ambiguous about the type of diabetes studied were excluded. Articles were included, on the other hand, if the particular type of diabetes was not indicated in the title, keywords or abstract but the publication mentioned risk factors like obesity that are more specific to type 2.

Topics relating to SDG6: Clean water and sanitation

The data are presented in Table F3 of the statistical annex, which is available online, along with the full dataset.

Sustainable withdrawal and supply of freshwater

This dataset includes all articles pertaining to the sustainable governance, management and policy of withdrawal, supply and use of freshwater. Freshwater sources include groundwater, lakes and rivers.

Water harvesting

This dataset includes articles on techniques for harvesting precipitation, stormwater, runoff and fog. It also includes articles on rainfed agriculture and irrigation.

Desalination

This dataset includes articles related to desalination technologies, such as osmosis, membrane distillation, solar distillation and micro- and nanofiltration.

Wastewater treatment, recycling and re-use

This dataset includes articles on the management and treatment of wastewater, sewage water, polluted water and greywater, as well as re-use and recycling techniques.

National integrated water resource management

This dataset includes articles pertaining to the management of water resources and their allocation for domestic, agricultural and industrial use. It includes articles on policies and laws pertaining to water use and allocation (e.g. reservoir management), as well as system modelling (e.g. water use, reservoir or water quality) to enable strategic decision-making and the optimization of processes. Articles about management of water distribution networks or the smart water grid are also part of the dataset. Some articles may analyse the situation at local level, since the use of the key word 'national level' served, above all, to exclude articles on transboundary water resource management, as this topic is treated separately.

Transboundary water resource management

This dataset includes articles on governance, management, co-operation, law, allocation and sharing of transboundary and international freshwater resources, as well as related conflicts and disputes. Water resources analysed encompass aquifers, groundwater, lakes and rivers.

Topics relating to SDG7: Affordable and clean energy

The data are presented in Table F4 of the statistical annex, which is available online, along with the full dataset.

Cleaner fossil fuel technology

This dataset includes articles on clean coal technology, integrated carbon capture and storage technology, carbon dioxide (CO_2) capture from fossil fuel power plants, desulfurization of flue gases, the clean production of synthetic fuel, refinery processes to limit contaminants emitted, the clean synthesis of methanol and the conversion of CO_2 to hydrocarbons or fuel.

Photovoltaics

This dataset includes articles relating to stationary hosts, namely, solar cells, solar panels and photovoltaic cells. It also includes larger-scale projects, such as solar farms. Articles about developing materials specifically for solar cells are included.

Hydropower

This dataset includes articles relating to both large and small hydropower projects. It includes articles on the different types of dams but also on other hydropower methods (e.g. tidal and wave energy) and the development of specific hydropower station parts like hydroturbines.

Biofuels and biomass

This dataset includes articles on the main biofuels, such as bioethanol, biomethane, biobutanol and biocrude, as well as their production in biorefineries. Articles on the processing of biomass to create biofuels are also included. Little content was found on the hydraulic head of such processes.

Wind turbine technologies

This dataset includes articles about wind turbines and wind farms, with a focus on both their conception and their use. All main types of wind turbines are included, such as horizontal-axis, vertical-axis, offshore and floating wind turbines. Articles about the main types of generator used in the turbines are also part of this dataset, examples being doubly fed induction and Savonius.

Nuclear fusion

This dataset includes articles relating to studies of nuclear fusion for energy purposes, as well as articles relating to the development of materials, parts and methods to build nuclear fusion reactors. It voluntarily excludes articles focused on nuclear fusion reactions in stars because those are mostly unrelated to energy production on Earth.

Geothermal energy

This dataset contains articles on geothermal energy, the generation of electricity from geothermal energy and different techniques for harnessing this energy, most notably, enhanced geothermal systems and hydraulic stimulation. A small number of articles about the safe digging of wells to access this energy is also included.

Hydrogen energy

This dataset contains articles on most aspects of hydrogen energy harnessing. Machines and systems designed to have hydrogen as their source of power, the industrial production and storage of hydrogen and hydrogen-based fuel cells (i.e. most fuel cells) are all included.

Smart-grid technology

This dataset contains articles on smart grid technology and techniques and protocols to enable smarter, more reliable electrical networks. Articles are included on new devices such as smart power meters and grid-friendly appliances, as well as on protocols related to distributed grids and microgrids, such as islanding detection. Articles relating to the cybersecurity of smart grids are also included in the dataset. Furthermore, articles related to new challenges for the grid, such as electric vehicle (EV) charging on a national scale, are included because many of those articles advocate 'vehicle-to-grid' energy transfers and real-time pricing when mentioning EV charging. An effort was made to exclude papers specific to EVs that do not address the electrical grid. Some articles about renewable energy sources are part of the dataset but they only concern the integration of these new power sources in the grid.

Topics relating to SDG9: Infrastructure, industrialization and innovation

The data are presented in Table F5 of the statistical annex, which is available online, along with the full dataset.

Carbon pricing

This dataset includes articles related to all aspects of the taxation and trading of carbon and other greenhouse gas emissions (also known as cap and trade): the carbon market, carbon allowance, carbon credits, carbon trading, carbon pricing and carbon taxation.

Eco-industrial waste management

This dataset includes articles on the safe treatment and disposal of waste chemicals and other industrial wastes, such as those from the construction, pharmaceutical and electronic industries, reactive waste and wastewater treatment. Articles are also included on the minimization of waste production, such as zero-waste manufacturing and waste-to-energy processes. Articles about remediation techniques for soils contaminated with industrial wastes (heavy metals, petroleum, etc.) are included but these exclude nuclear waste, since it is treated separately.

Radioactive waste management

This dataset contains articles about radioactive waste management, storage and disposal. Articles about methods such as safe geological disposal, transmutation and vitrification of nuclear waste are included.

Eco-alternatives to plastics

This dataset includes articles studying the synthesis, effects and potential applications of ecologically friendly alternatives to traditional fossil fuel-derived plastics. Articles on biodegradable plastics or alternatives based on cellulose, starch or banana fibres are included. Any articles containing terms such as 'sustainable alternative to plastics' or 'environmentally degradable plastics' are also included.

Eco-construction materials

This dataset includes articles relating to 'eco-building', 'green building', 'sustainable construction', 'ecological housing', 'sustainable architecture' and other permutations of those terms. There is also a focus on low-energy and zero-energy buildings, sustainable building certifications (e.g. Leadership in Energy and Environmental Design [LEED] or Passivhaus) and environmentally friendly materials, sustainable materials and recycled materials. Articles on retrofitting buildings and infrastructure are included. The dataset is primarily focused on passive construction practices and materials, rather than active methods such as photovoltaic panels or wind generator integration, covered in other datasets (see Table F4). Sustainability here is almost always defined from an environmental standpoint but some articles are also focused on economic or social sustainability of buildings or infrastructure. For example, an article might tackle the health implications of using sustainable alternatives to traditional materials or designs for the occupants, or buildings that can better withstand the extreme conditions found in poorer countries or anticipated conditions relating to climate change.

Greater battery efficiency

This dataset contains articles on battery efficiency, new battery technologies and battery design and optimization. Since most articles are written to report improvements on batteries, most rechargeable battery types were included as keywords in the dataset, with the greatest contributions coming from more recent battery types such as Li-ion. Articles concerning the recycling of spent batteries and battery integration in the electrical grid are also part of the dataset.

Sustainable transportation

This database focuses on the development and use of sustainable means of transportation, including electric vehicles, solar vehicles, plug-in hybrid vehicles and hydrogen vehicles. Most articles indexed focus on cars but planes, boats, trains and trucks are all included. The dataset includes articles on self-driving cars and associated technologies, such as autonomous intersection management, lane detection and advanced driver assistance. This is because self-driving cars would achieve better fuel economy (or energy economy, if electric) and have the potential to reduce the number of road accidents. Articles are included on other methods of sustainable transportation and policies, such as carsharing, public bicycles, car-free cities and congestion pricing. 'Sustainable transport' is used as a keyword but, with the exception of publications using the terms 'solar boat' and 'electric boat', shipping is not explicitly included in this dataset.

Topics relating to SDG13: Climate action

The data are presented in Table F6 of the statistical annex, which is available online, along with the full dataset.

National and urban greenhouse gas emissions

This dataset includes articles on the measurement, assessment and estimation of greenhouse gas emissions at the national and urban levels. It includes articles on national and urban emission sources, such as agricultural, industrial and urban sources, including transportation.

Carbon capture and storage

This dataset includes articles on carbon capture techniques, such as pre-, post- and oxyfuel combustion, direct air capture and carbon sequestration, which is the capture and storage of carbon through natural (biological, chemical and physical) processes that include photosynthesis, oceanic and geological capture. The publications refer to both natural sequestration and human attempts to enhance sequestration. This dataset is restricted to gaseous carbon; it excludes solid-state carbon, namely particulate matter.

Local impact of climate-related hazards and disasters

This dataset focuses on articles assessing climate hazards and climate-related disasters and the impact of these on small and vulnerable communities. Examples are coastal erosion, sea level rise, droughts, floods and extreme weather events.

New technologies to protect from climate-related hazards

This dataset focuses on any technology, innovation or tool that has the potential to mitigate the impact of climate change or climate-related disasters on communities. This includes publications on ways of building more resilient infrastructure, as well as those on better modelling tools to predict the risk of occurrence and consequences of catastrophic events, enabling better preparedness.

Local disaster risk reduction strategies

This dataset focuses on planning processes and techniques that help to reduce the risk faced by local and vulnerable communities with respect to climate change and related disasters, such as drought, flooding, extreme storms and wildfires. The publications included consider communities' strategies, preparation and mitigation efforts as well as their capacity to recover and be resilient.

Climate-ready crops

This dataset includes articles related to agriculture and crops that are tolerant of and resilient to salinity, flooding, drought and other climate-related stressors.

Topics relating to SDG14: Life below water

The data are presented in Table F7 of the statistical annex, which is available online, along with the full dataset.

Coastal eutrophication

This dataset includes all articles on coastal eutrophication, phosphorus removal and pollution, algal blooms, water nutrient dynamics and specific harmful algal species.

Floating plastic debris in the ocean

This dataset includes articles related to plastic debris and plastic pollution in the ocean. Search terms cover plastic and plastic types in aquatic environments, plastic consumption by marine wildlife, as well as the identification of plastics and byproducts in marine and coastal environments and food chains.

Ocean acidification

This dataset includes all articles on ocean acidification and seawater acidification, including those assessing the impact on species.

Sustainably manage marine tourism

This dataset includes articles on tourism management and the effects of tourism on oceanic or coastal zones. Articles often focus on the environmental impact but the socio-economic impact is also covered. Whenever there was a focus on sustainability, cruises, diving, recreational fishing and sea animal tourism (watching, swimming and diving) were all specifically included as search terms.

Sustainably manage fisheries and aquaculture

This dataset includes articles on fish farming and aquaculture, as well as commercial fishing as it relates to sustainable resource management. The dataset was constructed using key terms related to sustainable fisheries and aquaculture, as well as species and techniques commonly associated with these practices. When aquatic species names were used, this was done in conjunction with terms related to sustainability, to limit the dataset to the topic as much as possible.

Ecosystem-based approaches in marine environments

This dataset contains articles relating to management of coastal zones using an ecosystem-based approach, meaning the ecosystem and its interactions are considered as a whole in management processes. The dataset includes articles on concepts like integrated coastal zone management, which is closely linked to ecosystem-based approaches. When the dataset was verified, no articles focusing on ecosystems in international waters were found.

Topics relating to SDG15: Life on land

The data are presented in Table F8 of the statistical annex, which is available online, along with the full dataset.

Sustainable use of terrestrial ecosystems

This broad dataset includes all papers on the conservation, preservation, restoration and sustainable use of terrestrial ecosystems: forest, habitat, land, wildlife, pond, freshwater, drylands, etc. It includes research on species richness, biodiversity, etc. It also includes articles on ecosystem services.

Status of terrestrial biodiversity

This dataset includes articles on species with a defined conservation status, as well as more fundamental research needed to monitor, predict and protect those species. For example, articles on causes of species endangerment, such as habitat loss, are included.

Minimize poaching and trafficking of protected species

This dataset includes articles on the Convention on the International Trade of Endangered Species (CITES), wildlife trade, poaching, illegal fishing, illegal harvesting, wildlife forensics, etc.

Tackle invasive alien species

This dataset includes all articles related to biological invasions, alien species, invasive plants, etc. Basic research on invasive potential has also been included.

Use of ecosystem-based approaches in protected areas on land

This dataset includes articles on ecosystem-based management of public and private land. An effort was made to exclude marine and freshwater ecosystems, to avoid overlap with other selected topics.

Extent of water-related ecosystems

This dataset includes articles measuring the extent of water-related ecosystems (excluding oceans) such as wetlands, rivers, estuaries, lakes, aquifers, swamps, fens, peatlands, marshes, mangroves and artificial water bodies. It involves articles either measuring or improving delineation protocols of such systems, as well as estimating the water quantity involved. Articles on disappearing systems are included, as well as articles on the inventory of water-based ecosystems. Some articles on ecosystems related to either the water quality or quantity found in ecosystems such as forests and mountains are also included when these are directly related to water.

Socio-ecological impact of terrestrial protected areas

This dataset covers conservation planning, reserve selection and reserve design, as well as the prioritization of protected areas and the future habitat suitability under climate change. The negative and positive impact of protected areas on humans, such as human–wildlife conflict or valuation of ecosystem services, are also targeted.

Annex 5: Technical and methodological note on the Statistical annex

METHODOLOGICAL NOTE

Bibliometric data

Publication data have been compiled for UNESCO by Science-Metrix from the Scopus (Elsevier) database, excluding Arts, Humanities and Social Sciences.

Publications listed under the People's Republic of China do not include those for its Special Administrative Regions of Hong Kong and Macao, or the Taiwan Province of China.

Economic data

Data on economic indicators, such as gross domestic product (GDP) and purchasing power parity (PPP), are based on the World Bank's economic data release of July 2020: http://data.worldbank.org/products/wdi (see the note on the cut-off date).

It should be noted that, since 2014, the UNESCO Institute for Statistics has used data on total general government expenditure (all sectors) from the International Monetary Fund's World Economic Outlook database as the denominator for its indicator, expenditure on education as a percentage of total government expenditure. For more information about the change in methodology, please visit: www.uis.unesco.org/education

Education data

The UNESCO Institute for Statistics (UIS) compiles education statistics, from official administrative sources at the national level, through its annual Survey of Formal Education. The objective of this survey is to provide internationally comparable data on key aspects of education systems, such as access, participation, progression and completion, as well as the associated human and financial resources dedicated to them.

In addition to this survey, these data are also collected through a joint data collection mechanism involving UIS, the Organisation for Economic Co-operation and Development (OECD) and the Eurostat. The data collected are used to monitor and report on international development goals related to education, including the education goal of *The 2030 Agenda for Sustainable Development*.

For more information on the UIS Survey of Formal Education, please visit: http://uis.unesco.org/en/uis-questionnaires

GERD

R&D expenditure in US dollars at the purchasing power parity (PPP$) rate is presented in Table B1 using constant 2005 prices.

Innovation data

The UNESCO Institute for Statistics collected data on innovation within the manufacturing industry every two years up to 2017 through its innovation data collection. In addition, the institute obtained innovation data directly from Eurostat and the African Science, Technology and Innovation Indicators (ASTII) Initiative of the African Union/NEPAD Planning and Coordinating Agency for countries which participate in the data collections of these organizations. With a few exceptions, innovation data refer to a three-year reference period that varies from one country to another. The data collected are featured in the institute's international database at: http://data.uis.unesco.org

Population data

Population data were obtained from the World Bank's World Development Indicators (accessed September 2020). For Eritrea, data were available only for 2011 and this value was used for all years, which could lead to a slight overestimation of the number of publications per million inhabitants for later years. Population data for the Cook Islands and Niue were obtained directly from the national statistical service. Niue provided data for only 2011 and 2017; the remaining years were estimated based on the growth between 2011 and 2017.

Research and experimental development (R&D) data

The UNESCO Institute for Statistics collects data on resources devoted to research and experimental development through its R&D statistics survey. In addition, it obtains data directly from the OECD, Eurostat, the Ibero-American and Inter-American Network on Science and Technology Indicators (RICYT) and the African Science, Technology and Innovation Indicators (ASTII) Initiative of the African Union/NEPAD Planning and Coordinating Agency for countries which participate in the data collections of these organizations. The data collected can be found at: http://data.uis.unesco.org

Data obtained from the OECD are based on the OECD's Research and Development Statistics database released in April 2020. Data obtained from Eurostat are based on the Eurostat Science and Technology database, as of April 2020. Data received from RICYT are as of July 2020. Data obtained from ASTII are based on the African Innovation Outlook III (2019).

Cut-off date for data in the statistical annex and chapters

R&D and economic data presented in the regional/individual country chapters may not always correspond to the data given in the statistical annex or in Chapter 1. The reason for this is that the underlying economic data used to calculate R&D indicators are based on the World Bank's economic data release of July 2020 for the annex and Chapter 1, whereas data in the other chapters might come from a different release of economic data by the World Bank.

TECHNICAL NOTE

Bibliographic data

Number of papers: this is the number of peer-reviewed scientific publications (e.g. articles, reviews and notes only) indexed in the Scopus database. Publications are assigned to countries according to the author affiliation address on the publications. Double counting is avoided at both the national and regional levels. For instance, a paper co-authored by two researchers from Italy and one author from France is counted only once for France and once for Italy but also once for Europe and once for the world.

The regions were defined using a blend of geographic and economic groupings based on status in 2019. For example, the United Kingdom was included in the European Union region due to its membership in the period under study (2011 to 2019). The total for the People's Republic of China excludes its Special Administrative Regions of Hong Kong and Macao and the Taiwan Province of China.

Number of international collaborations: this is the number of publications involving authors from at least two different countries, counted once for each entity. For instance, if 13 authors from one organization collaborate with 1 author from another, that article counts as a co-publication, counting once for each organization. For the computation of international collaboration, territories were considered to be part of their respective mainland countries. Thus, collaboration between Guadeloupe and France would not be considered as international co-authorship. The international collaboration rate of a country is simply a measure of how many articles are co-published with international partners as a proportion of the given country's total output.

Average of relative citations: this is an indicator of the scientific impact of papers produced by a given entity (the world, a country, an institution) relative to the world average (i.e. the expected number of citations, set as 1). Citation scores for conference papers are not included in the citation analysis because citation patterns associated with them tend to distort impact measurements.

Field classification of publications: see Annex 2 for the broad fields of science used to prepare statistics at the level of scientific disciplines: agriculture, fisheries and forestry; animal and plant biology; built environment and design; chemistry; cross-cutting strategic technologies; engineering; environmental sciences; geosciences; health sciences; ICTs, mathematics and statistics; and physics and astronomy.

Education data

The fields of study are defined as per the International Standard Classification of Education: *Fields of Education and Training* (ISCED-F 2013). Agriculture (includes agriculture, forestry, fisheries and veterinary); natural sciences (includes biological, environmental, geological and physical sciences), mathematics and statistics; engineering (includes engineering, manufacturing and construction); ICTs (information and communication technologies); health (health and welfare); social sciences and journalism; business, administration and law; and arts and humanities (contains the ISCED-F 2013 broad fields of arts and humanities, education and services).

Table 1: **Levels of tertiary education**

UNESCO Science Report term	ISCED 2011 equivalent
Short cycle	ISCED level 5 – short-cycle tertiary education
Bachelor's or equivalent	ISCED level 6 – bachelor's or equivalent level
Master's or equivalent	ISCED level 7 – master's or equivalent level
PhD or equivalent	ISCED level 8 – doctoral or equivalent level

International classifications for the type of degree received follow the International Standard Classification of Education (ISCED 2011) of level of education (Table 1).

Data on internationally mobile students that are collected by the UNESCO Institute for Statistics (UIS), OECD and Eurostat encompass students who are pursuing a tertiary degree and thus exclude students on exchange programmes. Data on inbound internationally mobile students reported by host countries are used by the UIS to estimate the number of outbound internationally mobile students from a given country. As not all host countries report disaggregated data on internationally mobile students by country of origin, the number of outbound students from a given country may thus be underestimated.

R&D data

The definitions and classifications used to collect R&D data are based on the *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development* (OECD). Some of the key definitions related to R&D data are presented in the glossary of the present report.

Two types of R&D indicator are usually compiled: data on R&D personnel measure researchers, technicians & equivalent staff directly involved in R&D, as well as other support staff; data on R&D expenditure measure the total cost of carrying out the R&D activity concerned, including indirect support.

Global and regional estimates for R&D expenditure and researchers presented in Chapter 1 are calculated based on country-level data obtained from the UNESCO Institute for Statistics, accessed in August 2020.

Patent data

Number of granted patents: this is the number of granted patents indexed in the PATSTAT database for five patent offices, namely the US Patent and Trademark Office (USPTO), European Patent Office (EPO), Japanese Patent Office (JPO), Korean Intellectual Property Office (KIPO) and State Intellectual Property Office of the People's Republic of China (CNIPA). Patents are assigned to countries according to the country of the inventors on the applications. Double counting is avoided at both the national and regional levels. For instance, a patent application submitted by two inventors from Italy and one inventor from France is counted