

# Executive summary

Artificial intelligence (AI) and robotics are major breakthrough technologies that are transforming the economy and society. To understand and anticipate this transformation, policy makers must first understand what these technologies can and cannot do. The OECD launched the *Artificial Intelligence and the Future of Skills* project to develop a programme that could assess the capabilities of AI and robotics and their impact on education and work. This report represents the first step in developing the methodological approach of the project. It reviews existing taxonomies and tests in psychology and computer science, and discusses their strengths, weaknesses and applicability for assessing machine capabilities.

---

## **Assessing AI and robotics capabilities is a necessary foundation for understanding their implications for education, work and the larger society.**

---

An ongoing programme of assessment for AI and robotics will add a crucial component to the OECD's set of international comparative measures that help policy makers understand human skills. The Programme for International Student Assessment (PISA) describes the link between the education system and the development of human skills, while the Programme for the International Assessment of Adult Competencies (PIAAC) links those skills to work and other key adult roles. A programme for assessing AI and robotics capabilities will relate human skills to these pivotal technologies, thereby providing a bridge from AI and robotics to their implications for education and work, and the resulting social transformations in the decades to come.

---

## **There are a number of taxonomies and tests of human skills. These provide different perspectives and opportunities for understanding AI capabilities.**

---

Taxonomies stemming from the cognitive psychology literature are hierarchical models of broad cognitive abilities, such as fluid intelligence, general memory/learning, visual and auditory perception, assessed by factor analysis of cognitive ability tests. These tests have been widely used and validated for assessing human skills.

Research interest in social and emotional skills is growing and their testing is advancing. These skills are focused on individuals' personality, temperament, attitudes, integrity and personal interaction. Recent research considers not just individual abilities but also collective ones. This emerging literature studies the factors of "collective intelligence" and is developing tests to measure them.

Education research has also contributed to defining and shaping the understanding of human skills. This domain focuses on subject-specific knowledge (e.g. in mathematics, biology and history), basic skills such as literacy and numeracy, and more complex transversal skills such as problem solving, collaboration, creativity, digital competence and global competence. A wide range of tests is available from international and national large-scale educational assessments.

---

## Skills can be linked to work tasks and occupations, and measured through complex vocational tests.

---

Another major area – industrial-organisational psychology – links abilities to tasks specific to particular occupations. The resulting comprehensive occupation taxonomies classify occupations by work tasks, and the required skills, knowledge and competences. The most widely used classifications are the Occupational Network (O\*NET) database of the US Department of Labor and the European classification of skills, competences, qualifications and occupations (ESCO). Assessments in this domain comprise a variety of vocational and occupational tests.

---

## Healthy human adults share some basic skills that AI systems do not have.

---

Many taxonomies for assessing skills overlook ubiquitous low-level or basic cognitive skills. These are rarely assessed in human adults because there are few meaningful individual differences in the absence of severe disability. However, AI systems do not necessarily have these skills (e.g. navigating in a complex physical environment, understanding basic language or knowing basic rules of the world). Taxonomies and assessments for these skills are found in the fields of animal cognition, child development and neuropsychology. A recently emerging field assesses basic (low-level) skills of AI systems drawing on these fields of psychology.

---

## Evaluating AI and robotics systems is challenging and applying human tests can be misleading.

---

AI assessment focuses on functional components of intelligent mechanisms, such as knowledge representation, reasoning, perception, navigation and natural language processing. These are strongly linked to the underlying technique used by the mechanism. Many components overlap with the ability categories developed in psychology for humans, but the match is not exact. In addition, many capabilities that AI is developing – such as language identification and the generation of realistic images – are not well covered by human skill taxonomies or tests.

Moreover, the design of human tests takes for granted that the test takers all share basic features of human intelligence, which might be radically different from AI. For example, integrating basic skills, such as natural language understanding and object recognition, is easy for humans. However, most AI systems are trained to perform a specific narrow task, but they are not (or are rarely) able to integrate and apply these to perform a different type of task. This makes it difficult to generalise from an AI system's performance on a specific human ability test to an underlying AI skill, let alone infer general intelligence.

---

## Different types of empirical assessments can gauge AI capabilities, but these are scattered and not systematic.

---

A multitude of benchmarks and competitions assess and compare AI systems empirically. However, these have not yet been systematically classified. Increasingly, more institutions carry out rigorous evaluation campaigns to assess the capabilities of AI and robotics systems. These include the evaluation of individual functions, i.e. self-contained units of capability, such as self-localisation. They also include evaluation of complete tasks that constitute a meaningful activity, such as autonomous driving and text summarisation. Evaluation of AI systems is particularly well developed in certain areas, such as language understanding. Machine translation, in particular, is a field that holds many lessons for assessing AI.

---

**A systematic assessment of AI demands a comprehensive framework that covers all human skills necessary for work and life.**

---

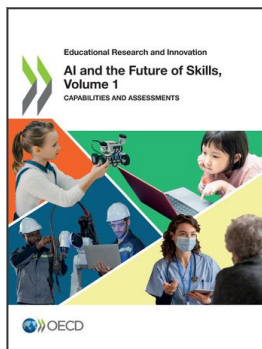
Providing valid, reliable and meaningful measures of AI and robotic capabilities requires a comprehensive approach that brings together different research traditions and complementary methodologies. The goal should be to address the full range of relevant human capabilities; the extra capabilities needed to consider for AI (because they are difficult for AI and often neglected in lists of human skills); and the full range of valued tasks that appear in education, work and daily life.

---

**A robust methodology involves understanding how AI and robotics systems are assessed and bringing together different assessment approaches.**

---

A multidisciplinary approach needs theoretical underpinning that considers the challenges linked to assessing AI and robotic capabilities with regard to human skills. The different disciplinary approaches can be organised across two dimensions. One relates to whether skill taxonomies and tests measure *primarily human or primarily AI capabilities*. The second dimension is whether they measure *single (isolated) capabilities or complex tasks* that require multiple capabilities. Future systematic assessment of AI capabilities should bring together different assessments along these two dimensions and skilfully integrate their potential to draw valid implications for the future of work and education.



**From:**  
**AI and the Future of Skills, Volume 1**  
Capabilities and Assessments

**Access the complete publication at:**  
<https://doi.org/10.1787/5ee71f34-en>

**Please cite this chapter as:**

OECD (2021), "Executive summary", in *AI and the Future of Skills, Volume 1: Capabilities and Assessments*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/568d57ea-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

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>.