How can we better measure the demand for AI and other skills on the labour market?

ai:conomics policybrief

April 2025

How can we better measure the demand for AI and other skills on the labour market?

Simon Janssen, Christina Langer, Markus Nagler, Nicholas Rounding, Michael Stops and Simon Wiederhold

1. Introduction

A large body of research literature shows that technological change has a significant impact on labour markets, as modern digital technologies are changing the demand for certain skills. On the one hand, new technologies can replace some human activities. On the other hand, they can create or complement new activities (Acemoglu et al., 2015; Acemoglu & Restrepo, 2018, 2019, 2020). With the proliferation of artificial intelligence (AI) in recent years, certain questions are becoming increasingly important in public debate and research: Is the demand for AI skills also growing on the German labour market? Does the increasing demand for AI skills mean that other skills - among low, medium and highly qualified workers - are less in demand? The aim of this research project is to create a reliable data basis in order to be able to answer such questions in a more informed way in the future.

Developments in generative AI, particularly tools such as ChatGPT, have significantly intensified the discussion about the impact of AI on the labour market, both in academia and in public debate and policy. While computers and software have transformed the world of work by performing routine tasks more precisely and efficiently, modern AI systems can now take on complex, non-routine tasks without relying on detailed instructions or repetitive rules (Brynjolfsson et al., 2025). As a result, many are optimistic about the productive potential of this new technology. Others, however, fear that AI could disrupt labour markets.

In the course of the intensive scientific and public debate on AI, there is a growing body of literature that deals with the effects of AI on labour markets. These initially focus on specific occupations such as call centre workers (Brynjolfsson et al., 2025, Dijksman et al., 2024), consultants (Dell' et al., 2023), writers or developers (Peng et al., 2023). However, a major challenge is to measure how the demand for and supply of skills has changed in the wake of the emergence of AI.

Some recent studies have measured the impact of AI by using data from online job adverts (Bessen et al, 2023; Gonschor and Storm, 2023; Babina et al, 2024, 2023; Alekseeva et al, 2021; Acemoglu et al, 2022). Unlike survey or other traditional data, job advertisements are available in near real-time and provide a forward-looking, largely objective measure of labour demand. However, in many countries, job vacancy data is difficult to link to labour market data at company or even employee level. As a result, most studies either focus only on the demand for AI skills while ignoring the supply, or they limit themselves to capturing the supply of AI skills only at a highly aggregated level, such as in specific industries or large regional labour markets. In such aggregated data sets, however, it is hardly possible to determine the extent to which changes in the employment of individual companies can be attributed to the increasing demand for AI skills. For example, it remains unclear whether companies that are increasingly demanding

AI skills tend to employ highly or low-skilled workers or whether employment develops differently in different occupations within companies. However, the investigation of such correlations is essential in order to ultimately understand the influence of AI on the labour market.

Some notable studies from Europe have addressed these challenges by taking advantage of the fact that public employment agencies in some European countries support companies in publishing job advertisements as part of their employer services (Damelang et al., 2024; Mueller et al., 2024). Researchers can link information from these job advertisements with labour market data from social security registers, as both social security registers and job advertisements are integrated into the administrative processes of the same public service providers. For the German labour market and for the early phase of the use of AI from 2015 to 2019, Peede & Stops (2024) merged data from the JOBBÖRSE of the Federal Employment Agency with administrative establishment data (the establishment history panel) via a provided establishment identifier. They investigate how the demand for AI skills influences the demand for labour at establishment level. They find that demand for AI skills is associated with a small decrease in demand for other technical skills. They find no effects on overall employment in the company, but slightly higher employment growth in jobs for highly qualified workers.

Companies are not required to use the employer services of public employment agencies. This also applies to the job websites operated by the Federal Employment Agency. Peede and Stops (2024) and Stops et al. (2025) propose representativeness analyses based on comparisons with the IAB Job Vacancy Survey. For the job data from the Federal Employment Agency, for example, they generally come to the conclusion that although these do not reach all the jobs identified by the IAB job survey in terms of quantity, they come very close to the distribution, e.g. by job level or economic sector, in the IAB job survey.

However, the extent to which very specific job advertisements, e.g. for highly qualified workers in sectors such as IT, consulting or the pharmaceutical industry, are sufficiently representative in the data from public employment agencies and whether other data sources for job advertisements can close any gaps in representation is still an open question. This also has implications for research into the impact of AI, as it can be assumed that modern AI technologies will change the demand for highly qualified labour in these sectors in particular.

There are several potential, differently accessible private data sources for Germany that are based on web crawling and whose properties are currently being analysed step by step. These include job data from Textkernel, on the basis of which, for example, the IW Cologne conducts analyses (see e.g. Büchel et al. 2023), an unnamed provider whose data is being examined in more detail by the RWI (Gonschor and Storm, 2023) and the data from the company Lightcast, which is certainly best known in the international scientific community and is also being tested for initial studies in Germany (Alipour et al., 2021).

A team of researchers from IAB Nuremberg, ROA Maastricht, FAU Nuremberg, IWH Halle and Stanford University have joined forces to clarify the outlined questions on representativeness and close any data gaps for research. The researchers are developing a comprehensive dataset of online job adverts obtained from the company Lightcast, which will be linked to social security data on German employees.

2. Vacancy data

In our research project, we use the extensive online job adverts of the company Lighcast. Lighcast searches all relevant online job boards, including the job boards of the Federal Employment Agency, as well as over 200 individual company websites. The data collection process involves the use of algorithms to cleanse and organise the information. Lightcast further categorises the ads according to factors such as occupation, industry, company location and other criteria (Alipour et al., 2021).

2.1 Register data of the Federal Employment Agency

We merge the Lightcast job data with the German social security register data, which contains information on German employees and their companies. More specifically, our analysis is based on the Integrated Employment Biographies (IEB) of the Federal Employment Agency. The IEB contains comprehensive social security data for Germany from 1975 to 2022, covering both employees subject to social security contributions and recipients of unemployment benefits. It excludes students, military personnel, civil servants, the self-employed and people who leave the labour market completely. For each employee, the IEB records the wage, duration of employment and various demographic and occupational characteristics. Unique identifiers for individuals and establishments allow us to track employees and companies over time. We supplement the IEB data with the Establishment History Panel (BHP), which provides company-level information such as size, median wage and sector for establishments with at least one employee covered by social security as of 30 June each year.

2.2 Data quality and data linkage

In contrast to the job data managed and provided by the employment agencies, the Lightcast job data is compiled from many different websites. On the one hand, this approach promises a comprehensive overview of the structure of vacancies on the German labour market. On the other hand, the job data from Lightcast does not contain any information that allows the companies and organisations that publish the vacancies to be precisely identified. Instead, Lightcast has to rely on the information provided with the texts of the vacancies. As a rule, the job texts contain information about the names and locations of the companies and it is therefore possible to make an assumption about which professions the companies are hiring for and what their requirements are.

The following paragraphs describe three challenges in linking online job advertisement data with social security data:

1. *Data protection:* Information on company names and locations is highly sensitive and requires special data protection regulations. In principle, researchers are not permitted to access information on the names and locations of the companies in the social security data. Therefore, the merging of the various data sources had to be carried out by a third party who is ultimately not involved in the data analysis. This ensured that we, as researchers, only use the final data product in a pseudonymised version that does not allow us to identify individual companies or employees
2. *Cross-company job adverts:* Job adverts have many different structures. While some job adverts relate to exactly one vacant position in a specific company and location, others may relate to many vacancies at different locations of a large company. Nevertheless, the majority of job adverts can be precisely assigned to their respective companies. Nevertheless, there are job adverts that are assigned to several companies at different locations or in different sectors.
3. *Information quality:* Not all job vacancy texts contain the same information about companies or the company location. While some provide the exact name and location of a company, others only contain aggregated information at district or county level. There are also job adverts that do not contain any information about company locations. In addition, the quality of information on job vacancies varies considerably. Some job adverts contain comprehensive information, including the full legal name and sector of the company, while other job adverts do not contain this information. For example, in some cases the legal form is fully advertised, e.g., "limited liability company", while others use abbreviations such as "GmbH". Furthermore, on some websites, capitalisation is completely disregarded (e.g., "gmbh") and longer company names are often (un)intentionally shortened. This problem is further complicated by the fact that there are many international companies with foreign names and foreign legal forms. Finally, job adverts often differ in their comprehensiveness and depth of detail.

In recent years, the Ai:conomics team has worked to overcome these difficulties and minimise their impact on the quality of the final product. To do this, we have used various accurate and probabilistic methods to, for example, remove differences in company names, remove duplicates, and accurately identify companies using regional and industry information. The work was based on a high level of expertise in data work, statistical software programmes and manual data cleansing.

3. Preliminary results of the matched data

Overall, the data available to us includes job advertisements from around 1.2 of the approximately 3.4 million companies in Germany. Of these companies, around 50 per cent have no employees subject to social security contributions and are therefore excluded from the data basis. As a preliminary result, we have so far identified around 400,000 (33.5%) 1.2 million companies in the register data (Figure 1). In comparison, previous studies that have attempted to link the Lightcast data to listed US companies have been able to link a third of the job adverts but not the companies. Unfortunately, there is no information on the number of linked companies in the literature to date.

Figure 1. Identified companies in register data

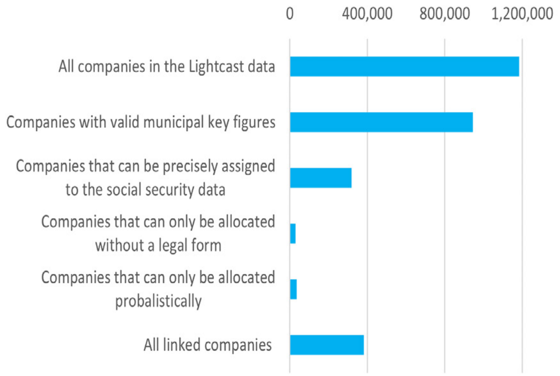
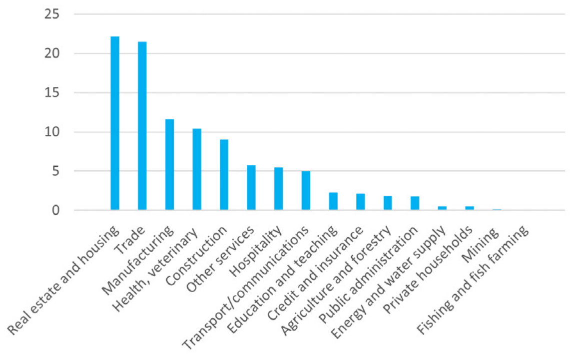


Figure 2 illustrates the distribution of matched firms across various industries. The figure shows that the majority of these firms are concentrated in the real estate, housing, and trade sectors, followed by manufacturing and health. Unpublished results already show a tendency for industries with a higher level of highly skilled labour to search for employees on online job data.

Figure 2. Share of matched firms by sector



4. Conclusions and outlook

This policy brief provides an update on the process of creating a new dataset that will combine information from online job vacancies with administrative register data on employees with social insurance. We describe the data linkage process in detail and outline the research gaps that our research team will be addressing in the near future with this innovative dataset. We also present preliminary findings that the ai:conomics team and its partners at FAU Nuremberg, IWH Halle and Stanford University have achieved in recent years. In the coming years, we plan to tap into the research potential of this data set to gain new insights into the spread of artificial intelligence on the German labour market, among other things.

Literature

Alekseeva, L., Azar, J., Gin'e, M., Samila, S., Taska, B., 2021. The demand for AI skills in the labour market. Labour Economics 71, 1-27.

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. Journal of Labor Economics, 40(S1), S293-S340.

Acemoglu, D., Gancia, G., & Zilibotti, F. (2015). Offshoring and directed technical change. American Economic Journal: Macroeconomics, 7(3), 84-122.

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. In American Economic Review (Vol. 108, Issue 6, pp. 1488-1542). American Economic Association.

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labour. In Journal of Economic Perspectives (Vol. 33, Issue 2, pp. 3-30). American Economic Association.

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labour Markets.

Alipour, J.-V. ;, Langer, C. ;, & O'kane, L. (2021). Is working from here to stay? A look at 35 million job ads (Vol. 22, Issue 06).

Babina, T., Fedyk, A., He, A.X., Hodson, J., 2024. Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics 151, 103745.

Babina, T., Fedyk, A., He, A.X., Hodson, J., 2023. Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition. in: Basu et al. (2024).

Büchel, Jan; Engler, Jan Felix; Mertens, Armin (2023): KI-Einsatzbereiche in Deutschland. Eine Analyse von KI-Stellenanzeigen, Gutachten im Projekt „Entwicklung und Messung der Digitalisierung der Wirtschaft am Standort Deutschand" des Bundesministeriums für Wirtschaft und Klimaschutz, Köln.

Bessen, J., Cockburn, I., Hunt, J. (2023). Is distance from innovation a barrier to the adoption of artificial intelligence? Unpublished manuscript.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work. The Quarterly Journal of Economics, qjae044.

Damelang, A., Rückel, A.-K., & Stops, M. (2024). 3|2024 Gender Stereotyping in the Labour Market: A Descriptive Analysis of Almost One Million Job Ads across 710 Occupations and Occupational Positions. IAB Discussion.

Dell', F., Saran, A., Mcfowland, R. E., Krayer, L., Mollick, E., Candelon, F., Lifshitz-Assaf, H., Lakhani, K. R., & Kellogg, K. C. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality.

Dijksman, S., Eijkenboom, D., Fouarge, D., Fregin, M.-C., Graus, E., Janssen, S., Levels, M., Montizaan, R., Steens, S., & Stops, M. (2024). Künstliche Intelligenz in der Arbeitswelt: Wie KI-unterstütztes Training die Gespräche im Kundenservice verändert. Web publication/site, Institut für Arbeitsmarkt und Berufsforschung (IAB).

Gonschor, M., Storm, E. (2023). The diffusion of artificial intelligence: New evidence from German online job vacancy data. Unpublished manuscript.

Mueller, A. I., Osterwalder, D., Zweimüller, J., & Kettemann, A. (2024). Vacancy Durations and Entry Wages: Evidence from Linked Vacancy-Employer-Employee Data. Review of Economic Studies, 91(3), 1807-1841.

Peede, Lennert & Michael Stops (2024): Artificial intelligence technologies, skills demand and employment: evidence from linked job ads data. (IAB Discussion Paper 15/2024), Nuremberg, 62 pp.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The Impact of AI on Developer Productivity: Evidence from GitHub Copilot.

Stops, Michael, Asma Ahmed EzElDin, Marie-Christine Heckel, Pascal Heß, Hauke Klevinghaus, Verena Malfertheiner, Lina-Jeanette Metzger & Lennert Peede (2025): „IAB-Kompetenz-Kompass" Bereitstellung von Stellentext-Daten als strukturierter Datensatz. (IAB-Forschungsbericht 01/2025), Nürnberg, 259 S.