**Automation risk in the EU labour market**

**Analysing jobs at risk of machine displacement based on online job vacancy data**

Jiri Branka, Konstantinos Pouliakas ([[1]](#footnote-1)) and Vladimir Kvetan

European Centre for the Development of Vocational Training (Cedefop)

**Keywords:** automation, AI, machine learning, risk, skills, tasks, technologies, EU

**Introduction**

Not long before the coronavirus outbreak popular fears about artificial intelligence (AI) algorithms and smart machines resulting in a jobless society were widespread (Brynjolfsson and McAfee, 2014; Ford, 2015). While a 2013 University of Oxford study cautioned that about half of all jobs in advanced economies may become extinct due to advancing machine learning methods (Frey and Osborne, 2013), subsequent studies deconstructing jobs according to their task composition tended to dispel such fears of rampant job destruction (Pouliakas, 2018).

Concerns about the Covid-19 crisis accentuating automation in labour markets have nonetheless resurfaced. The pandemic and associated social distancing measures have accelerated the incentives of companies and societies to adopt new digital and data-driven technologies. And indeed, most notable automation episodes in history have always tended to spike following major economic crises (Frey, 2019). Early predictions that Covid-19 will have a positive automation effect may however prove to be false. Occupations identified as high risk of Covid-19 exposure and social distancing disruption (Pouliakas and Branka, 2020) have been found, for instance, to [correlate weakly with those facing higher automation risk](https://www.nber.org/papers/w27249.pdf) (Chernoff and Warman, 2020). Many of the occupations and sectors mostly affected by Covid-19 are typically in the service sector and heavily reliant on interpersonal skills (e.g. hospitality, leisure, retail), which are relatively less susceptible to replacement by AI technologies.

The above findings highlight that in-depth understanding of which skills and job tasks may be displaced by AI and other digital technologies, especially within the context of the Covid-19 shock, and which occupations are more inclined to automation is crucial for the formulation of preventive upskilling and reskilling policies. Indeed, the need to design effective reskilling programmes that can enable individuals and firms to make the transition to a digital economy has been highlighted as a key action as part of the European Commission’s *New Skills Agenda*[[2]](#footnote-2).

**Value added of study**

Previous studies focussed on estimation of automation risk in labour markets have relied on representative survey data from individual workers, such as the OECD’s PIAAC survey (Artnz et al., 2017; Nedelkoska and Quintini, 2018) and the Cedefop European skills and jobs survey (ESJS) (Pouliakas, 2018). Such studies have sought to detect the latent relationship between the risk of automation and a limited set of broad skill groups and tasks. Such analyses have enriched the original occupational estimates of Frey and Osborne by adopting a task- or skill-needs approach. However, their main conclusions regarding the skills and tasks with higher probability of automation have been relatively generic and predictions are made at higher (2-digit) occupational level, due to data constraints associated with conventional labour market surveys.

The current study utilises instead a novel *big dataset* which contains information on the skills and work activities required by EU employers, enabling in-depth understanding of the detailed skills and task profiles of narrow occupational groups that may be rendered automatable by machine learning algorithms. Specifically, the research utilises the rich data collected as part of Cedefop’s analysis of online job advertisements (OJAs) in all EU countries, which has manifested in the so-called *Online Vacancy Analysis Tool for Europe* (Skills-OVATE) (Cedefop, 2019). Such data facilitate investigation of the specific types of skills and work activities that are associated with jobs that could be replaced by smart machines in the near future. In addition, the data permits the detection of particular ‘task-technology-skills’ bundles which may render themselves immutable to the impact of emerging digital technologies.

**Research questions**

This study aims to provide answers to the following research questions:

* Which occupations are characterised by a high probability of automation? What share of occupational groups in the EU labour market are prone to displacement by emerging digital technologies?
* Which detailed skills are automatable and which not?
* What clusters of ‘skills-work activities-technologies’ in EU jobs are characterised by higher automatability?
* Which groups of occupations should be targeted by employment and up-skilling/reskilling policies?

**Methodology**

This paper utilises detailed data on occupations (at 4-digit level), skills, work activities and technologies from Cedefop’s OVATE. For the purposes of the analysis we deploy a new version of OVATE that has been classified using the O\*NET hierarchical structure. Specifically, all detailed skills collected as part of the OJA have been allocated to the following O\*NET taxonomy: **Abilities** (e.g. originality, fluency of ideas, oral expression etc.); **Knowledge** (e.g. language, computers and electronics, sales and marketing, personnel and human resources, administration and management etc.); **Technology** (e.g. office suite software, web platform development, object or component oriented development software etc.); **Skills** (e.g. complex problem solving, time management, programming, management of financial resources, etc.); **Work activities** (e.g. interacting with computers, organising, planning and prioritising work, thinking creatively, assisting and caring for others, communicating with persons outside organisation, handling and moving objects, controlling machines and processes etc.); **Work styles** (e.g. adaptability, cooperation, initiative, dependability etc.).[[3]](#footnote-3)

Each skill term is mapped to 4-digit ISCO occupational groups. We use the 3rd level of the hierarchy, corresponding to a total of 142 relatively broad ‘skill clusters’. We calculate for each skill cluster the ‘skills intensity’ per 4-digit occupation group:

$S^{I}= \frac{S\_{io}}{S\_{o}}$ [1]

where the skills intensity (*SI*) is computed as the ratio of the advertised skill cluster, *i*, over the total number of posted vacancies in each 4-digit occupation (*o*).

To calculate the risk of automation for all EU occupations, we first match the assessment of automatability based on the opinions of machine learning experts as surveyed by Frey and Osborne (2013). Namely, we match to the OVATE data the Frey and Osborne ‘training dataset’ of 85 occupations that are deemed automatable or not, given the absence of a set of characteristics (‘engineering bottlenecks’) that may prevent them from machine replacement. Following a necessary process of dimensionality reduction of the big data at hand, specifically the use of a stepwise regression approach, we subsequently estimate a reduced form model of the latent relationship between automatability and the ‘skills-work-activities-technologies’ of occupations as demanded by EU employers.

More formally, we estimate a linear probability or logit model that detects the latent relationship between the “true” automatability of occupations, as extracted from the FO training data, and the OVATE list of required skills-tasks-technologies, as follows:

$P\left(y\_{ao}^{\*}=1\right)=β\_{0o}+β\_{1}XS\_{o}+β\_{2}WA\_{o}+β\_{3}T\_{o}+u\_{o}$ [2]

where $y\in \{0,1\}$ is a vector of the occupational automatability assessment, *S* is a matrix of employers’ skill-requirements as calculated in equation (1), *WA* refers to occupational work activities and *T* are demanded occupational technologies. The coefficients $β$ are estimated on a pooled sample of 403 occupations from all EU countries*.*

Having estimated the latent relationship (2), the coefficients of the model are subsequently applied to all other 4-digit occupations to obtain an out-of-sample prediction of the occupational risk of automation. Given the predicted automatability score of occupations, the analysis subsequently investigates which ‘skills-work activities-technologies’ bundles are associated with occupations at high or low risk of automation using clustering (unsupervised learning) techniques.

**Preliminary findings**

Our analysis of the big data available from EU employers’ online job advertisements reveals that the core skills insulated from automation include creativity, leadership, organisational and interpersonal communication skills (Figure 1). Interaction with digital devices is also a key trait of occupations that have lower automation risk, one which has assumed greater significance in the Covid-era given the growing need for workers to carry out their jobs remotely. By contrast, skills that are associated with mechanical knowledge and physical or monitoring activities and objects in the workplace are more prone to higher automation risk.

The analysis further reveals that assemblers, numerical clerks, refuse and other elementary workers, metal and machinery workers as well as drivers and mobile plant operators are some of the occupations facing highest risk of displacement. This contrasts with personal services and personal care workers, teaching, health and ICT professionals, who have low automatability. Overall, the clustering analysis illustrates that a core group of occupations facing the highest risk of automation is characterised by similar qualities, namely they involve work activities such as ‘evaluating standards’, ‘operating vehicles’, ‘monitoring processes’ and the use of technologies that include ‘information retrieval’, ‘application server software’ and analytical and scientific software’.

Overall, the paper’s conclusions provide useful input for the effective design of upskilling and reskilling policies that can facilitate adjustment of individuals, firms and economies to the automation dynamics of new and emerging digital technologies.

Figure 1. Job profiles and automation risk

 

*Note*: The list of skills and work activities are obtained from Cedefop’s European database of online job advertisements ([Skills OVATE](https://www.cedefop.europa.eu/en/events-and-projects/projects/skills-online-job-vacancies)); occupations are distinguished into automatable and non-automatable after matching Frey and Osborne (2013) training dataset to 4-digit OVATE occupations.

**Key references**

Arntz, M., Gregory, T. and Zierahn, U. (2017). “Revisiting the Risk of Automation”. *Economics Letters*, Vol. 159, pp. 157-160.

Brynjolfsson, E., and McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York: WW Norton & Company.

Cedefop (2019). *The online job vacancy market in the EU - driving forces and emerging trends*. Available online: <https://www.cedefop.europa.eu/files/5572_en.pdf>

Chernoff, A.W. and Warman, C. (2020), Covid-19 and implications for automation, NBER WP 27249 <http://www.nber.org/papers/w27249>

Ford, M. (2015). *Rise of the robots: technology and the threat of a jobless future*, Basic books: New York.

Frey, C. and Osborne, M. (2013). *The future of employment: how susceptible are jobs to computerisation*, University of Oxford.

Nedelkoska, L. and G. Quintini (2018). “Automation, skills use and training”, *OECD Social, Employment and Migration Working Papers*, No. 202, OECD Publishing, Paris.

Pouliakas, K. (2018). “Automation risk in the EU labour market: a skill-needs approach”, in Hogarth, T. (ed) *Economy, employment and skills: European, regional and global perspectives in an age of uncertainty*, Rome.

Pouliakas, K. and Branka, J. (2020). *EU jobs at highest risk of Covid-19 social distancing: Is the pandemic exacerbating the labour market divide?* Luxembourg: Publications Office of the European Union. Cedefop working paper; No 1. <http://data.europa.eu/doi/10.2801/968483>

1. () Contact information: Cedefop, Europe 123, 57001 Thermi, Thessaloniki, Greece; Tel: +030 490 292; email: konstantinos.pouliakas@cedefop.europa.eu [↑](#footnote-ref-1)
2. () <https://ec.europa.eu/social/main.jsp?catId=1223&langId=en> [↑](#footnote-ref-2)
3. () Within each broad ‘skill cluster’ lies a range of individual, more detailed, skills e.g. ‘interacting with computers’ is comprised of ‘use a computer’ / ‘use Microsoft office’ / ‘use office systems’ / ‘use spreadsheets’ etc. [↑](#footnote-ref-3)