANETI vacancies (count)*	*ANETI vacancies (rate)*	*Population (count)*	*Population (rate)*
Tunis	25,375	42.8%	12,846	52.0%	9987	11.4%	1056247	14.2%
Sousse	5,949	10.0%	2,090	8.5%	11557	13.2%	547403	7.4%
Ben Arous	4,593	7.8%	1,146	4.6%	6324	7.2%	573712	7.7%
Sfax	3,655	6.2%	1,311	5.3%	5531	6.3%	595728	8.0%
Nabeul	3,429	5.8%	871	3.5%	10037	11.4%	535970	7.2%
Monastir	3,357	5.7%	545	2.2%	9771	11.1%	548828	7.4%
Ariana	3,123	5.3%	1,723	7.0%	3667	4.2%	511655	6.9%
Bizerte	1,045	1.8%	372	1.5%	6174	7.0%	370757	5.0%
Manouba	790	1.3%	305	1.2%	1407	1.6%	288186	3.9%
Zaghouan	566	1.0%	121	0.5%	3690	4.2%	77394	1.0%
Mahdia	452	0.8%	126	0.5%	2385	2.7%	187804	2.5%
Medenine	407	0.7%	163	0.7%	2203	2.5%	377240	5.1%
Kairouan	383	0.6%	114	0.5%	1457	1.7%	201531	2.7%
Gabes	288	0.5%	94	0.4%	785	0.9%	262771	3.5%
Gafsa	163	0.3%	34	0.1%	2055	2.3%	250012	3.4%
Jendouba	94	0.2%	31	0.1%	1120	1.3%	122996	1.7%
Beja	79	0.1%	27	0.1%	1123	1.3%	133595	1.8%
Siliana	63	0.1%	25	0.1%	886	1.0%	94551	1.3%
Tozeur	52	0.1%	10	0.0%	1296	1.5%	75685	1.0%
Kasserine	48	0.1%	5	0.0%	1974	2.3%	191346	2.6%
Kef	37	0.1%	8	0.0%	868	1.0%	137290	1.8%
Sidi Bouzid	33	0.1%	7	0.0%	975	1.1%	116428	1.6%
Tataouine	30	0.1%	9	0.0%	1614	1.8%	95031	1.3%
Kebili	13	0.0%	1	0.0%	815	0.9%	84879	1.1%

Note: The table presents the geographic distribution of online ads in our data, vacancies from the National Employment Agency of Tunisia (ANETI) and population. The count columns (2, 4, 6, and 8) present absolute counts; the rate columns (3, 5, 7, and 9) present the share as a percentage of the total in that column (i.e., online ads as a percentage of all ads, N=59,254; digital online ads as a percentage of all ads mentioning at least one digital skill, N = 24,718; ANETI vacancies as percentage of the total, N = 87,701; and population as percentage of the overall Tunisian population, N = 7,437,039). ANETI vacancy data from the National Institute of Statistics (2019b). Population data from the last census (National Institute of Statistics, 2014). For 5,320 ads in our dataset, location data is missing.

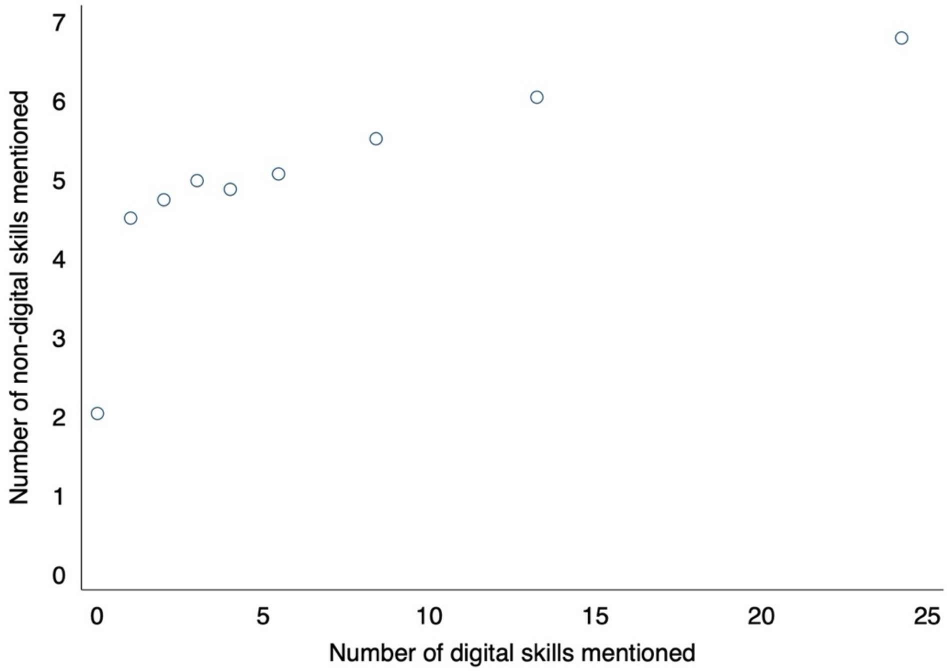
Figure A.1: Job ads used to identify a starting set of digital skills

First ad Job title: "Développeur .net Senior" Description: "Le groupe est leader des applications de la carte à puce (Pétrole, Collectivités, Loisirs et Monétique) en France. Fort d'une croissance importante, la filiale en Tunisie souhaite renforcer son effectif IT par un **Développeur .NET** senior Principales missions : Développement de toutes les fonctionnalités techniques. Être force de proposition et contribuer dans les choix techniques. Respect des bonnes pratiques de **codage**. Contribution dans l'encadrement et la formation des nouvelles recrues. Compétences : Maîtrise parfaites des environnements de développement .Net Une très bonne maîtrise des langages de **programmation** : **C#** Bonnes connaissances des technologies Web **front end** : **HTML**, **Javascript** et **CSS** (**Bootstrap** et **JQuery**) Maîtrise du langage **SQL** Maîtrise de l'**orienté objets** et des patterns les plus importants. Bon niveau de communication en français à l'oral et à l'écrit Capacité à encadrer des profils juniors Capacité à estimer et analyser les tâches Capacité à faire des POC et proposer des solutions techniques talents."
Second ad Job title: "Responsable Marketing Digital / E-commerce" Description: "Société commerciale, sise à Dende-Manouba, désire recruter un chargé de marketing et de **communication digital** dont la mission: Valoriser l'image de la société, assurer sa présence sur le web et générer des visites de prospects qualifiés sur le site internet et les réseaux sociaux; Produire l'ensemble des contenus numériques et multimédias (rédactionnels, visuels, photos, vidéos…) et les diffuser avec une approche cross-canal (Réseaux sociaux, sites web, mini-sites, emailings, …); Mise en place de stratégie marketing, développement et la gestion du site web; Animer les réseaux sociaux et faire grandir les communautés de la société; Exigences de l'emploi De **formation multimédia** ou **infographie**, avec une spécialisation en **e-marketing**/**e-communication**/marketing stratégique."
Third ad Job title: "Webmarketer" Description: "Connaissances : Passionné(e) d'internet, vous avez une forte culture du web et du digital Maîtrise des techniques de **SEO** / **SEA** : Google est votre ami ! Maîtrise des **outils Google** (**Google Analytics**…) Parfaite maîtrise du Français / bonnes connaissances en Anglais, Connaissances du langage **HTML**, Connaissances de base en **PHP** / **MySQL**. Expérience significative requise en **webmarketing** / **référencement** Formation exigée : Bac + 2/3 min et métiers du web (Webmarketing/Web/Communication…) Expérience exigée : 1 à 2 ans min" [...]
Fourth ad Job title: "Ingénieur développeur JS / MEAN Stack" Description: "Exigences: - Expérience dans le **développement frontend** avec **Angular** 2 => 9 - Connaissances **CSS** / **SCSS** - Expérience dans le **développement backend** avec **NodeJS** / **Loopback** ou frameworks **Express JS** équivalents / **Spring** - Connaissances basiques de **MongoDB** / **Mysql** - Connaissances basiques en **Data science** - Bonne communication en français et en anglais - Discipline et autonomie - Capacité à faire des recherches sur Google / Stack Overflow / lire la documentation / explorer le code source - 0-3 années d'expérience (Débutant / junior / confirmé) - Bac + 5 ou équivalent en **ingénierie informatique** / **logicielle** / **réseaux** ou équivalents."

Fifth ad
Job title: “Ingénieur développement”
Description: “Compétences techniques exigées Maîtrise de la suite **BI SAP Business Objects**. Connaissances des **outils de visualisation de données** (**QLIK**) Maîtrise d'un ou de plusieurs des langages suivants : **Spark**, **Scala**, **R**, **Python**, **Java**. Connaissance des **outils ETL** (**Talend**, **ABINITIO**) et **entrepôts de données**.. Connaissance des outils et technologies **Big Data**: **Hadoop**, **Hive**, **OOZIE**, **ELK**. Expérience avec les bases de données relationnelles (**Oracle**, **MySQL**) et **NoSQL** (**Mongo**, **Cassandra**). Bonnes connaissances en administration **système Linux**.”

Note: This figure presents the five ads used to identify a ‘starting set’ of digital skills. The ads are presented in French as originally published. The starting set of digital skills is highlighted in bold.

Figure A.2: Relationship between the number of digital and non-digital skills mentioned



Note: The figure shows a binned scatter plot of the relationship between the number of non-digital and digital skills mentioned in ads. The plot is constructed by grouping the data on the number of digital skills into equal-sized bins and portraying the means for the number of digital (x-axis) and non-digital (y-axis) skills within each bin. (N=59,254)