NETWORK
LISER

Discussion Paper Series

IZA DP No. 18415 February 2026

AI-Powered Skill Classification: Mapping Technology Intensity in the German Labor Market

Sabrina Genz
University Utrecht, IAB Nuremberg and Instituut Gak

Terry Gregory
Luxembourg Institute of Socio-Economic Research (LISER), ZEW Mannheim and IZA@LISER

Florian Lehmer
IAB Nuremberg

IZA NETWORK @ LISER
ISSN: 2365-9793

AI-Powered Skill Classification: Mapping Technology Intensity in the German Labor Market*

Abstract

The rapid evolution of technology is reshaping labor markets by altering skill demands and job profiles. This paper introduces a novel skill-based measure of occupational technology intensity – the Occupational Technology Skill Share (OTSS) – that distinguishes between manual, digital, and frontier technologies, including artificial intelligence (AI). Using natural language processing, generative AI, and supervised machine learning, we develop an AI-powered skill classification that enriches occupation-linked skill labels with standardized GenAI-generated descriptions and structured indicators of technological content, enabling transparent classification by technology intensity. We compute OTSS for all occupations in the German labor market. For the average worker in 2023, manual technologies account for the largest share of skill content (42%), followed by digital (38%) and frontier technologies (20%). Frontier technologies remain concentrated in specialized occupations, while digital technologies are widespread. Linking these measures to administrative data from 2012–2023 shows a broad shift from manual and digital toward frontier skills across occupations, and reveals a non-linear, U-shaped relationship between changes in frontier skill intensity and employment growth.

JEL classification

JEL, J21, J24, O33

Keywords

artificial intelligence, digitalization, skills, employment growth

Corresponding author

Terry Gregory

terry.gregory@liser.lu

* We gratefully acknowledge helpful comments by seminar participants at the RWI Text-to-Data workshop, LISER Data science lunch seminar, ELMI conference and seminar participants at Utrecht School of Economics. We thank Britta Matthes for insights into the German skill data, and the editors Matthias Parey and Monica Costa Dias for their helpful guidance. We remain solely responsible for any remaining errors and impressions. Sabrina Genz gratefully acknowledges funding from Instituut Gak. Terry Gregory gratefully acknowledges funding from the Data Science and Simulation (DSS) grant at LISER.

1 Introduction

The rapid advancement of technology is reshaping labor markets, demanding new skills and transforming entire occupations. While digital technologies like computers and industrial robots have long been integrated into various occupations, frontier technologies – such as artificial intelligence (AI) – are emerging as key drivers of recent change. Understanding how frontier skills spread across jobs is crucial for ensuring businesses and workers can successfully adapt and for designing effective workforce policies targeting training and facilitating labor market transitions. However, measuring the presence and spread of frontier technology skills in jobs remains challenging due to their complexity and evolving nature.

To overcome this challenge, we propose a multi-step approach to rank occupations by technology intensity, distinguishing between manual, digital, and frontier skills. Using an occupation-level expert database with rich information on occupational skills, we apply natural language processing (NLP) and generative AI (GenAI), and train a machine learning model that categorizes job skills by technology level. This allows us to calculate technology intensities of jobs – measured by the Occupational Technology Skill Shares (OTSS) – and rank occupations according to their technological advancement. Linking these measures to occupational employment data, we analyze the diffusion and development of digital and frontier skills in the German labor market for the most recent period of technological progress (2012-2023). The OTSS data supporting the findings of this study are publicly available for use by researchers, policy-makers, and practitioners.[1]

In particular, we exploit the German BERUFENET expert database,[2] a rich data set containing 8,775 distinct job skills (skill titles) in either 2012 or 2023 that we can link to 4,647 unique occupations using the occupation-skill matrix provided by the German Federal Employment Agency. The job skills are updated regularly by experts and reflect the current state of typically required skills and activities performed to conduct a specific occupation. For instance, as of 2023, the occupation *data scientist* contains 36 core and additional job skills such as "applied computer science", "data lake", "project management", "programming", "machine learning", and "IT coordination". Based on the list of job skill titles, we follow a supervised machine learning approach to train an algorithm to classify job skills into three technology classes. We categorize skills into manual skills (e.g., social therapy, group dancing, tile laying, metalwork), digital skills (e.g., customer relationship management, accounting software, enterprise resource planning systems) and

[1]The OTSS data underlying this study – including occupation-level manual, digital, and frontier technology skill shares for 2012 and 2023 – are publicly available via GitHub: `https://github.com/otss-data/otss`.

[2]The most current version of the database can be accessed online via https://berufenet.arbeitsagentur.de/. See Dengler et al. (2014) for a detailed documentation of the database.

frontier skills (e.g., machine learning, augmented reality, collaborative robots) while also differentiating between production and service work contexts. Specifically, we employ a neural network classifier[3] that draws on three sources of information: standardized skill descriptions generated by the generative AI model (GenAI), represented numerically; structured indicators of technological content also generated by GenAI; and metadata from the BERUFENET system (skill groups and areas). Our approach does not hand over the classification itself to the GenAI tool, but rather uses GenAI to generate more context for each skill title for the training of the supervised machine learning model. We call this an AI-powered skill classification, as it is empowered by AI but still makes the classification process transparent and reproducible. We provide evidence on the value of our AI-powered approach by reporting a systematic comparison of alternative classification strategies (e.g. rule-based text-mining and classifiers trained on embeddings of raw skill titles only), suggesting that the additional semantic context generated by the LLM is important for identifying advanced technology skills that are not explicitly named in titles. As a further external validation, we compare our measures with individual-level survey data on technology at the workplace.

Our results indicate strong predictive power of the classifier in assigning job skills to technology classes and levels of technological intensity. Using the occupation–skill matrix, we aggregate classified skills to the occupational level and compute employment-weighted averages, thereby characterizing the OTSS of the average worker in 2023 in the economy. From this perspective, manual technologies account for the largest share of implemented skill content (42%), followed by digital technologies (38%) and frontier technologies (20%). Frontier skills are particularly prevalent in technical, scientific, and IT-related occupations, whereas digital skills are broadly distributed across a wide range of occupational domains, including many service occupations. This pattern highlights that while frontier technologies remain highly concentrated, digital technologies constitute the dominant and most pervasive form of technological exposure in the contemporary workforce.

In a next step, we link OTSS to administrative employment data to examine (i) the most prevalent frontier-OTSS occupations at the German labor market in 2023 and (ii) to document the evolution of frontier OTSS among employed workers in Germany from 2012 to 2023. When accounting for the employment size of occupations at the German labor market in 2023, the most common frontier-OTSS occupations are predominantly concentrated in technical, engineering, and IT-related domains. By contrast, occupations with high digital OTSS span a wider range of occupational domains such as office clerks, sales occupations, and occupations in business administration and financial services. One

[3]A neural network is a machine learning algorithm that learns non-linear relationships between input variables and target categories. We use a Multi-Layer Perceptron (MLP), one of the simplest network architectures.

distinct difference between the most common frontier-OTSS and digital-OTSS occupations is that, on average, digital-OTSS occupations employ more individuals than the most common frontier-OTSS occupations. Furthermore, we document shifts in the OTSS within occupations between 2012 and 2023. We document a widespread decline in the relative share of manual and digital OTSS within occupations, facilitating a substantial rise in frontier OTSS across nearly all occupations. Lastly, we relate employment growth to these changes in OTSS within occupations. Our analysis reveals a U-shaped pattern: employment growth is highest in occupations with either low or high OTSS. Overall, these findings highlight an ongoing shift in the German labor market toward advanced technology skills.

We contribute primarily to two main literature strands. First, we contribute to the literature measuring technology intensity in occupations relying on the task-based approach or exposure indices. The task-based approach computes Routine Task Intensity (RTI) of jobs based on occupational descriptions such as O*NET for the US (Autor and Dorn, 2013) or BERUFENET for Germany (Dengler et al., 2014) to identify automation-prone jobs. Exposure indices estimate forward-looking automation risks (Frey and Osborne, 2017; Arntz et al., 2017; Grienberger et al., 2024). Several AI-specific indices predict the exposure to AI advancements: Felten et al. (2018) quantify AI's impact by machine learning capabilities, Webb (2020) links AI, robotics, and software to job tasks, and Brynjolfsson et al. (2018) introduce the suitability for machine learning (SML) index, showing most jobs mix automation-prone and complementary tasks. Other studies create indices linking AI research intensity to occupations leveraging performance benchmarks from AI research (Tolan et al., 2021). Recent studies create an index measuring exposure to large language models, such as ChatGPT (Engberg et al., 2024). While these previous studies use occupational task data to construct forward-looking exposure measures, we leverage granular and up-to-date skill content of jobs to calculate technology intensity within detailed occupations, distinguishing between manual, digital, and frontier skills. By focusing on observed skill content rather than potential exposure, our approach provides a complementary and more direct perspective on how advanced technologies are currently implemented in jobs. Moreover, our novel methodology is flexible and can be applied in other countries with detailed occupational databases or even in other contexts leveraging rich textual data such as job postings.

Second, we contribute to a literature strand that focuses on how the demand for labor and skills is changing in response to the expanding capacities of new technologies.[4] Task-based models show that digitalization substitutes routine tasks while complementing

[4]A related strand of the literature examines the relationship between the skill intensities and wages. Empirical studies have documented positive wage premia for ICT skills (Krueger, 1993; Autor et al., 2003; Falck et al., 2021) and digital skills more broadly (Deming and Noray, 2020; Wiederhold and Langer, 2023). Evidence on the wage effects of AI-related skills is more limited but emerging, with recent contributions including (Babina et al., 2024; Acemoglu et al., 2022; Engberg et al., 2023).

cognitive work (Autor et al., 2003), increasing demand for both low- and higher-skilled workers and contributing to job polarization (Autor et al., 2003; Goos et al., 2014). The impact of automation varies: robots reduce employment and wages in exposed sectors (Acemoglu and Restrepo, 2020), but boost productivity without major job losses elsewhere (Graetz and Michaels, 2018) in the US and are even totally offset by employment gains in the service sector in Germany (Dauth et al., 2021). AI increasingly augments high-skill jobs rather than replacing them (Felten et al., 2019), though Webb (2020) finds AI disrupts cognitive occupations more than manual ones. Job posting data confirms rising demand for AI and digital skills, reshaping demand for specific skills rather than eliminating jobs (Acemoglu et al., 2022; Babina et al., 2024). Other studies show that automation and AI primarily transform tasks within existing jobs (Spitz-Oener, 2006; Gathmann et al., 2024; Consoli et al., 2023). While existing studies emphasize the potential shift toward more complex and advanced skills driven by rapid technological change, they mostly rely on task-based or forward-looking exposure measures. Our study provides direct, skill-based evidence on the diffusion of frontier technologies across occupations in the German labor market. By linking changes in occupational skill requirements to administrative employment data, we further document how the expansion of frontier skills relates to employment changes across all jobs between 2012 and 2023.

The article proceeds as follows. Section 2 introduces our concept of skills and describes our micro-level data relating digital and frontier skills to granular occupations and administrative employment records. Section 3 outlines our approach to identifying digital and frontier skills using a supervised machine learning approach. Section 4 explains how we construct the share of digital and frontier skills for each occupation and compares our novel measure to survey data on technology usage at the workplace. Section 5 documents the prevalence of digital and frontier skills in the German labor market in 2023 and illustrates the evolution of these skills since 2012. Section 6 concludes.

2 Data and concepts of skills and technology

In the following section, we describe our conceptual framework for distinguishing skills by level of technology, which provides the basis for our skill classification in section 3. Furthermore, we describe the data that we use to train our classifier in predicting the technology class of each skill in our data. The primary analysis relies on cross-sectional data from 2023 as described below, while Section 5 extends the analysis over time by incorporating analogous data from 2012 to examine the evolution of skill requirements and employment patterns.

Conceptual framework for skills by level of technology One big disadvantage of existing data sets is that they do not allow for distinguishing jobs by their level of technology use.[5] Up to now, it is therefore not possible to assess the number of jobs that require frontier technology skills such as those related to artificial intelligence. To overcome this gap, we exploit the richness of the expert online database BERUFENET, which provides a granular list of skills required in jobs.[6] Building on the framework used by other studies for classifying work equipment (Genz et al., 2021; Arntz et al., 2024, 2025a,b), we categorize job skills into the three different groups shown in Table 1, distinguishing between service and production contexts.[7]

At the most basic level, *manual* skills are those required to operate work processes without the support of digital technology. These skills are typical of the First and Second Industrial Revolutions (before the Digital Revolution).[8] In production settings, this includes the use of manually controlled equipment (cell (6)), such as rust removal, tile laying, or handcrafted metalwork. In service settings, manual skills correspond to work equipment that is not IT-supported (cell (5)), such as voice teaching, group activities, or analogue communication tools. For simplicity, we refer to these as service tools below.

The next category, *digital* skills, refers to those that enable workers to interact with and control digital technologies that support or indirectly manage the work process. In service occupations, this includes skills such as using customer relationship management software, operating enterprise resource planning systems, or using reservation software (cell (3)). In production contexts, digital skills include operating computer numerical control systems, fire alarm technologies, or computer-aided design tools (cell (4)). These skills reflect the computerization wave of the Third Industrial Revolution (First Digital Revolution) that started in the 1970s and enabled IT-based automation of specific sub-processes.

The highest level, *frontier* skills, refers to the Fourth Industrial Revolution (Second Digital Revolution) since the late 2000s. In service occupations, frontier skills include advanced capabilities such as applying machine learning techniques, operating AI-powered chatbots, or conducting fraud detection (cell (1)). In production settings, they involve working with self-controlled technologies such as collaborative robots, digital twin systems,

[5] An exception are surveys on actual technology use at the individual level, such as Arntz et al. (2025a). Because large-scale administrative or household data lack such measures, most empirical studies rely on occupational exposure indices constructed from external sources (e.g., Felten et al. (2018, 2019); Webb (2020)) or on job-posting based indicators of AI adoption (e.g., Engberg et al. (2024)).

[6] While O*NET contains broad skill descriptors, these rarely reference specific technologies. Although O*NET also includes a separate "Tools and Technology" component, these entries are not part of the skill taxonomy and remain less detailed than the occupation-specific skill text in BERUFENET. This richer textual information is essential for our machine-learning classification of technology levels. Moreover, BERUFENET corresponds directly to the German occupational structure used in our analysis.

[7] With production, we mean job, task, or work environment related to factories, blue-collar work, engineering, manufacturing, machinery, transport or logistics. All other jobs are service-related, including office work environments and services like administration, healthcare, education, retail, hospitality, finance, and information technology.

[8] The First Industrial Revolution (starting around 1800) marks the transition from hand production to the widespread use of machines powered by water and steam. The Second Industrial Revolution (starting in the late 19th century) saw the introduction of electricity-powered mass production and assembly lines.

Table 1: Skill levels by technology intensity and examples

Skill Level	Service	Production
Frontier	**(1) IT-integrated**	**(2) Self-controlled**
Skills related to technologies that perform work processes autonomously	Machine Learning Fraud detection Chatbots	Collaborative robots Digital twin technology Autonomous driving systems
Digital	**(3) IT-supported**	**(4) Indirectly controlled**
Skills related to technologies where humans are indirectly involved in work processes	Customer relationship management Reservation systems Enterprise resource planning	Computer numerical control Fire alarm systems Computer-aided design
Manual	**(5) Not IT-supported**	**(6) Manually controlled**
Skills related to technologies where humans conduct work processes manually	Voice teaching Group dancing Social therapy	Rust removal Tile laying Roof and facade metalwork

Notes: Building on the framework introduced in Genz et al. (2021), we classify skills according to three technology levels: manual, digital, and frontier technologies. Numbers (1)–(6) label the six conceptual categories discussed in the text.

or autonomous driving technologies (cell (2)). These are skills that allow workers to engage with technologies capable of autonomously performing complex work processes.

Note that this taxonomy classifies all occupational skills into one of the three groups; the "manual" category therefore also functions as a residual group that absorbs skills not involving digital or frontier technologies, such as social, communicative, or managerial skills, which are outside the scope of our technology-oriented classification.

The purpose of this taxonomy is not to offer a comprehensive typology of all occupational skills, but rather to distinguish skills that require digital or frontier technologies from those that do not. This focus allows us to construct a meaningful measure of the *technology intensity* of occupations. At the same time, the classification does not capture non-technological skill dimensions – such as social, managerial, or creative skills – which remain grouped in the residual manual category.

Expert database on job skills BERUFENET is a publicly accessible online platform provided by the German Federal Employment Agency, offering comprehensive information on all recognized occupations in Germany.[9] Primarily designed to support career guidance, it is widely used by vocational counselors and job placement officers during job placement procedures of job seekers as well as by the general public for occupational orientation. In 2023, the database contains detailed descriptions of 4,647 unique occupations (German: Einzelberuf), corresponding to the 8-digit level of the German Classification of Occupations

[9]The database can be accessed online via https://berufenet.arbeitsagentur.de/

(KldB2010),[10] including information on, among others, occupational job *competencies* – only the titles – describing the specific tasks, tools, and technologies used in each occupation. For simplicity, we refer to these competencies as *skills* or skill titles throughout the paper. Content is maintained and regularly updated by a dedicated expert team in response to update requests stemming from the Federal Employment Agency's advisory activities.[11]

For this paper, we focus on the job skills provided for each of the 4,647 unique occupations. In 2012 and 2023, the database contains 9,031 unique valid skills that are assigned to the unique occupations either as *core skills*, *additional skills* and *further skills*.[12] *Core skills* refer to essential skills that are indispensable for performing the duties of a given job. *Additional skills* denote skills that may be beneficial for carrying out the duties but are not strictly necessary for employment in the job. Note that a skill may be a *core* skill in one job, while being an *additional* (or *further*) skill in another job. For instance, the occupation *data scientist* contains 22 core job skills such as "applied computer science", "data lake", "data warehouse", "programming", and "project management" as well as 14 additional skills including "business intelligence", "IT coordination", and "machine learning" (see Table 2). When we construct our technology measures, we will stick to skills listed as core or additional in jobs, as *further skills* mainly consist of extensive, internal reference lists (e.g. drawn from related occupations) and do not reflect the central skill requirements of a job.

While we cannot weight skills by their relative importance, the extensive skill lists provided in BERUFENET contain substantial variation. Occupations differ widely in the number and mix of technology-related skills they include. For example, the occupation of data scientist contains 36 skills, of which 26 can be later classified as frontier (see Section 3). Such differences in composition provide informative variation for identifying technology-skill intensity even without explicit importance weights.

In the BERUFENET database, skills are organized using the Digital Competence Code (DKZ), a hierarchical classification system that assigns each skill to increasingly specific semantic categories. A schematic illustration of this structure in the DKZ is provided in Appendix Figure A.1. For example, the skills "Zoom" and "Microsoft Teams" both belong to the *skill field* (6-digit) "Office/Communication Software". This skill field is part of the *skill area* (4-digit) "Software – Commercial Application", which in turn falls under the

[10]For details on the structure of the German Classification of Occupations see Matthes and Paulus (2013).

[11]Due to its completeness and timeliness, BERUFENET has been used in prior research to derive occupational task profiles (Dengler et al., 2014) and assess the substitutability of occupations in the context of digitalization and automation (Grienberger et al., 2024). For a detailed description of the structured process by a dedicated team of experts specializing in education and vocational training to keep the BERUFENET database up to date, see Dengler and Matthes (2018).

[12]We ignore artist entries that capture specific works in the skill group *Media, Arts, Design* (e.g. named operas), which are explicitly labelled in the database and can therefore be easily dropped, see Appendix Table A.1.

Table 2: Skills for the occupation *data scientist*

Core	**Additional**
Applied computer science	Business administration
Content recommendation systems	Business informatics
Continuous integration	Business intelligence
Data analysis	High-performance computing (HPC/parallel computers)
Data architecture	IT coordination
Data lake	IT organization
Data modeling	Knowledge management systems
Data protection	Machine learning
Data transfer, data preparation	Market research
Data warehouse	Modeling, simulation (IT)
Database administration, management and organization	Network administration, management, organization
E-commerce, e-business	System support, system administration, system management
Information and communication management	Teaching (university)
Internet, intranet technology	User consulting, user support (IT)
Mathematics	
Programming	
Project management	
Research, information procurement	
Security systems (IT), data security	
Software technology, software engineering	
Statistics	
System software (development, programming, analysis)	

Notes: The table shows the competencies for the occupation data scientist (KldB2010 8-digit=43104132) as of 2023. See https://web.arbeitsagentur.de/berufenet/beruf/129987#taetigkeit_kompetenzen_kompetenzen for the list of skills required for a data scientist as of today.

broader *skill group* (2-digit) “IT, Data Processing, Computers”. This multi-level structure provides a consistent semantic context, ensuring that related digital tools and software applications are grouped together while remaining distinguishable from other domains such as “financial accounting software” or “technical application software”.

We follow Dengler et al. (2014) and drop skills assigned to the skill group related to types of work (“K 12 - Arbeits-, Einsatzformen”), workplace setting (“K 16 - Arbeitsorte”), and industries (“K 17 - Branchen”). Examples of such skills include “supermarket” (workplace) and “traineeship” (work form). As these categories do not convey information about the actual work activities, they are not relevant for our analysis and are therefore omitted. These account for 256 skills that we drop, ending up with 8,775 skills in either 2012 or 2023 (see Appendix Table A.1).

Beyond its hierarchical DKZ structure, the BERUFENET database also provides frequently associated search terms for each individual skill, offering additional semantic context that is not captured by the classification alone. For example, the skill *Chatbots* is linked to search terms such as “AI agents”, “Chatbots”, “LLM – Large Language Model”, and “Generative AI”. These related terms help users understand the broader conceptual environment of the skill and improve navigation within the BERUFENET system by guiding them toward occupations and skill profiles connected to these terms. The bottom layer of Appendix Figure A.1 illustrates how such search terms extend the hierarchical DKZ representation for individual skills.

We use the skill groups (2-digit), skill areas (4-digit), and search terms to enhance the interpretative accuracy in translation and description generation tasks using GenAI (next paragraph). For instance, the skill *tango* may refer to either a dance style or a control system for scientific instrumentation. Supplementing the input with the search terms provided by the system, as well as skill group and area, yields the additional contextual keywords “media”, “art”, “design”, and “dance”, which help disambiguate the meaning, guiding the model toward the intended interpretation of *tango* as a dance style. The skill groups and areas are also used as input variables in the classification procedure to capture general differences (Section 3).

Although the BERUFENET database offers extensive coverage of occupational skills, it provides only short skill titles without standardized definitions. This makes it difficult to understand the technological content of a skill based on the title alone. To address this, we use a large language model to generate short, consistent descriptions for each skill, which we call LLM skill descriptions. These LLM descriptions add the semantic detail needed to interpret skills correctly and also improve the performance of our keyword-based indicators. Without this additional context, many digital or frontier technologies – such as cloud-based environments like Microsoft Azure or specific SAP modules – would not be identifiable from titles alone (see discussion in Section 3). The LLM-generated descriptions therefore supply essential semantic information that BERUFENET does not provide.

Administrative data on occupations We use the IAB Employment History (BeH, Beschaeftigtenhistorik V10.08.00-202312) to analyze employment at the 5-digit level of German Classification of Occupations (KldB2010). The BeH is a comprehensive administrative dataset maintained by the German Federal Employment Agency, covering employment spells from 1975 through 2023 for all individuals in jobs subject to social insurance contributions. It is based on mandatory employer notifications submitted annually, as well as employment registrations and deregistrations required for social security purposes. Consequently, the dataset excludes civil servants, self-employed, military personnel, and individuals in phased retirement programs.

Occupations are coded using the 5-digit KldB2010, where the first four digits describe specific tasks and required skills, and the fifth digit indicates job complexity (e.g., helper or expert level). In addition to occupational data, we utilize establishment-level information on the registered sector of employment. Our sample consists of all individuals aged 16 to 65 in regular employment in 2023, resulting in a total of 30,361,893 individuals.

When linking our detailed skill information from BERUFENET to the administrative employment data, we are constrained to the 5-digit KldB2010 classification, which comprises 1,297 unique occupations, as employment data at a more granular occupational level are not available.

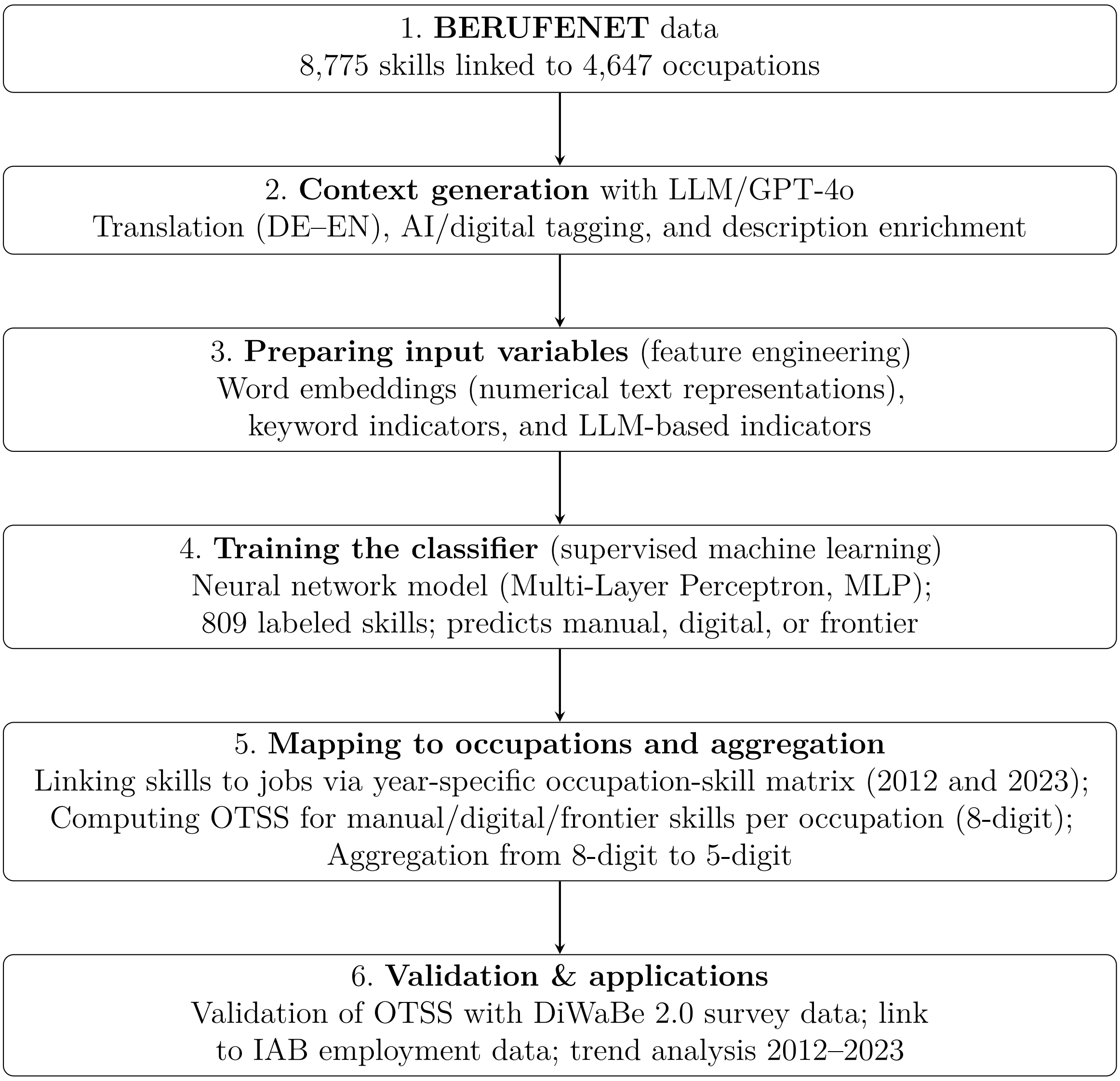
3 Classifying job skills with machine learning

The aim of this paper is to rank occupations by their technology skill intensity using detailed job skills from the BERUFENET database. To this end, we classify each job skill into one of the three technology classes – manual, digital, or frontier – defined in Table 1 and differentiate by service and production work contexts. We adopt a supervised learning approach by training a neural network classifier to predict the technology class of each skill based on a set of input variables (called *features* in machine learning) described below. Figure 1 provides an overview of the full workflow, summarizing how the data sources, contextual enrichment, and machine learning steps are connected in the skill classification pipeline that produces the OTSS. The detailed workflow is presented in the following discussion. We conduct these steps for both years 2012 and 2023.

Generating additional context using large language models While BERUFENET provides the official list of skills associated with each occupation (*Step 1* in Figure 1), it contains only short skill titles without standardized definitions. To make these skill titles usable for text-based machine-learning methods, we employ an LLM to enrich the raw BERUFENET entries with additional, standardized context (*Step 2*) including (i) a concise *LLM description* for each skill and (ii) structured indicators of its technological content (digital, AI, Industry 4.0, frontier, IT-integrated or self-controlled). These supplement keyword-based indicators and BERUFENET metadata (skill group and area) in the set of input variables (features) used by the classifier.

Specifically, we use OpenAI's large language model *gpt-4o* to iterate over the 8,775 skills. As a first step, each skill is translated from German to English, providing the LLM with contextual information (search terms, skill group, and skill area). For each translated skill, the LLM is then prompted to generate additional information along the following dimensions: (i) whether the skill is directly related to digital technology; (ii) if digital, we ask for a concise description of the technology-related skill, if not, a general description of the skill is requested (*LLM description*); whether the skill is related to (iii) frontier technology; (iv) technologies related to Industry 4.0; (v) artificial intelligence; (vi) self-controlled production technology or IT-integrated office and communication technology. A final prompt (vii) asks for an assessment of whether the skill is related to production jobs or tasks. Prompts are executed within a simulated chat session to preserve contextual coherence with the initial response and are initiated with system roles targeted to the

Figure 1: Overview of the AI-powered skill classification pipeline and OTSS construction



Notes: The figure summarizes the workflow from data collection and contextual enrichment to supervised machine learning classification of skills, construction of OTSS, and validation for 2012 and 2023.

subsequent tasks, such as a translator or job expert. We ask for a short explanation, as the accuracy of the model improves, when it has to think about a justification. The full prompt together with the system roles and context prompts are detailed in Appendix Table A.2.

Importantly, our final classification does *not* rely solely on the LLMs classification into manual, digital, or frontier skill categories. Instead, it generates consistent and interpretable inputs that serve as input features for the supervised learning model, which then learns how to predict the technology class based on the input. In this way, the LLM enriches the raw BERUFENET titles by producing detailed descriptions and answering targeted questions that translate each skill into structured indicators. This workflow combines the strengths of the two data sources in a complementary manner. BERUFENET provides the mapping of skills to occupations, which remains unchanged, while the LLM supplies semantic detail and additional indicators. This enrichment enables the supervised model to distinguish between skills that cannot be differentiated on the basis of titles alone.

Preparing input variables In *Step 3*, we construct the input variables for the classifier, starting with LLM descriptions. These descriptions provide additional context beyond the short skill titles available in BERUFENET, clarifying what each skill actually involves. For instance, relying only on the raw text of a skill title might place all Microsoft products in the same group because of the shared word "Microsoft". However, *Microsoft Windows* and *Microsoft Azure* differ strongly with respect to technology: whereas Windows is a digital technology, Azure is a cloud computing environment for AI and big-data applications – and therefore classified as a frontier technology. By using the LLM descriptions, we can distinguish such differences more accurately.

Before analyzing the LLM descriptions, we apply standard text-preprocessing steps that prepare the text for quantitative analysis. These include splitting the text into individual units (*tokenization*), removing very frequent function words that carry little meaning (*stop-word removal*), and reducing different forms of a word to a common base form (*lemmatization* or *stemming*). These procedures ensure that the classifier captures the semantic content of the descriptions rather than linguistic variation. We then convert the cleaned text into numerical representations that reflect semantic similarity (*word embeddings*) using the Word2Vec model.[13] Besides the text-based information, we include a number of binary indicators that capture whether the LLM classified a skill as digital, frontier, AI-related, Industry 4.0-related, self-controlled vs. IT-integrated, and production-related (see Section 2). By *production*, we mean any job, task, or work environment related to factories, blue-collar work, engineering, manufacturing, machinery, transport, or logistics.

Keywords. We also add variables indicating the presence and number of specific keywords in the LLM-generated skill descriptions that are associated with digital and frontier technologies. To do so, we construct manually curated dictionaries comprising 64 digital-technology keywords (e.g., information system, software, Windows, CNC, computer) and 238 frontier-technology keywords (e.g., 3D printing, artificial intelligence, machine learning, augmented reality, sensors, smart). These dictionaries were developed through an iterative process: starting from our conceptual definitions of digital and frontier technologies (see Table 1), we reviewed BERUFENET skill titles together with the enriched LLM descriptions and collected terms that consistently appeared in connection with the corresponding technologies. We refined the lists through repeated testing, removing overly

[13]Word embeddings map words or short texts into a high-dimensional vector space in which semantically similar terms receive similar numerical representations, enabling the classifier to capture semantic structure in the descriptions.

broad or ambiguous terms and adding terms where initial classifications suggested under-coverage. Based on the cleaned LLM descriptions, we count the occurrences of these keywords per skill. The distribution of frontier-technology keyword matches is summarized in Appendix Table A.3. From this, we derive binary indicators for the presence of digital and frontier keywords, as well as count variables capturing the number of matched terms in each category.

In addition to text-based features and LLM-generated indicators, we include categorical metadata from BERUFENET. Specifically, we use dummy variables for the official BERUFENET skill areas (and, where applicable, skill groups), which capture the occupational structure not contained in the text descriptions.

Overall, this yields a comprehensive, data-driven set of input variables from which the model learns the relative importance of each component for predicting a skill's technology class. Table 3 provides illustrative observations of how the LLM-generated indicators are transformed into the structured information used by the classifier.

Training the neural network classifier In *Step 4*, we classify job skills into technology categories using a *neural network model* (technically, a *Multi-Layer Perceptron*, MLP) implemented in Python's `scikit-learn` library.[14] This supervised learning approach estimates patterns between the input variables X (describing each skill) and the target variable y (its technology class). In our setting, the target variable distinguishes between three categories: *manually controlled (manual)*, *indirectly controlled (digital)*, and *self-controlled (frontier)* skills in production contexts, or equivalently, *non-IT supported (manual)*, *IT-supported (digital)*, and *IT-integrated (frontier)* skills in service contexts.

Compared with a conventional logistic regression, the neural network can capture more complex, non-linear relationships by including one or more intermediate (*hidden*) layers between the input and output. The construction of the input variables X is described in the following subsection.

[14] A Multi-Layer Perceptron is a simple type of neural network that learns non-linear relationships between input variables and an outcome through an iterative training process known as backpropagation.

Table 3: Examples of GenAI-generated skill descriptions and technology indicators

		GenAI-generated indicators						
Original skill title (BERUFENET)	Skill description generated with GenAI	Digital	Frontier	AI	Ind. 4.0	IT-int.	Self-con.	Prod.
Speechwriting	Speechwriting is the skill of crafting speeches that effectively communicate a message to an audience. It involves understanding the purpose of the speech, the audience's expectations, and the speaker's style and voice. A speechwriter must be adept at organizing ideas, using persuasive language, and incorporating rhetorical devices to engage listeners. This competency requires strong research skills to gather relevant information and the ability to tailor content to fit the context and occasion. Additionally, speechwriting often involves collaboration with the speaker to ensure the speech aligns with their personal delivery style and objectives.	0	0	0	0	0	0	0
Project planning software MS Project (MS Office)	Project planning software MS Project (MS Office) is a digital tool used for project management that allows users to plan, schedule, and manage projects effectively. As of 2023, it provides features such as task assignment, resource allocation, progress tracking, and timeline visualization through Gantt charts. MS Project integrates with other Microsoft Office applications, enabling seamless data sharing and collaboration. It supports project managers in organizing tasks, setting deadlines, and monitoring project performance to ensure timely and efficient project completion.	1	0	0	0	1	0	0
Machine learning	Machine Learning is a subset of artificial intelligence that involves the use of algorithms and statistical models to enable computers to improve their performance on a specific task through experience and data. In 2023, it was directly related to digital technology as it relied heavily on electronic devices such as computers and servers to process large datasets, develop predictive models, and automate decision-making processes. Machine Learning applications included natural language processing, image recognition, and recommendation systems, all of which utilized software and hardware components to analyze and interpret data efficiently.	1	1	1	1	1	0	0
Autonomous driving systems	Autonomous Driving Systems is a digital technology that involves the use of electronic devices and software to enable vehicles to navigate and operate without human intervention. In 2023, this technology integrated various components such as sensors, cameras, radar, and LIDAR to perceive the environment, along with advanced algorithms and machine learning models to process data and make driving decisions. These systems were capable of tasks like lane keeping, adaptive cruise control, and traffic sign recognition, contributing to the development of self-driving cars that could perform certain driving functions autonomously under specific conditions.	1	1	1	1	0	1	1

Note: The table illustrates how generative AI (GenAI) enriches short skill titles by generating standardized textual descriptions and structured indicators capturing technological content. Indicators denote whether a skill is digital, frontier, AI-related, associated with Industry 4.0 (Ind. 4.0), IT-integrated (IT-int.), self-controlled (Self-con.), or related to production contexts (Prod.). For simplicity, the table reports only a subset of the inputs used in the supervised classification model.

Training. To train the classifier, we manually label 809 out of 8,775 job skills according to the three technology classes, distinguishing between service and production contexts, i.e. overall six groups. Labeling is based on expert judgment supported by the LLM-generated descriptions and adheres to the framework introduced in Section 2. While assigning skills to technology categories inevitably involves some degree of judgment, we restrict the labeled training set to skills whose technological content can be classified with relatively high confidence in order to ensure a high-quality GroundTruth dataset, but make sure we include skills from all broader skill groups (see Appendix Table A.4 for the distribution of labeled skills across skill groups). To give a few illustrative examples drawn from different occupations and skill groups, manual skills include *sealing (roof)*, *pipe construction*, and *wood sorting* in production, and *infants and toddlers care*, *art education*, and *social therapy* in services. Digital skills include indirectly controlled production skills such as *fire alarm technology* and *CAD application Nemetschek Allplan*, as well as IT-supported service skills such as *email program Pegasus Mail* and *web content management system (CMS) Mambo / Joomla*. Frontier skills relate to technologies that perform work processes autonomously, including self-controlled production technologies such as *autonomous transport systems*, *LiDAR technologies*, and *driver assistance systems*, and IT-integrated service technologies such as *Microsoft Azure cloud computing*, *computer science*, and *programmatic advertising*.

Using this labeled subset, we train the MLP classifier to predict the technology classes based on the feature set. For model evaluation, we randomly split the labeled data into a training and a test set of equal size, setting aside 50% of the labeled skills (405 observations) as a held-out test set that is not used during training.

Evaluation. We evaluate model performance using the 405 manually labeled job skills in the test set by comparing the model's predictions with the true class labels. Appendix Table A.5 reports standard measures of prediction quality commonly used in classification tasks: precision, recall, and the combined F1-score. Precision measures the share of skills predicted to belong to a given technology class that are correctly classified, that is, for which the predicted class matches the true label. Recall measures the share of skills that truly belong to a given technology class that are correctly identified by the model. The F1-score combines precision and recall into a single balanced measure that penalizes imbalances between the two and is high only when both precision and recall are high. Overall, the model performs strongly: high precision values indicate that the model makes few incorrect positive predictions (false positives), while high recall values show that it successfully identifies most relevant skills (i.e. the model makes few false negatives). The macro-averaged F1-score is close to 0.8, indicating a strong and well-balanced classification performance across technology classes.

To provide evidence on the value of the multi-step approach more generally, Table 4 compares alternative classification strategies. Panel A benchmarks each method against the full set of human-labeled skills, which serve as the ground truth in this exercise. We focus on the 3-class scheme pooling service and production skills within manual, digital and frontier. The results show that applying text-mining directly to short skill titles performs substantially worse than applying the same rules to LLM-generated descriptions, particularly for frontier technologies. This demonstrates that the additional semantic context provided by the LLM is essential for identifying advanced technologies that are not explicitly named in the raw titles. Moving beyond rule-based methods, an AI-powered classifier trained on text embeddings of skill titles already performs markedly better. However, the highest agreement with human labels is achieved by the full AI-powered classifier, which combines text embeddings of LLM-generated skill descriptions with additional structured, non-embedding features.

Table 4: Performance of text-mining and AI-powered classifiers (3-class scheme)

	Accuracy (1)	Cohen's κ (2)	Recall (Manual) (3)	Recall (Digital) (4)	Recall (Frontier) (5)
Panel A: Rule-based text-mining					
Textmining (titles)	0.637	0.434	0.993	0.420	0.418
Textmining (LLM-generated descriptions)	0.760	0.639	0.827	0.871	0.552
Panel B: AI-powered classification (MLP)					
AI-powered classifier (title embeddings only)	0.860	0.790	0.876	0.837	0.866
AI-powered classifier (descriptions + structured features)	0.932	0.897	0.971	0.920	0.895

Notes: The table reports agreement between alternative classification strategies and human-labeled technology classes using a three-class scheme (Manual, Digital, Frontier). Accuracy denotes the share of correctly classified observations. Cohen's κ measures agreement adjusted for chance. Recall is reported separately for each technology class.

Feature importance analysis for the full AI-powered classifier (Appendix Figure A.2) assesses the contribution of each input by measuring the decline in predictive performance under permutation. Features related to the skill descriptions are aggregated to one block. The results show that textual embeddings derived from the LLM-generated skill descriptions dominate the model's predictions, accounting for by far the largest drop in performance. All other features play a secondary role: among the structured inputs, skill-area dummies contribute the most, followed by indicators related to production and frontier technologies and the presence of corresponding keywords. Variables capturing digital and AI-related content add modest predictive value, while Industry 4.0, IT-integrated, and self-controlled technology indicators exhibit near-zero or slightly negative importance, suggesting limited incremental contribution once textual information is taken into account.

Prediction. Using the trained classifier, we predict the technology class for each unique skill in the full dataset and assign it to one of the three technology classes: frontier, digital, or manual. Figure A.3 in Appendix A.3 illustrates the distribution of these predicted classifications for the skills that appear as either core or additional in any occupation in 2023.

Examples of skills identified as either frontier or digital are presented in Appendix Table A.6. In line with existing classification approaches, skills directly associated with AI, such as "augmented reality", "machine learning", and "data lake", are categorized as frontier skills. However, our classification also captures frontier technologies more broadly, encompassing skills such as "maintenance and repair robots", "computer-assisted surgery", and "networked production systems". Representative digital skills include "bookkeeping", "CNC programming" and "office management".

4 Constructing the Occupational Technology Skill Share (OTSS)

This section implements *Step 5* of the skill-classification pipeline depicted in Figure 1 by linking the skill-level technology classification developed in Section 3 to (8-digit) occupations using the year-specific occupation–skill matrices provided by the German Federal Employment Agency.[15] This allows us to quantify the technology intensity of jobs and show how digital and frontier technologies are transforming job profiles. From the 8,775 classified skills, we focus only on skills that ever appear as core or additional skills in any occupation – in 2023 this amounts to 5,065 (see full skill selection steps in Appendix A.1 for 2012 and 2023). To validate the resulting occupational skill intensity measures, we compare aggregated values in occupation groups (2-digit occupations) with survey data on actual technology use at the workplace, confirming the reliability of our approach. The details are discussed below.

Occupational skill intensities Formally, combining the skill-level technology classification with the occupation–skill matrix allows us to compute, for each unique 8-digit occupation, the share of skills belonging to technology class i in occupation j, which we refer to as the Occupational Technology Skill Share (OTSS):

$$\mathrm{OTSS}_{ij} = \frac{\text{Number of skills assigned to technology class } i \text{ and occupation } j}{\text{Total number of assigned skills in occupation } j} \tag{1}$$

where $i = \{manual, digital, frontier\}$. That is, the numerator counts the number of skills associated with occupation j that are classified (based on the model predictions) as belonging to technology class i, while the denominator captures the total number of skills

[15]The occupation–skill matrices used throughout this section corresponds to the versions available as of 19 December 2023 and 12 December 2012.

linked to occupation j. For each occupation, we calculate the proportion of required skills associated with each technology class relative to the total number of assigned skills.

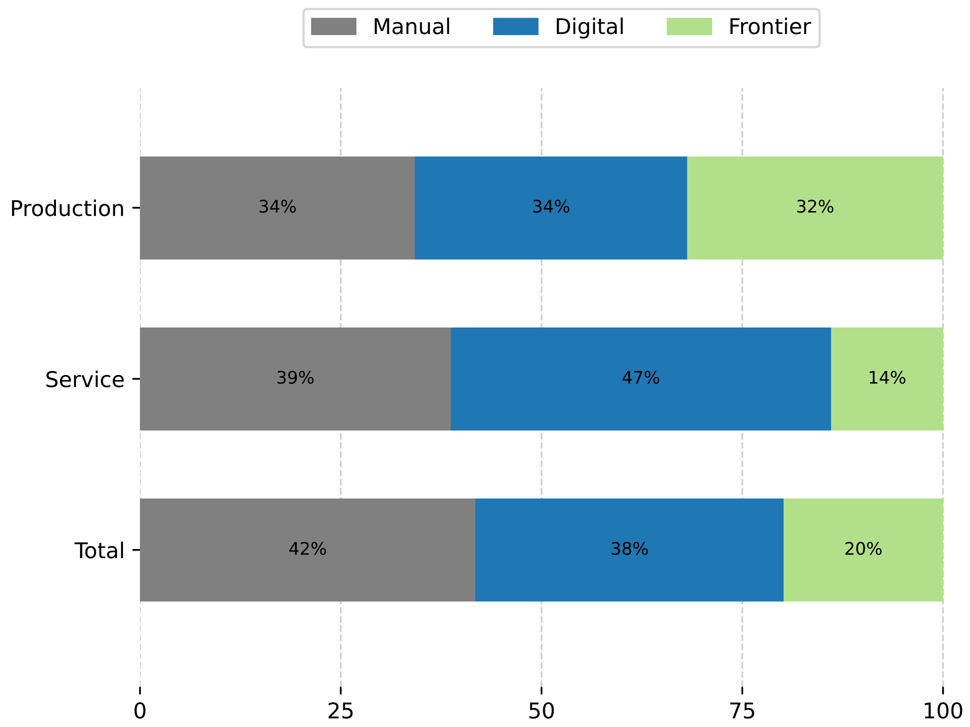
Returning to the example of the *data scientist* for which Table 2 displays all skills. Based on our classification, the resulting OTSS values are: 72% frontier skills, 28% digital skills and no manual skills, placing the occupation in the highest percentile of the frontier skill distribution. In total, 26 skills are classified as frontier skills (e.g., 'high-performance-computing', 'data lake' and 'machine learning') and 10 skills are classified as digital skills (e.g., 'project management', 'IT coordination' and 'system management'). A key advantage of our skill classification is its broader coverage of cutting-edge frontier technologies beyond artificial intelligence alone. As a result, occupations without a primary AI or IT focus can rank highly in terms of frontier skill intensity. One such example is *advanced driver assistance systems / autonomous driving engineer* (KldB2010 27184136), with a skill composition of 88% frontier skills and 12% digital skills. Frontier skills of this occupation include, for instance, 'data analysis', '5G technology', and 'machine-to-machine-communication'.

An example of an occupation with one of the highest shares of digital skills is the *customer information specialist* (KldB2010 71401147), whose skill profile comprises 8% frontier skills, 92% digital skills and no manual skills. Representative digital skills for this occupation include, for instance, 'customer data management', 'mail processing', and 'managing files, documents'.

Occupational skill composition To characterize the technological content of work in the economy, we construct employment-weighted averages of our occupational skill shares described above. Because employment counts are not observed at the 8-digit level, we combine the skill measures with employment information available at the 5-digit occupational level using a two-step procedure. First, we collapse skill shares from the 8-digit to the 5-digit occupational level by taking simple averages across all 8-digit occupations belonging to a given 5-digit occupation. This step implicitly assumes that employment is evenly distributed across 8-digit sub-occupations within each 5-digit occupation. Second, we aggregate these 5-digit skill measures to the economy level using employment weights.

Figure 2 illustrates the resulting employment-weighted OTSS for the average worker in 2023. The figure displays three stacked horizontal bars corresponding to overall technology type, production-related technology skills, and service-related technology skills. Within each bar, gray segments represent manual skills, blue segments represent digital skills, and green segments represent frontier skills. All bars are normalized to sum to 100 percent. The figure is based on the universe of 5,065 distinct classified skills that appear as either core or additional skills in occupations existing in 2023. Across all skills, 20% of the skill content of work is frontier-related, 38% is digital, and 42% is manual. The distribution of technology intensity differs markedly between production and service contexts. Production-

Figure 2: Average occupational OTSS in 2023 (employment-weighted)



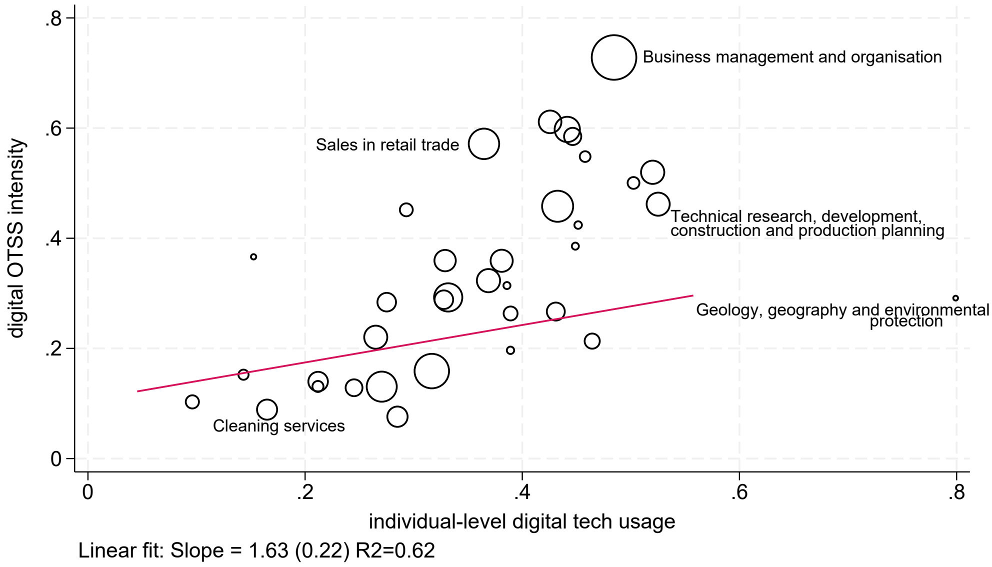
Notes: The figure shows the employment-weighted OTSS of an average worker in the economy. OTSS values are constructed at the 8-digit occupational level and aggregated using employment data at the 5-digit level. The bars report the distribution of manual, digital, and frontier skills *Overall* as well as by technologies related to *Production* and *Service*. All bars are normalized to sum to 100 percent. The analysis is based on the 5,065 distinct skills that appear as either core or additional skills in at least one occupation existing in 2023. We also show the unweighted version of this figure which represents the skill composition of the average occupation in the German economy in Appendix Figure A.4.

related tasks exhibit a substantially higher frontier OTSS (32%) than services (14%), indicating greater reliance on highly advanced or self-controlled technologies. By contrast, service-related tasks are dominated by digital skills (47%) and display comparatively lower frontier OTSS, suggesting that digitalization in services is more often incremental rather than frontier-driven.

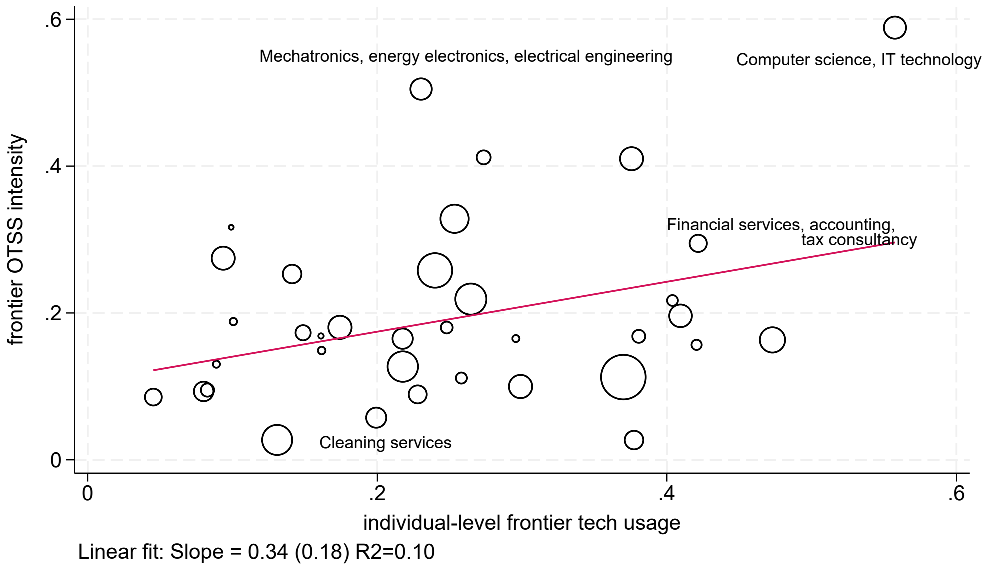
Comparison to survey data on technology use at the workplace As the final step of the empirical workflow illustrated in Figure 1 (*Step 6*), we validate our occupational OTSS against survey-based evidence on actual workplace technology use and apply them in combination with administrative employment data. Our digital and frontier OTSS constructed from BERUFENET reflect expert assessments of which skills are required for each occupation. To test whether these measures capture actual workplace technology use, we compare them with individual-level survey data on workplace technology usage at the

Figure 3: Comparing OTSS and self-reported workplace technology usage

(a) Digital skills



(b) Frontier skills



Notes: Panel (a) plots our digital OTSS as of 2023 against the self-reported digital technology usage shares at the workplace at the 2-digit KldB2010. Panel (b) our frontier OTSS as of 2023 against the self-reported frontier technology usage shares at the workplace at the 2-digit KldB2010. The size of the circles denotes the employment levels in each 2-digit KldB2010 occupation, which are also used as weights in the linear fit regression.

occupational level.

The DiWaBe 2.0 dataset comprises responses from approximately 9,800 employees in Germany, who were interviewed by telephone about their use of non-computerized (manual), digital, and intelligently connected (frontier) technologies at their workplaces, following a technology framework aligned with ours (explained in Table 1). Details of the survey data and the comparison between both measures are provided in Appendix Section A.5. Figure 3 displays the correlations between the two data sources at the 2-digit KldB2010 level: The upper panel shows the relationship between survey-based digital technology use and our digital skill intensity measure, while the lower panel presents the corresponding correlation for frontier technologies.

Despite differences in data sources and construction methodologies, both measures display statistically significant correlations with the survey data. The frontier OTSS correlates positively with the individual-level measure of frontier technology usage at the workplace (correlation coefficient 0.309, p-value<0.10). In the lower panel of Figure 3, occupations in computer science and IT technology lie in the upper-right part of the distribution, indicating agreement between both measures for these highly frontier-intensive occupations. This positive correlation provides supportive evidence that our measure of frontier skill requirements captures meaningful variation in the actual use of frontier technologies in the workplace.

A substantially stronger relationship (correlation coefficient 0.786, p-value<0.01) emerges when comparing our digital OTSS to the individual-level survey measure. The upper panel of Figure 3 shows a strong alignment between digital OTSS and reported usage of digital work equipment across occupations, with particularly high values observed in business management and organisation. This strong correlation provides further evidence for the reliability and consistency of our classification.

Overall, the occupational-level correlations suggest that our measures of digital and frontier skill intensity effectively capture the actual use of these technologies in the workplace. In the following section, we move beyond their use as aggregation weights and employ administrative employment data to analyze how digital and frontier skill intensities relate to employment levels and growth in the German labor market.

5 Frontier jobs in the German labor market

Building on the employment-weighted OTSS measure constructed above, this section uses granular administrative employment data to map the distribution of frontier and digital skills across the German labor market in 2023. We then compute the OTSS based on skill data from 2012 to trace the evolution of occupational skill requirements over the past decade. Finally, we relate changes in digital and frontier OTSS between 2012 and 2023 to employment dynamics in order to assess how employment growth varies with shifts in occupational skill composition.

Frontier occupations on the German labor market This subsection focuses on frontier-intensive occupations and combines the frontier OTSS with employment counts at the 5-digit KldB2010 level to identify which technologically advanced occupations are also of substantial labor market relevance. Table 5 presents the most common high-frontier-skill occupations (upper panel) and high-digital-skill occupations (lower panel), ranked by employment.

Table 5: Most common high-OTSS occupations by technology type (employment-weighted)

KldB2010	Occupation title	# employees	OTSS (%)		
5-digit			Manual	Digital	Frontier
*Most common occupations with highest **frontier technology** skill share*					
27104	Technical R&D occupations (highly complex)	218,216	8	34	58
43414	Software developers (highly complex)	203,338	5	31	64
26212	Construction electricians (skilled)	177,929	12	35	53
43224	IT application consultants (highly complex)	114,505	9	33	58
26312	ICT specialists (skilled)	103,412	9	38	54
26252	Industrial electrical equipment technicians (skilled)	97,482	8	37	55
43102	Computer science occupations (skilled)	92,890	4	35	61
43343	IT system administrators (complex)	91,830	4	40	56
43103	Computer science occupations (complex)	64,271	3	25	72
43223	IT application consultants (complex)	54,648	7	37	55
*Most common occupations with highest **digital technology** skill share*					
71402	Office clerks and secretaries (skilled)	1,487,732	11	82	7
71302	Business administration & management occupations (skilled)	771,916	14	80	6
72213	Accountants (complex)	230,517	24	69	8
71104	Managing directors & executive board members	216,197	16	65	19
71401	Office clerks and secretaries (un-/semi-skilled)	213,301	10	81	9
61122	Sales occupations, excl. ICT (skilled)	208,308	14	66	20
72302	Tax consultants (skilled)	110,901	24	76	0
61212	Wholesale & foreign trade management assistants (skilled)	103,368	14	74	12
61113	Purchasing specialists (complex)	103,115	7	70	22
71403	Office clerks and secretaries (complex)	100,246	13	79	8

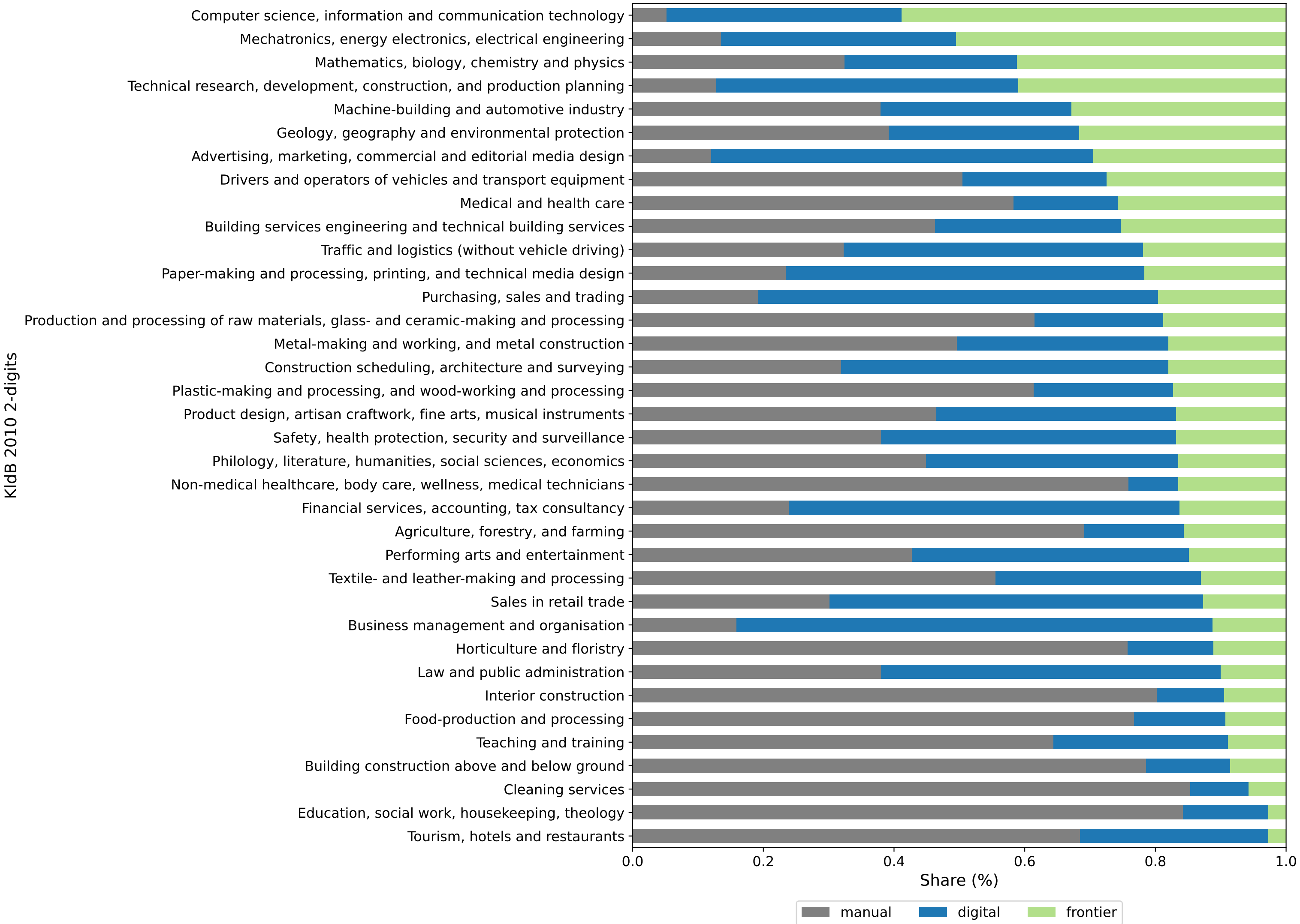
Notes: To identify the ten most common occupations among those with high frontier OTSS (upper panel) and digital OTSS (lower panel), we restrict attention to the 100 occupations with the highest frontier or digital OTSS at the KldB2010 5-digit level and rank them by employment size in descending order. Official english titles shortened for better readability.

The most common frontier occupations in Germany, based on employment size, are predominantly concentrated in technical, engineering and IT-related domains. The two most prevalent frontier-intensive occupations are technical research and development occupations with a frontier OTSS of 58% and occupations in software development with a frontier OTSS of 64%, followed by electricians in construction and occupations in IT-applications. Other prominent occupations include ICT specialists as well as computer science occupations and IT system administrations. These occupations employ between roughly 55,000 and over 218,000 individuals. Notably, these occupations tend to involve both skilled and (highly) complex tasks, as indicated by the fifth digit of their occupation code being either two, three or four, corresponding to skilled, complex or highly complex work requirements.

The lower panel of Table 5 shows that the most common digital occupations in Germany span a wider range of occupational domains and complexity levels, and tend to employ more individuals on average than the most common frontier occupations. At the top are office clerks and secretaries, with approximately 1.5 million employed individuals, followed by occupations in business administration (772,000 individuals) and occupations in accounting (230,000 individuals). Other prominent occupations include managing directors and executive board members, followed by further groups of office clerks and secretaries and sales occupations. Each of these most common digital occupations engages more than 100,000 individuals, highlighting the significant scale of digitally intensive roles in the German labor market. In contrast, the more limited presence of frontier-intensive occupations may reflect the relatively slower diffusion of frontier technologies up to 2023.

Figure 4 shows the employment-weighted average OTSS for manual (grey), digital (blue) and frontier (green) skills across occupational groups (2-digit KldB2010 occupations), ranked by the frontier OTSS. The highest share of frontier skills is consistently found among information technology and engineering occupations in Germany, in line with the more granular findings at the 5-digit KldB2010 presented in Table 5. In particular, the occupational group with the highest frontier OTSS are occupations in computer science, information and communication technology, with a frontier skill share of 59%.
This occupational group also relies strongly on digital skills, with 36%, while manual skills play only a minor role. This distribution underscores the centrality of digital skills in the IT domain, while also suggesting that many applications remain rooted in established digital technologies rather than involving cutting-edge technologies. Closely following are occupations in mechatronics, energy electronics and electrical engineering, with 50% of skills classified as frontier. These jobs are also digitally intensive (36%) and less manually intensive (14%). Scientific occupations in mathematics, biology, chemistry, and physics also show a high frontier OTSS (41%), though they stand out with a comparatively high manual component (32%). This combination likely reflects the continued importance of hands-on experimental and laboratory work alongside digital and computational tasks in scientific professions.

Figure 4: Ranking occupational groups by OTSS (KldB2010, 2-digit, employment-weighted)



Notes: The figure reports employment-weighted average OTSS of manual, digital, and frontier skills by occupational group at the 2-digit KldB2010 level. Aggregation is based on 5-digit KldB2010 occupations, with weights given by their employment levels. The employment numbers and the OTSS values by 2-digit KldB2010 level are also reported in Appendix Table A.7.

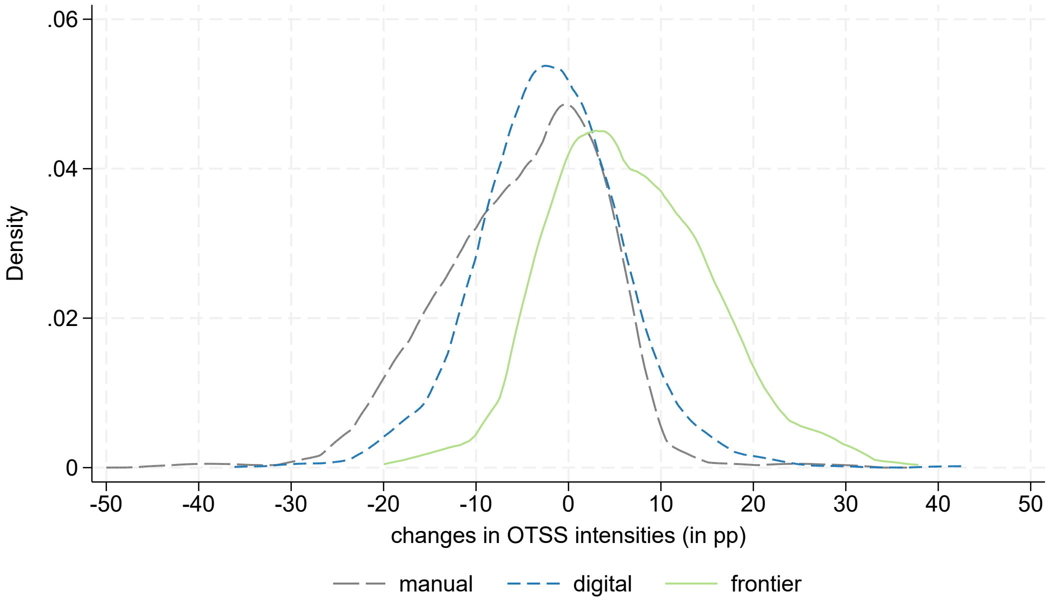
In contrast, several main occupational groups that are high in digital skill intensity rank considerably lower in terms of frontier skills. This pattern is particularly evident in administrative, commercial and service occupations, which often involve routine, desk-based tasks that do not require frontier technologies. Two examples for such occupational groups are business management and organisation occupations and financial services, accounting, and tax consultancy occupations. These two occupational groups are also among the largest in the German labor market, employing approximately 3.9 million and 1.3 million individuals, respectively (see Appendix Table A.7). At the opposite end of the spectrum, several craft and personal service occupations exhibit a dominant share of manual skills, with only minimal involvement of digital and frontier technologies. For example, cleaning service occupations contain 85% manual skills, placing them at the bottom of the ranking when occupational groups are sorted by frontier skill intensity.

Evolution of OTSS from 2012 to 2023 To illustrate how occupational skill requirements have evolved over time, we construct changes in technology skill intensities within each occupation from 2012 to 2023. For this purpose, we use the skill data from 2012 and compute the OTSS following the same methodology described in Section 3 and Section 4. Note that observed changes in OTSS at the 5-digit level may not only reflect changes in skill shares within occupations. Part of the shift may also arise from the emergence of new or redefined occupations (on the 8-digit level) by 2023. As a result, changes in the distribution of skill intensities over time capture both within-occupation change and compositional effects driven by occupational entry and updating.

Figure 5 presents kernel density estimates of employment-weighted changes in OTSS across 5-digit KldB2010 occupations between 2012 and 2023 for manual (gray), digital (blue), and frontier technologies (green). The distributions reveal a pronounced shift within 5-digit KldB2010 occupations: Manual OTSS is distributed clearly below zero, indicating a widespread decline in the relative share of manual skills within occupations (-4.8pp on average). Digital OTSS displays a modest leftward shift, suggesting a slight overall decline in its relative importance (-1.9pp on average). Frontier OTSS, in turn, exhibits a more positive distribution, pointing to a substantial rise in its relevance across most occupations (+6.7pp on average). Among the 1,181 5-digit KldB2010 occupations that existed in both years, the vast majority shows increasing frontier shares (914 occupations), while only few remain unchanged (53) or display declining frontier shares (214). These declines are concentrated mainly among IT-related occupations which already exhibited very high frontier OTSS in 2012. Overall, the evidence reflects a structural transformation in occupational skill profiles, marked by declining reliance on manual skills and a pronounced shift toward frontier skills, which is in line with ongoing technological advancement and workplace innovation.

Additional evidence of these changes is provided in Appendix Figure A.7, which

Figure 5: Distribution of OTSS changes 2012-2023 (employment-weighted)



Notes: The figure displays employment-weighted Epanechnikov kernel densities of the percentage point changes in the OTSS within 5-digit KldB2010 occupations.

compares the levels of occupational skill intensities in 2012 and 2023. The appendix figure shows that the decline in manual skills is concentrated among occupations with initially low to medium manual intensity, while highly manual-intensive occupations see only a slight decline. By contrast, frontier skill intensities exhibit a broad rightward shift across the entire distribution, indicating a widespread diffusion of frontier technologies rather than changes limited to already frontier-intensive occupations.

Changes in frontier skill intensity and employment growth To better understand how evolving skill requirements impact labor market dynamics, we examine the association between changes in OTSS and employment growth across occupations. A natural question in this context is whether increases in frontier OTSS reflect technologies that augment labor demand or substitute for it. Our approach does not impose any structural assumption on this: the classification identifies which skills are related to frontier technologies, but it does not determine how these technologies affect labor. We therefore interpret the relationship between changes in frontier OTSS and employment growth as reduced-form evidence on how occupations adjust when technology-related skills expand. A positive association would be consistent with augmenting effects, whereas a negative association would point toward substitution. The quadratic specification we estimate allows for both patterns to appear simultaneously, for example, if modest increases in frontier OTSS are complementary but very large increases are associated with substitution pressures.

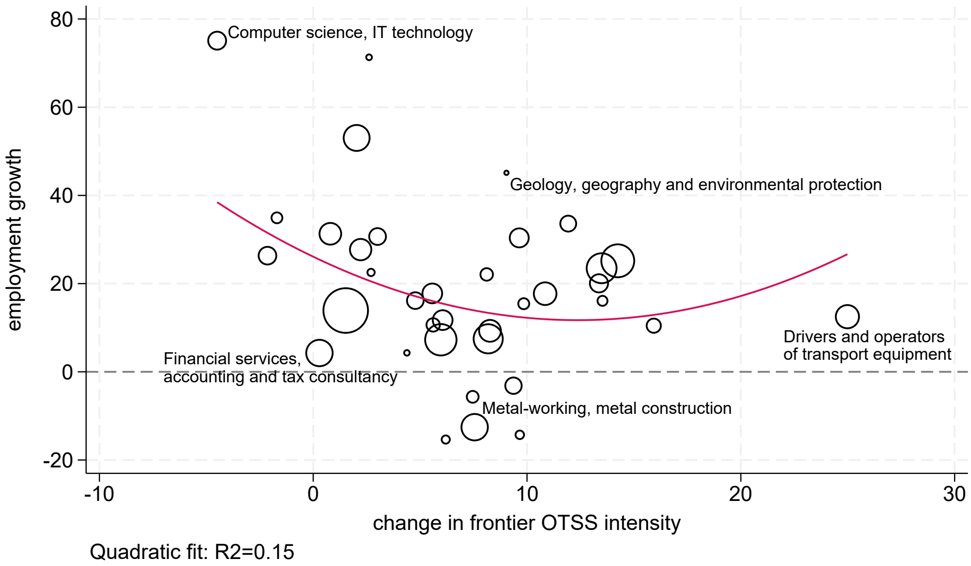
Figure 6 explores this relationship by plotting employment growth against changes in frontier OTSS between 2012 and 2023 for each occupational group at the 2-digit KldB2010 level. The figure reveals a polarization pattern: employment growth tends to be strongest among occupations at the extremes of the distribution, those with either very small or very large changes in frontier OTSS, while occupations with moderate changes in frontier OTSS tend to exhibit weaker, sometimes stagnant or even declining employment growth.

One example of an occupation group that experiences a large increase in frontier OTSS alongside moderate employment growth is the group *drivers and operators of vehicles and transport equipment*. For these occupations, the frontier OTSS increased from 2% in 2012 to 27% in 2023. Examples include *transport equipment operators* (KldB2010 52531100) and *forestry machine operators* (KldB2010 52512104). These jobs have experienced an increase in frontier OTSS driven by the introduction of advanced machinery, navigation, and automation technologies: for instance, frontier OTSS among transport equipment operators rose from 0% to 31%, while forestry machine operators saw an increase from about 6% to 26% between 2012 and 2023.

The strongest employment growth, however, is observed in *computer science and IT occupations*, which expanded from roughly 550,000 workers in 2012 to nearly 1 million in 2023. Despite this strong employment growth, these occupations display a moderate decrease in frontier OTSS. This decline does not reflect a decline in frontier technologies but rather a compositional effect driven by additional skills required in the occupations. Examples include *IT consultants* (KldB2010 43224101) and *network service technician* (KldB2010 43313100). In these two occupations, the number of frontier skills remains stable or increases slightly, while the total number of skills expands strongly due to the addition of digital skills. For instance, frontier skills for IT consultants increased from 15 to 16 skills between 2012 and 2023, yet total skills grew from 20 to 26 skills, leading to a decline in the frontier OTSS from 74% to 62%. Also, new occupations have entered with a relatively high digital skill shares, such a *business continuity manager* (KldB2010 43384111) or *digitalization management clerk* (KldB2010 43112131).

Overall, this section shows that combining our OTSS measures with employment data provides a nuanced picture of how frontier and digital technologies evolve in the German labour market. Frontier skills remain predominantly concentrated in technical and scientific occupations, while digital skills are more broadly distributed across sectors. Over the past decade, the expansion of frontier skill requirements has been accompanied by polarised employment growth: occupations with very small changes and very large increases in frontier OTSS experienced the strongest employment gains, whereas occupations with intermediate changes recorded weaker or even negative growth. This non-linear relationship is also reflected in a simple regression framework relating employment growth to changes in frontier OTSS using a linear and a quadratic term. Across the full sample as well as

Figure 6: Employment change against frontier OTSS (employment-weighted)



Notes: The figure plots the percentage change in employment against the percentage point changes in our emplyoment-weighted OTSS measure at the 2-digit KldB2010. The size of the circles denotes the employment numbers in each 2-digit KldB2010 occupation in the initial year 2012.

across subgroups by age, gender and qualification, the point estimates display a negative coefficient on the linear term and a positive coefficient on the squared term (see Table A.8), consistent with the U-shaped pattern visible in Figure 6. In the full sample regression, neither the linear nor the quadratic term is statistically significant at conventional levels. In contrast, the linear and squared terms are statistically significant in several subgroup regressions, in particular for younger (16-34) and older (45+) workers, for men and women, for workers without vocational training and for workers with a university degree. While purely descriptive, these results suggests that the observed polarisation does not appear to be driven by a specific demographic or qualification group, but represents a more general feature of occupational adjustment to frontier technologies. These trends highlight the ongoing structural transformation of job profiles and the growing relevance of frontier technologies for future workforce development.

6 Conclusion

This study introduces a novel, skill-based approach to measuring technological change in the German labor market, enabling a direct and granular assessment of how digital and frontier technologies reshape occupational skill requirements and employment dynamics. Our research proposes a new method to classify job skills into manual, digital, and frontier technology skills using generative AI and machine learning techniques in combination with rich textual data from the German BERUFENET database. We refer to this approach as an AI-powered skill classification: generative AI is used to enrich and contextualize skill descriptions, while the classification itself is performed by a supervised learning model, ensuring transparency, reproducibility, and systematic validation against alternative approaches. The classification allows us to construct new measures of occupational technology skill intensity – Occupational Technology Skill Shares (OTSS) – , which we link to administrative labor market data to capture the depth and direction of occupational transformation in terms of digital and frontier skills between 2012 and 2023.

Aggregating these measures using employment weights provides a worker-centered perspective on technology exposure. In 2023, manual OTSS accounts for the largest share of implemented skill content in the German workforce (42%), followed by digital OTSS (38%) and frontier OTSS (20%). Frontier skills are predominantly concentrated in technical, engineering, and IT-related occupations, while digital skills are more broadly distributed across a wider range of occupational domains, particularly administrative, commercial, and managerial occupations. However, over the past decade, frontier skills have increased substantially within almost all occupations, suggesting a structural shift in occupational requirements in the German labor market. At the same time, manual skill requirements have declined markedly, while digital skill requirements have remained broadly stable or slightly decreased on average, potentially reflecting the diminishing relevance of less-advanced digital skills in the face of rapid technological progress. The relationship between shifts in skill demand and employment growth is characterized by polarized pattern, where employment growth is relatively stronger in occupations experiencing large increases in frontier OTSS, as well as in occupations with comparatively limited changes in frontier OTSS. These patterns should be interpreted as descriptive evidence on changing skill requirements rather than as direct evidence on the employment or productivity effects of frontier technologies, which may differ across stages of adoption and over time.

For researchers and policymakers, our paper offers a novel, fine-grained measure of technological skill intensity to monitor and study shifts in occupational skill requirements across the labor market. Our methodology is adaptable beyond the German context and can be readily applied in other countries with detailed occupational databases – such as O*NET in the United States, or comparable European resources – to monitor technological transformation in near real time. Beyond its application to official occupational databases, the same classification approach can potentially be used to analyze the skills demanded in job advertisements, providing a timely indicator of changing employer needs and emerging technology trends. Because the enrichment is created through targeted prompts, the same framework can be adapted to other research questions, for example to identify "green" skills, without altering the underlying occupational classification system. By linking expert-based skill assessments and AI-powered classification, our methodology opens new avenues for studying skill mismatches, tracing the diffusion of new technologies, and supporting targeted education and training policy in an era of rapid technological transformation, while remaining flexible enough to be extended to other domains.

Finally, while our approach does not take a stand on whether frontier technologies augment or substitute labor, the empirical patterns provide reduced-form insights into how occupations adjust when technology-related skills expand. The non-linear association between changes in frontier-skill intensity and employment growth suggests heterogeneous adjustment paths, with some occupations experiencing complementarities and others showing signs of substitution. Future work could combine our measure with task-level or firm-level data to distinguish these mechanisms more structurally.

References

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

Acemoglu, D. and Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6):2188–2244.

Arntz, M., Baum, M., Brüll, E., Dorau, R., Hartwig, M., Lehmer, F., Matthes, B., Meyer, S.-C., Schlenker, O., Tisch, A., and Wischniewski, S. (2025a). Digital Transformation and the Changing World of Work (DiWaBe 2.0): A Data Source for Research on Artificial Intelligence and Other Technologies in the Workplace. *Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAUA) Report*, (Projectnumber: F 2573, F 2601):1–43.

Arntz, M., Böhm, M. J., Graetz, G., Gregory, T., Lehmer, F., and Lipowski, C. (2025b). Firm-level technology adoption in times of crisis. IZA Discussion Paper 17846, IZA Institute of Labor Economics. JEL Codes: O33, E22, E32, J23.

Arntz, M., Genz, S., Gregory, T., Lehmer, F., and Zierahn-Weilage, U. (2024). Frontier technology adopters and the aggregate decline of routine jobs. IZA Discussion Paper 16740, IZA – Institute of Labor Economics, Bonn.

Arntz, M., Gregory, T., and Zierahn, U. (2017). Revisiting the risk of automation. *Economics Letters*, 159:157–160.

Autor, D., Levy, F., and Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.

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

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? In *AEA Papers and Proceedings*, volume 108, pages 43–47.

Consoli, D., Marin, G., Rentocchini, F., and Vona, F. (2023). Routinization, within-occupation task changes and long-run employment dynamics. *Research Policy*, 52(1):104658.

Dauth, W., Findeisen, S., Suedekum, J., and Woessner, N. (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association*, 19(6):3104–3153.

Deming, D. J. and Noray, K. (2020). Earnings dynamics, changing job skills, and stem careers. *The Quarterly Journal of Economics*, 135(4):1965–2005.

Dengler, K. and Matthes, B. (2018). The Impacts of Digital Transformation on the Labour Market: Substitution Potentials of Cccupations in Germany. *Technological Forecasting and Social Change*, 137:304–316.

Dengler, K., Matthes, B., and Paulus, W. (2014). Occupational Tasks in the German Labour Market - An Alternative Measurement on the Basis of an Expert Database. *FDZ-Methodenreport*, 2014(12):1–36.

Engberg, E., Görg, H., Lodefalk, M., Javed, F., Längkvist, M., Monteiro, N. P., Nordås, H. K., Pulito, G., Schroeder, S., and Tang, A. (2024). Ai unboxed and jobs: A novel measure and firm-level evidence from three countries. Technical report, RF Berlin-CReAM Discussion Paper Series.

Engberg, E., Koch, M., Lodefalk, M., and Schroeder, S. (2023). Artificial intelligence, tasks, skills and wages: Worker-level evidence from germany. Technical report, Working Paper.

Falck, O., Heimisch-Roecker, A., and Wiederhold, S. (2021). Returns to ICT skills. *Research Policy*, 50(7):104064.

Felten, E. W., Raj, M., and Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, 108:54–57.

Felten, E. W., Raj, M., and Seamans, R. (2019). The occupational impact of artificial intelligence: Labor, skills, and polarization. *NYU Stern School of Business*.

Frey, C. B. and Osborne, M. A. (2017). The future of employment: how susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280.

Gathmann, C., Grimm, F., and Winkler, E. (2024). Ai, task changes in jobs, and worker reallocation. *CESifo Working Paper Series*, No. 11585.

Genz, S., Gregory, T., Janser, M., Lehmer, F., and Matthes, B. (2021). How Do Workers Adjust When Firms Adopt New Technologies? IZA Discussion Paper No. 14626, Institute of Labor Economics (IZA), Bonn.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.

Graetz, G. and Michaels, G. (2018). Robots at Work. *The Review of Economics and Statistics*, 100(5):753–768.

Grienberger, K., Matthes, B., and Paulus, W. (2024). Folgen des technologischen wandels fuer den arbeitsmarkt: Vor allem hochqualifizierte bekommen die digitalisierung verstaerkt zu spueren. *IAB-Kurzbericht*, (05):1–8.

Krueger, A. B. (1993). How computers have changed the wage structure: evidence from microdata, 1984–1989. *The Quarterly Journal of Economics*, 108(1):33–60.

Matthes, B. and Paulus, W. (2013). The german classification of occupations 2010: structure, coding and conversion table. *FDZ-Methodenreport*, (08):1–34.

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.

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., and Gómez, E. (2021). Measuring the occupational impact of ai: tasks, cognitive abilities and ai benchmarks. *Journal of Artificial Intelligence Research*, 71:191–236.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. Working Paper, Accessed at: https://www.michaelwebb.co/webb_ai.pdf.

Wiederhold, S. and Langer, C. (2023). The value of early-career skills. *Kiel, Hamburg: ZBW-Leibniz Information Centre for Economics*.

Online Appendix

A.1 BerufeNet skills database

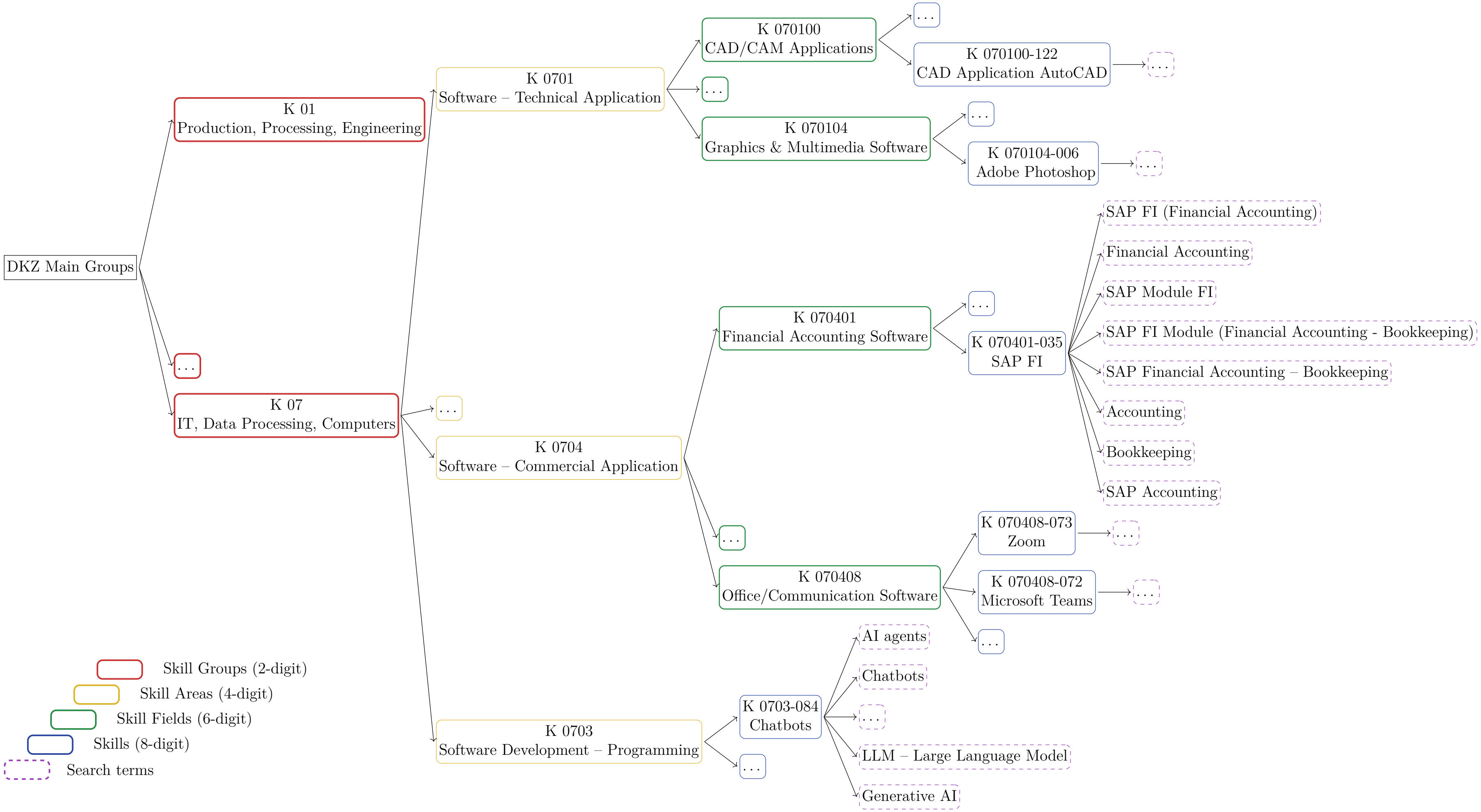
Table A.1: Selection steps of skills

Step (1)	Description (2)	2012 (3)	2023 (4)	2012 or 2023 (5)
0	Raw administrative skills (no restrictions)	17,747	21,139	21,258
1	Valid skills (excluding artists)	7,937	9,019	9,031
2	Exclude non-skill skill groups (K12, K16, K17)	7,681	8,764	**8,775**
3	Exclude skills without technology classification	7,547	8,477	–
4	Link classified skills to occupations; retain core or additional skills	**4,218**	**5,065**	–

Skill selection steps

- **Step 0** reports the full administrative list of skills by year, prior to any conceptual or analytical restrictions.
- **Step 1** applies the core definition of a valid skill, excluding artists entries in skill group *Media, Arts, Design* such as named operas (labeled in data with `kuenstler = "j"`).
- **Step 2** removes skill groups that capture contextual or structural characteristics (Arbeits- und Einsatzformen, Arbeitsorte, Branchen) rather than actual skills. This defines the conceptually clean skill universe used for classification.
- **Step 3** enforces the requirement that a skill must have a valid technology classification – some skills cannot be classified if the LLM things the skill does not exist.
- **Step 4** restricts the classified skill sets to skills that appear as either core or additional skills in any occupation using the year-specific occupation-skill matrix.

Figure A.1: Demonstration of the hierarchy within the digital competency classification code (DKZ) in BerufeNet



Notes: As of 2023, the hierarchy graph provides a representative excerpt of the DKZ (Digital Competence Code) structure. For reasons of readability and space constraints, the figure does not display the complete set of skill groups, skill areas, and skill fields. As of 2023, excluding artists (kuenstler ≠ "j"), the DKZ taxonomy comprised 17 skill groups (2-digit), 117 skill areas (4-digit), 196 skill fields (6-digit), 9,019 individual skills, and 47,007 unique search terms.

A.2 Generated input features

A.2.1 Prompts for generating additional context on skills

Table A.2: Prompts and system roles for LLM

Prompt name	**System role**	**Prompt**	**System role of the context prompts**	**Name of context prompts**
prompt_translation	You are an AI assistant specialized in German-English translation.	Give me the English translation of the job competency {competency}, which is related to the following search words: {keywords}. Your answer should consist only of the translated job competency, not the search words.	N/A	N/A
prompt_digital	You are an AI assistant specialized in classifying job competencies based on my definitions.	Digital technology refers to electronic devices or systems that create, store, and process data. Examples include computers, smartphones, the internet, software applications or other devices that involve any digital components. Is the job competency {translation} directly related to any digital technology as it existed in {year}, regardless of its level of development or integration at that time? Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Respond only with either ’yes’ or ’no’, followed by a concise explanation. If the technology did not exist in {year}, respond with ’Did not exist in {year}’ without any further answer.	You are an AI assistant specialized in German-English translation.	prompt_translation
prompt_description	You are an AI assistant specialized in describing job competencies.	if digital: Provide a description, in running text, of the digital technology directly related to the job competency {translation}. Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Answer in the format: {translation} is.. if not digital: Provide a description, in running text, of the job competency {translation}. Answer in the format: translation is..	You are an AI assistant specialized in classifying job competencies based on my definitions.	prompt_digital
prompt_frontier	You are an AI assistant specialized in classifying job competencies based on my definitions.	Cutting-edge technology refers to technological devices, techniques or achievements that are at the frontier of their field, even if they are not yet widely implemented or mature. Is {translation} directly related to a cutting-edge technology? Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Respond only with either ’yes’ or ’no’, followed by a concise explanation.	You are an AI assistant specialized in classifying job competencies based on my definitions.	prompt_digital
prompt_industry4	You are an AI assistant specialized in classifying job competencies based on my definitions.	Industry 4.0 refers to the fourth industrial revolution, characterized by the integration of digital technologies into both manufacturing and service industries. This revolution builds on previous industrial advancements, taking them to a new level of automation, interconnectivity, and real-time data exchange. In manufacturing, Industry 4.0 includes technologies such as the Internet of Things, artificial intelligence, and smart factories. In service industries, it involves the use of AI, big data analytics, and cloud computing. Is {translation} directly related to industry 4.0 technologies? Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Respond only with either ’yes’ or ’no’, followed by a concise explanation.	You are an AI assistant specialized in classifying job competencies based on my definitions.	prompt_digital
prompt_ai	You are an AI assistant specialized in classifying job competencies based on my definitions.	Is {translation} directly related to artificial intelligence? Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Respond only with either ’yes’ or ’no’, followed by a concise explanation.	You are an AI assistant specialized in classifying job competencies based on my definitions.	prompt_digital
prompt_self_controlled	You are an AI assistant specialized in classifying job competencies based on my definitions.	Self-controlled production technology and IT-integrated office and communication technology perform work processes largely automatically and autonomously. Examples for self-controlled production technology are smart factories, cyber-physical systems, or the internet of things. Examples for IT-integrated office and communication technology are the use of big data, cloud computing or online markets. Is {translation} directly related to either self-controlled production technology or IT-integrated office and communication technology? Assume what the technology was able to do in {year}, not what it does currently or will do in the future. Respond only with either ’yes, self-controlled’, ’yes, IT-integrated’ or ’no’, followed by a concise explanation.	You are an AI assistant specialized in classifying job competencies based on my definitions.	prompt_digital
prompt_production	You are an AI assistant specialized in classifying job competencies based on my definitions.	Given this description, is {translation} related to production jobs or tasks? By ’production’ we mean any job, task, or work environment related to factories, blue-collar work, engineering, manufacturing, machinary, transport or logistics. Respond only with either ’yes’ or ’no’, followed by a concise explanation.	1) You are an AI assistant specialized in German-English translation. 2) You are an AI assistant specialized in describing job competencies based on my definitions.	prompt_translation, prompt_description

Notes: For the primary analysis the reference year is {year}==2023. For the analysis of changes in skill intensities over time presented in Section 5, we repeat the procedure with the occupation-skill input data from 2012 and use the reference year {year}==2011. To improve the interpretative accuracy of the translation prompt, we use the broader skill groups, skill areas and search terms as additional input features, referred to as {keywords}.

A.2.2 Keyword matches

Table A.3: Frequency of number of keywords found

Number of keywords found	Frequency (frontier technology)	Frequency (digital technology)
0	5,989	4,570
1	1,280	629
2	753	552
3	493	1,016
4	252	974
5	154	702
6	67	345
7	24	152
8	11	64
9	6	18
10	1	7
11	1	2

Note: The table reports the frequency distribution of the number of detected technology-related keywords per skill description. Keyword counts are based on predefined dictionaries for frontier and digital technologies and are computed over the full set of skill descriptions. Each observation corresponds to one skill; frequencies therefore sum to the total number of skills in the dataset. All keyword counts are derived from LLM-generated skill descriptions as of 2023.

A.3 Classification

Table A.4: Distribution of training observations by skill group

Code	Skill Group	Training Count	Total Count	Share (%)
K00	Agriculture, Forestry, Horticulture	26	204	12.7
K01	Production, Manufacturing, Technology	107	1,707	6.3
K02	Construction, Architecture	32	314	10.2
K03	Business, Administration	49	696	7.0
K04	Transport, Logistics	27	226	11.9
K05	Hospitality, Tourism	30	166	18.1
K06	Services	28	299	9.4
K07	IT, Computing	214	1,613	13.3
K08	Science, Research, Development	39	379	10.3
K09	Social Services, Education, Health, Sports	75	954	7.9
K10	Media, Arts, Design	56	931	6.0
K11	Language Skills	19	103	18.4
K13	Product and Merchandise Knowledge	65	818	7.9
K14	Licenses, Certifications, Driving Permits	42	365	11.5
Total		**809**	**8,775**	

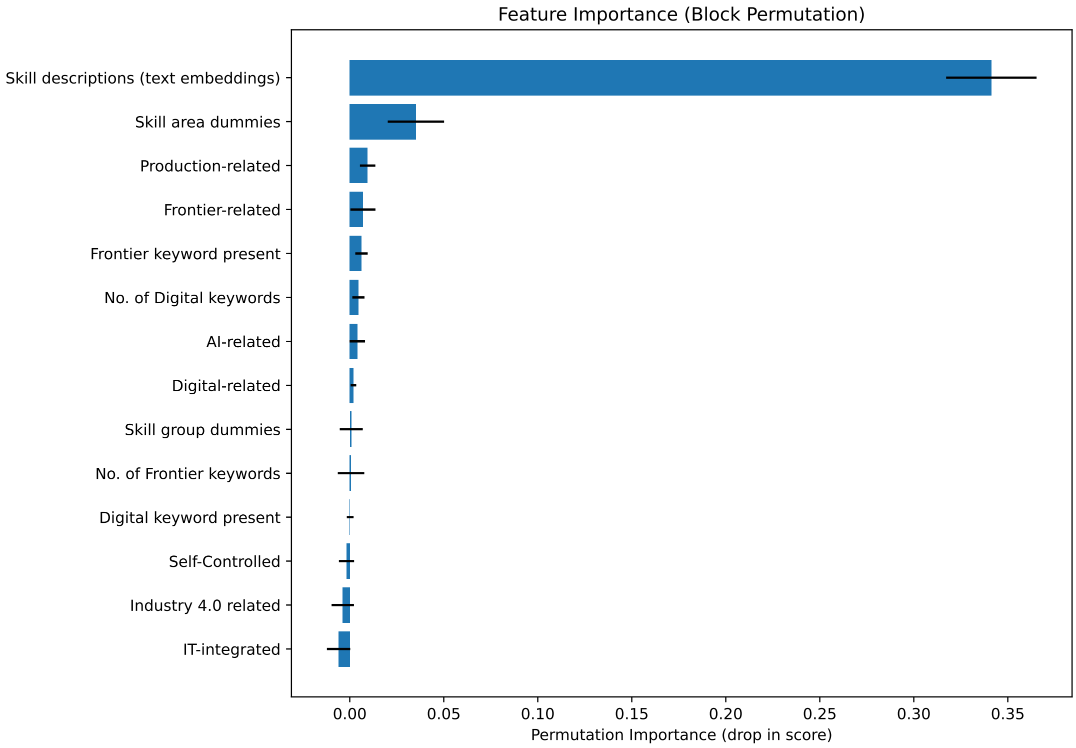
Note: The table reports the distribution of observations across skill groups in the training sample and the full dataset. The column *Share (%)* denotes the percentage of training observations relative to the total number of observations within each skill group.

Table A.5: Evaluation of the skill classifier's performance on the test set

Skill technology class	precision (1)	recall (2)	F1-score (3)	support (4)
Production:				
1-Manually Controlled *(manual)*	0.88	0.84	0.86	76
2-Indirectly Controlled *(digital)*	0.63	0.69	0.66	48
3-Self-Controlled *(frontier)*	0.73	0.73	0.73	51
Service:				
1-Non-IT Supported *(manual)*	0.91	0.90	0.91	81
2-IT-Supported *(digital)*	0.75	0.82	0.78	79
3-IT-Integrated *(frontier)*	0.79	0.70	0.74	70
Accuracy			0.79	405
Macro Avg	0.78	0.78	0.78	405
Weighted Avg	0.80	0.79	0.79	405

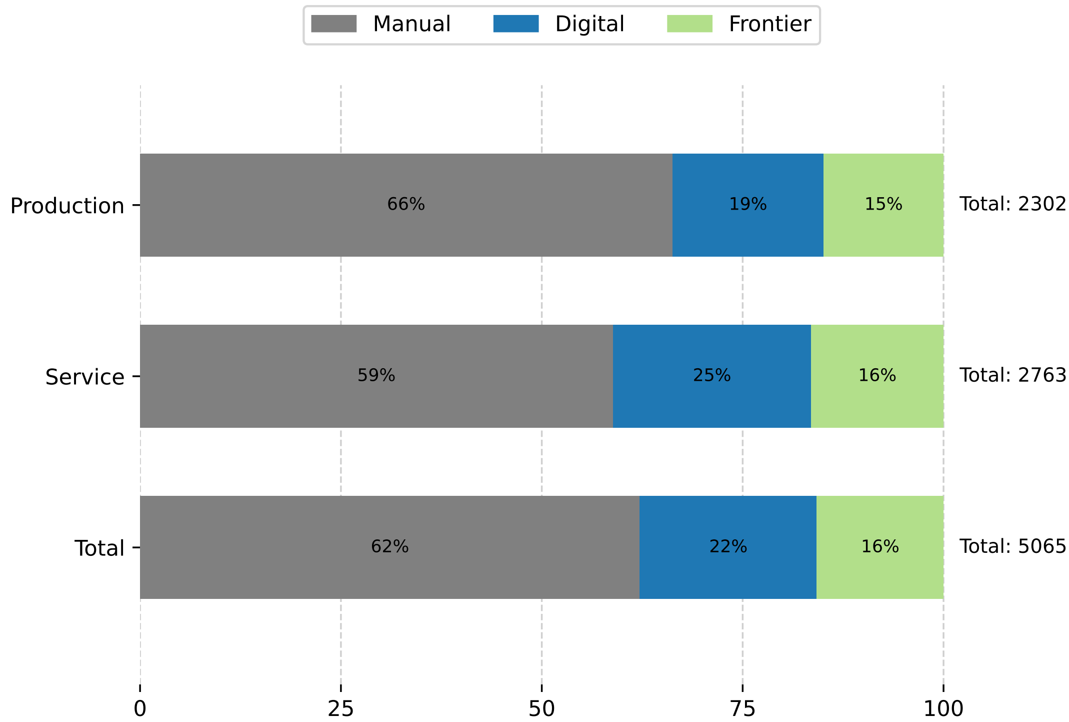
Notes: Table reports precision, recall, and F1-score for each skill class as predicted by the multi-layer perceptron (MLP) classifier, using a manually labeled test set of job skills. Results are shown separately for production and service contexts. Model performance is evaluated on a random 50% split of the labeled data; macro and weighted averages are reported across all classes.

Figure A.2: Feature importance



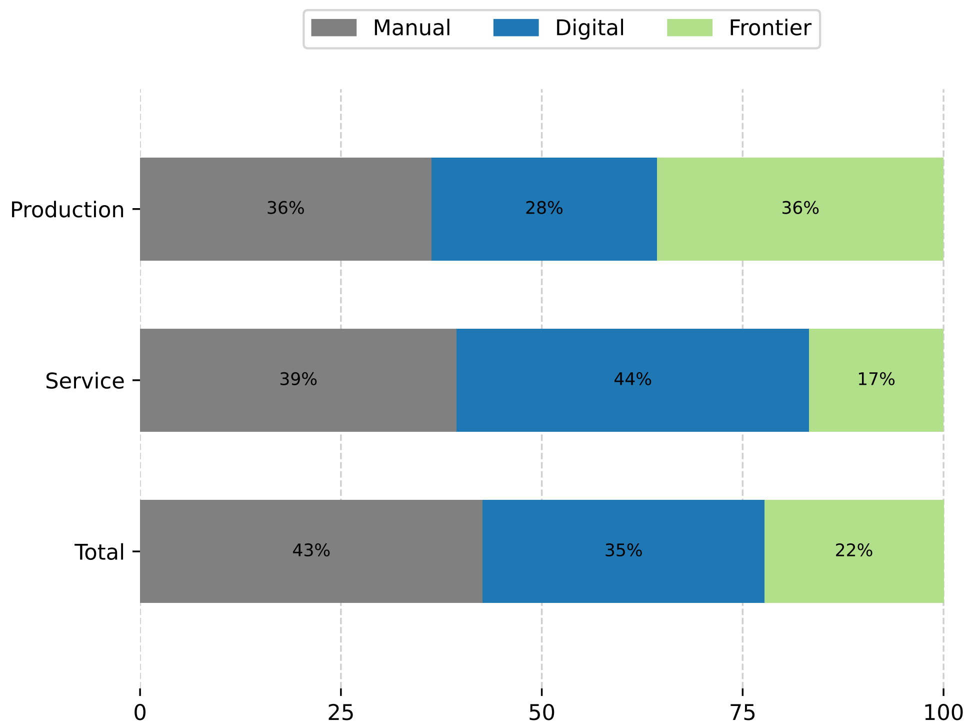
Notes: Feature importance was estimated using permutation importance on the test set (n=405), which measures the decrease in model performance when each feature's values are randomly shuffled. Higher values indicate a greater contribution to classification accuracy. Results are based on the trained MLPClassifier described in Section 3.

Figure A.3: Classification of unique job skills in BERUFENET as of 2023



Notes: The figure shows the distribution of skills classified as manual, digital, or frontier technologies, based on the 5,065 skills that appear as either core and additional skills in any occupation that existed in 2023.

Figure A.4: Average occupational OTSS as of 2023 (unweighted)



Notes: The figure shows the average unweighted OTSS for classified manual, digital, or frontier skills across all occupations. The numbers represent the skill composition of the average occupation in the German economy. The shares are calculated based on the 5,065 skills that appear as either core or additional skills in any occupation that existed in 2023.

A.4 Skill ranking

Table A.6: Examples of frontier and digital skills

Competency	# occupations
Frontier technology *competencies*	
Data lake	188
Maintenance and repair robots	67
Modelling, simulation (IT)	51
3D printing	38
Augmented reality	33
Machine learning	23
Computer-assisted surgery	20
Networked production systems	14
Autonomous driving systems	10
Collaborative robots (cobots)	3
Digital technology *competencies*	
Bookkeeping, accounting	506
Office and administrative work	332
CNC knowledge, CNC programming	248
Office organisation, office management	77
Payroll accounting, payroll processing	61
Personnel administration	48
Digital printing	44
Operating CNC machines	15
Scanning	13
Creating press releases	10

Notes: The table displays examples of competencies that are classified as frontier technology competencies (Panel A) and digital technology competencies (Panel B). The second column gives the number of distinct occupations at the 8-digit level of the KldB2010, which contain these competencies.

A.5 Survey data on technology use at the workplace

To bolster confidence in our digital and frontier skill measures, we compare our OTSS measures to individual-level survey data on technology usage at the workplace. For this purpose, we exploit the DiWaBe 2.0 survey, which is a representative cross-sectional individual-level survey conducted in 2024. The survey encompasses approximately 9,800 employees subject to social insurance contributions in Germany (see Arntz et al. (2025a) for a detailed description). The survey is designed to assess the impact of technological change, particularly AI and other frontier technologies, on the workplace of individual workers. Respondents were asked to think about their workplace and estimate the share of their working time spent using three categories of work equipment: non-computerized (equivalent to our manual technology classification), computerized (equivalent to our digital technology classification) and intelligently networked (equivalent to our frontier technology classification). Each respondent's occupation is identified at the 5-digit level of the KldB2010 classification.

A key limitation of the individual-level survey data is the relatively small number of observations available at detailed occupations. Only about two-thirds of respondents consented to link their responses with administrative employment records, resulting in 6,816 observations spread across the 1,297 occupations at the 5-digit KldB2010. Therefore, we need to aggregate the DiWaBe information to the more aggregated 2-digit KldB2010 occupational level to yield a representative picture of technology usage at the workplace. In contrast, the expert-based assessment of required skills for each occupation from the BERUFENET is available at the highly granular 8-digit KldB2010. For validation purposes, we also aggregate our OTSS measures to the 2-digit KldB2010 level, enabling comparisons across 36 occupational main groups (German: 'Berufshauptgruppen').

Finally, the two data sources differ in terms of temporal coverage. The DiWaBe survey is a cross-sectional survey which captures a snapshot of the self-reported technology usage at the workplace in 2024. One major advantage of the expert database BERUFENT is that it is continuously updated (see Section 2) and therefore captures variations of required skills within granular occupations at the 8-digit KldB2010.

For these reasons, we regard our technology skill intensity measures based on the BERUFENET database as more reliable and suitable for the main empirical analysis. Nonetheless, we regard the survey-based DiWaBe measures as a valuable external benchmark to validate the accuracy and relevance of our constructed technology skill indicators.

A.6 Occupational group ranking by frontier OTSS

Table A.7: Ranking of occupational groups by frontier OTSS (KldB2010, 2-digit, employment-weighted)

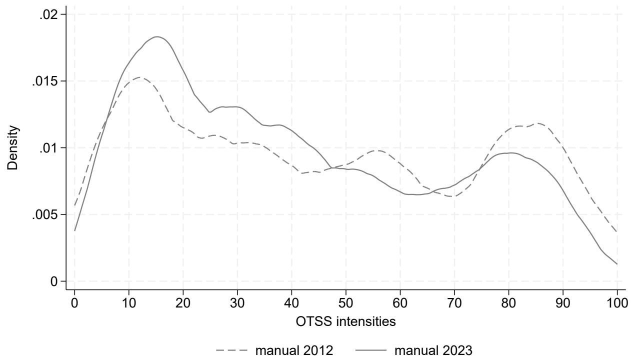
KldB2010			OTSS (in %):		
2-digit	Occupational group title	# employees	Manual	Digital	Frontier
43	Computer science, information and communication technology	972,643	5	36	59
26	Mechatronics, energy electronics and electrical engineering	905,073	14	36	50
41	Mathematics, biology, chemistry and physics	369,272	32	26	41
27	Technical research and development, construction, and production planning and scheduling	1,052,093	13	46	41
25	Technical occupations in machine-building and automotive industry	1,577,055	38	29	33
42	Geology, geography and environmental protection	46,370	39	29	32
92	Occupations in advertising and marketing, in commercial and editorial media design	588,677	12	58	29
52	Drivers and operators of vehicles and transport equipment	1,059,063	51	22	27
81	Medical and health care occupations	2,342,064	58	16	26
34	Building services engineering and technical building services	680,234	46	28	25
51	Traffic and logistics (without vehicle driving)	1,913,836	32	46	22
23	Paper-making and -processing, printing, and in technical media design	226,753	23	55	22
61	Purchasing, sales and trading	1,011,215	19	61	20
21	Production and processing of raw materials, glass- and ceramic-making and -processing	105,108	62	20	19
24	Metal-making and -working, and in metal construction	1,063,401	50	32	18
31	Construction scheduling, architecture and surveying	272,367	32	50	18
22	Plastic-making and -processing, and wood-working and -processing	452,045	61	21	17
93	Product design, artisan craftwork, fine arts and the making of musical instruments	55,380	46	37	17
53	Safety and health protection, security and surveillance	331,262	38	45	17
91	Philology, literature, humanities, social sciences, and economics	98,366	45	39	17
82	Non-medical healthcare, body care, wellness and medical technicians	815,633	76	8	17
72	Financial services, accounting and tax consultancy	1,269,118	24	60	16
11	Agriculture, forestry, and farming	197,905	69	15	16
94	Performing arts and entertainment	114,356	43	42	15
28	Textile- and leather-making and -processing	98,714	56	31	13
62	Sales occupations in retail trade	1,838,517	30	57	13
71	Business management and organisation	3,932,383	16	73	11
12	Horticulture and floristry	241,764	76	13	11
73	Law and public administration	1,072,363	38	52	10
33	Interior construction	345,231	80	10	10
29	Food-production and -processing	751,048	77	14	9
84	Teaching and training	639,191	64	27	9
32	Building construction above and below ground	556,969	79	13	9
54	Cleaning services	801,062	85	9	6
83	Education and social work, housekeeping, and theology	1,789,019	84	13	3
63	Tourism, hotels and restaurants	686,599	68	29	3

Notes: We only take non-missing KldB2010 5digit occupations for calculating the means and exclude armed forces. Occupations are ranked according to their frontier OTSS. This table displays the numbers used to generate Figure 4 in the main text.

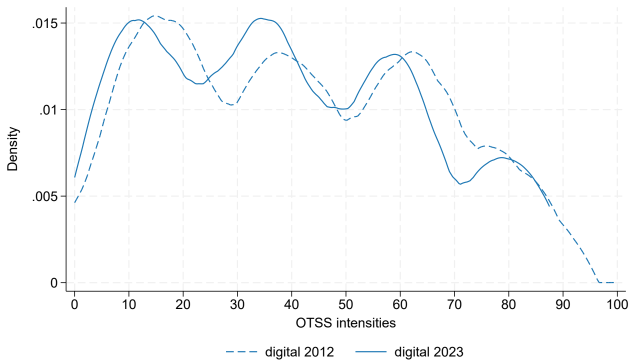
A.7 Distribution of occupational skill intensities

Figure A.5: Distribution of OTSS in 2012 and 2023 (employment-weighted)

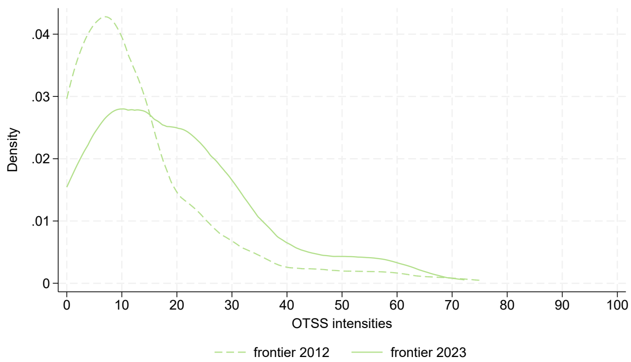
(a) Manual OTSS



(b) Digital OTSS



(c) Frontier OTSS



Notes: The figure plots employment-weighted kernel density estimates of OTSS at the 5-digit KldB2010 level in 2012 (dashed lines) and 2023 (solid lines). Panel (a) shows the distribution of manual OTSS, Panel (b) digital OTSS, and Panel (c) frontier OTSS. Employment weights are year-specific, i.e. skill intensities in 2012 are weighted by employment in 2012, and those in 2023 by employment in 2023.

A.8 Heterogeneous responses to changes in frontier skill intensity

Table A.8: Heterogeneous employment responses to changes in frontier OTSS

	All	Age group			Gender		Education		
		16–34	**35–44**	**45–65**	**Men**	**Women**	**No Voc. Training**	**Voc. Training**	**University Degree**
$\Delta OTSS^{frontier}$	-2.33	-4.42**	-1.68	-3.23**	-3.01**	-3.64*	-5.99***	-1.23	-3.79***
	(1.40)	(1.93)	(1.54)	(1.51)	(1.38)	(1.83)	(1.92)	(1.26)	(1.27)
$(\Delta OTSS^{frontier})^2$	9.40	22.71**	5.93	11.63*	10.50*	23.16**	40.38***	4.17	24.33***
	(6.18)	(9.55)	(6.74)	(6.32)	(5.45)	(9.83)	(7.78)	(4.82)	(8.20)
Observations	36	36	36	36	36	36	35	36	36
R^2	0.15	0.25	0.08	0.21	0.23	0.15	0.22	0.05	0.31

Notes: Each column reports coefficients from a separate regression of employment growth between 2012 and 2023 on the change in (employment-weighted) frontier skill intensity between 2012–2023 ($\Delta OTSS^{frontier}$) and its square, estimated at the 2-digit KldB 2010 occupation level. Column *All* reproduces the baseline specification underlying Figure 6. Age, gender, and education groups are based on administrative BeH records. Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.