

6

Implications of evolving AI capabilities for employment and education

This chapter summarises the results of the assessment of artificial intelligence (AI) capabilities in literacy and numeracy and discusses their implications for policy. The chapter first considers likely impacts of developing computer capabilities on employment. For this purpose, it analyses the use of literacy and numeracy at work and the proficiency of workers who use these skills on a daily basis. It then discusses the education implications of AI advancements. In particular, it highlights the need for developing skills in the population that are beyond those of AI technology. The chapter also touches upon the importance of supplying workers with diverse skills, including digital skills, to help them cope with occupational changes resulting from the use of technology.

The preceding chapters described artificial intelligence (AI) capabilities in literacy and numeracy assessed in 2016 and 2021 using expert evaluations with the OECD Survey for Adult Skills (PIAAC). This chapter summarises the results and discusses their implications for policy. It considers likely impacts of developing computer capabilities on employment by looking at the use of literacy and numeracy skills at work. The chapter also discusses the education implications of AI advancements. This focuses on the need for developing skills in the population that are beyond those of technology and the importance of diversifying the skill set of people to enable them to compete, but also to work together, with AI.

Summary of results

The study asked 11 computer experts to rate the capacity of AI to solve the questions of PIAAC's literacy and numeracy tests. Four experts in mathematical reasoning of AI provided additional ratings in numeracy. AI's likely performance on the tests was determined by looking at the majority expert opinion on each question.

Assessment of current AI capabilities in literacy and numeracy

According to most experts, AI can perform high on the PIAAC literacy test. It can solve most of the easiest questions, which typically involve locating information in short texts and identifying basic vocabulary. It can also master many of the harder questions, which require understanding rhetorical structures and navigating across larger chunks of text to formulate responses (see Chapter 4). Overall, AI is expected to solve between 80-83% of all literacy questions, depending on how the majority response of experts is calculated (see Table 6.1).

Table 6.1. Summary of AI and adults' performance in PIAAC

Share of PIAAC questions that AI can answer correctly according to the majority of experts and probability of successfully completing items of adults at different proficiency

	Literacy	Numeracy
AI measures		
Yes/No	80%	66%
Weighted	81%	67%
Weighted + Maybe	83%	73%
Quality of AI measures		
Agreement	78% average majority	62% average majority
Uncertainty	20% Maybe/Don't know responses	12% Maybe/Don't know responses
Adults' performance		
Average adults	50%	57%
Level 2 adults	41%	52%
Level 3 adults	67%	74%
Level 4 adults	86%	90%

Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

This evaluation rests on high consensus among experts. The group judgements of whether AI can solve each PIAAC literacy question were supported by 78% of experts, across questions, on average (see Table 6.1). However, many votes were excluded from the group response as they were less informative of AI's potential outcome on the PIAAC test. These were the Maybe- and Don't know-answers to the question

of whether AI can correctly solve an item. These responses amounted to 20% of all responses in the literacy assessment (see Table 6.1).

The analyses set out in Chapter 4 compared AI literacy performance to that of adults at varying proficiency levels. PIAAC assesses respondents' proficiency and questions' difficulty on the same levels, going from low (Level 1 and below) to high proficiency/difficulty (Levels 4-5). Respondents with proficiency at a given level have a 67% chance of correctly completing the items at this level, higher chance of success at lower levels of difficulty and lower chances of success at more difficult levels (see Chapter 3).

According to the evaluation of experts, AI can perform similar to or better than Level 3 adults at all levels of question difficulty in literacy (see Figure 4.6, Chapter 4). This is also indicated by the overall success rate of AI on the literacy test, estimated at 80%, which is between that of Level 3 and Level 4 adults (see Table 6.1). This suggests that AI can potentially outperform a large proportion of the population on the PIAAC literacy test. Across the OECD countries that participated in PIAAC, on average, 35% of adults are proficient at Level 3 and 54% score below this level; only 10% of adults perform better than Level 3 in literacy (OECD, 2019, p. 44^[4]).

AI performs less well in numeracy – according to the 15 experts who completed the numeracy assessment. Following their majority vote, AI can answer around two-thirds of the easier and intermediate numeracy questions of PIAAC, and less than half of the hardest questions (see Chapter 4). This amounts to an overall success rate of 66-73% on the entire numeracy test, depending on the type of aggregate measure (see Table 6.1).

AI's estimated success rate in numeracy is beyond that of Level 2 adults (see Table 6.1). However, AI could not outperform these adults at each level of question difficulty, as shown in Chapter 4. According to most experts, AI would score similar to low-performing adults at Level 1 and below. The estimated performance at Level 2 of question difficulty is close to that of Level 2 adults. At Level 3 and above, AI's outcome expected by experts corresponds to adults' proficiency at Level 3.

These results should be interpreted with caution, given the high disagreement among experts in the numeracy domain. Two groups of experts with opposing opinions emerged: five experts rated AI's potential performance on most of the numeracy questions low, and five believed that AI could answer most questions. As a result, thin majorities determined AI's capacity to solve the numeracy questions. On average, across questions, the majority opinion on AI performance was supported by 62% of experts (see Table 6.1). This may have produced arbitrary results since AI's success on single questions was often decided by a difference of only one vote.

The disagreement among experts relates to ambiguity about the generality of the evaluated AI. Some experts imagined narrow AI solutions for separate PIAAC questions. Others considered general AI systems that can reason mathematically and process all types of numeracy questions simultaneously, including similar questions that are not part of the test. These considerations affected experts' evaluations: the latter experts gave lower ratings than the former. However, the group discussion revealed much agreement behind this divergence in ratings. Both groups seemed to agree the test can be solved by developing a number of narrow AI solutions, while an AI with general mathematical abilities is still out of reach of current technology.

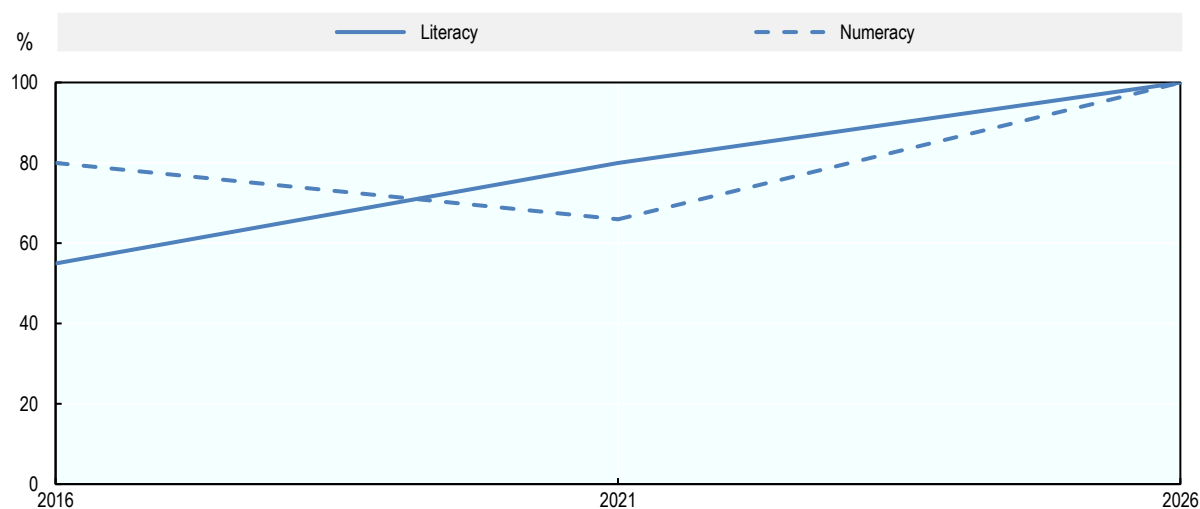
Development of AI literacy and numeracy capabilities over time

This study follows up a pilot assessment of AI capabilities with PIAAC conducted in 2016 (Elliott, 2017^[5]). The pilot study asked 11 computer scientists to rate AI capabilities with respect to PIAAC's literacy, numeracy and problem-solving tests. The assessment approach used in the follow-up study is comparable to that of the pilot.

The comparison of both assessments reveals a considerable improvement in AI's literacy capabilities since 2016 (see Figure 6.1). In 2016, experts assessed the potential success rate of AI on the literacy test at 55%. By contrast, experts in 2021 expected AI to solve 80% of the test correctly. Chapter 5 showed that expected AI literacy performance increased at all levels of question of difficulty: from 71% to 93% at Level 2, from 48% to 68% at Level 3, and from 20% to 70% at Level 4-5, while performance at Level 1 and below remained at 100%.

Figure 6.1. Success rate of AI in PIAAC according to experts' assessments

Percentage of PIAAC questions that current AI and AI in 2026 can answer correctly according to most experts



StatLink  <https://stat.link/brmyfj>

These results reflect the technological developments in AI in the period since the pilot assessment. The introduction of large pre-trained language models in 2018 has pushed the state of the art in natural language processing (NLP) forward considerably. These models are trained once, on a large corpus of data, and then used as base models for developing NLP systems for specific tasks and domains. Their success lies in the huge amounts of training data used, as well as in the application of advanced architectures, such as Transformers (Russell and Norvig, 2021^[6]). The latter allow the models to capture relations over longer paragraphs and to “learn” the meaning of words in context.

Given these technological advancements and the heavy investment and research in NLP (see also Box 6.1), experts judged that AI's literacy capabilities will continue to develop. They expect that AI will be able to solve all literacy questions in PIAAC by 2026.

The numeracy assessment produced implausible results of declining AI numeracy capabilities over time (see Figure 6.1). This may have to do with methodological issues: disagreement among experts in the domain or varying interpretations of the rating exercise across studies. Another reason may relate to the more limited scope of research on AI's mathematical capabilities on problems such as PIAAC at the time of the pilot assessment. This lack of information may have led experts to overestimate AI's likely performance on the numeracy test in 2016.

Box 6.1. ChatGPT as an example of AI literacy capabilities

In November 2022 – roughly one year after the 11 core experts completed the online assessment, and three months after the four experts in mathematical reasoning re-assessed AI in numeracy – ChatGPT was released. ChatGPT is an AI chatbot developed by OpenAI, a prominent AI research laboratory. Its ability to answer diverse questions and interact in a human-like manner attracted huge public attention. Next to mimicking conversation, ChatGPT performs a variety of tasks, such as composing poetry, music and essays, or writing and debugging code. It demonstrated for the first time to the broader public what state-of-the-art language models are capable of.

ChatGPT relies on a somewhat upgraded version of GPT-3, the model that the experts often considered in their evaluation of AI capabilities. GPT-3 “learns” language in a self-supervised manner, by predicting tokens in a sentence based on the surrounding text. By contrast, ChatGPT is built upon the InstructGPT models of the GPT-3.5 series, which are trained to follow instructions using reinforcement learning from human feedback. Specifically, GPT-3 is fine-tuned with data containing human-written demonstrations of the desired output behaviour. Subsequently, the model is provided with alternative responses ranked by human trainers. The ranks act as reward signals to train the model to predict which outputs humans would prefer. This makes the model better at following a user’s intent (Ouyang et al., 2022^[7]).

However, the model still has important limitations. It can produce plausible-sounding but incorrect responses (OpenAI, 2023^[8]). It can generate toxic or biased content and respond to harmful or inappropriate instructions, although it has been trained to refuse such requests. In addition, the model often fails to respond to incomplete inquiries by asking clarifying questions.

The discussion with experts suggested the numeracy capabilities of AI are unlikely to have changed much between 2016 and 2021. During the period, constructing mathematical models from tasks that require general knowledge and are expressed in language or images has received less research attention. The interest and investment by companies were also limited, focused on specific areas of mathematical reasoning, such as verifying software.

However, this has recently begun to change. In 2021, benchmarks for mathematical reasoning of AI were released (Hendrycks et al., 2021^[9]; Cobbe et al., 2021^[10]). These allow researchers to train and test models in solving mathematical problems of various kinds. In addition, multimodal models that process information in different formats have received more attention (Lindström and Abraham, 2022^[11]). These models are particularly relevant for solving the types of numeracy questions contained in PIAAC since the questions use formats as diverse as images, graphs, tables and text. These trends led experts to expect that AI will advance considerably in numeracy and could solve all PIAAC numeracy questions by 2026 (see Figure 6.1).

In sum, AI’s performance in literacy estimated by experts increased by 45% between 2016 and 2021 and is expected to increase by a further 20% by 2026. AI’s potential performance on the PIAAC’s numeracy test is unlikely to have changed much between 2016 and 2021. However, experts expect it to reach a ceiling by 2026. For comparison, the literacy and numeracy skills of adults change much more slowly (see Chapter 2). Over 13 to 18 years – between the 1990s and 2010s – the distribution of literacy skills in the adult population and the working population has, on average, changed marginally across 19 countries and economies. The same goes for numeracy skills compared over five to nine years across seven countries.

Policy implications of evolving AI capabilities in literacy and numeracy

Fast-evolving AI capabilities in key skill domains raise questions about whether AI will substitute workers in jobs and what implications this would have for education systems. Before turning to these questions, several limitations of this analysis for formulating clear policy conclusions should be discussed.

First, the analysis focuses on the technological capabilities of AI and not on AI deployment in the economy. Whether and how evolving AI technology in the domains of NLP and mathematical reasoning will be adopted in the workplace depends on many factors. Among these are the costs of the technologies, capital investment, regulation and social acceptance (Manyika et al., 2017^[12]).

Second, the study assesses AI capabilities in only two skill domains – literacy and numeracy. However, workers use a variety of skills to perform a variety of tasks in occupations. An assessment of AI capabilities across the full range of skills used in the workplace will be needed to determine the exact impacts of AI on employment.

Third, the comparison of AI and human performance in PIAAC should not imply that AI can carry out all kinds of everyday literacy and numeracy tasks as flexibly as adults at a corresponding level of proficiency. In fact, some experts criticised that education tests applied on AI do not necessarily capture the general underlying capabilities that would allow for performing a wide range of similar tasks (as they do when applied on humans). However, this problem – known as overfitting to the test – is common to all tests for evaluating AI. The study attempted to decrease the risk of overfitting by providing experts with more information on the underlying skills that PIAAC is supposed to measure (see Chapter 3).

Despite the limitations of the analysis, it is safe to conclude that the rapid development of AI capabilities laid out here will have important effects on employment, and in particular, on the employment of workers with low literacy and numeracy skills. This may rebound on education systems as they will be increasingly expected to equip people with the skills needed to work in the digitised economy.

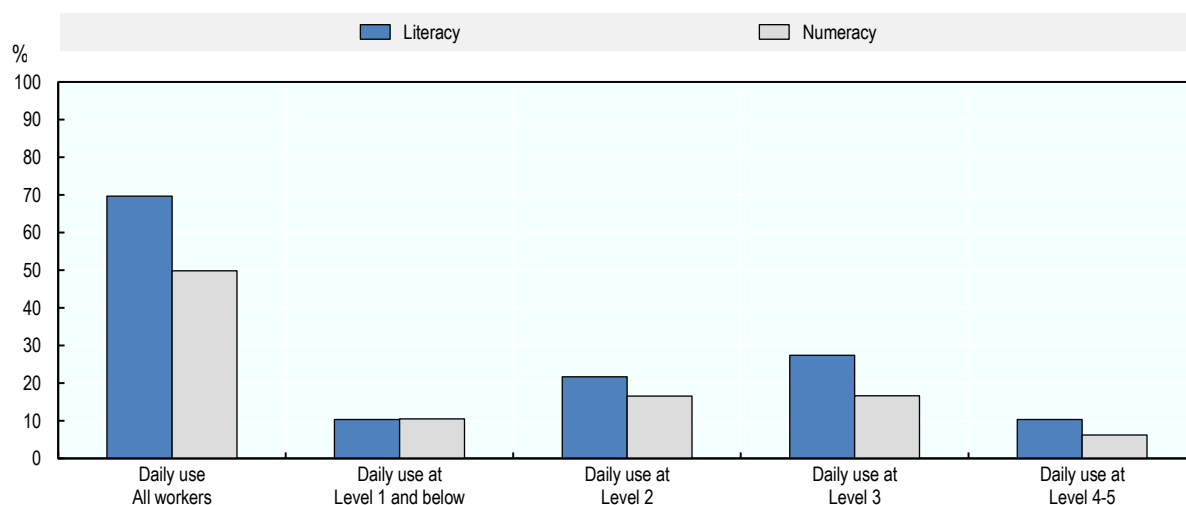
Implications for employment

How evolving AI will affect employment depends not only on how AI capabilities compare to human skills, but also on how skills are used in the economy. AI would substantially affect demand for workers if it can reproduce those skills that are in high demand in the economy. The study chose to assess AI with PIAAC precisely because this test measures key information-processing competencies that are an integral part of work.

Figure 6.2 shows that 70% of workers use literacy skills on a daily basis at work. These workers can be affected by advancing AI capabilities if their literacy proficiency is comparable to or below that of computers. According to computer experts, AI performance in literacy is beyond proficiency Level 3 in PIAAC. Approximately 27% of workers use literacy daily at Level 3 of proficiency. Another 32% perform literacy tasks on a daily basis, having a literacy proficiency below Level 3. Taken together, AI could affect the literacy-related tasks of 59% of the workforce.


Figure 6.2. Percentage of workers at different proficiency levels who use literacy and numeracy on a daily basis at work

Percentage shares from all workers



Note: The use of literacy skills includes reading books; professional journals or publications; manuals or reference materials; diagrams, maps or schematics; financial statements; newspapers or magazines; directions or instructions; letters, memos or mails. The use of numeracy skills includes the use of advanced math or statistics; preparing charts, graphs or tables; use of simple algebra or formulae; calculating costs or budgets; using or calculating fractions or percentages; using a calculator. The bars show the percentage shares of all workers who report performing at least one of these practices daily at work.

Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/zdks8g>

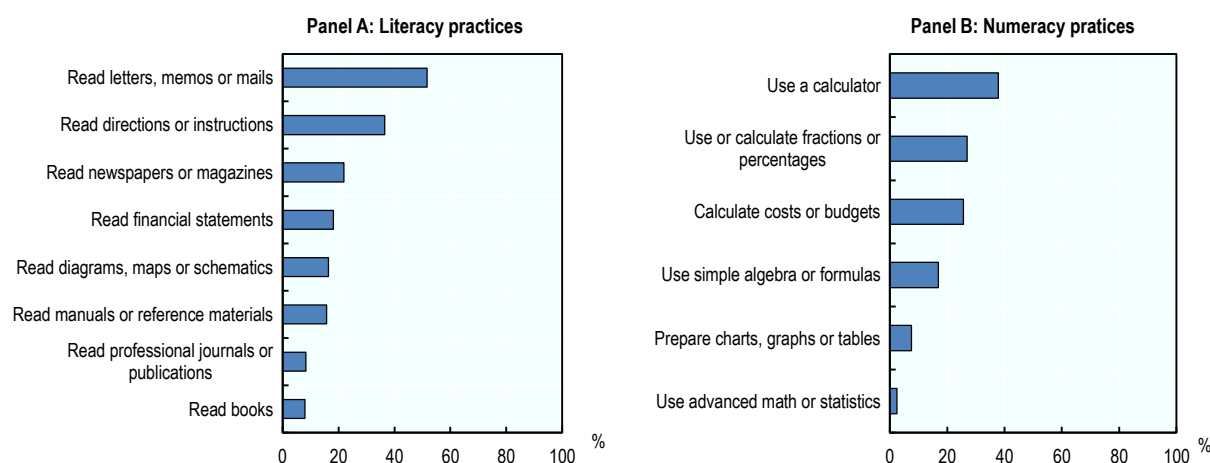
Similarly, Figure 6.2 shows that 50% of workers perform numeracy tasks daily at work. The AI numeracy capabilities assessed by experts are beyond those of adults at proficiency Level 2 on most PIAAC questions and close to those of Level 3 adults on some questions. Across the 39 countries and economies, on average, 27% of the workforce uses numeracy on a daily basis at or below Level 2 proficiency. A share of 44% uses numeracy at Level 3 or below that level. AI can negatively affect the employment of these workers if numeracy tasks constitute a substantial part of their daily work.

AI's impact on employment also depends on the difficulty of tasks performed in occupations. As this study shows, AI performs better on literacy and numeracy tasks that are easier for humans and worse on tasks that are difficult for humans. Thus, AI is more likely to affect workers with easier tasks – independent of their skill proficiency.


Figure 6.3 indicates that more workers perform easy than hard tasks at work. Across the 39 countries participating in PIAAC, on average, 52% of workers read memos or mails, 37% read directions and instructions, and 22% read newspapers and magazines each day. A smaller share of the workforce reads longer text such as professional journals (8%) or books (8%) each day. Similarly, simple numeracy skills are used more broadly than complex numeracy skills. Between 26% and 38% of workers, across all countries, on average, calculate costs or budgets, use a calculator, fractions or percentages daily at work. By contrast, only 3% use advanced math and statistics, 8% prepare charts and graphs, and 17% use simple algebra or formulae each day. This shows that literacy and numeracy tasks that are potentially easy for AI are more prevalent in the economy, even though the workers that perform them may be more proficient than computers.

Figure 6.3. Daily use of literacy and numeracy practices at work

Percentage of workers reporting using a practice on a daily basis



Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/8epnr0>

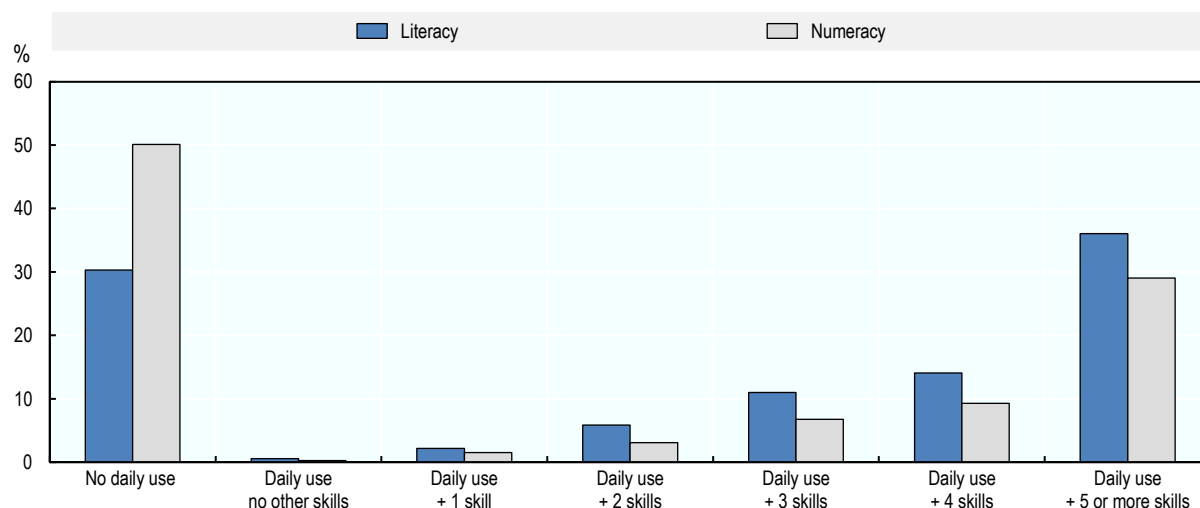
AI's potential to automate jobs further depends on the skill mix jobs require. Jobs that involve a diverse set of skills are more sheltered from automation as it is less likely that AI reproduces many and different skills of workers at once. If AI capabilities come close to reproducing some skills in a rich set of skill requirements, workers will still be needed for other skills. By contrast, workers who use only one or a few skills at work intensively can be completely replaced by machines with the corresponding capability.

Next to literacy and numeracy skills, the Survey of Adult Skills (PIAAC) collects information on the performance of a variety of practices at work (OECD, 2013^[13]). These include, for example, the frequency of writing documents, solving complex problems, working with a computer or instructing, teaching or training people. Based on this information, Figure 6.4 explores how workers combine literacy and numeracy with the following generic skills at work: writing, digital skills, problem solving, learning at work, influencing skills, co-operative skills, organising skills and physical skills.¹ Although these skills do not cover all possible skills used in the workplace, they give a glimpse of how workers use skills in concert in their jobs.

Figure 6.4 shows the percentage shares of all workers who do not use literacy or numeracy on a daily basis, and who use these skills daily alone or in combination with other generic skills. The results show that the biggest shares of workers with daily use of literacy and numeracy use a diverse skill mix at work. Across all countries participating in PIAAC, on average, 36% of all workers combine literacy and 29% of workers use numeracy with at least five other generic skills in their daily work routines. By contrast, less than 1% of the workforce reports using literacy or numeracy with none of the other generic skills daily. However, some workers combine literacy and numeracy with only a few other skills – 20% of workers use literacy and 12% use numeracy with up to three other skills at work.

Figure 6.4. Daily use of literacy and numeracy at work together with other skills

Percentage shares from all workers



Note: Daily use of literacy and numeracy together with the following skills: writing, digital skills, problem solving, learning at work, influencing skills, co-operative skills, organising skills and physical skills. See note 1 at the end of this chapter.

Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/2prxu9>

The above results suggest that advances in AI with regard to literacy and numeracy can negatively impact employment since literacy and numeracy skills are widely used at work. This holds particularly for the employment of workers who use these skills at a proficiency below that of machines, who perform easy tasks manageable for AI or who use only a few other skills intensively at work. These workers constitute a considerable share of the workforce.

However, AI may change the nature of work in ways that do not affect the aggregate demand for labour. A standard economic view is that long-term effects of AI on employment may be positive due to increases in productivity. AI is expected to improve productivity in companies by performing particular tasks faster and more accurately than human workers. This would leave the latter time to concentrate on more important tasks that may include creativity, management or critical thinking. This, in turn, would enable companies to produce more at a lower cost. Lower costs are expected to increase the demand for the products. This would boost labour demand both in the AI-using companies and in other companies connected to the value chain (OECD, 2019^[14]).

In addition, AI is expected to create demand for new tasks – tasks related to adoption and use of machines in the workplace. In future, more workers will be needed for producing data, developing AI applications, operating AI systems and analysing their outputs. An OECD study that analyses job-postings data for 2012-18 in four countries – Canada, Singapore, the United Kingdom and the United States – shows increasing demand for AI-related skills (Squicciarini and Nachtigall, 2021^[15]). In the United States, for example, the total number of AI-related job openings increased from around 20 000 in 2012 to almost 150 000 in 2018. In particular, skills related to data mining and classification, NLP and deep learning are more often advertised on line.

Moreover, AI can lead to the emergence of completely new occupations and industries by enabling the creation of new products and services. This technology has been seen as an “invention of a method of

invention” (Cockburn, Henderson and Stern, 2018^[16]). This means it is expected to accelerate the process of innovation at an unprecedented rate. What makes AI a trigger of innovation is its wide applicability: learning algorithms have many potential new uses in a variety of sectors and occupations. In addition, AI is increasingly used in science to formulate hypotheses, search and systemise information, or identify hidden patterns in high-dimensional data (Bianchini, Müller and Pelletier, 2022^[17]). This can facilitate scientific discovery and the creation of novelty.

How advancing AI will reshape work and the demand for skills remains an open question. What is certain is that workers will need new skills to meet future demands – skills that allow them to compete and work together with AI. This raises questions about the role of education in preparing people for the future.

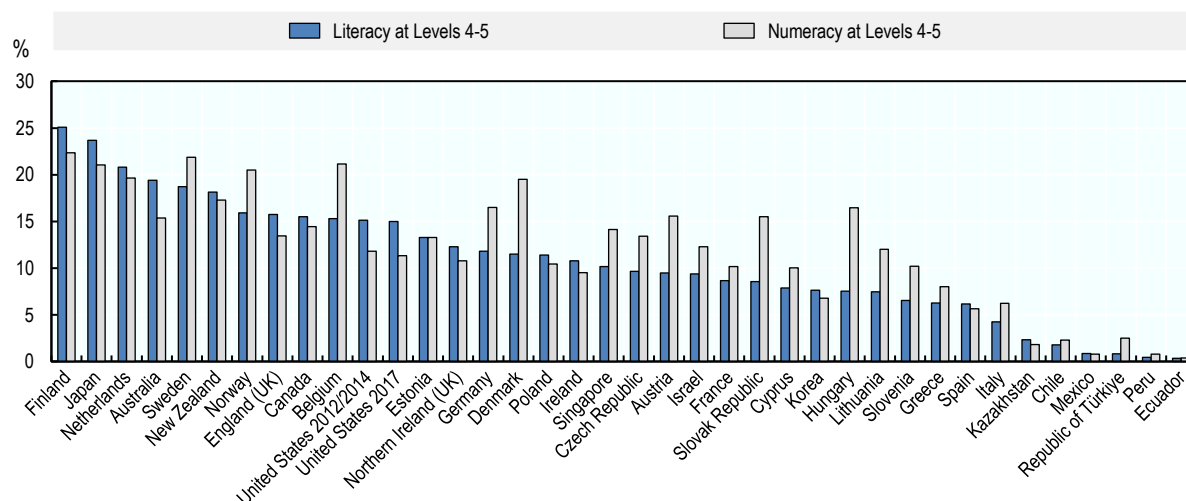
Implications for education

Technological change pressures education systems to supply the economy with a suitably skilled workforce. As one likely response, education would attempt to increase the skill level of the workforce beyond that of computers. In the domains of literacy and numeracy, this would mean to lift up the working population to the highest proficiency levels – Levels 4 and 5. This proficiency would enable workers to understand, interpret and critically evaluate complex texts and multiple types of mathematical information. Developing such skills is not only relevant for outperforming AI in reading and mathematical tasks. Much more importantly, strong literacy and numeracy skills build the foundation for developing other higher-order skills, such as analytic reasoning and learning-to-learn skills. They also ease the access to new knowledge and know-how (OECD, 2013^[13]).

Chapter 2 showed the foundation skills of the working population have not changed substantially in the past decades. Of course, future efforts to improve the literacy and numeracy skills of workers could be more successful. In particular, countries with high proportions of highly proficient adults in their workforce can serve as examples of good practice. Other countries can extract formative lessons and borrow promising policies from the comparison with these high performers to increase their skills stock.

Figure 6.5 shows the shares of workers with literacy and numeracy proficiency at Levels 4-5 across the 39 countries and economies that participated in PIAAC. The best ranking country in literacy – Finland – has 25% of adults at literacy proficiency Levels 4-5, followed by Japan with 24% and the Netherlands with 21%. In numeracy, the three top ranking countries – Finland, Sweden and Belgium – have between 21-22% of workers at Levels 4-5. This shows that even the best performers to date cannot supply more than a quarter of their workforce with the literacy and numeracy skills needed to outperform AI. For median performers these shares are much smaller – 10% of workers at Level 4-5 in literacy (Singapore) and 12% in numeracy (Lithuania). These countries would have to double the shares of highly proficient workers in literacy and numeracy to reach their best-performing peers.

Figure 6.5. Proportion of workers with high literacy and numeracy proficiency



Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/6km89z>

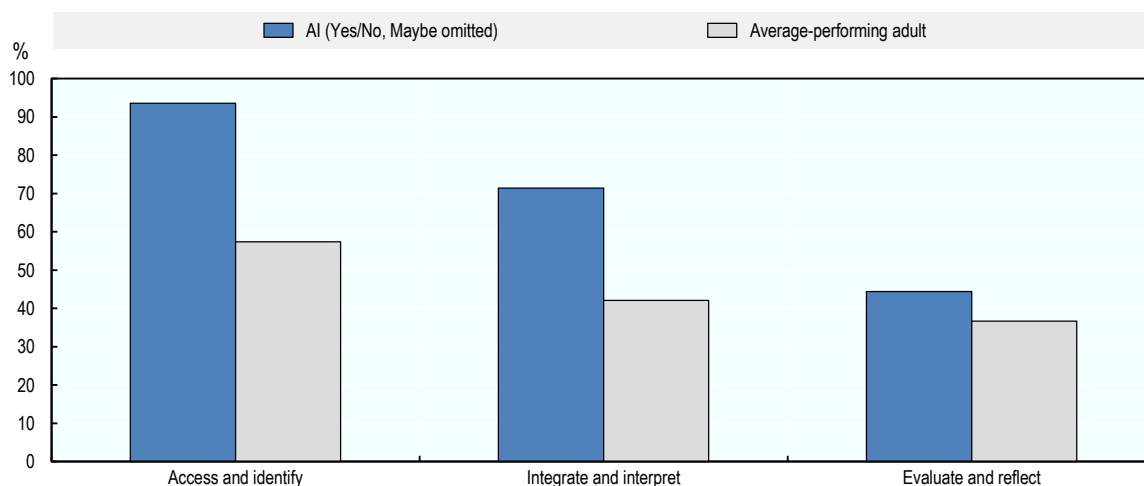
Another goal of education providers could be to trigger those components of foundation skills that prove hard for AI. As shown in Chapter 3, literacy and numeracy are complex constructs. Literacy, for example, involves the use of three cognitive strategies. Individuals should be able *to access and identify* information in texts; *to integrate and interpret* the relations between parts of the text, such as those of cause/effect or problem/solution; and *to evaluate and reflect* on information from texts using own knowledge or ideas (OECD, 2012^[18]). Figure 6.6 shows that not all of these literacy sub-skills are easy for AI. According to experts' evaluations, AI is expected to solve 94% of the questions that require accessing and identifying information and 71% of the questions that involve integrating and interpreting relations in the text. Expected performance on questions containing evaluation and reflection is lower, at 44%.

These findings reflect the technological developments in NLP. As Chapter 2 showed, state-of-the-art AI excels on Question-Answering tasks, such as those of the Stanford Question Answering Dataset (SQuAD) (Rajpurkar et al., 2016^[19]; Rajpurkar, Jia and Liang, 2018^[20]) and the General Language Understanding Evaluation (GLUE) benchmark (Wang et al., 2018^[21]; Wang et al., 2019^[22]). These benchmarks test the ability of systems to answer questions related to texts by accessing/identifying the information containing the right answer. Progress was also registered in natural language inference (NLI) (Storks, Gao and Chai, 2019^[23]). This is the task of “understanding” the relationship between sentences, which comes close to the “integrate and interpret” tasks in PIAAC. By contrast, AI still struggles with language tasks that require logical reasoning and common knowledge (see, for example Yu et al. (2020^[24])). This could explain the low expert ratings on PIAAC questions that require evaluating and reflecting on texts.

However, evaluation and reflection on information in texts is also more challenging for humans. An average respondent in PIAAC has a 37% probability of successfully completing questions in this category, compared to a 57% probability of success on questions requiring the cognitive strategy “access and identify” and 43% on questions involving the “integrate and interpret” strategy. Strengthening people’s ability to evaluate and reflect on texts would not only give them an important advantage over machines. This skill would also enable them to cope with the information overload of the digital age and determine the accuracy and credibility of sources against the background of spreading fake news and misinformation.

Figure 6.6. Literacy performance of AI and average adults by cognitive strategy required in PIAAC questions

Share of literacy questions that AI can answer correctly according to the majority of experts compared to the probability of successfully completing items of average-performing adults



Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/2ckahz>

In Figure 6.1, projections to 2026 suggest that AI systems will likely soon be able to perform the full range of literacy and numeracy tasks on PIAAC. If this is correct, then the objective for education may need to change substantially. With more capable systems, even high proficiency in literacy and numeracy may no longer be sufficient to allow people to compete with AI. In that context, it seems more plausible that adults will begin to work regularly with AI systems for performing literacy and numeracy tasks. AI systems may help them carry out the tasks more effectively than they could do on their own. As a result, the focus of education may need to shift towards teaching students how to use AI systems effectively.

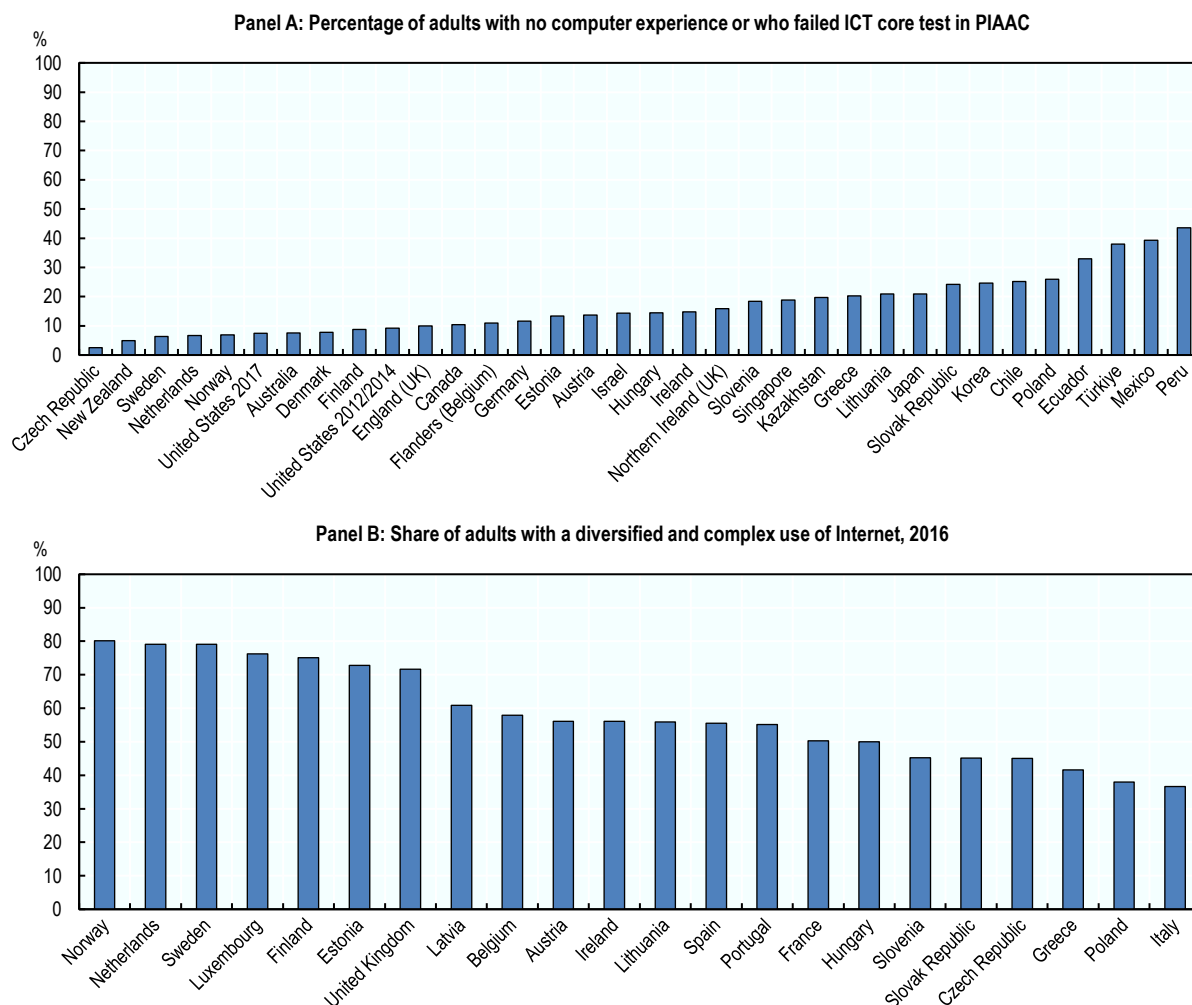
Education systems should also seek to strengthen the digital skills of individuals. These skills would help individuals meet the demands of increasingly digitised workplaces and seize the opportunities brought about by technological advances. Figure 6.7 shows two indicators of the availability of digital competencies in the population (OECD, 2019^[25]). The first is the share of adults who are insufficiently familiar with computers. These adults either reported having no prior computer experience in PIAAC or could not perform basic computer tasks (e.g., using a mouse or scrolling through a webpage) to take part in PIAAC's computer-based assessment (OECD, 2019^[4]). The second indicator is the share of adults with diversified and complex use of the Internet. It is based on previous analysis of the OECD of data from the European Community Survey on Information and Communication Technologies (ICT) Usage in Households and by Individuals and covers fewer countries (OECD, 2019^[25]).

The figure shows that both indicators vary widely across countries. Among countries with available data, Norway, the Netherlands and Sweden have around 80% of the population equipped with skills allowing for diverse and complex Internet use (Panel B). In these countries, as well as in New Zealand and the Czech Republic, less than 7% of the population cannot work with computers (Panel A). By contrast, in Greece and Poland, around 40% of the population can perform many and complex online activities, and one-fifth and one-quarter, respectively, cannot use computers at all. In Peru, the share of adults who cannot use computers exceeds 40%. These latter countries must up-skill large proportions of their adult population to


meet the skill needs emerging from technological change. Another possibility is that the lack of digital skills in these countries slows the spread of new technologies in their economies. This could have negative effects on competitiveness, productivity, innovation and, eventually, on employment.

The current ICT skills of adults shown in Figure 6.7 reflect massive change over the past four decades. Although formal data are not available, in 1980 – before the widespread adoption of computers, the Internet and smartphones – most adults in all countries would likely have failed the ICT core test in PIAAC. They would also have likely said they made no use of the early Internet.

Figure 6.7. Digital skills of adults



Source: Adapted from OECD (2019^[41]), *Skills Matter: Additional Results from the Survey of Adult Skills*, Figure 2.15, <https://doi.org/10.1787/1f029d8f-en>, and OECD (2019^[25]), *OECD Skills Outlook 2019: Thriving in a Digital World*, Figure 4.16, <https://doi.org/10.1787/df80bc12-en>.

StatLink  <https://stat.link/xryga5>

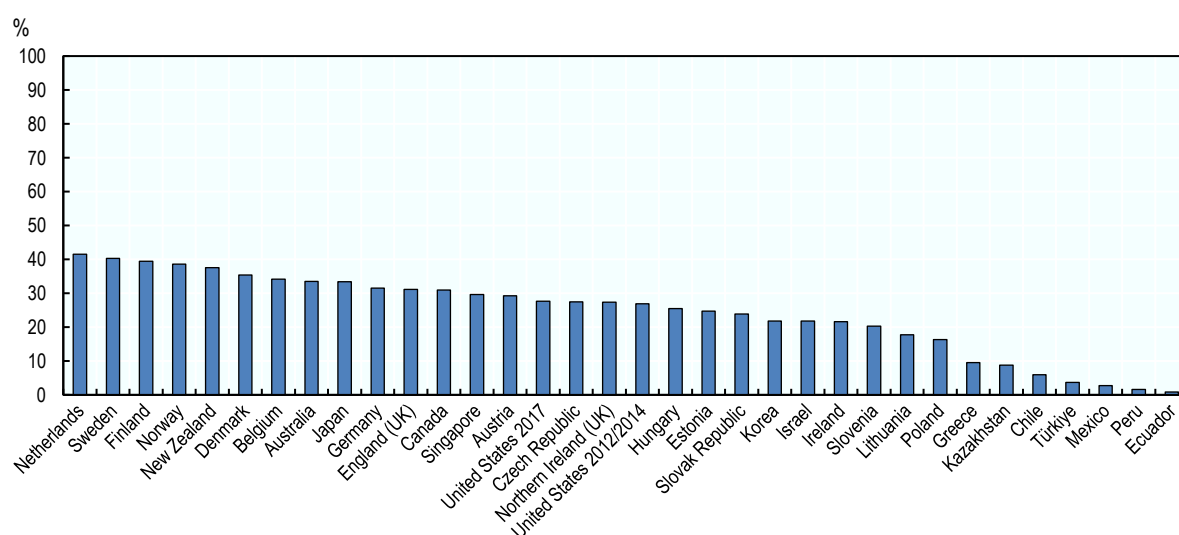
As argued above, the use of diverse skills at work can shelter workers from automation. Therefore, education systems should aim at equipping people with a well-rounded skill set. This would enable people

to adapt to potential changes in their occupations induced by technology. It would also ease their mobility between occupations since diverse skills apply in different work contexts.

Figure 6.8 shows the proportion of working adults with solid skills in three key areas – literacy, numeracy and problem-solving in technology-rich environments.² Concretely, the figure shows the proportion of workers with literacy and numeracy skills at Level 3 or above and problem-solving skills at Level 2 or above (see also OECD (2019^[25])). At 42%, the Netherlands has the highest proportion of working adults with strong skills in all three domains. However, in nine of the participating countries, the share of workers with a well-balanced skill mix is lower than 20%.

Figure 6.8. Proportion of workers with a well-balanced skill set

Proportion of workers with literacy and numeracy skills at Level 3 and above and problem-solving skills in technology-rich environments at Level 2 and above



Source: OECD (2012^[1]; 2015^[2]; 2018^[3]), *Survey of Adult Skills (PIAAC) databases*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023).

StatLink  <https://stat.link/lfn861>

In sum, AI advances in key cognitive skills will likely pose a challenge for education. Many education systems will initially need to up-skill substantial parts of the population to help them keep up with improving AI capabilities in literacy and numeracy. As AI capabilities in cognitive areas continue to improve, education systems may need to substantially shift their approach to focus far more on working with powerful AI systems that have high levels of literacy and numeracy skills. In addition, education will be increasingly expected to strengthen many other skills, including digital skills, to help people develop strong, diverse skill sets. Such skill sets can help people avoid the risks and benefit instead from the opportunities of the AI revolution.

A new approach to assessing AI

This study provided an example of how AI capabilities improve with respect to two key cognitive skills of humans – literacy and numeracy. These skills were selected because they are of key importance at work and in everyday life, and the foundation for acquiring additional skills and knowledge (OECD, 2013^[13]).

Yet, these skills are hard to cultivate, compared to others seen as necessary in the digitised future, such as digital skills. Over the last few decades, literacy and numeracy skills have not changed substantially in most countries, while it did not take long for most people to learn to use a computer or the Internet.

The study showed that AI has developed strong capabilities in literacy and numeracy. According to experts, these capabilities are likely to improve further over the next five years. This raises questions about the possible impacts of advancing AI – about how it will change the ways key skills are used in the workplace and taught in education. Ultimately, to understand how AI will affect future skill use and skill needs, the assessment of AI capabilities should go beyond the general cognitive skills addressed in PIAAC. This would require information on the full range of skills used in occupations and on the proficiency of people with respect to these skills.

This exploratory project is part of a bigger effort by the OECD to assess AI. The AI and the Future of Skills (AIFS) project is developing a comprehensive and authoritative approach to regularly measuring AI capabilities and comparing them to human skills. The capability measures will cover various skill domains that are crucial for work and important in education.

Expert ratings of AI on education tests are an important tool in this approach. Over the past years, the project has repeated and extended the gathering of these expert judgements. For example, the project explored the use of a large-scale expert survey to assess potential AI performance on the PISA science test. It also collected expert judgement on whether AI can perform occupational tests from vocational training and education.

More recently, the AIFS project started to use information from direct tests of AI systems. These include benchmarks, competitions and formal evaluation campaigns that apply AI techniques directly to various kinds of tasks, producing success or failure. The project is developing an approach to inventorying and selecting high-quality direct tests for the assessment. It is working to develop an approach to synthesise the information from such evaluation tests into indicators of AI performance that are understandable and policy relevant.

To help policy makers understand the implications of the AI measures, the project will link them to existing taxonomies of occupational tasks (e.g. ESCO (European Commission, n.d.^[26]), O*NET (National Center for O*NET Development, n.d.^[27])). These taxonomies provide a way of systematically considering the range of skills needed for performing work tasks and the way these different skills are brought together in occupations.

Moreover, the project will map the AI performance measures to information about the skill proficiency of workers. With the rapid development of AI across a wide range of skill areas, such an approach can systematically identify which skills will likely become obsolete and which may become more significant for work and in education.

The project's first methodology report described its initial work (OECD, 2021^[28]). Successive volumes in this series will describe the development of the set of AI measures and the project's explorations in their use. Armed with this information, policy makers can better understand the implications of AI for education and work.

References

- Bianchini, S., M. Müller and P. Pelletier (2022), “Artificial intelligence in science: An emerging general method of invention”, *Research Policy*, Vol. 51/10, p. 104604, <https://doi.org/10.1016/j.respol.2022.104604>. [17]
- Cobbe, K. et al. (2021), “Training Verifiers to Solve Math Word Problems”. [10]
- Cockburn, I., R. Henderson and S. Stern (2018), *The Impact of Artificial Intelligence on Innovation*, National Bureau of Economic Research, Cambridge, MA, <https://doi.org/10.3386/w24449>. [16]
- Elliott, S. (2017), *Computers and the Future of Skill Demand*, Educational Research and Innovation, OECD Publishing, Paris, <https://doi.org/10.1787/9789264284395-en>. [5]
- European Commission (n.d.), *The ESCO Classification*, <https://esco.ec.europa.eu/en/classification> (accessed on 24 February 2023). [26]
- Hendrycks, D. et al. (2021), “Measuring Mathematical Problem Solving With the MATH Dataset”. [9]
- Lindström, A. and S. Abraham (2022), “CLEVR-Math: A Dataset for Compositional Language, Visual and Mathematical Reasoning”. [11]
- Manyika, J. et al. (2017), *A Future That Works: Automation, Employment, and Productivity*, McKinsey Global Institute (MGI). [12]
- National Center for O*NET Development (n.d.), *O*NET 27.2 Database*, <https://www.onetcenter.org/database.html> (accessed on 24 February 2023). [27]
- OECD (2021), *AI and the Future of Skills, Volume 1: Capabilities and Assessments*, Educational Research and Innovation, OECD Publishing, Paris, <https://doi.org/10.1787/5ee71f34-en>. [28]
- OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, <https://doi.org/10.1787/eedfee77-en>. [14]
- OECD (2019), *OECD Skills Outlook 2019: Thriving in a Digital World*, OECD Publishing, Paris, <https://doi.org/10.1787/df80bc12-en>. [25]
- OECD (2019), *Skills Matter: Additional Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/1f029d8f-en>. [4]
- OECD (2018), *Survey of Adult Skills (PIAAC) database*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023). [3]
- OECD (2015), *Survey of Adult Skills (PIAAC) database*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023). [2]
- OECD (2013), *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264204256-en>. [13]
- OECD (2012), *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264128859-en>. [18]

- OECD (2012), *Survey of Adult Skills (PIAAC) database*, <http://www.oecd.org/skills/piaac/publicdataandanalysis/> (accessed on 23 January 2023). [1]
- OpenAI (2023), *Introducing ChatGPT*, <https://openai.com/blog/chatgpt> (accessed on 23 February 2023). [8]
- Ouyang, L. et al. (2022), “Training language models to follow instructions with human feedback”. [7]
- Rajpurkar, P., R. Jia and P. Liang (2018), “Know What You Don’t Know: Unanswerable Questions for SQuAD”. [20]
- Rajpurkar, P. et al. (2016), “SQuAD: 100,000+ Questions for Machine Comprehension of Text”. [19]
- Russell, S. and P. Norvig (2021), *Artificial Intelligence: A Modern Approach*, Pearson. [6]
- Squicciarini, M. and H. Nachtigall (2021), “Demand for AI skills in jobs: Evidence from online job postings”, *OECD Science, Technology and Industry Working Papers*, No. 2021/03, OECD Publishing, Paris, <https://doi.org/10.1787/3ed32d94-en>. [15]
- Storks, S., Q. Gao and J. Chai (2019), “Recent Advances in Natural Language Inference: A Survey of Benchmarks, Resources, and Approaches”. [23]
- Wang, A. et al. (2019), “SuperGLUE: A Stickier Benchmark for General-Purpose Language Understanding Systems”. [22]
- Wang, A. et al. (2018), “GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding”. [21]
- Yu, W. et al. (2020), “ReClor: A Reading Comprehension Dataset Requiring Logical Reasoning”. [24]

Notes

¹ The use of each skill is assessed with a number of variables from PIAAC: *writing* – frequency of writing letters, memos or mails; articles; reports; or of filling in forms; *digital skills* – frequency of using the Internet for mail; for finding work-related information; for conducting transactions; frequency of using spreadsheets; Microsoft Word; programming languages; or online real-time discussions; *problem solving* – frequency of solving complex problems at work; *learning at work* – frequency of learning from co-workers/supervisors; of learning-by-doing; keeping up to date; *influencing skills* – frequency of teaching people; giving presentations; selling; advising people; influencing people; negotiating with people; *co-operative skills* – co-operating with co-workers more than half of the time; *organising skills* – planning others’ activities; *physical skills* – working physically for longer time. A skill is used daily if the respondent reports daily use of at least one of the activities used to measure the skill.

² See note 1 in Chapter 1.



From:

Is Education Losing the Race with Technology? AI's Progress in Maths and Reading

Access the complete publication at:

<https://doi.org/10.1787/73105f99-en>

Please cite this chapter as:

OECD (2023), "Implications of evolving AI capabilities for employment and education", in *Is Education Losing the Race with Technology?: AI's Progress in Maths and Reading*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/ed919621-en>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.