

Chapter 9

DIGITAL INNOVATION

KEY FINDINGS

- Around one-third of patents owned in OECD countries are related to information and communication technologies. This share has fallen over the last decade but has increased markedly in the People's Republic of China (hereafter "China"), India and the Russian Federation.
- China has rapidly increased its contributions to the science underpinning digital technologies, overtaking the United States in the volume of contributions to computer science journals. However, the United States remains ahead in terms of highly cited work.
- Two-thirds of academic authors create new data or code as part of their published scientific work. However, barriers to sharing limit re-usability. Additionally, the technologies used and the resulting benefits vary greatly between scientific fields and across countries.
- The research community is using unique, persistent and pervasive international digital identification (UPPI) systems as one way to join relevant data held in different digital systems (e.g. those for funding applications and published outputs). However, only widespread adoption and use of these UPPIs can maximise their benefits.
- Digital technologies are playing a direct role in efforts to manage the COVID-19 pandemic and find a vaccine. In particular, artificial intelligence (AI) and associated technologies such as machine learning are finding innovative applications to a wide array of challenges driven by COVID-19. However, such applications work by identifying patterns in data. They require large amounts of data to find these patterns; the outputs are only as good as the training data.
- Open and collaborative approaches will allow the widest pool of researchers possible to access the tools and data needed. In this way, they can devise innovative uses of AI and maximise the chances of finding effective containment measures and treatments. Innovative incentives, such as research prizes and hackathons, can also help focus resources on this pressing societal challenge.

Introduction

Digital technologies are both a key area of research and innovation and themselves a key foundation for developments in research and innovation. This chapter looks at recent trends in innovation in digital technologies before examining how digitalisation and data are profoundly impacting the science, research and innovation that help drive those technological developments. Subsequently, it looks at how digital technologies, and particularly artificial intelligence (AI), are helping search for ways to manage and treat the COVID-19 pandemic. Finally, the chapter touches upon the increasing role of digital technologies in managing national science and innovation systems and policy.

Around one-third of patents owned in OECD countries are related to information and communication technologies (ICTs). This share has fallen over the last decade but has increased markedly in China, India and the Russian Federation. These three countries have moved from being mainly specialised in ICT manufacturing and software production to other parts of the value chain, including product and component design.

China has also rapidly increased its contributions to the science underpinning digital technologies. Indeed, it has overtaken the United States in the volume of contributions to computer science journals. However, when citation levels – which provide one indication of impact – are considered, the United States remains ahead.

Digital technologies and data are increasingly shaping and facilitating scientific research. Scientists are generally positive about the impacts of digitalisation on their work, with digital technologies facilitating science across borders, collaboration and efficiency. On average, two-thirds of academic authors create new data or code as part of their published scientific work. However, barriers to sharing limit re-usability. Furthermore, the technologies used and the resulting benefits vary greatly between scientific fields and across countries. Policy makers can help identify and mainstream best practices across disciplines.

The research community is overcoming the challenges of joining up relevant data held in different digital systems (e.g. those for funding applications and published outputs). This helps reduce reporting burdens on scientists and to better understand and monitor national and international science and innovation systems. Establishing unique, persistent and pervasive international digital identification (UPPI) systems for researchers and research organisations is one example of these practices. However, only widespread adoption and use of these UPPIs can maximise their benefits. Policy makers can drive uptake by promoting their use in interactions such as research funding applications and research outputs, including journal articles and academic papers.

Digital technologies are also playing a direct role in efforts to manage the COVID-19 pandemic and find a vaccine. In particular, AI and associated technologies such as machine learning are finding innovative applications to a wide array of challenges driven by COVID-19. However, such applications work by identifying patterns in data. They require large amounts of data to find these patterns; the outputs are only as good as the training data.

Open and collaborative approaches will allow the widest pool of researchers possible to access the tools and data needed. In this way, they can devise innovative uses of AI and maximise the chances of finding effective containment measures and treatments. Innovative incentives, such as research prizes and hackathons, can also help focus resources on this pressing societal challenge.

Innovation in digital technologies

Rapid innovation can be readily observed in many digital products in daily use. For example, smartphones and the networks they rely on are moving to implement 5G technology despite 4G (LTE) networks only beginning commercial rollout a decade ago. At the same time, online email and video streaming services are implementing increasingly sophisticated features underpinned by machine learning and AI. These advances culminate from a vast array of research and innovation activities.

Patents are often used to protect ICT-related technologies in relevant areas. These include high-speed networks, mobile communication, digital security, sensor and device networks, high-speed computing, large-capacity and high-speed storage, large-capacity information analysis, cognition and meaning understanding, human interface, and imaging and sound technologies (Inaba and Squicciarini, 2017^[1]). Importantly, patent protection is granted only for a product or a process that brings a novel technical solution. As such, looking at the volumes of such patents granted can offer one indication of the extent of innovation in ICT-related technologies.

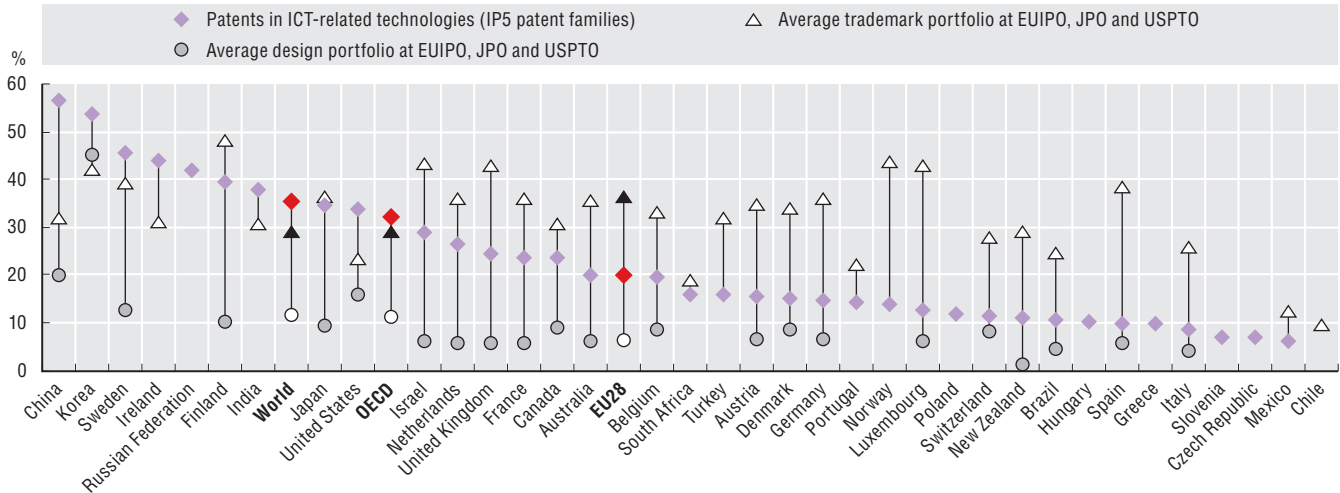
Over 2014-17, ICT-related technologies accounted for about one-third of all IP5 patent families filed by owners in OECD countries (Figure 9.1). These data refer to filings in at least two intellectual property offices, including at least one of the five largest offices globally. This represents a decrease on the share observed a decade earlier (37%). In contrast, the ICT-related share of IP5 patent families owned in China increased by one-fifth. This makes China's IP5 patent portfolio the most specialised in ICT. In the Russian Federation, India and Portugal, the share of patents related to ICT more than doubled. Meanwhile, it increased by almost two-thirds in Ireland, aided by several large technology companies establishing operations there.

Design patents protect the “look and feel of products”. A significant portion of these patents can relate to ICT product designs. ICT designs, for example, comprise almost half of the average design portfolio held by Korean firms across the European Union Intellectual Property Office, the Japan Patent Office and the United States Patent and Trade Office. For other countries, these average shares are much lower. However, they still reach 10% to 20% in China, Sweden, Finland and the United States, showing the importance of ICT product design. In comparison to 2004-07, ICT designs in 2014-17 maintained their share, relative to designs in general in the US market (+0.1 percentage point). In contrast, they declined as a share of all design filings in Europe (-0.8 percentage points) and in Japan (-2.5 percentage points).

Meanwhile, China has moved beyond ICT manufacture to aspects of design. It doubled its share of ICT design patents filed in the United States (from 13% to 26%). It also increased its share of ICT designs registered in Japan by almost one-third (to 21%). Finally, it maintained its registered design share in European markets (16%).

Figure 9.1. ICT-related patents, trademarks and designs, 2014-17

As a percentage of total IP5 patent families or total trademarks and total design patents at the EUIPO, JPO and USPTO, by country of ownership



Notes: Patents protect technological inventions (i.e. products or processes providing new ways of doing something or new technological solutions to problems). IP5 patent families are patents filed in at least two offices worldwide, including one of the five largest IP offices: the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the US Patent and Trademark Office (USPTO) and the National Intellectual Property Administration of the People’s Republic of China (NIPA). Data refer to IP5 families, by filing date, according to the applicants’ residence using fractional counts. Patents in ICT are identified using the list of IPC codes in Inaba and Squicciarini (2017^[1]). Only economies with more than 250 patent families in the periods considered are included. Data for 2016 and 2017 are incomplete. StatLink contains more data.

Source: OECD based on OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats> (accessed in March 2020).

StatLink <https://doi.org/10.1787/888934192433>

The share of trademarks that are ICT-related and registered by organisations in OECD countries grew in all markets considered. The highest increase was observed in 2014-17 in the European market (up six percentage points to 37% from 2004 to 2007). There was similar growth in the US market (up five percentage points to 24%). Trademarks filed in the Japanese market also increased significantly (up 23 percentage points to 36%).

Overall, OECD countries seem to move progressively towards ICT IP bundling strategies. Such approaches place relatively more emphasis on the look and feel of products and on extracting value from branding. The reverse is true in countries such as China, India and the Russian Federation, which appear to be pursuing technological catch-up strategies, and to be protecting their products through designs and brands (OECD, 2017^[2]).

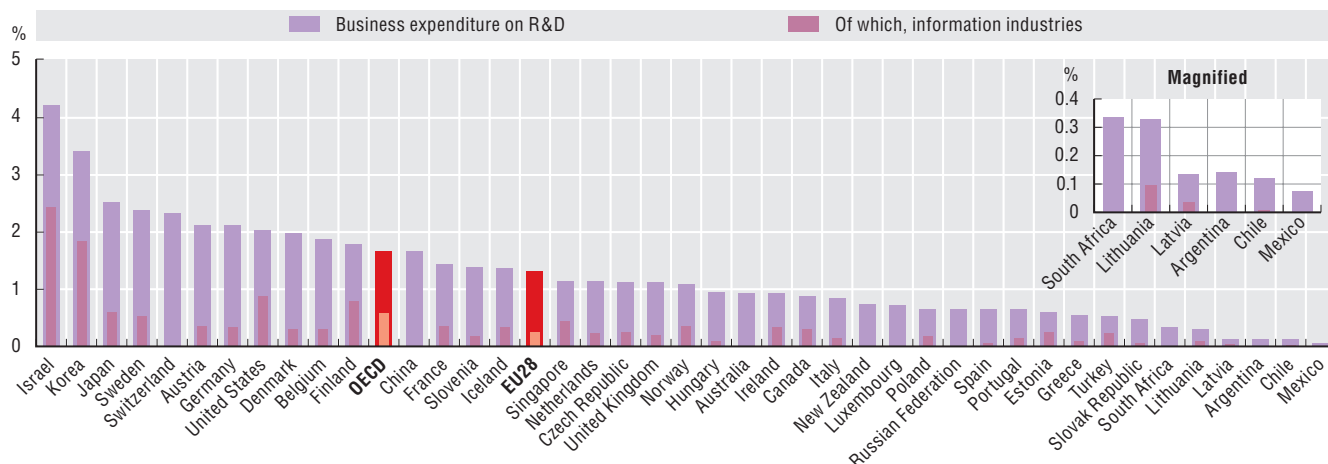
It is generally accepted that computer programs should be protected by copyright, whereas apparatus using computer software or software-related inventions should be protected by patent (WIPO, n.d.^[3]). As such, copyright is relevant in protecting certain elements related to digital technologies – notably software code. Even free and open-source licences rely on copyright law to enforce their terms. However, copyright protection is formality-free in the 178 countries party to the Berne Convention for the Protection of Literary and Artistic Works. This means that protection does not depend on compliance with formalities such as registration or deposit of copies. As such, there are no structured registers or databases for copyrights analogous to those for patents from which to derive similar statistics. Nevertheless, the OECD has started developing experimental approaches to measure contributions to open source software (OECD, 2019^[4]).

In addition, patent applications have increased in-step with the digital transformation. This has increased both the volume and complexity of patent examinations. This, in turn, has led to longer lags between application and any eventual granting of patents (WIPO, 2019^[5]). This may particularly create frictions in the area of digital technology and especially in ICTs, given the speed and complexity of technological development.

Research and development (R&D) is important in driving these advances in digital technologies. Businesses undertake the majority of R&D. The “information industries” – which comprise producers of ICT goods and services, as well as producers of digital content – contribute strongly in Israel and Korea. These countries have particularly high business R&D intensity (R&D expenditure as a share of gross domestic product), with the information industries accounting for over half of this (Figure 9.2). Firms in the information industries also undertake over 40% of all business R&D in Finland, Estonia, Turkey and the United States, confirming the knowledge-intensive nature of these industries.

Figure 9.2. Business R&D expenditure, total and information industries, 2017

As a percentage of gross domestic product



Notes: R&D = research and development. “Information industries” comprise ISIC Rev.4 divisions: “Computer, electronic and optical products” (26), “Publishing, audiovisual and broadcasting activities” (58 to 60), “Telecommunications” (61) and “IT and other information services” (62 to 63). Zone estimates (OECD and EU28) correspond to member countries’ R&D intensity averages weighted by GDP in purchasing power parity. For information industries, they exclude countries where no data are available: Australia, Mexico, Luxembourg, New Zealand and Switzerland for the OECD aggregate, and Bulgaria, Croatia, Cyprus,¹ Luxembourg and Malta for EU28. Data on total business expenditure on R&D (BERD) refer to 2017 except for South Africa (2016). Information industries values relate to the same reference year where possible or are based on shares for the most recent available year: Chile (2015), Korea (2015) and the United Kingdom (2016). Not available for Australia, China, Luxembourg, Mexico, New Zealand, the Russian Federation, South Africa and Switzerland. For Singapore, data that refer to the information industry component is estimated based on 2013 shares.

1. Note by Turkey

The information in this document with reference to “Cyprus” relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Turkey recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Turkey shall preserve its position concerning the “Cyprus issue”.

Note by all the European Union Member States of the OECD and the European Union

The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

Source: OECD based on OECD, ANBERD (database), <http://oe.cd/anberd> and OECD, Main Science and Technology Indicators (database), <http://oe.cd/msti> (accessed in March 2020).

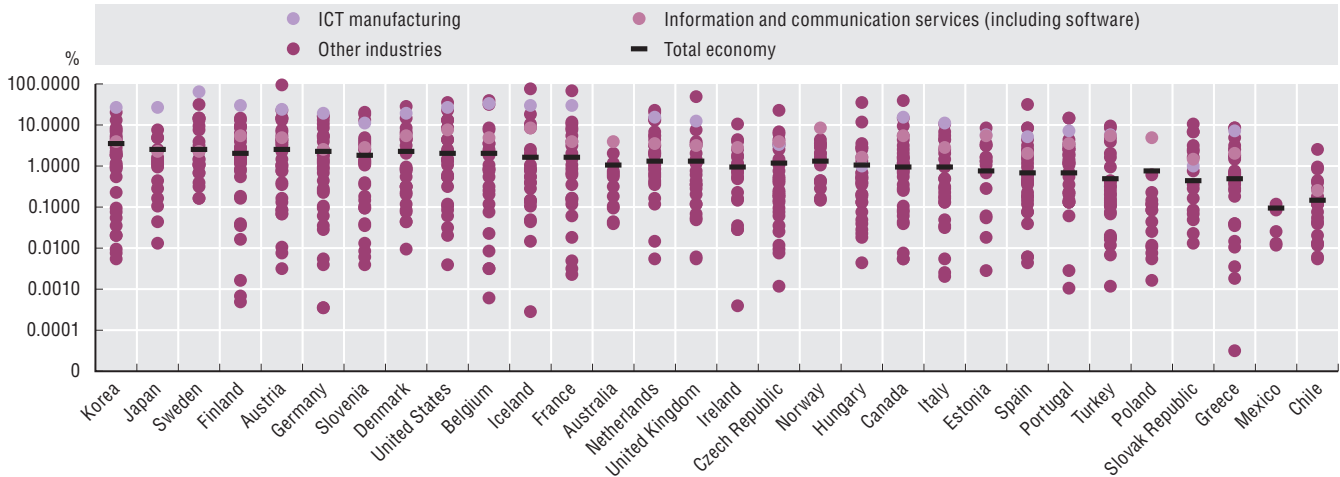
StatLink <https://doi.org/10.1787/888934192452>

Software development and R&D are closely intertwined (OECD, 2015_[6]; OECD/Eurostat, 2018_[4]). Firms mainly producing software, a component of the wider information and communication services sector,¹ are among the most R&D-intensive firms across most countries (Figure 9.3). Similarly, the ICT manufacturing industry is above the average R&D intensity in all the OECD countries presented. In general, both ICT manufacturing and information and communication services report higher than average incidence of innovation activities more broadly² (Figure 9.4).

Advances in scientific knowledge underpin developments in a wide range of digital technologies and techniques. The field of computer science, which contributes towards advances in areas such as machine learning and AI, is just one example. Over the last decade, China almost trebled its contribution to computer science journals. In so doing, it overtook the United States in the production of scientific documents in this field. However, China’s share in the world’s top-cited documents (top 10%, normalised by type of document and field) is close to 9%, remaining well below the United States at 15%. The share of highly cited papers published by authors based in China has nonetheless more than doubled since 2008 (Figure 9.5).

Figure 9.3. R&D intensity of ICT and other industries, 2016

As a percentage of gross value added in each industry, log scale



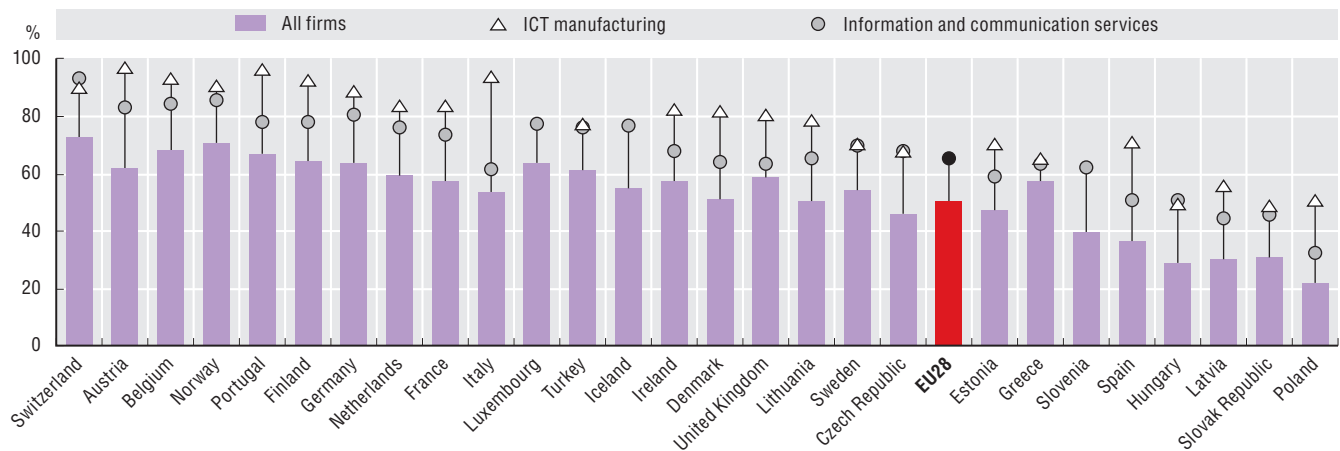
Notes: ICT = information and communication technology. For Korea, Sweden, Germany, Denmark, Island, Spain, Portugal and Turkey, data refer to 2015. For Chile and Canada, data refer to 2014. For Poland and Norway, data refer to 2017. Figures are based on estimates of business R&D by industry reported on a main activity basis according to ISIC Rev.4. R&D intensity has been calculated for each industry, where both R&D and Value Added (VA) data were available. These ratios are sensitive to the statistical units used in both frameworks. A broader discussion is available in Galindo-Rueda and Verger (2016^[7]). In particular, national practices differ in respect to the treatment of large and complex multi-activity enterprises and those firms specialised in providing R&D services. StatLink contains more data.

Source: OECD based on OECD, ANBERD (database), <http://oe.cd/anberd>, OECD, STAN (database), <http://oe.cd/stan>, OECD, National Accounts (database) and OECD, Research and Development Statistics (database), <http://oe.cd/rds> (accessed in March 2020).

StatLink <https://doi.org/10.1787/888934192471>

Figure 9.4. Businesses that have either introduced an innovation or have any kind of innovation activity, 2016

As a percentage of all businesses in the relevant sector



Notes: ICT = information and communication technology. ICT manufacturing is unavailable for Luxembourg, Iceland, EU28 and Slovenia. “All firms” refers to “Innovation core activities”, defined by European Commission ruling 995/2012, to comprise NACE activities: B (Industry), C (Manufacturing), D (Electricity, gas, steam and air conditioning supply), E (Construction), G46 (Wholesale trade except motor vehicles), H (Transportation and storage), J (Information and communication), K (Financial and insurance services), M71 (Architectural and engineering activities; technical testing and analysis), M72 (Scientific research and development), M73 (Advertising, market research). “ICT manufacturing” refers to “Manufacture of computer, electronic and optical products”.

Source: OECD based on Eurostat, Community Innovation Survey (database) (accessed in March 2020).

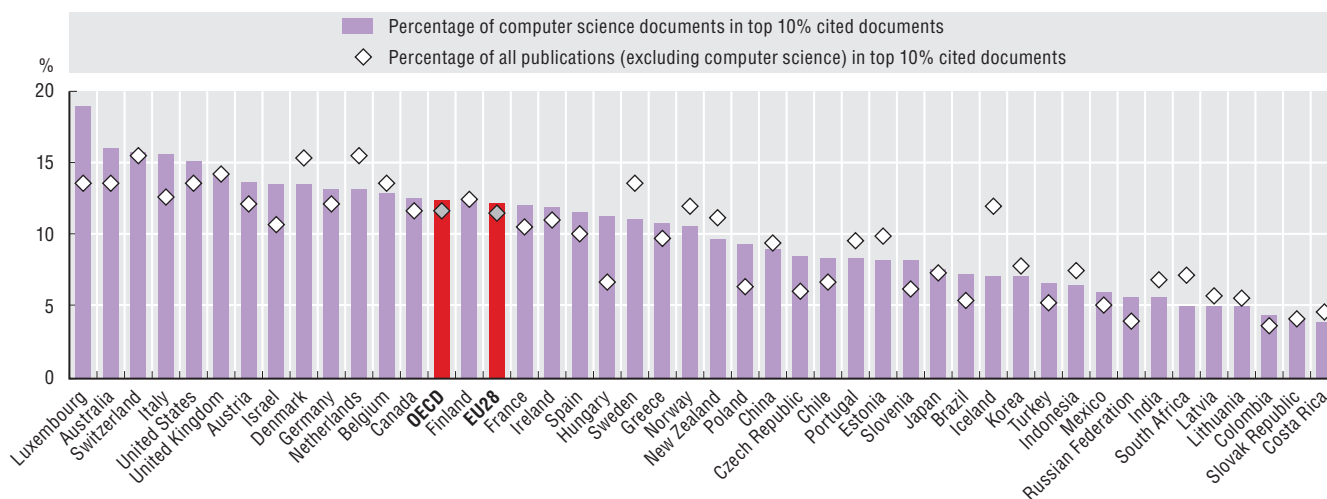
StatLink <https://doi.org/10.1787/888934192490>

In countries including Hungary, the Russian Federation, Poland, the Czech Republic, India and Brazil, scientific research in the computer science field has a much higher citation rate than overall scientific production. In 2018, over 16% of computer science publications by Australia-based authors featured among the world’s top 10% cited scientific documents. This figure reaches 19% for Luxembourg, although based upon a much smaller number of scientific outputs.

Scientific research is advancing digital technologies and techniques in other ways. The related field of AI is experiencing especially intensive activity, along with blockchain technology and work to advance quantum computing (Chapter 11).

Figure 9.5. Top 10% most-cited documents in computer science, 2018

Percentage of domestic documents (fractional counts) in the top 10% citation-ranked documents



Notes: Computer science publications consist of citeable documents (articles, conference proceedings and reviews) featured in journals specialising in this field. “Top-cited publications” are the 10% most-cited papers normalised by scientific field and type of document. Instead of counting a publication repeatedly if two or more countries contribute to it, fractional counting distributes such publication across contributors so that all publications have the same equal weight.

Sources: OECD calculations based on Elsevier, *Scopus Custom Data*, Version 1.2018 and 2018 Scimago Journal Rank from the Scopus journal title list (accessed in March 2020).

StatLink <https://doi.org/10.1787/888934192509>

The digitalisation of science and innovation

Building on the discussion of how science and innovation are driving digitalisation, this section looks at how digitalisation is impacting the processes and practices of science and innovation.³

Like almost all other activities, science, technology and innovation are going digital. Digital transformation is a multifaceted phenomenon that is impacting innovation in all sectors of the economy. Digital technologies have enabled the creation of completely new digital products and services and the enhancement of others with digital features. Production processes are also subject to substantial change, with new modes of human-to-machine interaction (OECD, 2020_[8]).

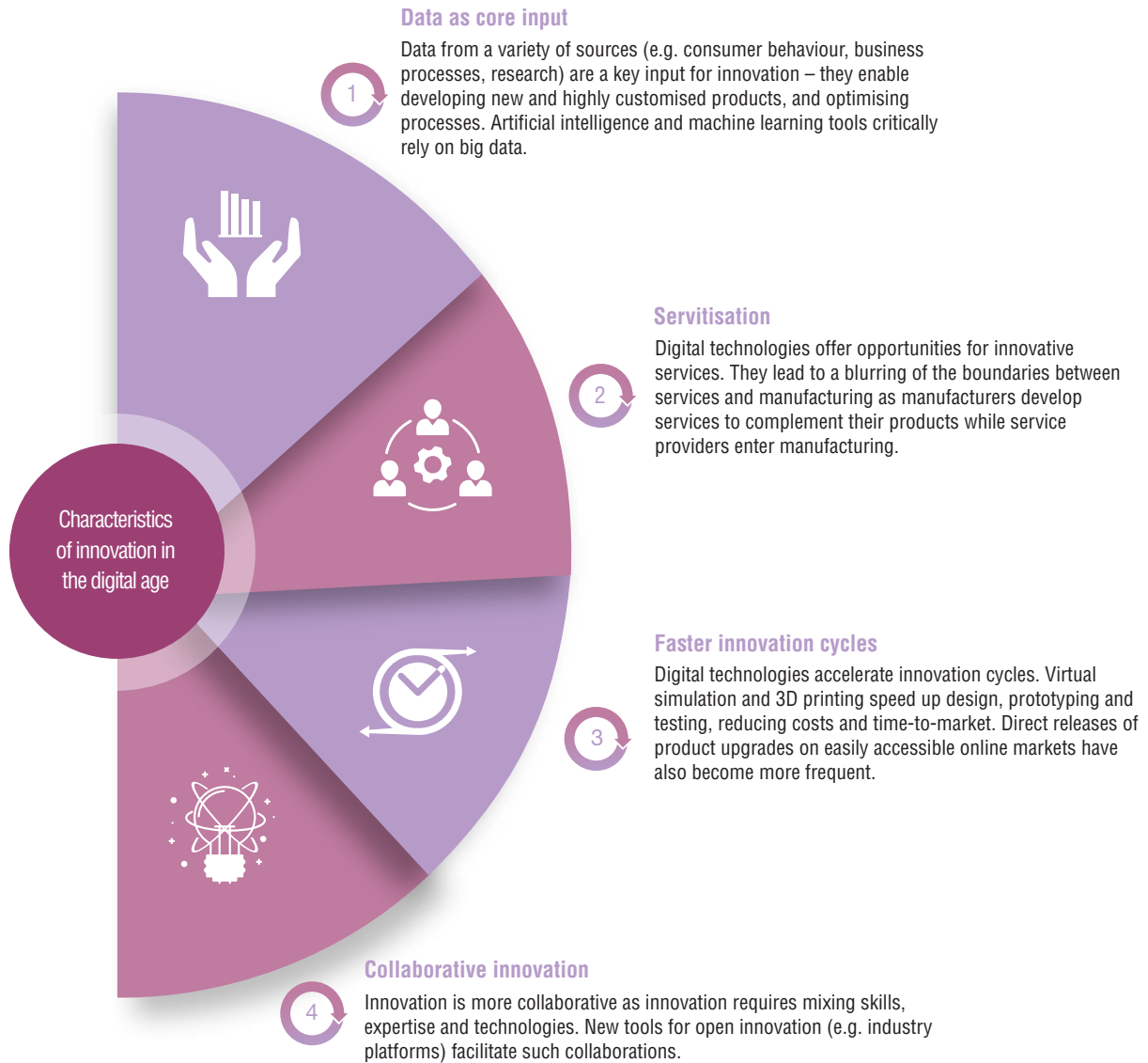
New opportunities are emerging across innovation processes from research to development to commercialisation. Researchers are using big data analytics and large-scale computerised experiments. Developers are exploiting new techniques of simulation and prototyping. Meanwhile, the use of market platforms is aiding commercialisation.

Since industries differ in their products and processes, their structures and in how they innovate, the impacts of digitalisation⁴ on innovation are also likely to differ. For instance, products produced by primary sectors such as food or mining remain largely unchanged. Conversely, the media, music and gaming industries have almost completely digitised their product and service offering. Nevertheless, the production processes for products such as food and minerals can now be increasingly digitalised. Another example is the wide deployment of robots in the automotive industry, while automation remains at early stages in sectors such as agriculture and retail. In many industries, online platforms – enabled by data and digital systems – are changing the way economic agents interact and how markets work.

Digital technologies have lowered information-related production costs and increased the “fluidity” of innovative products. Digitised knowledge (i.e. knowledge that takes the form of data) and information

can circulate and be reproduced, shared or manipulated instantaneously anywhere by any number of actors. As a consequence of changes in costs and fluidity, four trends affect innovation practices across all sectors of the economy in the digital age, as summarised in Figure 9.6.

Figure 9.6. Characteristics of innovation in the digital age



Source: OECD (2019^[9]), *Digital Innovation: Seizing Policy Opportunities*, <https://dx.doi.org/10.1787/a298dc87-en>.

Given the considerable variety in sectors' products and processes, digital technologies (e.g. AI, Internet of Things, drones, virtual reality, 3D printing) will create varying opportunities for innovation, including the following (OECD, 2020^[8]):

- Digitalisation of final products and services.** Some industries have almost completely digitised their products over past decades (e.g. the media, music and gaming industries). Others by their nature remain mainly physical, such as food and consumer products. Many industries present a mix of digital and physical components in their final products, with the digital elements often becoming progressively more important. In the automotive industry, vehicles increasingly integrate digital features.
- Digitalisation of business processes.** Digitalisation may affect sectors' business processes differently. It depends on the nature of the activities and the characteristics of production (e.g. whether it involves the assembly of physical products, if the sector has long supply chains, etc.). In particular, digital technologies offer opportunities for digitalisation (and automation) of production processes; for interconnecting supply chains; and, for improving interactions with the final consumer.

- **Creating new digitally enabled markets and business models.** New markets or market segments enabled by digital technologies, often adjacent to traditional sectors, have been created over recent years. E-commerce, car-sharing services and FinTech services are well-known examples. While new business models are emerging across the economy, the scale and disruption potential of these trends vary across sectors. In some cases, those business models may displace traditional ones (e.g. travel agencies). In other cases, the two models may co-exist and expand the product or service offering (e.g. brick-and-mortar existing simultaneously with online retail stores).

Links between digitalisation and innovation in business

Although the ways in which innovation responds to and influences digitalisation can be mediated by R&D and invention, they are different concepts. The *Oslo Manual* definition of an innovation (OECD, 2018^[10]) refers to a new or improved product or process (or combination of both). It must differ significantly from a unit's previous products and processes and be available to potential users or brought into use by the unit. Importantly, innovation requires that implementation takes place – moving beyond the realm of ideas and inventions. At a minimum, the innovation has to be new to the organisation in question. Thus, this is a broad concept that also encompasses the diffusion of digital technologies where this involves a significant change from the viewpoint of the business adopting them.

Data from business innovation surveys show the information services industry⁵ generally exhibits the greatest rates of reported innovation (e.g. 75% in the case of France). This may partly reflect relatively higher rates of obsolescence for certain types of digital technologies (e.g. hand-held devices). Such rates of obsolescence drive more rapid innovation cycles (as highlighted in Figure 9.6).

Digital innovations can be found in any sector. They comprise product or process innovations that incorporate ICTs and also innovations that rely significantly on ICTs for their development or implementation. A wide range of business process innovations can entail fundamental changes in the organisation's ICT functions and their interaction with other business functions and the products delivered.

Figure 9.6 also notes that data are now a core element of the innovative process. The *Oslo Manual* recognises developing data and software as a potential innovation activity. Data accumulation by companies can entail significant direct or indirect costs.

One recent OECD-Statistics Canada study examined patterns of advanced technology use and business practices (ATBPs) among Canadian firms using the Statistics Canada 2014 Survey of Advanced Technology. Mapping ATBP portfolios through factor analysis has helped reveal seven main categories of ATBP specialisation (Galindo-Rueda, Verger and Ouellet, 2020^[11]). These categories are logistics software technologies; management practices and tools; automated production process technologies; geomatics and geospatial technologies; bio-and-environmental technologies; software and infrastructure-as-a-service; and additive and micro manufacturing technologies. The data indicate a strong complementarity between management practices and production, and adoption of logistics technologies.

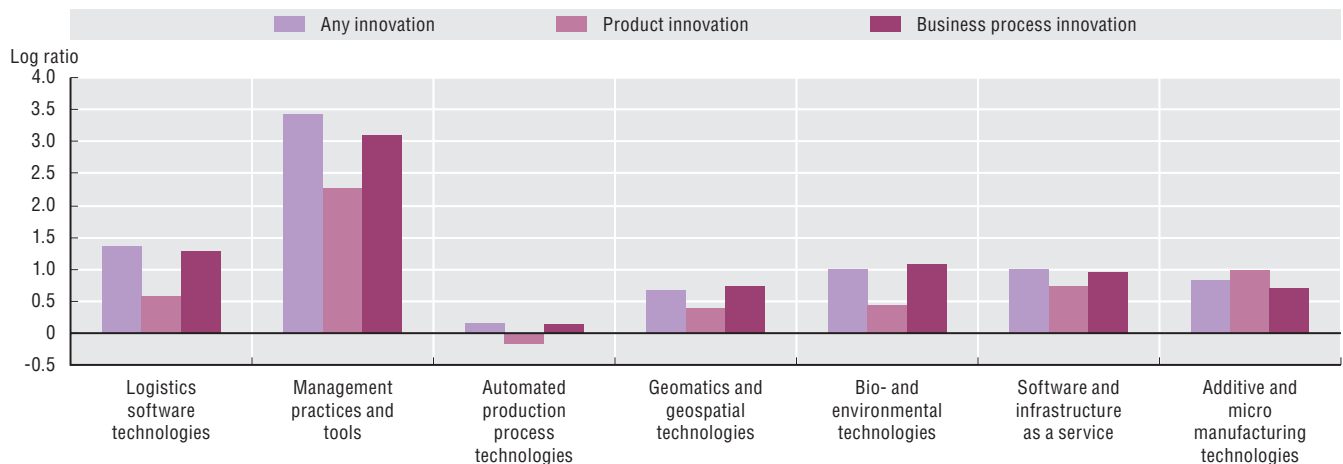
The study also found that innovation is highly correlated with the use of certain business practices and advanced technologies (Figure 9.7). Regression results suggest that using advanced technologies doubles the odds of reporting innovations. The results also indicate complementarity between technology and management in explaining innovation. A positive relationship is also found between the development of technologies and innovation, especially for products, pointing at the advantages of being lead adopters.

Digital technology as an enabler of scientific research

Research is a key driver of technological developments and a foundation for product and process innovations. Digitalisation is changing the ways research is conducted and disseminated – both in businesses and in other organisations such as universities. The International Survey of Scientific Authors (ISSA) asked a global sample of scientists how digitalisation is impacting their work (Bello and Galindo-Rueda, 2020^[12]). It assessed whether digital tools make scientists more productive; to what extent they rely on big data analytics, or share data and source codes developed through their research; and to what degree they rely on a digital identity and presence to communicate their research. Preliminary results reveal contrasting patterns of digitalisation across fields.

Figure 9.7. The link between innovation and the adoption of technology and business practices, Canada, 2014

Estimated log odds ratios of reporting an innovation between technology and/or practice users and non-users



Note: Estimates control for technology development activity, country of ultimate ownership control, and business size and industry.

Source: Galindo-Rueda, Verger and Ouellet (2020_[11]), "Patterns of innovation, advanced technology use and business practices in Canadian firms", <https://doi.org/10.1787/6856ab8c-en>.

StatLink  <https://doi.org/10.1787/888934192528>

Use of advanced digital tools, including big data analytics, is a defining feature of the computer sciences, followed by multidisciplinary research, mathematics, earth and materials sciences, and engineering (Figure 9.8). The life sciences (with the exception of pharmaceuticals) and the physical sciences (other than engineering) report the largest relative efforts to make data and/or code usable by others. Digital productivity tools have much broader adoption. Interestingly, the fields making less use of advanced digital and data/code dissemination tools – namely those in the social sciences, arts and humanities – are more likely to engage in activities that enhance their digital presence and external communication (e.g. use of social media).

Younger scientists are more likely to engage in all dimensions of digital behaviour, just as ICT-use surveys find younger individuals generally make greater use of digital technologies. Women scientists are less likely than their male counterparts to use and develop advanced digital tools. However, they are more likely to engage in enhancing their digital presence, identity and communication. Scientific authors working in the business sector are also more likely than those in other sectors to use advanced digital tools linked to big data. At the same time, they are less likely to engage in data/code dissemination activities and online presence and communication. By contrast, authors in the higher education sector use digital productivity tools more (with most tools asked about in the survey relating to academic tasks), and also had a greater online presence and use of digital communication tools (Bello and Galindo-Rueda, 2020_[12]).

Research paradigms and digitalisation

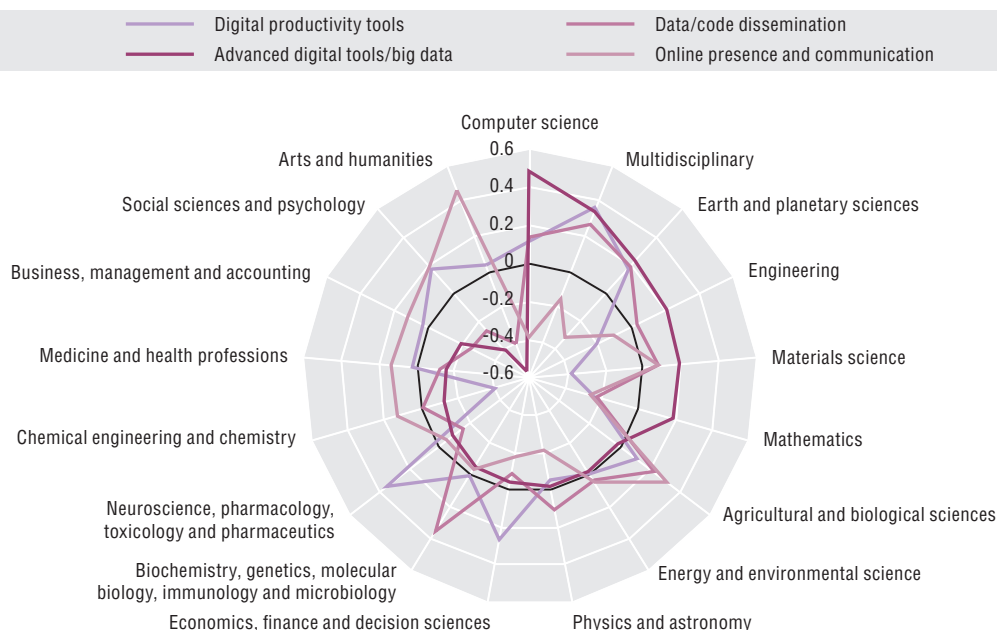
Since digital tools can transform how scientific research is conducted, ISSA survey respondents were asked to describe their scientific research work with respect to the use of theory, simulations, empirical non-experimental and experimental activity. Scientific research practices correlate with digital practices in complex ways. Researchers engaged in computational and modelling work (37% of the sample) are most likely to use advanced digital tools. However, they are also less likely to engage in online presence and communication activities. Together with researchers involved in experimental work (49%), they are also the most likely to engage in data and code dissemination practices, for example through online platforms such as GitHub.

Those reporting work on gathering information (37%) are surprisingly not among those most likely to disseminate data and code. This suggests considerable scope for digitalisation of their data diffusion activities. Among this group, the use of digital productivity tools is nonetheless high. Those involved

in theoretical work (46%) tend to make limited use of most digital practices. The incidence of digital practices among those undertaking empirical, non-experimental work (45%) is most common in the social sciences. It is relatively constrained in terms of data/code dissemination (creating a challenge for replicability) and advanced digital tools.

Figure 9.8. Patterns of digitalisation in science across fields, 2018

Average standardised factor scores for four different facets of digitalisation



Notes: This is an experimental indicator. Figures refer to the weighted average of four standardised factor scores representing latent digitalisation indicators within each science field. Sampling weights adjusted by non-responses are used in the weighting procedure. The factor analysis is based on the responses by scientists to questions relating to the use of digital tools or adoption of digitally enabled practices. The resulting four factors have been interpreted and labelled based on how strongly they correlate with the survey-based underlying variables. Factor scores are estimated in units of standard deviations from their means and represent a person's relative position on a latent factor compared to the rest of the individuals. How to read: Computer science's highest score for the factor representing use of advanced digital tools (grey line) represents high relative intensity on this facet. Conversely, a low relative intensity is seen on the digital facet representing online presence and communication (dotted line) for scientists in this area.

Source: Bello and Galindo-Rueda (2020_[12]), "Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)", <https://doi.org/10.1787/1b06c47c-en>.

StatLink <https://doi.org/10.1787/888934192547>

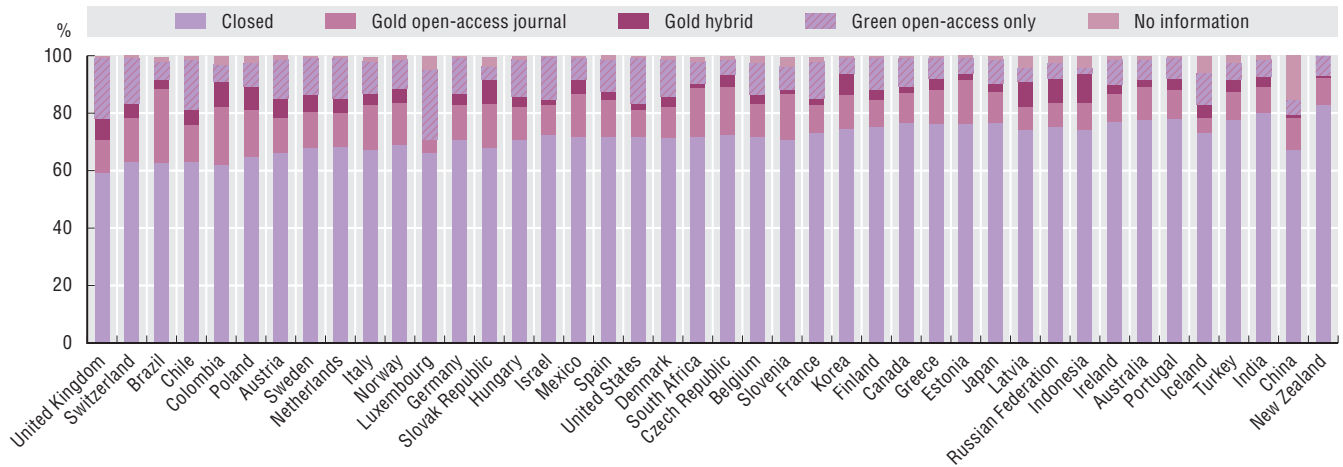
Digitalisation can be a key enabler of "open science" practices. For example, digitalisation can help reduce transaction costs; promote data re-use; increase rigour and reproducibility; and decrease redundant research. Broadening access to scientific publications, data and code is at the heart of open science so that potential benefits are spread as widely as possible (OECD, 2015_[13]). Interest is growing in monitoring the use of such practices (Gold et al., 2018_[14]).

Access to scientific research articles plays an important role in the diffusion of scientific knowledge. Digital technology facilitates the sharing of scientific knowledge to promote its use for further research and innovation. Digital technology can also be used to rapidly query whether a large number of research outputs published on line are available openly. This approach reveals that 60% to 80% of content published in 2016 was, one year later, only available to readers via subscription or payment of a fee (Figure 9.9).

Journal-based open access (usually termed "gold" OA) is particularly noticeable in Brazil, as well as in many other Latin American economies. Repository-based OA (also known as "green" OA) is especially important for authors based in the United Kingdom. About 5% of authors appear to be paying a fee to make their papers publicly available in traditional subscription journals (also known as "gold hybrid" OA).

Figure 9.9. Open access of scientific documents, 2017

As a percentage of a random sample of 100 000 documents published in 2016, by country of affiliation



Notes: This is an experimental indicator based on an automated query of a random (non-stratified) sample of 100 000 citable documents (articles, reviews and conference proceedings) published in 2016 and indexed in the Scopus database, with valid DOIs associated to them (more than 90% of cases). The open-access status of the documents has been assessed using the R wrapper for the oaDOI API produced by ImpactStory, an open source website that aims to help researchers explore and share the online impact of their research. The API returns information on the different mechanisms by which to access legal copies of each document. StatLink contains more data.

Sources: OECD (2017^[2]), OECD Science, Technology and Industry Scoreboard 2017: The Digital Transformation, <https://doi.org/10.1787/9789264268821-en>.

StatLink  <https://doi.org/10.1787/888934192566>

Citations provide one indicator of the “impact” of scientific work. It might be expected that open-access publications would be more likely to be cited due to ease of access. However, bibliometric analysis confirms previous findings of a mixed picture (OECD, 2015^[15]; Boselli and Galindo-Rueda, 2016^[16]), as not all forms of OA appear to confer a citation advantage. Results from the ISSA confirm that authors of documents in gold OA journals tend to report significantly lower earnings. This points to strong and self-reinforcing prestige effects that are dissociated from dissemination objectives in the digital era (Fyfe et al., 2017^[17]). Nevertheless, evidence points to OA increasingly becoming the norm. This provides just one illustration of how the “promise” of digital advances can come up against challenges arising from entrenched behaviours and ways of working.

Open access to data and code

With data as a core element of innovative activities (Figure 9.6), measuring and understanding access to data and code are also important for mapping open science practices. The latest ISSA study goes beyond considering only the access status of publications to examine the accessibility of the code and data developed as part of the published research. The study shows that, on average, two-thirds of respondents create new data, code, or both as part of their published scientific work (Figure 9.10).

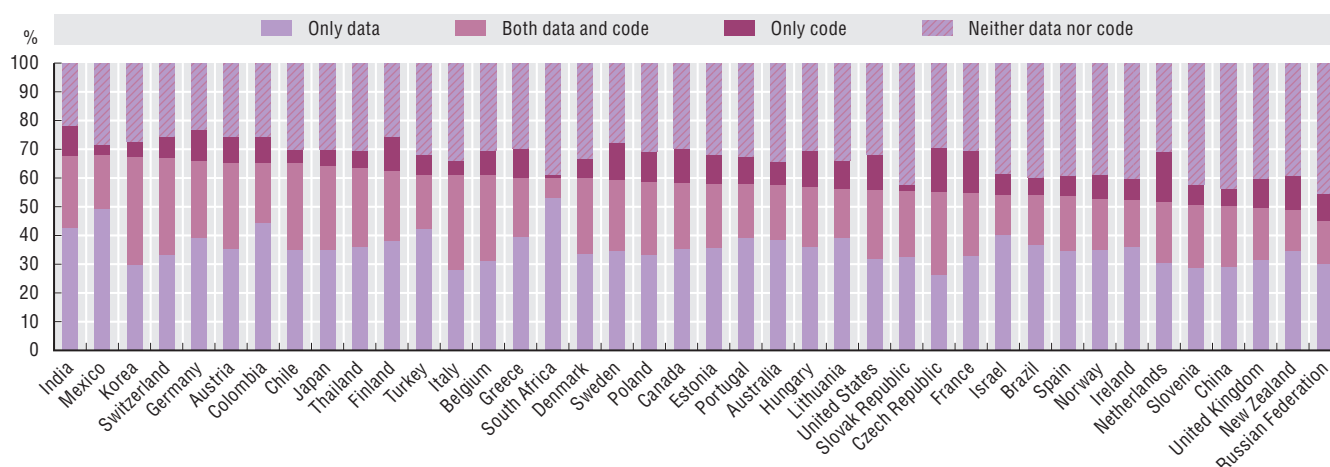
The use of repositories for data archiving and dissemination seems to be most common among respondents in the life sciences. Informal data or code sharing among peers seems to be the main way researchers in all fields make data available to others. Nevertheless, the publication of research data or code does not automatically imply that other researchers can easily use and re-use them. Barriers include access costs or challenges such as not knowing the coding language or software used. Standard mechanisms for requesting and securing data access appear uncommon across all disciplines, being used by fewer than 30% of respondents when sharing data or code. Likewise, only about 10% of respondents applied a data usage licence to their data. Re-usability of data seems to be supported mainly through the provision of detailed metadata, especially in the physical sciences and engineering. Compliance with standards that facilitate data combination across sources is more common in health and life science but less so in physical sciences and engineering.

In all fields, authors tend to report several barriers to access of scientific outputs. These include formal sharing requirements set by publishers, funders or the respondent’s organisation; or intellectual property protection (Bello and Galindo-Rueda, 2020^[12]). Career objects and peer expectations were

reported as driving enhanced access. While capabilities for managing disclosure and sharing do not seem to be limiting factors, dissemination costs in terms of time and money are deemed strong barriers. Privacy and ethical considerations also tend to limit access to scientific outputs in health sciences.

Figure 9.10. Scientific production resulting in new data or code, by country of residence, 2018

Weighted percentage of respondents with outputs in each category



Notes: Weighted estimates adjusted for sample design and nonresponse. Only countries with more than 70 observations are included. StatLink contains more data.

Source: OECD calculations based on OECD International Survey of Scientific Authors (ISSA), <http://oe.cd/issa>.

StatLink  <https://doi.org/10.1787/888934192585>

Open access to scientific documents, data and code are increasingly important components of a wider shift towards “open innovation” based on knowledge assets both within and outside the organisation. Co-operation is a key way to source this knowledge to generate new ideas and bring them quickly into use. At the same time, organisations exploit their own ideas, as well as innovations of other entities. In this context, academic research occupies a major place (OECD, 2008_[18]). Open innovation involves leveraging the collective and collaborative potential of institutions and individuals with different or unrelated backgrounds. They come together to contribute towards a common goal or project. This can lead to co-creating new products, processes and business models – with digital technology often a key component.

Governments can use digital technologies to support the open innovation ecosystem. For example, the Infocomm Media Development Authority (IMDA) of Singapore maintains a digital Open Innovation Platform. It hosts digital challenges set by enterprises, trade associations and non-profit organisations seeking to solve business challenges and societal problems. The platform facilitates a vibrant community of over 8 000 registered “solvers” across various geographies and skillsets. In this way, it can crowdsource innovative solutions through a highly structured process. This is complemented by a physical PIXEL Innovation Lab that gathers entrepreneurs and innovative enterprises to work on building cutting-edge digital technologies. Additionally, IMDA has a regulatory sandbox for businesses and their data partners to explore and pilot innovative uses of data. On the one hand, the sandbox reduces uncertainty for businesses. On the other, it allows the regulator to learn of new developments in industry and assess the need for policy action to support data innovation.

Scientists’ perspectives on digitalisation and its impacts

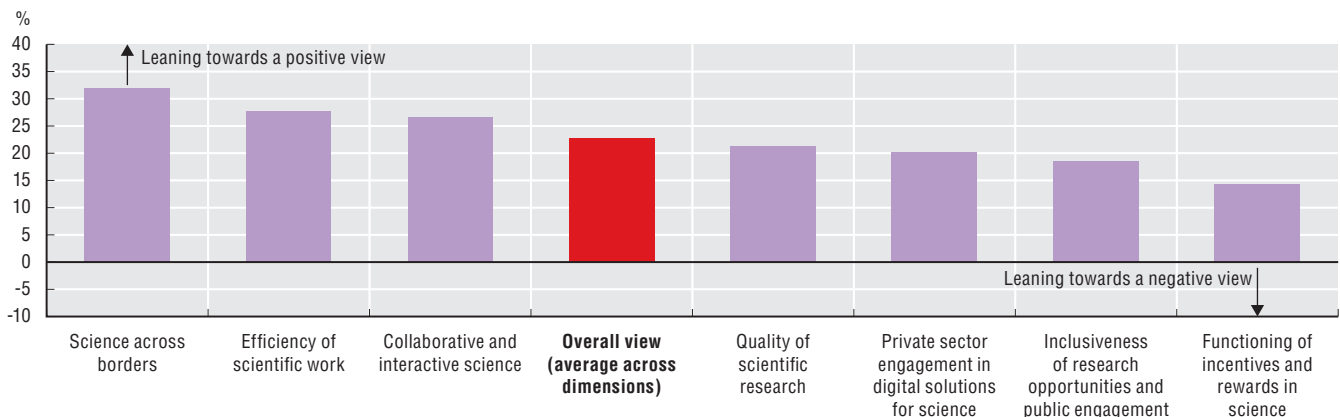
How do scientists themselves view the digital transformation of scientific research and its impacts? Evidence from the 2018 ISSA study suggests that scientists are on average positive across several dimensions (Figure 9.11). Many respondents feel that digitalisation has positive potential to promote collaboration, particularly across borders, and improve the efficiency of science. While remaining positive, scientists appear less optimistic regarding the potential impact of digitalisation on the system of incentives and rewards. Specifically, they are concerned about being “rated” on the basis of their digital “footprint”, such as their publications and citations, as well as downloads of their work. They

also have reservations about whether digitalisation can bring scientific communities and scientists together with the public (inclusiveness). Finally, they sometimes question the role of the private sector in providing digital solutions to assist their work. Younger authors are generally more positive than older peers, except regarding the impacts of digitalisation on the incentive system; this may reflect concerns about their future careers.

Across countries, the average sentiment towards the impacts of digitalisation (Figure 9.12) appears consistent overall with results from broader population surveys on attitudes towards science and technology (OECD, 2015_[19]). Scientists in emerging and transition economies appear to be more positive on average towards the impacts of digitalisation on science. The position of scientists in the most R&D-intensive European economies is more reserved, while still positive in the main. These results do not imply that scientists are by and large dismissive of the potential pitfalls of digitalisation. A significant minority of respondents tended to agree with “negative” statements about the impacts of digitalisation on science. They were concerned, for example, about the promotion of hypothesis-free research in computationally intensive data-driven science. For these respondents, digitalisation could also accentuate divides in research between those with advanced digital competences and those without. It could also encourage a celebrity culture in science, premature diffusion of findings and individual exposure to pressure groups. Digitalisation could also lead to use of readily available but inappropriate indicators for monitoring and incentivising research. Finally, authors agreed with the statement that digitalisation could concentrate workflows and data in the hands of a few companies providing digital tools.

Figure 9.11. Scientific authors’ views on potential impacts of the digitalisation of science, 2018

Average sentiment towards “positive” digitalisation scenarios, as percentage deviation from the mid-range of possible responses



Notes: This is an experimental indicator. Survey respondents were asked to rate opposing scenarios on different dimensions from (1 = fully agree with a negative view) to (10 = fully agree with a positive view). For interpretability, weighted average scores on each dimension and the general summary view (weighted average across dimensions) are presented as percentage deviations from the midpoint. This means, for example, that with respect to the subject of “Science across borders”, respondents are on average 32% oriented towards the positive outcome, relative to the neutral perspective. Weighted average scores consider the sample design and non-response.

Source: OECD (2019_[4]), *Measuring the Digital Transformation: A Roadmap for the Future*, <https://dx.doi.org/10.1787/9789264311992-en>.

StatLink <https://doi.org/10.1787/888934192604>

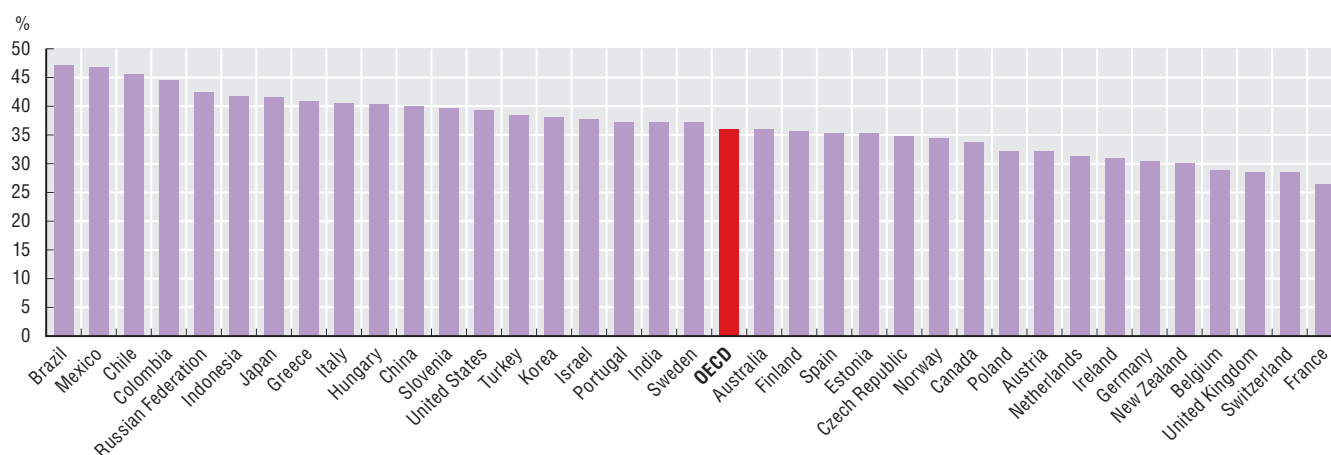
Digital technology in research against COVID-19

Digital technologies are playing a direct role in efforts to manage the COVID-19 pandemic and find a vaccine. In particular, AI and associated technologies such as machine learning are finding innovative applications to a wide array of COVID-19 driven challenges.⁶

Before the world was even aware of the threat posed by COVID-19, AI systems had detected the outbreak of an unknown type of pneumonia in China. As the outbreak turned into a global pandemic, AI tools and technologies can support policy makers, the medical community and society at large to manage every stage of the crisis and its aftermath (detection, prevention, response, recovery) and to accelerate research (Chapter 11).

Figure 9.12. Scientific authors' views on the digitalisation of science, by country of residence, 2018

Average sentiment towards a "positive" digitalisation scenarios, as percentage deviation from the mid-range of possible responses



Notes: This is an experimental indicator. Cross-country comparisons should be interpreted with caution as the population of corresponding scientific authors is not uniformly representative of their scientific community. Economies with fewer than 75 survey responses not shown. Average scores are weighted and consider the sample design and non-response.

Source: Bello and Galindo-Rueda (2020^[12]), "Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)", <https://doi.org/10.1787/1b06c47c-en>.

StatLink  <https://doi.org/10.1787/888934192623>

AI powering research to understand and treat COVID-19

AI tools and techniques can rapidly analyse large volumes of research data. In this way, they can help the medical community and policy makers understand the COVID-19 virus and accelerate research on treatments. AI text and data mining tools can uncover the history of the virus along with transmission, diagnostics and management measures, as well as lessons from previous epidemics. For instance, several institutions are using AI techniques such as deep learning models to help identify candidate drugs or treatments that might treat COVID-19. This helps narrow the list of potential candidates for further investigation by scientists, making the research process more efficient and effective. Meanwhile, DeepMind and several other organisations have used deep learning to predict the structure of proteins associated with SARS-CoV-2, the virus that causes COVID-19.

Access to data and computing power are key inputs for this process. Collaborative initiatives are helping make relevant datasets on epidemiology, bioinformatics and molecular modelling accessible to researchers. For example, the COVID-19 Open Research Dataset Challenge by the US government and partner organisations has made available more than 29 000 academic research articles for coronavirus and COVID-19. Technology companies such as IBM, Amazon, Google and Microsoft are making computing power available; individuals are donating computer processing power (e.g. Folding@home); and public-private efforts have emerged like the COVID-19 High Performance Computing Consortium and AI for Health.

Facebook has used its massive AI-powered computational infrastructure to generate mobility data sets that inform researchers and public health experts about how populations are responding to physical distancing measures. This complements previous efforts including a partnership with the Center for International Earth Science Information Network at Columbia University. This collaboration used state-of-the-art computer vision techniques to identify buildings from publicly accessible mapping services to create highly accurate population datasets (Herdağdelen et al., 3 June 2020^[20]; Bonafilia et al., 2 April 2019^[21]).

Innovative incentives, including prizes, open source collaborations and hackathons, are also helping accelerate research on AI-driven solutions to the pandemic. For example, the United Kingdom's CoronaHack – AI vs. COVID-19 seeks ideas from businesses, data scientists and biomedical researchers on using AI to control and manage the pandemic.

AI helping to detect, diagnose and prevent the spread of the coronavirus

AI can also help detect, diagnose and prevent the spread of the virus. Algorithms that identify patterns and anomalies are already working to detect and predict the spread of COVID-19. Meanwhile, image recognition systems are speeding up medical diagnosis:

- AI-powered early warning systems can help detect epidemiological patterns by mining mainstream news, online content and other information channels in multiple languages to provide early warnings. This can complement syndromic surveillance and other health care networks and data flows (e.g. World Health Organization Early Warning System, Bluedot).
- AI tools can help identify virus transmission chains and monitor broader economic impacts. In several cases, AI technologies have demonstrated their potential to infer epidemiological data more rapidly than traditional reporting of health data. Institutions such as Johns Hopkins University and the OECD (OECD.AI)⁷ have also made available interactive dashboards that track spread of the virus through live news and real-time data on confirmed coronavirus cases, recoveries and deaths.
- Rapid diagnosis is key to limiting contagion and understanding the way COVID spreads. Applied to images and symptom data, AI could help rapidly diagnose COVID-19 cases. Attention must be given to collecting data representative of the whole population to ensure scalability and accuracy.

Limiting contagion is a priority in all countries and AI applications are also helping to slow spread of the virus:

- A number of countries are using AI technology in population surveillance to monitor COVID-19 cases. In Korea, for example, algorithms use geolocation data, surveillance-camera footage and credit card records to trace coronavirus patients. China assigns a colour code (red, yellow or green) to each person indicating contagion risk using cell phone software. Machine-learning models use travel, payment and communications data to predict the location of the next outbreak and inform border checks. Meanwhile, search engines and social media are helping track the disease in real time.
- Many countries, including Austria, China, Israel, Poland, Singapore and Korea have set up contact tracing systems to identify possible infection routes. Israel, for example, used geolocation data to identify people coming into close contact with known virus carriers. It then sent text messages directing them to isolate themselves immediately.
- AI is identifying, finding and contacting vulnerable, high-risk individuals. For example, Medical Home Network, a Chicago-based non-profit, has an AI platform to identify Medicaid patients most at risk from COVID-19 based on risk of respiratory complications and social isolation.
- Semi-autonomous robots and drones are responding to immediate needs in hospitals. They deliver food, medications and equipment; clean and sterilise; and aid doctors and nurses.

AI technologies have great potential to help policy makers and the health community develop ways to slow the spread of COVID-19 and to aid the search for treatments, including for vaccines. Multidisciplinary and multi-stakeholder co-operation and data exchange both nationally and internationally can boost this contribution. The AI community, medical community, developers and policy makers, for example, can formulate the problem, identify relevant data and open datasets, share tools and train models.

However, AI is not a silver bullet. AI systems based on machine-learning work by identifying patterns in data, and require large amounts of data to find these patterns. The outputs are only as good as the training data. In some cases, diagnostic claims have been called into question and some chatbots have given different responses to questions on symptoms (OECD, 2020_[22]). This further emphasises the general point that the data used to train AI need to have the appropriate qualities to draw robust and generalised conclusions – including being designed to avoid biases related to race and gender.

Nevertheless open and collaborative approaches will allow the widest pool of researchers possible to access the tools and data needed to devise innovative uses of AI and maximise the chances of finding effective containment measures and treatments.

The digitalisation of science and innovation policy

As well as profoundly affecting science, research, and innovation processes, digitalisation is also beginning to impact the way in which policy is made in these areas.⁸ Scientific research and innovation increasingly leave digital “footprints”, in datasets that are becoming ever larger, more complex and

available at higher speed. At the same time, technological advances – in machine learning and natural language processing, for example – are opening new analytical possibilities.

Science, technology and innovation (STI) can harness the power of digitalisation to link and analyse datasets covering diverse areas of policy activity and impact. For example, Digital Science and Innovation Policy (DSIP) initiatives already experiment with semantic technologies. On the one hand, they link datasets with AI to support big data analytics. On the other, they link datasets with interactive visualisation and dashboards to promote data use in the policy process. Other policy areas and government services can similarly benefit from data and digitalisation (Chapter 4).

An overarching aim is to increase the effectiveness of national research and innovation ecosystems. In particular, data linking and synchronisation across digital systems can help optimise administrative workflows to reduce reporting burdens. It can support performance monitoring and management. Finally, it can provide anticipatory intelligence to identify needs for support or policy interventions. The insights gained support improved policy formulation and design in various ways.

Figure 9.14 provides a stylised conceptual view of a DSIP initiative and its main components. All of these elements interact in ways reflecting each country's institutional set-up. The main elements consist of various input data sources that feed into a “data cycle” enabled by interoperability standards. These standards include unique, persistent and pervasive identifiers (UPPIs). DSIP systems can perform various functions catering to various users' needs.

Data are predominantly sourced from administrative databases held by funding agencies (e.g. databases of grant awards) and organisations that perform research, development and innovation (RD&I). These include current research information systems in universities, and proprietary bibliometric and patent databases. Some DSIP systems have grown out of these databases. Through integration with external platforms or development of add-on services, they have evolved into infrastructures that can deliver comprehensive data analysis on RD&I activities. Other systems have been established from the ground up. Several DSIP systems harvest data from the web to build a picture of the incidence and impacts of science and innovation activities. Web sources include, but are not limited to, company websites and social media.

DSIP infrastructures can increase the scope, granularity, verifiability, communicability, flexibility and timeliness of policy analyses. They can lead to the development of new STI indicators (Bauer and Suerdem, 2016^[23]), the assessment of innovation gaps (Kong et al., 2017^[24]), strengthened technology foresight (Kayser and Blind, 2017^[25]) and the identification of leading experts and organisations (Shapira and Youtie, 2006^[26]; Johnson, Fernholz and Fosci, 2016^[27]; Gibson et al., 2018^[28]). Furthermore, in some countries, researchers and policy makers have started to experiment with natural language processing and machine learning. They are using it to track emerging research topics and technologies (Wolfram, 2016^[29]; Mateos-Garcia, 6 April 2017^[30]) and to support RD&I decisions and investments (Yoon and Kim, 2012^[31]; Yoon, Park and Kim, 2013^[32]).

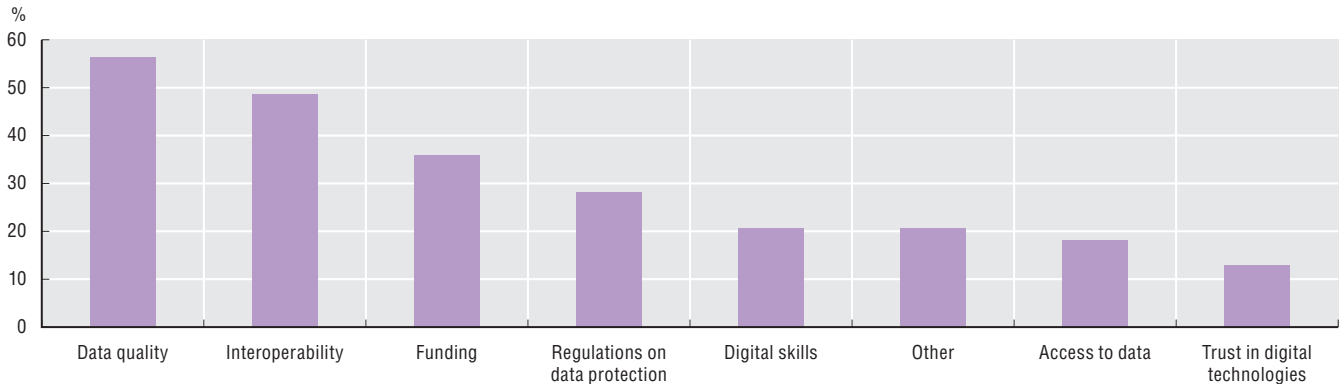
Realising the potential of DSIP involves overcoming several possible barriers. The OECD DSIP survey received responses from 39 initiatives in the OECD countries and partner economies. Drawing on these responses, DSIP administrators identified data quality, interoperability, sustainable funding and data protection regulations as the biggest challenges facing their initiatives.

Other challenges cited less often were access to data, the availability of digital skills and trust in digital technologies. Policy makers wishing to promote DSIP face further systemic challenges. These include overseeing fragmented DSIP efforts and multiple (often weakly co-ordinated) initiatives; ensuring responsible use of data generated for other purposes; and balancing the benefits and risks of private-sector involvement in providing DSIP data, components and services.

Data interoperability, in particular, is a challenge to which digital tools may help to provide a solution. Research and innovation activities, by their nature, are shaped by a large number of actors. As a result, data on the incidence and impacts of research and innovation are dispersed across a variety of public and private databases and the web. Harvesting these datasets from external sources requires the development of common data formats and other interoperability enablers including, but not limited to, application programming interfaces (APIs), ontologies, protocols and UPPIs for RD&I actors.

Figure 9.13. Main challenges facing DSIP initiatives, 2018

Percentage of surveyed DSIP systems



Note: Questionnaire respondents could select more than one challenge facing their DSIP initiatives.

Source: OECD (2018^[33]), *OECD Science, Technology and Innovation Outlook 2018: Adapting to Technological and Societal Disruption*, https://doi.org/10.1787/sti_in_outlook-2018-en.

StatLink  <https://doi.org/10.1787/888934192642>

An integrated and interoperable system can considerably reduce the reporting and compliance burden on RD&I actors, freeing up time and money for research and innovation. In addition, it allows quicker, cheaper and more accurate data matching that can, in turn, enable cheaper, more timely and more detailed insights. This can allow for more responsive and tailored policy design. Furthermore, the gradual emergence of internationally recognised identifiers makes it easier to track the impacts of research and innovation activities across borders and map international partnerships.

However, interoperability issues raise important questions. On a technical level, policy makers must ask what kind of digital system can make existing and new data interoperable. On a semantic level, they must grapple with metadata and language issues. With respect to governance, they must reflect on how all stakeholders can be aligned to agree upon an interoperability system. A specific issue concerns the role and effectiveness of data standards, particularly in a mixed ecosystem containing both legacy and new systems.

Many DSIP systems use national identifications (IDs) – e.g. business registration and social security numbers – as well as country-specific IDs for researchers. Nevertheless, attempts are being made to establish international standards and vocabularies to improve the international interoperability of DSIP infrastructures. These include UPPIs, which assign a standardised code unique to each RD&I entity (e.g. researcher, research organisation, funder, project or outputs such as publications). These are designed to be persistent over time and pervasive across various datasets.

Some UPPIs exist as an integral part of, or support for, commercial products. These include publication/citation databases, research information systems and supply-chain-management services. Others exist solely to provide a system of identifiers for wide adoption and use. Open Researcher and Contributor ID (ORCID), for example, aims to resolve name ambiguity in scientific research through a digital register of unique identifiers and basic associated identity information for individual researchers. Registers often incorporate links to a wide range of further information. For example, ORCID records allow details of education, employment, funding and research works to be added manually or brought in by linking to other systems, including Scopus and ResearcherID.

As a UPPI system gains traction there may be a “network effect”, whereby each additional registrant increases the value of the system to all users. Eventually, the UPPI system may become an expected way for entities to unambiguously identify each other. This results in strong incentives for those not yet registered to join.

Besides UPPIs, APIs have become a standard for enabling machine-to-machine interactions and data exchanges. Several countries have started to proliferate APIs across the whole landscape of government websites and databases, improving data re-use. Improvements in access to administrative datasets have positive impacts on the functionality and reliability of the results of analyses delivered by DSIP systems.

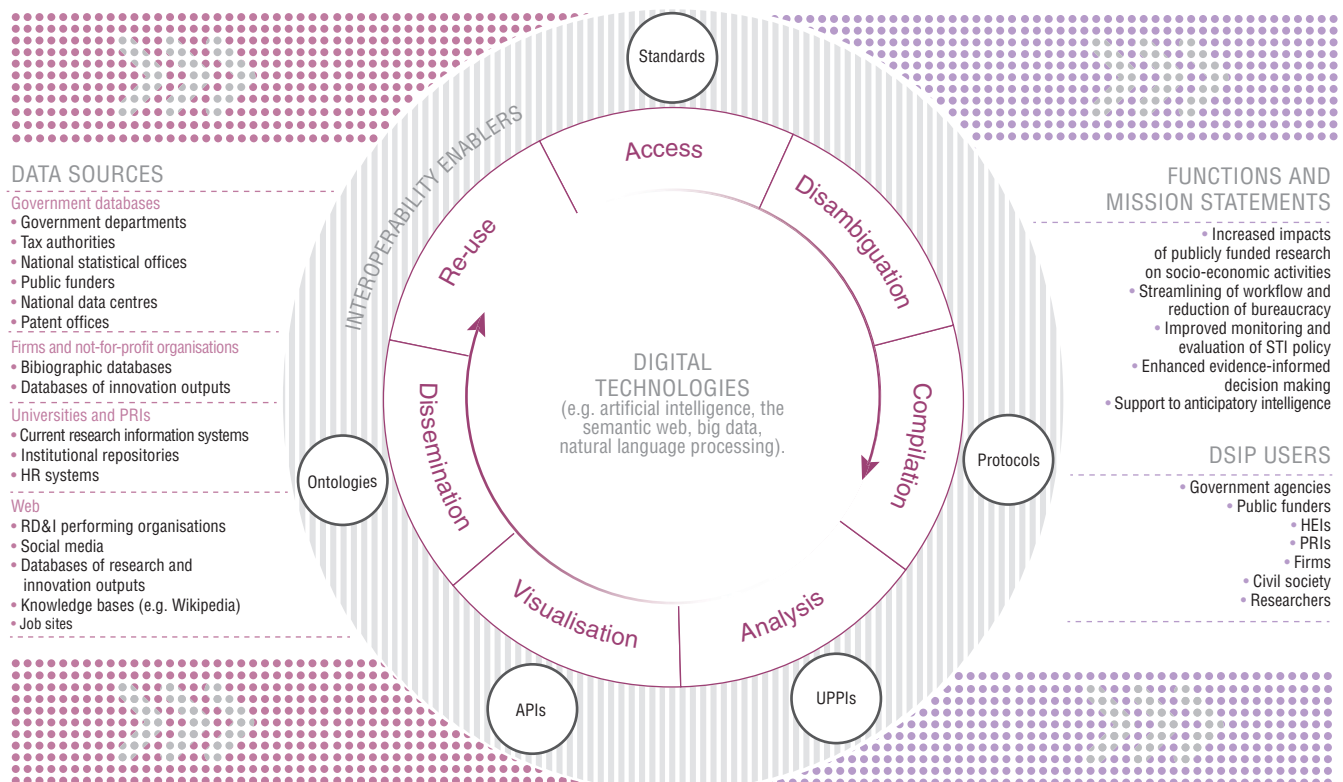
Aside from government agencies and other public funders, R&DI-performing organisations store a significant share of research and innovation data. However, these often have different formats and structures – even for the same type of information. The Common European Research Information Format and metadata formats by Consortia Advancing Standards in Research Administration Information were originally designed to serve the needs of higher education institutions in data management. Some DSIP systems use them to harvest curated data from research institutes and directly apply them in analysis.

The digital transformation of STI policy and its evidence base is still in its early stages. This means policy makers can actively shape DSIP ecosystems to fit their needs. This will require strategic co-operation through interagency co-ordination and sharing of resources (such as standard digital identifiers), and a coherent policy framework for data sharing and re-use in the public sector. Since several government ministries and agencies formulate science and innovation policy, DSIP ecosystems should be founded on the principles of co-design, co-creation and co-governance (OECD, 2018_[33]).

Interoperability remains a major hurdle, despite the recent proliferation of identifiers, standards and protocols. There is the potential opportunity for policy makers to influence the development of international UPPI systems. Key issues are target populations, information captured, compatibility with statistical systems, governance systems and especially adoption both by entities and by potential users. International efforts related to data documentation and the development of standards for metadata could be consolidated to improve data interoperability.

Governments can usefully co-operate with the private and not-for-profit sectors in developing and operating DSIP systems. However, they should ensure public data remain outside of “walled gardens” and open for others to readily access and re-use. They should also avoid vendor lock-ins, deploying systems that are open and agile. In a fast-changing environment, this will provide governments with greater flexibility to adopt new technologies. It will also allow them to incorporate unexploited data sources in their DSIP systems to realise benefits for RD&I actors.

Figure 9.14. A stylised conceptual view of common main components of a DSIP initiative



Note: DSIP = digital science and innovation policy; STI = science, technology and innovation; HEI = higher education institution; PRI = public research institution; API = application programming interface; UPPI = unique, persistent and pervasive identifier; HR = human resources; RD&I = research, development and innovation.

Source: OECD (2018_[33]), OECD Science, Technology and Innovation Outlook 2018: Adapting to Technological and Societal Disruption, https://doi.org/10.1787/sti_in_outlook-2018-en.

References

- Bauer, M. and A. Suerdem (2016), “Relating science culture and innovation”, presentation at the OECD blue sky meeting on science and innovation indicators, Ghent, 19-21 September 2016, https://www.oecd.org/sti/097%20-%20OECD%20Paper%20attitudes%20and%20innovation_v4.0_MB.pdf. [23]
- Bello, M. and F. Galindo-Rueda (2020), “Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2)”, *OECD Science, Technology and Industry Working Papers*, No. 2020/3, OECD Publishing, Paris, <https://doi.org/10.1787/1b06c47c-en>. [12]
- Bonafilia, D. et al. (2 April 2019), “Mapping the world to help aid workers, with weakly, semi-supervised learning”, *Computer Visions - ML Applications* blog, <https://ai.facebook.com/blog/mapping-the-world-to-help-aid-workers-with-weakly-semi-supervised-learning>. [21]
- Boselli, B. and F. Galindo-Rueda (2016), “Drivers and Implications of Scientific Open Access Publishing: Findings from a Pilot OECD International Survey of Scientific Authors”, *OECD Science, Technology and Industry Policy Papers*, No. 33, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5jlz70k0bx-en>. [16]
- Fyfe, A. et al. (2017), “Untangling academic publishing: A history of the relationship between commercial interests, academic prestige and the circulation of research”, *Briefing Paper*, University of St. Andrews, <https://doi.org/10.5281/zenodo.546100>. [17]
- Galindo-Rueda, F. and F. Verger (2016), “OECD Taxonomy of Economic Activities Based on R&D Intensity”, *OECD Science, Technology and Industry Working Papers*, No. 2016/4, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5jlv73sqqp8r-en>. [7]
- Galindo-Rueda, F., F. Verger and S. Ouellet (2020), “Patterns of innovation, advanced technology use and business practices in Canadian firms”, *OECD Science, Technology and Industry Working Papers*, No. 2020/02, OECD Publishing, Paris, <https://doi.org/10.1787/6856ab8c-en>. [11]
- Gibson, E. et al. (2018), “Technology foresight: A bibliometric analysis to identify leading and emerging methods”, *Foresight and STI Governance*, Vol. 12/1, pp. 6-24, <https://foresight-journal.hse.ru/en/2018-12-1/217552882.html>. [28]
- Gold, E. et al. (2018), “An open toolkit for tracking open science partnership implementation and impact [version 1; not peer-reviewed]”, *Gates Open Research* 2/54, <https://doi.org/10.21955/gatesopenres.1114891.1>. [14]
- Herdağdelen, A. et al. (3 June 2020), “Protecting privacy in Facebook mobility data during the COVID-19 response”, *Facebook Research* blog, <https://research.fb.com/blog/2020/06/protecting-privacy-in-facebook-mobility-data-during-the-covid-19-response/>. [20]
- Inaba, T. and M. Squicciarini (2017), “ICT: A new taxonomy based on the international patent classification”, *OECD Science, Technology and Industry Working Papers*, No. 2017/01, OECD Publishing, Paris, <https://doi.org/10.1787/18151965>. [1]
- Johnson, R., O. Fernholz and M. Fosci (2016), “Text and data mining in higher education and public research: An analysis of case studies from the United Kingdom and France”, report commissioned by the Association des directeurs et personnels de direction des bibliothèques universitaires et de la documentation, Paris, <https://adbu.fr/competplug/uploads/2016/12/TDM-in-Public-Research-Final-Report-11-Dec-16.pdf>. [27]
- Kayser, V. and K. Blind (2017), “Extending the knowledge base of foresight: The contribution of text mining”, *Technological Forecasting and Social Change*, Vol. 116, pp. 208-215. [25]
- Kong, D. et al. (2017), “Using the data mining method to assess the innovation gap: A case of industrial robotics in a catching-up country”, *Technological Forecasting and Social Change*, Vol. 119, pp. 80-97. [24]
- Mateos-Garcia, J. (2017), “We are building a formidable system for measuring science – but what about innovation?”, *Nesta* blog, 26 July, <http://www.nesta.org.uk/blog/we-are-building-a-formidable-system-for-measuring-science-but-what-about-innovation/>. [30]
- OECD (2020), “OECD Competition Assessment Toolkit”, webpage, <https://www.oecd.org/competition/assessment-toolkit.htm> (accessed on 21 October 2020). [22]
- OECD (2020), *Protecting Online Consumers During the Covid-19 Crisis*, webpage, <http://www.oecd.org/coronavirus/policy-responses/protecting-online-consumers-during-the-covid-19-crisis-2ce7353c/> (accessed on 21 October 2020). [8]
- OECD (2019), *Digital Innovation: Seizing Policy Opportunities*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/a298dc87-en>. [9]
- OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, <https://doi.org/10.1787/276aaca8-en>. [34]
- OECD (2019), *Measuring the Digital Transformation: A Roadmap for the Future*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264311992-en>. [4]

- OECD (2018), “Enhancing product recall effectiveness: OECD background report”, *OECD Science, Technology and Industry Policy Papers*, No. 58, OECD Publishing, Paris, <https://doi.org/10.1787/ef71935c-en>. [10]
- OECD (2018), *OECD Science, Technology and Innovation Outlook 2018: Adapting to Technological and Societal Disruption*, OECD Publishing, Paris, https://doi.org/10.1787/sti_in_outlook-2018-en. [33]
- OECD (2017), *OECD Science, Technology and Industry Scoreboard 2017: The Digital Transformation*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264268821-en>. [2]
- OECD (2015), “Assessing government initiatives on public sector information: A review of the OECD Council Recommendation”, *OECD Digital Economy Papers*, No. 248, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js04dr9l47j-en>. [13]
- OECD (2015), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264239012-en>. [6]
- OECD (2015), “Making Open Science a Reality”, *OECD Science, Technology and Industry Policy Papers*, No. 25, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5jrs2f963zs1-en>. [15]
- OECD (2015), *OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society*, OECD Publishing, Paris, https://doi.org/10.1787/sti_scoreboard-2015-en. [19]
- OECD (2008), *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD, Paris, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0362>. [18]
- Shapira, P. and J. Youtie (2006), “Measures for knowledge-based economic development: Introducing data mining techniques to economic developers in the state of Georgia and the US South”, *Technological Forecasting and Social Change*, Vol. 73/8, pp. 950-965. [26]
- United Nations (2008), *International Standard Industrial Classification of all Economic Activities (ISIC), Rev. 4*, Statistical Papers, Series M, No. 4, Rev. 4, Department of Statistical and Economic Affairs, United Nations, New York, https://unstats.un.org/unsd/publication/seriesM/seriesm_4rev4e.pdf. [35]
- WIPO (2019), *World Intellectual Property Indicators 2019*, World Intellectual Property Organization, Geneva, https://www.wipo.int/edocs/pubdocs/en/wipo_pub_941_2019.pdf. [5]
- WIPO (n.d.), “Copyright protection of computer software”, webpage, <https://www.wipo.int/copyright/en/activities/software.html> (accessed on 21 October 2020). [3]
- Wolfram, D. (2016), “Natural synergies to support digital library research”, presentation at the joint workshop on bibliometric-enhanced information retrieval and natural language processing for digital libraries, 3 July, <https://www.aclweb.org/anthology/attachments/W16-1501.Presentation.pdf>. [29]
- Yoon, J. and K. Kim (2012), “Detecting signals of new technological opportunities using semantic patent analysis and outlier detection”, *Scientometrics*, Vol. 90/2, pp. 445-61. [31]
- Yoon, J., H. Park and K. Kim (2013), “Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis”, *Scientometrics*, Vol. 94/1, pp. 313-31. [32]

Notes

- Information and communication services comprises the following ISIC Rev.4 industries: Publishing activities; Motion picture, video and television programme production, sound recording and music publishing activities; Programming and broadcasting activities; Telecommunications; Computer programming, consultancy and related activities; Information service activities (United Nations, 2008_[35]). Software publishing is a subset of the first category but cannot be presented separately in the figure since many countries do not break these down.
- Innovation activities include all developmental, financial and commercial activities by a firm that are intended to result in a new or improved product or business process (or combination thereof) that differs significantly from the firm’s previous products or business processes and that has been introduced on the market or brought into use by the firm (OECD, 2018_[10]).
- This section draws upon Chapters 2 and 4 of OECD (2020_[8]) authored respectively by Fernando Galindo-Rueda and Dominique Guellec; and Caroline Paunov and Sandra Planes-Satorra.

4. Digitisation is the conversion of analogue data and processes into a machine-readable format. Digitalisation is the use of digital technologies and data, as well as interconnection, that results in new or changes to existing activities. Digital transformation refers to the economic and societal effects of digitisation and digitalisation (OECD, 2019^[34]).
5. Information service activities (ISIC Rev.4, division 63) comprises the following industries: Data processing, hosting and related activities; Web portals; News agency activities; Other information service activities not elsewhere classified (United Nations, 2008^[35]).
6. This section draws upon OECD (2020^[22]).
7. <https://www.oecd.ai/>.
8. This section draws upon Chapter 12 of OECD (2018^[33]) and Chapter 7 of OECD (2020^[8]), authored by Michael Keenan, Dmitry Plekhanov, Fernando Galindo-Rueda and Daniel Ker.



From:
OECD Digital Economy Outlook 2020

Access the complete publication at:
<https://doi.org/10.1787/bb167041-en>

Please cite this chapter as:

OECD (2020), "Digital innovation", in *OECD Digital Economy Outlook 2020*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/8e8f2750-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>.