

### Why are indicators on the digitalisation of scientific research needed?

Ministers from OECD countries and partners meeting at the OECD Ministerial Meeting held in Daejeon (Korea) in 2015 recognised in their joint declaration ([www.oecd.org/sti/daejeon-declaration-2015.htm](http://www.oecd.org/sti/daejeon-declaration-2015.htm)) that science, technology and innovation (STI) are being revolutionised by the rapid evolution of digital technologies. These technologies are changing the way in which scientists work, collaborate and publish; increasing the reliance on access to scientific data and publications; opening new avenues for public engagement and participation in science and innovation; facilitating the development of research co-operation between businesses and the public sector; and contributing to the transformation of innovation. The OECD was asked to monitor this transformation and invited to convene the international community working on STI data and indicators to develop new thinking and solutions for empirical evidence to guide policy. The 2016 OECD Blue Sky Forum (<http://oe.cd/blue-sky>) identified the digitalisation of STI both as a priority object of measurement and as a fundamental enabler of future statistical and analytical work (OECD, 2018).

### What are the challenges?

Tracing the resources invested in expanding knowledge of the possibilities of digitalisation and their outputs requires data infrastructures with a wider scope that provide high levels of detail. Research domains evolve over time and new paradigms emerge that draw upon existing research. Conventional practices and categories, though necessary, are not sufficient to track the digitalisation of research and its impact on digital transformation. New text-mining tools enable the extraction of relevant information from qualitative sources to produce indicators about the nature, method and purpose of research. Provided the underlying data are sufficiently rich and comprehensive and the use of the mining techniques is sound, it is possible to estimate with some accuracy the share of scientific production related to topics such as AI or the percentage of funding agency support for R&D projects that makes use of AI tools. However, the databases that can in principle support this type of analysis are often fragmented, difficult to gain access to, and challenging to combine and use.

Data users increasingly demand access to a very fine-grained level of detail that can be hard to reconcile with the preservation of confidentiality and, depending on the way in which data are used to inform decisions, may induce selective disclosure of information by individuals or organisations, if this is in their interest. Because project funding and career decisions are influenced by quantitative indicators, a roadmap for science measurement in the digital age has to address the potential trade-off between reconciling data availability and integrity arising from exploiting new sources. Another key challenge is to connect measures of science digitalisation to policy measures.

### Options for international action

Initiatives to leverage information captured by digital systems to measure both scientific research and its digital transformation has been supported by consortia of independent non-profit organisations, academia and business that promote the growing adoption of standards. These standards target digital objects that capture multiple dimensions of scientific activity, such as documents, projects, data and code, allowing them to be persistently identified and retrieved, and also enabling them to be interpreted, linked or attributed to individuals and their organisations. The ORCID identifier (<https://orcid.org>) is a salient example of the new data infrastructure about science and research, having attained high adoption rates among serving researchers, their organisations, and many publishers and research funders. Although the main motivation for STI data infrastructures with such features is to facilitate scientific and related management processes rather than statistics, these data can provide the basis for higher quality statistics and indicators. Administrative reporting requirements often set out the basis for what official surveys can collect from respondents. Different groups within the science and science management community can work to ensure the integrity of the information, but there are significant career and commercial interests at play, as well as co-ordination challenges.

Because of the current limitations of available sources of information about some rapidly evolving dimensions of science, the OECD has been experimenting with the development of a dedicated survey tool. The OECD International Study of Scientific Authors (ISSA) is a new global survey-based initiative that was first piloted in 2015 to explore the changing scientific publishing landscape; its 2018 edition examined how science is going digital. ISSA explores a range of key dimensions that are potentially relevant across all fields of science:

