# Foreword

Artificial intelligence (AI) has emerged as a significant area of development. Its integration into various sectors necessitates a comprehensive understanding of its capabilities, especially in relation to human skills. The AI and the Future of Skills (AIFS) project by the OECD's Centre for Education Research and Innovation (CERI) has undertaken this task, aiming to provide a methodological framework for assessing and comparing AI capabilities to human skills. This framework should provide a basis for informed discussions on AI's impact on education, work and society.

The project has undergone two phases of developing a rigorous approach to assessing AI's capabilities. The first phase focused on identifying relevant AI capabilities and the tests best suited to evaluate them. Leveraging insights from various fields including computer science, psychology and education, the project offered a multi-disciplinary perspective on the challenges and prospects of assessing AI.

The second phase, the focus of this report, further refines the methodology of the assessment. It encompasses a range of exploratory AI evaluations to identify most promising practices for systematically and periodically assessing AI. These explorations are threefold. First, by assessing AI capabilities with OECD's education tests using expert judgement, the project explored ways to understanding AI's progress in competencies that are traditionally human – competencies in reading, mathematics and science. Second, the project asked experts to rate AI on real-world occupational tasks, such as those encountered in nursing or product design, to provide critical insights into AI's application potential. By situating AI within these occupational contexts, we gain a clearer picture of its impending impact on the economy. Third, the project considered the vast and evolving benchmarks available in AI research that result from direct assessments of AI systems.

These methods, while promising, are not without their challenges. This report underscores the difficulties in solely relying on expert judgements to evaluate AI. While expert input is valuable, achieving consensus, particularly in novel domains, can be challenging. Moreover, the variability in AI applications and the intricacies of real-world tasks suggest the need for diverse evaluation metrics. Therefore, the project decided to integrate both expert judgements and direct AI measures in its subsequent phase to provide a thorough and balanced evaluation. This integrative approach aims to provide decision-makers with a nuanced understanding of AI's capabilities.

The next project phase intends to produce an integrated assessment framework for AI. This will contain a set of key AI indicators that can serve as reference points for various stakeholders. These indicators, informed by a combination of expert input and direct assessments, will offer guidance for policy formulation and implementation.

As AI continues to evolve, having a clear framework to understand its capabilities becomes crucial. The AIFS project's efforts contribute to this understanding, laying the groundwork for informed decisions in education and employment sectors. This work reflects OECD's commitment to producing rigorous, evidence-based insights that can inform decision-making in the context of AI's continued growth and integration into various sectors.
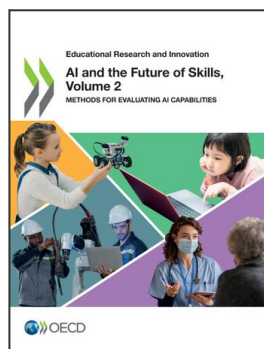
From:

# AI and the Future of Skills, Volume 2
Methods for Evaluating AI Capabilities

Access the complete publication at:
https://doi.org/10.1787/a9fe53cb-en