Chapter 25

Future Data and Knowledge Needs

© Shutterstock/kan_chana

Coordinating Lead Authors: Florence Daguitan (Tebtebba, Indigenous Peoples' International Centre for Policy Research and Education), Charles Mwangi (Global Learning and Observations to Benefit the Environment [GLOBE] Program), Michelle Tan (ADEC Innovations)

Lead Authors: Graeme Clark (University of New South Wales), Daniel Cooper (University of Oxford), James Donovan (ADEC Innovations), Sheryl Gutierrez (ADEC Innovations), Nina Kruglikova (University of Oxford), Pali Lehohla (Pan African Institute for Evidence), Joni Seager (Bentley University), William Sonntag (Group on Earth Observations)

GEO Fellow: Amit Patel (Planned Systems International, Inc.)

Executive summary

Citizen science is providing unprecedented opportunities for engaging the public in collecting and analysing vast amounts of environmental data (*well established*). The potential for massively dispersed teams of observers, coupled with new technologies such as smart sensors, mobile telephony, Internet and computing capabilities, is offering new approaches for research and engaging the public on environmental issues. As well as collecting large volumes of data, the advancement of new technologies has also enhanced the quality and veracity of the data collected. Key opportunities presented by citizen science include greater frequency of data from dispersed sources, the ability to address large knowledge and funding deficits, the ability to educate the public about environmental policy issues, and the use of local knowledge. {25.2.1}

Big data is one of the world's emerging valuable resources, shifting the landscape of environmental assessment at global, national and local scales (*well established*). Traditional processing techniques cannot handle the volume, velocity, variety and veracity of big data, demanding new algorithms, programming and statistical methods to derive information and draw evidence-based conclusions. There is enormous potential for advancing environmental knowledge if big data can be effectively harnessed and interrogated. {25.2.2}

Governments, organizations, academia and the private sector have initiatives seeking opportunities to tap the potential of big data for sustainability and development (*well established*). Current initiatives include the establishment of the United Nations Pulse Labs for pilot studies on big data, the formation of the United Nations Global Working Group on Big Data in monitoring the Sustainable Development Goals (SDGs), and the availability of repository sites and open data sources from multilateral organizations, research centres and government collaborations. Big data from web-based and geospatial mapping technologies, remote sensing and statistical visualization provide a basis for environmental assessment. {25.2.2}

Challenges for using big data in environmental assessments include its accessibility, quality, varying scale and context, and incomplete time series (*well established*). Despite efforts to generate globally acceptable and available big data, capacities are limited by resources and funding, especially in developing countries. Much real-time big data are controlled and held by the private sector, though many data products are made freely available for public good in a process known as data philanthropy. Recommendations for building a holistic system for big data include the establishment of leadership and data governance; collaborations among governments, institutions and the private sector; and institutionalizing legal frameworks with safeguards on information. {25.2.2}

Strengthening the ability to gather, interpret and use data for effective planning, policymaking, management and evaluation could provide countries with a comprehensive view of environmental impacts (*well established*). Governments and society need to adapt to the evolving data landscape, including the possible use of artificial intelligence to manage environmental concerns. Coping with the shift in the data landscape entails new information-technology skills and a holistic approach in utilizing emerging and existing data and knowledge tools. {25.3}

Traditional knowledge held by indigenous peoples and local communities is increasingly seen as a valuable resource for environmental assessment and sustainable development (*well established*). This revaluation is evidenced by the increase in discussions and studies on traditional knowledge, and its inclusion in global policy agreements. In order to address current and future challenges such as climate change, research suggests that the best approaches may be characterized by the coordination of modern science and technology with traditional knowledge. While cooperation between local and global communities and knowledge systems has proven to be successful for the health of individuals and the planet, certain challenges remain. {25.2.3}

25.1 Introduction

This chapter discusses emerging areas of environmental information and statistics, including citizen science, big data and traditional knowledge. It aims to summarize the gaps and opportunities for improving the environmental knowledge base.

The global landscape is changing, technology is advancing and more and more data are available. These new data sources will not override the need for traditional means of data collection but will provide additional opportunities for environmental monitoring and assessment. This chapter analyses these new and emerging means of data collection and presents a perspective for the future of environmental monitoring and assessment.

25.2 Emerging tools for environmental assessment

Citizen science, big data and traditional knowledge are not new sources of information; what is new is their regular and systematic use in environmental assessments. This section highlights some current experiences and the need to use these innovative sources of information to fill data gaps.

25.2.1 Citizen Science

Citizen science entails the engagement of volunteers in science and research. Volunteers are commonly involved in data collection, but can also be involved in initiating questions, designing projects, disseminating results and interpreting data (Blaney *et al*. 2016). Coupling Citizen Science with new emergent technologies is providing unprecedented opportunities for doing research and sensitization of the public on environmental issues (Newman *et al.* 2012, p. 298).

The possibility of tapping into a massive, dispersed team of observers in different regions of the world has created opportunities for collating and analyzing data at unprecedented spatial and temporal scales. Citizen Science projects have the potential to gather large amounts of scientific data but this is only helpful if data collected is utilized in one way or the other (Dickinson, Zuckerberg and Bonter 2010; Kim *et al.* 2011; Dickinson *et al.* 2012).

Citizen science has numerous benefits, the main one being the opportunity to collect data over wider spatial coverage and longer periods at lower cost. Additional benefits include the creation of jobs, increased scientific literacy, citizen engagement in local and environmental issues, cost effectiveness for governments and benefits to the environment being monitored. Citizen Science also allows the expertise of scientists to be brought to the public while at the same time exposing the scientists to the indigenous knowledge and expertise available within the local community (Conrad and Hilchey 2011; Blaney *et al.* 2016). Some of the key benefits of citizen science are highlighted in **Figure 21.1: Selected targets and their related clusters as examined in this chapter**.

The fields of astronomy and ornithology have led the charge for citizen science. In 1900, Frank Chapman, an ornithologist with the American Museum of Natural History initiated the Christmas Bird Count (CBC). This project has survived thanks to the enthusiasm of citizen scientists over the years and is currently being run by the National Audubon Society (Dickinson, Zuckerberg and Bonter 2010). Since then, there have been many citizen science projects over the years at local, regional and global scales, covering different areas of interest.

More recently, citizen science projects have included a wide variety of initiatives, ranging from building collaborative knowledge (e.g. Wikipedia, OpenStreetMap), volunteer computing (e.g. CitizenGrid, climateprediction.net), and pattern classification (e.g. Galaxy Zoo, eyewire), to the community collection of observations (e.g. bird counting, air sensor toolbox) (Mathieu *et al.* 2016).

Many environmental interests that transcend government boundaries, such as pollution and bird migration, have increased the engagement of citizen scientists to monitor these issues of concern. More innovative projects include the use of Google's reCAPTCHA, which has facilitated the digitization of books and millions of articles by turning words that cannot be read by computers into CAPTCHAs for people to solve (Conrad and Hilchey 2011; Google 2018).

There are two main approaches used in the organization of citizen science projects; top-down or bottom-up. These approaches are similar to the concepts in Chapter 10 on evaluation of policy effectiveness.

Figure 25.1: Some of the benefits of citizen science

Individual Citizen

- Learn observational and analytical skills
- Gain a better understanding of the natural world
- Job opportunities
- Capacity building

Governments

- Lower cost of data collection
- Wider spatial and temporal coverage of data
- Promote environmental stewardship

Communities

- Monitor the health of the environment
- Increased interaction of the community
- Promote environmental stewardship

Scientists and Researchers

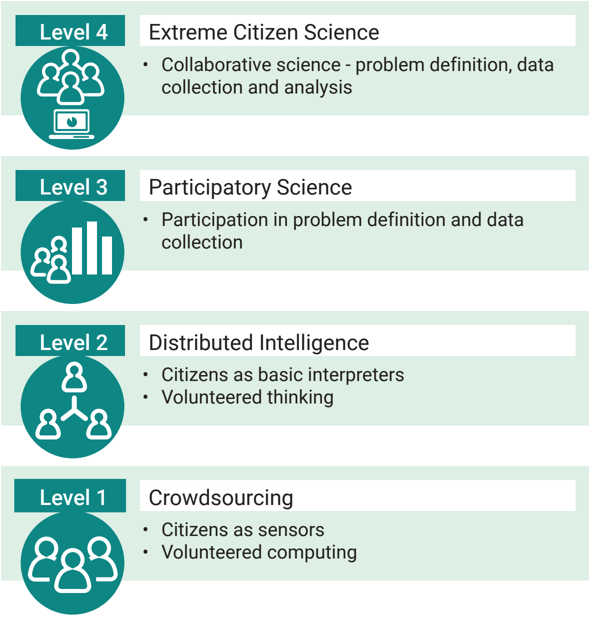
- Large numbers of participants reduce workload
- Scientists are able to build connection with community
- Teach people how to research

The top-down approach is mostly driven by scientists who train volunteers on the procedures and the research to be undertaken. Based on this approach, the volunteers play limited roles mostly in data collection. The bottom-up approach is driven by the community. More often than not, this is driven by the need of the community to understand or gather evidence of a concern. The community can then approach scientists for support and guidance during the process (Roelfsema *et al.* 2016; Shirk *et al.* 2012).

The level of engagement, skills and knowledge needed by volunteers to participate in citizen science projects varies depending on the scope of the research. Some projects require basic data collection knowledge requiring minimal or no training of volunteers, while others require intensive training (Haklay 2013; Shirk *et al.* 2012). **Figure 21.4: Future projections of global average crop yield (top left), crop production (top right), agricultural area (bottom left), and forest and other natural land area (bottom right)**, illustrates the various levels of engagement of volunteers in citizen science projects.

Citizen scientists can help to uncover critical information about our environment which could possibly take scientists years to discover by themselves. An example is illustrated by the infographic shown in **Figure 25.3** where rivers need a citizen science movement for monitoring, and how the collected data and findings are used to maintain ecosystem services and human wellbeing (Pottinger 2012). The figure also demonstrates a step-by-step procedure for conducting citizen science. This data collection and analysis procedure can be replicated across the Drivers (Chapter 2) and the various environmental themes (Chapters 5 to 9).

Figure 25.2: Levels of citizen science by increasing depth of the participation



Source: Haklay (2013).

Trends in citizen science

The technology revolution has heralded multiple novel ways of collecting, archiving, analyzing, and transmitting data. The emergence of the internet-of-things (IoT), miniaturized smart sensors with geo-location functions, ease of accessing internet and data as well as the potential of cloud storage and computing has expanded the possibilities and opportunities for data collection and analysis. This rapid advancement in technology coupled with greater exposure and sensitization of the public, have led to an explosion in uptake of projects based on citizen science (Mathieu *et al.* 2016).

The availability of internet and geographic information system (GIS)-enabled web applications has enabled citizen scientists to collect large volumes of geographically-referenced data and submit them electronically to centralized databases. An example of such a system is the Global Learning for Observation to Benefit the Environment (GLOBE) program, which uses students to collect environmental data and archive it in the GLOBE program database (Dickinson *et al.* 2012; GLOBE 2018).

The expansion in the use of smartphones, the possibility of digital photo validation of observations, and the capability of creating simple online data-entry systems is revolutionizing the process of initiating citizen science projects while ensuring data accuracy at minimal cost. Currently, it is now possible to create mobile phone apps for collecting different types of datasets and automatically geo-locating the data, using the in-built GPS receiver chip on most mobile phones (Dickinson, Zuckerberg and Bonter 2010; Dickinson *et al.* 2012).

Scientists are now increasingly using citizen scientists to collect geo-referenced *in-situ* data which can be used to support the calibration and validation of Earth Observation satellite data products. Citizen scientists are also involved in the interpretation and digitization of Earth Observation (EO) data sets (Mathieu *et al.* 2016; See *et al.* 2016). Tomnod is such an example of using crowdsourcing and citizen scientists to identify objects and places in satellite images. Tomnod was used in trying to locate the missing Malaysian Airlines flight MH370 aircraft using satellite imagery. Approximately 2.3 million Internet users submitted 18 million tags for over 745,000 satellite images clearly illustrating the potential of citizen science (Mazumdar *et al.* 2017).

Another example of the use of citizen scientists to validate satellite data is the partnership between NASA's Global Precipitation Measurement (GPM) satellite mission and the GLOBE program. The GLOBE program is an environmental

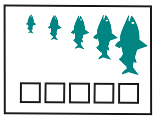
Figure 25.3: An example of citizen science that demonstrates how it is needed and can be replicated

Case Study: Why our rivers need a citizen science movement

Most of our decisions are based on incomplete or inadequate data/information - in the absence of professional scientists to fill these gaps, citizen science can step in to help uncover information and findings. This case study explores opportunities where citizen scientists have filled this gap.

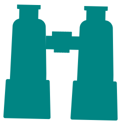
TOO MUCH WORK, TOO FEW SCIENTISTS

- There is too much research that needs to be done yet hardly enough scientists to undertake it all by themselves.
- Volunteers in the US state of Oregon are helping scientists survey 146 miles of streams by locating and counting salmon and native trout species and helping restore habitat.
- Hundreds of volunteers with The Nature Conservancy annually survey how much desert land is made wet by the San Pedro River; they cover more than 250 miles.



WHAT WE DON'T KNOW CAN HURT US

- Citizen science can be used to document basic information about a river system, as well as changes over time to its flow, sediment load, species and water quality.
- In China, the South-North Water Transfer Project was envisioned to have massive impacts on many key waterways, volunteers were recruited for a 4 year assessment of 10,000km of China's western rivers.

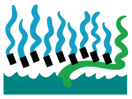


BECAUSE THEY'RE OUR RIVERS

- Volunteers for the Mystic River Watershed Association in the Eastern US, with support of scientists, take monthly samples at 15 locations along the river to monitor water quality.
- Advocacy based on their results has helped improve the river's cleanliness and has enabled the residents get involved in their natural environment through hands-on science.

UNCOVERING RIVER MYSTERIES

- Citizen scientists can fill in gaps in crucial baseline knowledge about a river's species or general health.
- The Mekong River in Southeast Asia supports several species of giant fish and very little is known about them. More information is needed on where they spawn, what natural cues drive them to spawn, population estimates, and maps of their life-cycle territory.
- A step-by-step guide on Mekong River citizen study illustrates how this can be realized.



Just 1 in 10 rivers now reach the sea

8/10 people depend on river resources

Just 2/3 of Earth's major rivers are dammed

STEP-BY-STEP CITIZEN SCIENCE GUIDE: CASE STUDY OF FISHING VILLAGERS DOCUMENTING MEKONG'S RIVER'S NATURAL WEALTH

IDENTIFY THE QUESTIONS YOU WANT TO ANSWER

- In 1994, Thailand built Pak Mun Dam on the largest tributary of the Mekong, destroying local fisheries and harming river-based communities.
- Information on local fisheries was scant.
- In 2001, the Thai government opened the dam's floodgates 1-year study of its impacts to fisheries.

FORM A RESEARCH TEAM

- South East Asia Rivers Network (SEARIN) and Assembly of the Poor teamed up to monitor the changes caused by the dam.
- Their innovative citizens' science research method, called Thai Bahn (Thai Villager) research, relied on local fishers to gather information.

DEVELOP A PLAN OF ACTION

ACTION PLAN

- Methods, areas of study, and research team members were all decided by the local villagers.
- SEARIN helped develop a plan of action, write up their findings and increase international awareness.

DOCUMENT YOUR FINDINGS

- The natural flows of the one-year trial period allowed people to resume traditional ways of life and eased resource conflicts among river communities.
- Local fish species not seen for eight years came back; researchers found a total of 156 fish species had returned to the Mun River.

ANALYZE YOUR DATA

- SEARIN helped create a report on the team's findings, in two languages.
- The report is considered one of the most thorough documentations of Mekong fisheries produced for that area.

SHARE YOUR FINDINGS, AND USE THEM FOR ACTION

- Thanks to this citizen science effort, the villagers succeeded in getting the Thai government to open the dam gates for four months each year to allow for fish migration.
- Subsequent governments have not implemeneted this agreement.
- The project has inspired many other citizen science projects to protect rivers in the region.

Source: International Rivers (2012).

educational program for primary and secondary schools, where students from schools across the world collect precipitation data using rain gauges as shown in **Figure 25.4**. The collected data, as well as data collected from other sources, is used by NASA to calibrate and validate GPM precipitation measurement data (United States National Aeronautics and Space Administration [NASA] 2018).

Figure 25.4: GLOBE Students in St. Scholastica Catholic School in Nairobi collecting and recording the amount of precipitation for the GPM Satellite Mission field campaign

Source: © GLOBE Program (Kenya).

Automated and autonomous equipment such as drones, remotely operated sensors, autonomous underwater vehicles (AUV's) and underwater gliders are predicted to play an increasing role in citizen science. These autonomous systems can be a primary source of data or complement data collected *in situ*, provide high resolution data nearly in real time, be deployed on a need basis and often enable access to remote or extreme locations such as observation of marine environments. In addition, they are low cost compared to satellites and are thus offering alternative and credible sources of EO data (Macauley and Brennan 2016; Garcia-Soto 2017).

Citizen science, as well as other data sources, contributes to Big Data collection and these huge volumes of data need processing. Numerous approaches have been explored to involve the huge numbers of citizen scientists to assist in analyzing these huge volumes of data, one of which is the development of game-like systems (gamification). Citizen participation in these games help to speed up the data analysis and allow science to advance more rapidly (Van Vliet and Moore 2016; Spitz *et al.* 2017; McCallum *et al.* 2018).

An example of gamification is Cropland Capture, a game version of the GeoWiki project, which engaged citizen scientists in global land cover research, helping researchers identify farmland around the world. The game managed to collect 4 million classifications from over 3,000 players identifying images with and without cropland present (See *et al.* 2013).

Table 25.1 shows some of the global and regional projects dedicated to citizen science.

The potential of citizen science should not be limited to engaging volunteers to collect and collate scientific data as illustrated in **Figure 25.5**. Citizen science can be used to sensitize and engage the community on issues related to their natural environment, to better understand them and allow them to take charge, and provide an avenue for showcasing the need

Table 25.1: A selection of citizen-science projects and websites

Programme	Region	Description	Website
UNEP Environment Live	Global	UN open access platform of global, regional and national environmental data	https://environmentlive.unep.org
SciStarter	Global	Aggregates information, video and blogs about citizen-science projects	www.scistarter.com
Data Observation Network for Earth	Global	Provides a framework to access data from multiple data sources (including citizen science data)	www.dataone.org
CitSci.org	Global	Provides tools for citizen scientists to guide them on the entire research process such as: process of initiating research projects, managing the process of data collection, and analysis	www.citsci.org
iSpot	Global	Website aimed at helping anyone identify anything in nature by connecting citizen scientists with experts in species identification	www.ispotnature.org
eBird	Global	Online database of bird observations with real-time data about bird distribution and abundance	www.ebird.org

Figure 25.5: Citizen scientists collecting environmental data

Source: © GLOBE Program (Kenya).

to maintain and conserve our ecosystems given the increasing pressures on the environment (Roelfsema *et al.* 2016).

Challenges of citizen science

Challenges in citizen science mostly revolve around three main issues: *organizational* issues, *data-collection* issues and *data-use* issues. At the organizational level, the challenges include the process of recruiting volunteers, motivating and providing incentives for their participation and ensuring sustainability of the initiative as well as funding. On data collection, the issues that arise include: data fragmentation, data representativeness, data quality (for example data intentionally flawed by the data collector) and/or lack of essential metadata. In data use, the challenges include: differences in protocols and standards, legal issues, data-privacy concerns and the question of allowing open access (Conrad and Hilchey 2011; Hochachka *et al.* 2012; Rotman *et al.* 2012; See *et al.* 2016)

Due to misunderstandings and lack of technical knowledge and skills to handle such data, concerns have emerged over the credibility, comparability, completeness of, and lack of metadata, as well as challenges in data access and sharing, and these have resulted in these data not being seriously considered by policy and decision makers. In most cases, perception of poor data quality, rather than the actual data quality and fitness for use, have influenced the value and use of citizen science data (University of the West of England, Science Communication Unit 2013; Storksdieck *et al.* 2016).

The key opportunities presented by citizen science, mainly include:

i. use of local knowledge;
ii. timely data from dispersed sources;
iii. capability to address large knowledge and funding deficits;
iv. ability to educate the public about environmental policy issues; and
v. enhance participatory democracy.

For citizen science to be widely accepted, there is a need for appropriate training and support for citizen science project coordinators and those that use the data that emerge from it. Careful design of citizen science projects and application of appropriate quality assurance methods, as illustrated in **Figure 25.3**, can ensure that the effort of citizen scientists is not wasted (University of the West of England, Science Communication Unit 2013; Storksdieck *et al.* 2016).

There are on-going initiatives, such as the Public Participation in Scientific Research (PPSR)-Core data model framework as illustrated in **Figure 25.6**, to establish data and metadata standards to facilitate international collaboration and improve data standardization, interoperability, integration, accessibility, and dissemination of citizen science data (Bowser *et al.* 2017). Citizen science has the potential to provide credible data to bridge the data gaps highlighted in Chapter 3 and to provide data to enable the monitoring of SDG environmental indicators.

25.2.2 Big data and data analytics

Big data can be defined as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze" (Manyika *et al.* 2011).

Data are one of the world's valuable resources, shifting the landscape of environmental assessment across global, national and local scales ("The world's most valuable resource is no longer oil, but data" 2017). From 1.8 zettabytes (1.8 trillion gigabytes) of data generated in 2011 (International Data Corporation [IDC] 2012), the total amount of data is expected to reach 40 zettabytes (40 trillion gigabytes) by 2020 (Dell EMC and IDC 2014). With this influx, traditional processing applications will be unable to cope with the quantity of data from multiple sources. Big data is characterized by the four Vs of large storage capacity (volume), speed at which data are generated and transmitted (velocity), the complexity of unstructured data types (variety), and the uncertainty of data sources (veracity) **(Figure 25.7)**. A fifth V (value) is achieved through the application of data analytics (International Business Machines [IBM] 2017).

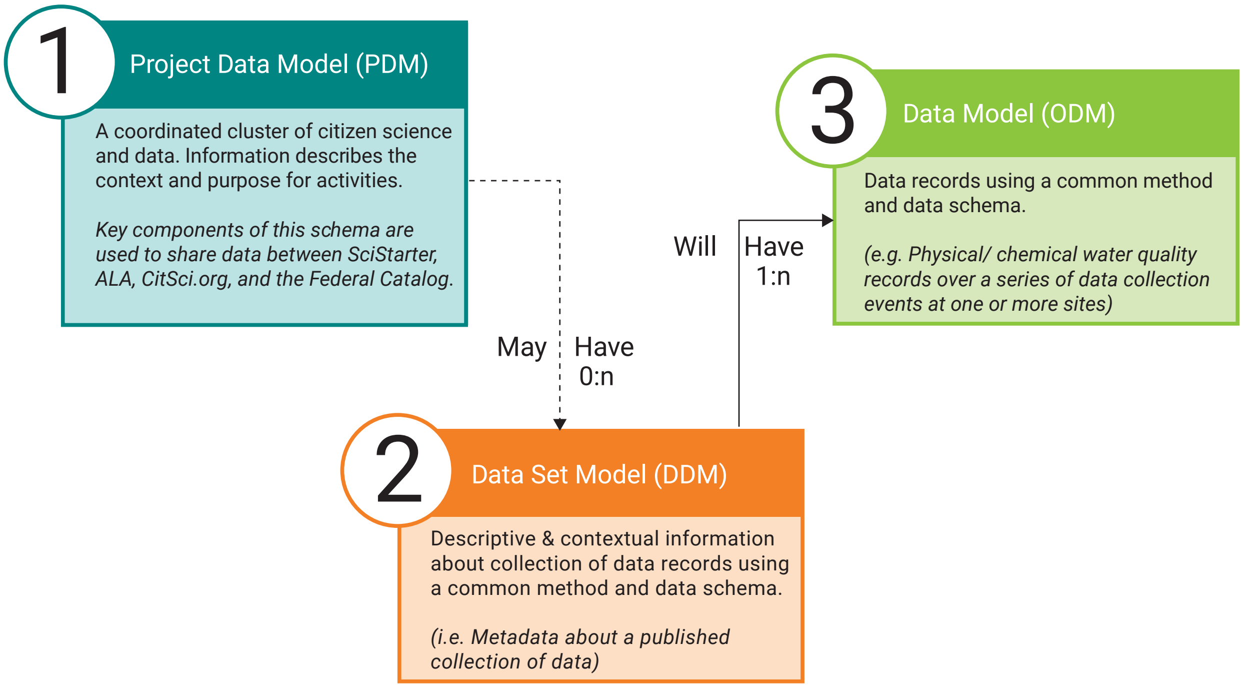
The science of data analytics is needed to create patterns from intricate data sets and find correlations (e.g. chemical pollution and locations in aerial photographs) by using algorithms, programming, and mechanical and statistical methods to draw evidence-based conclusions and obtain information that is useful for decision-making purposes (Monnappa 2017). Examples of insights drawn from big data analytics include those from projects in the United Nations Global Pulse initiative, such as:

i. urban dynamics drawn from mobile data used to improve transportation in Sao Paulo and Abidjan;
ii. campaign developments based on a survey of perceptions of HIV on social media; and
iii. and a support-services location plan based on the spatial epidemiology of Dengue fever (Kirkpatrick 2016).

Current trends and initiatives in big data
United Nations member states, in partnership with the academic and research communities, non-governmental organizations and the private sector, are seeking out innovations and looking for opportunities to tap into the optimum potential of big data for sustainability and development.

Innovation for public good
The United Nations Global Pulse initiative was founded in 2009 to progressively establish a global network of Pulse Labs to collect digital data for decision-making purposes (United Nations 2018a). Pulse Labs continue to innovate machines and to conduct pilot studies on the scalability of the capture and analytics of big data for sustainable development – some examples are presented in **Table 25.2.**

Figure 25.6: The PPSR-Core data-model framework



A coordinated cluster of citizen science and data. Information describes the context and purpose for activities.

Key components of this schema are used to share data between SciStarter, ALA, CitSci.org, and the Federal Catalog.

Data records using a common method and data schema.

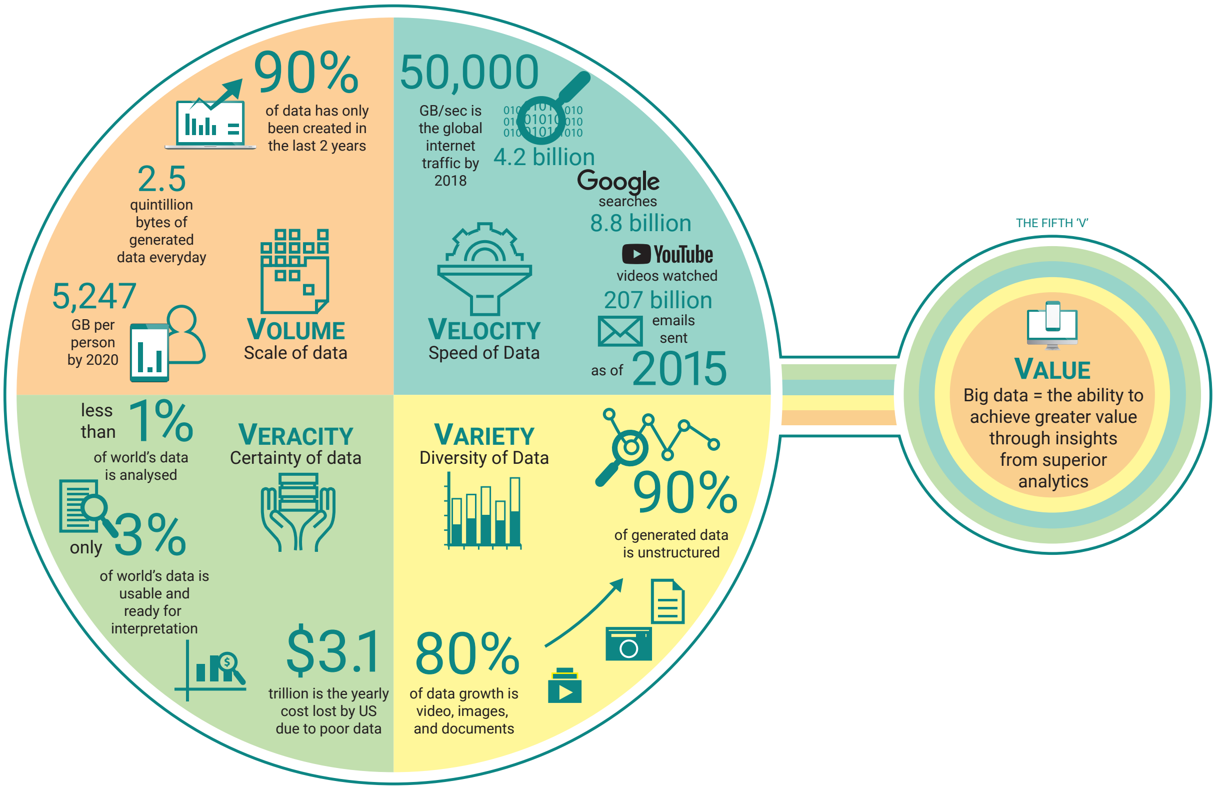
(e.g. Physical/ chemical water quality records over a series of data collection events at one or more sites)

Descriptive & contextual information about collection of data records using a common method and data schema.

(i.e. Metadata about a published collection of data)

Source: Bowser *et al.* (2017).

Figure 25.7: Characteristics of big data and the role of analytics



Source: Adapted and recreated the infographics of IBM, with information from World Bank (2016a), IBM (2017); IDC (2012); Harvard Business Review (2016).

Table 25.2: Pulse Lab research and studies

United Nations Global Pulse partner	Project description	Insights and results
Stellenbosch University Pulse Lab Kampala (2017)	Radio content analysis, prototype speech-to-text software that converts public radio content into categorized texts	Searchable topics of interest related to SDGs and development
Office of the United Nations High Commissioner for Refugees (UNHCR) Vacarelu (2017)	Understanding forced displacement of European refugees by utilizing Twitter data	Real-time social media monitoring system relevant to humanitarian actions
World Food Programme Webb and Usher (2017)	Determining the extent of drought in Indonesia, its impact on food market prices, and the resilience of affected areas through a vulnerability monitoring platform	Real-time information platform in support of climate-impacted populations
UNHCR Hoffman (2017)	Gaining insights on the displacement patterns from Libya to Italy and Malta, and the magnitude of rescue operations using vessel data	Revealed rescue activity patterns, capacity of rescue vessels, and patterns of distress signals. Optimized rescue operations by studying migration patterns in the Mediterranean

Source: Blog posts at United Nations Global Pulse (United Nations 2018b).

In recognition of the significance of big data for official statistics, the UN Statistical Commission (UNSC) established the UN Global Working Group (GWG) on Big Data in 2014 to tap the potential of big data in monitoring the SDGs. Various collaborations, research and projects addressing the quality, collection, accessibility, management and feasibility of big data have been developed over the years. These have included task-force teams focusing on the relevance of large volumes of information coming from mobile phones, satellite imagery, social media, virtual platforms and technological applications (United Nations 2018c).

Other initiatives of the United Nations include the UN Environment Live platform **(see Table 25.1)**, which in addition to a repository related to citizen science has data from official national, regional and global statistical and geospatial data series on different thematic areas (freshwater, forests, climate etc.). Another example is the Sustainable Development Goals Interface Ontology (SDGIO), developed by UN Environment to harmonize the relationships across different SDGs through a taxonomy and semantic framework, for SDG monitoring purposes (Jensen 2017).

Data collaboratives

Governments, leading technology companies, innovators, academia, research institutions and non-governmental organizations are convening to understand the challenges around big data and to search for development solutions through collaborative learning. The UNSC Global Working Group believes that a multifaceted approach to data collection, leading to timely delivery of trusted information, can be made possible through close relationships among the private and public sectors, including civil society (United Nations 2018c). Such dialogues open doors for the co-creation of more innovation hubs, allowing capacity-building and skills transfer from countries with more experience in big data to those entering the field more recently (e.g. Vacarelu 2017).

Open data access

Access to open data is essential to harnessing big data's potential for sustainability and development. The global non-profit network Open Knowledge International and the World Wide Web Foundation's Open Data Barometer promote open data as accessible, readily available and free of charge for universal use (World Wide Web Foundation 2017). Open access to valuable and timely data from the outputs of surveys, field experiments and scientific research provides a powerful resource for presenting the state of the environment, validating our knowledge of the anthropogenic climate-change impacts, and towards proposing feasible solutions.

Box 25.1: Examples of open-data systems

National Bioscience Database Center

National database of Japan, containing global life sciences data that can be readily accessed by users.

Air Now

A global database that presents the daily nationwide forecast of the Air Quality Index (AQI) of over 400 cities, providing visual representation of the air quality condition relevant to health interpretation.

Open Data for Business Assessment and Engagement Tool

Launched by World Bank in 2016 to establish the use of open government data for industrial or business purposes.

Estonia's X Road

Online e-government system, serving as a platform of data exchange among participating institutions and private companies, with 70 per cent composed of government agencies.

Source: Estonia Digital Society (n.d.); United States Environmental Protection Agency (n.d.); Japan Science and Technology Agency (2011); Manley (2015).

Box 25.1 presents a selection of open-data initiatives at global and national levels.

Environmental assessments and evaluation

Big data analytics enable illustrations of trends and progress over time (e.g. comparing current with historical data at a specific spatial site), reaches more end users beyond geographical boundaries, and allows a predictive analysis of the future using models and comparisons with historical data. Web-based and geospatial mapping technologies, remote sensing and statistical visualization provide a basis for analysis of environmental implications and raise issues on the scalability of data collected, as reflected in the sample of assessments and evaluations in **Box 25.2**. Big data from satellite imagery and sensors make the environmental indicators measurable (Uitto 2016).

The Data-Pop Alliance describes big data as a socio-technological phenomenon, emerging from a novel data ecosystem that defines the complexity of human behaviours and beliefs – generated and captured by digital devices, computational and analytical tools, and the active correspondence of communities (Anttila-Hughes *et al.* 2015). With this paradigm shift, technology has been used to survey public insights, leveraging big data to improve the environmental process.

Box 25.2: Examples of web-based and geospatial technologies using big data

Big data use for environmental Assessments and Evaluation

- ❖ **Participatory web-based GIS tool** used in the Strategic Environmental Assessment (SEA) process, complementing the traditional public consultation by developing a user-friendly and comprehensible system in Ireland
- ❖ Reporting of factors affecting **environmental litigation through GIS-generated graphics and location** (e.g. extent of pollution and potential contaminants)
- ❖ **Geospatial relations in Environmental Epidemiology,** investigation of disease in relation to the subject's location (e.g. disease mapping, cluster analyses and geographic correlation studies). Epidemiology studies increased from 43 in 1990 to 934 in 2014, based on the PubMed publication index.

Source: Gozales et.al. (2012); Rominger and Ikeda (2015).

Box 25.3: Comprehensive air-quality forecasting in India using big data

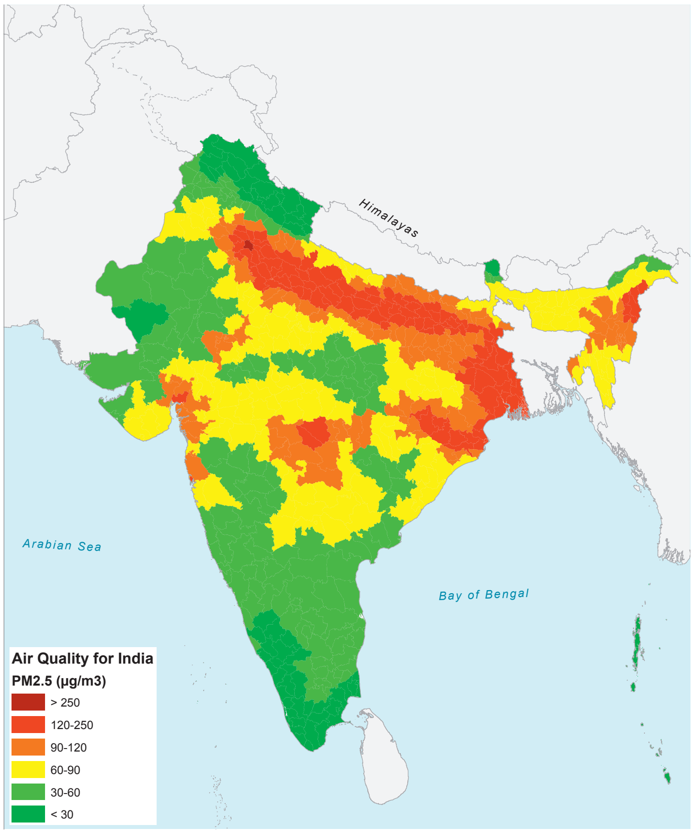
There are several hundred manuals and about a hundred regulatory air-monitoring stations in operation in India. These are limited to urban agglomerations, so second-tier towns and rural areas do not have access to any on-ground monitoring data. Before being able to manage and improve air quality effectively, citizens and policymakers need to know the status of air quality and to have information on the sources of pollution. The system, known as India air-quality forecasts, developed by *urbanemissions.info,* uses a modelling approach to predict, for the next three days, the estimated pollution levels and source contributions for all of the 640 districts in India. This is not a substitute for a robust monitoring system, but the estimates can be used to support informed decision-making while more monitoring capacity and systems are built.

While the methodology is continually improving, the key challenge in this approach is using a detailed emissions inventory and its spatial and temporal granularity. The programme currently uses information from official reports, academic publications, and survey analysis, and the following open data (dynamic feeds), which are updated every day.

- Remote-sensing satellite data (from NASA's Visible Infrared Imaging Radiometer Suite) for location of open fires, which when overlaid with land-use imagery, builds a dynamic emissions inventory for agricultural and forest fires at 1 km resolution.
- Meteorological data at 1 km resolution linked to emissions from multiple sectors. For example: (a) a surface temperature profile is used to trigger space heating in the residential sector (b) grids with precipitation over 1 mm/h are adjusted for lesser vehicle movement (c) grids with precipitation are adjusted for dust resuspension on the roads and dust at the construction sites (d) the dynamic calculations within the meteorological model estimation of likely dust storms, sea salt emissions, lightning, dry deposition rates and wet scavenging rates by grid.
- Google provides a wealth of information on traffic movement. Over cities, transit speeds are extracted at 1 km resolution, which is used as a proxy to dynamically allocate vehicle exhaust emissions and estimate road dust resuspension. For example, during peak times, if a grid shows speeds under 5 km/h, the emissions profile is adjusted to increase the exhaust emissions due to idling, and the road dust resuspension is zeroed.
- Google Earth imagery is used to generate spatial data on brick kilns, power plants, industrial zones, and mining and quarrying areas.
- Load dispatch centres across India report information on demand and supply by power grid, which is used as a proxy to dynamically adjust the use of diesel generator sets in cities.

These dynamic feeds are making air-quality forecasting more robust, allowing the model to capture trends, and they help to understand the source contributions better. Multiple microsensor networks are being tested and evaluated in several cities, promising to further improve the forecasting process and strengthen the on-ground data availability.

Figure 25.8: Forecasting air quality for Indian districts



Source: UrbanEmissions.info (2018).

Challenges in Big Data

Gaps in the collection, monitoring, analysis and interpretation of data identified in GEO-5 continue to challenge the reliability of Big Data as a tool in environmental assessment today in GEO-6. Issues include the accessibility, quality and sparsity of data for varying scales, contexts and time series (United Nations Environment Programme [UNEP] 2012). Constraints, generally encountered across the thematic areas covered in GEO-6, are summarized in **References**. These challenges on scope, privacy and potential for misinterpretation of data have not been sufficiently addressed over the years. Efforts to generate globally acceptable and available Big Data are pursued, but the actual capacities are limited by scarce resources and funding, especially in developing countries.

Although the private sector has been pioneering Big Data strategies, increasingly governments and global initiatives are exploring the benefits of Big Data for transparency, market analysis, research, education and environmental protection. In South-East Asia, six countries have formed an open-government partnership to advance their vision of scaling up public services and leveraging Big Data for development (Bhunia 2017).

Environmental agreements provide United Nations member states with guidelines on how to make environmental data publicly accessible, and provide open, geographically referenced data, together with opportunities for public participation in decision-making, multi-stakeholder involvement and the promotion of government transparency and accountability. These agreements include the Aarhus Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters (United Nations 1998) and the Kyiv *Protocol on Pollutant Release and Transfer Registers* (2003). The Data-Pop Alliance is a global coalition for people-centred data pioneered by the Harvard Humanitarian Initiative, Massachusetts Institute of Technology Media Lab, and the Overseas Development Institute. This alliance, funded by the United Nations Economic Commission for Latin America and the World Bank Group among others, creates a collaborative workspace for researchers, experts, practitioners and activists to overcome foreseen challenges on Big Data (i.e. technological biases, lack of access to an online knowledge-sharing facility, and limited technical capacity development) through research, capacity-building and community engagement (Data-Pop Alliance n.d.). A 2015 report by the alliance explores the opportunities for leveraging Big Data to monitor climate change hazards, mitigate the impacts, guide disaster response and increase the resilience of vulnerable countries (Anttila-Hughes *et al.* 2015).

Box 25.4: Some challenges of using Big Data

Big Data Challenges

Accessibility

Reluctance of governments and private sector from sharing information lead to data disintegration. Most data in the private domain are bounded by privacy and security limits-intellectual property rights, thus entailing a price for accessibility. Data sharing and availability are not only part of the legal issues, but of the political concern as well-like in China where the government place restrictions in the release of valuable environmental information.

Quality

There is no assurance in the reliability of data due to inconsistencies in the methodologies used by different countries and insufficient technical capability on data interpretation and analysis. Baseline studies are paralyzed by insufficient valuable data (e.g. biodiversity, hydrometeorological, wastewater treatment).

Sparsity

The availability of data varies across temporal and spatial scale. To fill the gaps, researchers utilize secondary information as proxy data and model estimates to fill the gaps, making global data incomparable.

Real-time Big Data are often controlled and held by the private sector (Kirkpatrick 2016). Therefore, collaborations are needed where both parties benefit without sacrificing the economic value of data, and at the same time maintain fair competition among businesses. The private sector has been providing the public sector, including research institutions and industry practitioners, with access to data through what Robert Kirkpatrick (2016) describes as data philanthropy. This collaboration has been in existence within the United Nations system. In pursuit of companies' contributions to SDGs, data scientists at the firms interpret private data for public good and well-being, which, in return, reduces the risks to business. Another form of collaboration is the public-private partnership where resources and capabilities on Big Data are shared between governments, National Statistical offices, research institutions, and the private sector, including leading technology and data companies across the globe. **Table 25.3** presents some examples of economic improvement achieved through public-private partnerships.

For Big Data to become an effective tool for environmental assessment and development, this emerging form of data and knowledge should be seen as a valuable asset. Big-data analytics involve not only compiling information but also creating a comprehensible view of the environment and its social attributes as a basis for proposing solutions and drafting policies. Factors that contribute to establishing a holistic data system include leadership and data governance, including the appointment of a chief data officer in national and local government agencies; partnerships among governments, institutions and the private sector; and institutionalizing legal frameworks with safeguards on information.

Table 25.3: Example public-private partnerships

Partnership	Project description	Source
Government of Nigeria and Cellulant	**Wide-scale mobile e-wallet system** that directly coordinates the distribution of seeds, and transfers subsidized fertilizers to farmers, thus streamlining public-service delivery.	World Bank (2016b), p. 94
Organisation for Economic Co-operation and Development and various governments	**Global earthquake model** that communicates earthquake risks through open-access to catastrophe models across the globe.	Thomas and McSharry (2015)
Willis Research Network and various partners	**Willis Research Network** provides and supports the scientific research and development of applications for universities, modelling companies, governments and non-governmental organizations	

25.2.3 Traditional knowledge

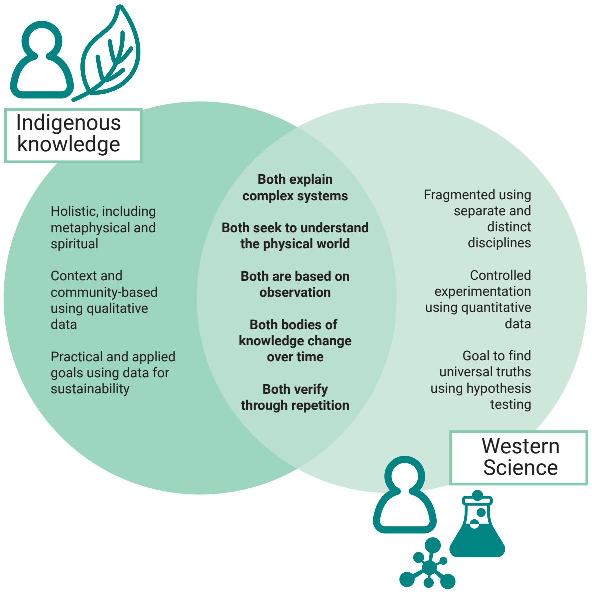
Many terms are used to describe the knowledge held by indigenous peoples and local communities. Some refer to the term traditional ecological knowledge, which is "a cumulative body of knowledge, practice, and belief, evolving by adaptive processes and handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and with their environment" (Berkes, Colding and Folke 2000, p. 1,252). Others prefer to use the terms indigenous knowledge, folk knowledge, local knowledge or traditional knowledge. There is no universally accepted definition of the diversity of expressions within this epistemic landscape; however, to include the widest understanding, this section uses the term traditional knowledge to include ecological, local, indigenous and folk knowledge.

According to Article 8j of the Convention on Biological Diversity, traditional knowledge includes cultural values, beliefs, rituals, community laws, local language and knowledge related to practical fields such as agriculture, fishing, hunting, medicine, horticulture, forestry and environmental management in general (Secretariat of the Convention on Biodiversity n.d.). The following definition, meanwhile, was published by the World Intellectual Property Organization (WIPO) in 2010 and remains the description given by the body (WIPO n.d.): "knowledge, know-how, skills and practices that are developed, sustained and passed on from generation to generation within a community, often forming part of its cultural or spiritual identity".

Perceived by some in the mainstream as superstitious and anecdotal, traditional knowledge has been historically marginalized. In the past 20 years, however, it has been acknowledged as a valuable resource for sustainable development. Its promotion and protection has been expressed in several United Nations agreements and agencies (e.g. Article 8(j) mentioned above, the United Nations University Institute for the Advanced Study of Sustainability, and WIPO). The foundational operating principles of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, established in 2012, include recognition and respect for the contribution of indigenous and local knowledge to the conservation and sustainable use of biodiversity. Moreover, the United Nations Convention to Combat Desertification (UNCCD) in its Decision 20/COP.12 adopted "improvement of knowledge dissemination including traditional knowledge, best practices, and success stories" (UNCCD 2016, p. 57).

An increasing amount of research concludes that traditional knowledge developed through direct interaction with local ecosystems is of equal value to that of Western scientific knowledge. Both knowledge systems have commonalities, but each has distinct features **(see Figure 25.9)** that complement each other to better understand the natural world (Agrawal 1995; Tsuji and Ho 2002).

Figure 25.9: Comparing indigenous/traditional knowledge and Western science



Source: Baker, Rayner and Wolowic (2011).

Traditional knowledge and Western scientific knowledge

GEO-5 called for coordination in the realm of knowledge-building, which implies that cooperation is necessary between Western scientists and the various holders of traditional knowledge. Recent progress has enabled the development of new tools and approaches in measurement, reporting and verification. Among others, these include the community-based monitoring and information systems, and the multiple evidence-base system, both of which foster knowledge co-creation between indigenous and Western systems, as an important way to help advance the recognition of the leadership role of indigenous peoples in stewarding their lands and waters (Raygorodetsky 2017). See more detail about these and other approaches in **Box 25.5** and **Table 25.4.**

These approaches combine traditional community monitoring systems with modern software (e.g. GIS, Google Maps, GPS, Microsoft Excel) and hardware (e.g. drones, remote sensing devices, trackers, smartphones, electronic tablets) to generate data and develop information on trends in ecosystems in order to assess development interventions. These modern tools and monitoring systems enable individuals and communities to decide on what actions to take by providing informed, improved and timely decision-making.

As traditional knowledge is based on *in situ* data generation, it is effectively used to give on-the-ground truth to Big Data, and like citizen science, could involve all sectors in the community. **Table 25.5** outlines new partnerships between communities and researchers that have produced innovative new approaches to the documentation and analysis of traditional knowledge.

Traditional knowledge and the SDGs

Indigenous peoples' engagement in the development of the 2030 Agenda for Sustainable Development (United Nations 2015) resulted in the inclusion of six references to

Box 25.5: Complementary uses of traditional knowledge and Western science

Community-based monitoring and information systems refer to initiatives by indigenous peoples and local communities/organizations to monitor their community's well-being and the state of their territories and natural resources, applying a mix of traditional knowledge and innovative tools and approaches (Ferrari, de Jong and Belohrad 2015).

The multiple evidence-base system generates new insights and innovations through complementarities between indigenous peoples, local communities and Western scientific knowledge systems. The system emphasizes that the evaluation of knowledge occurs primarily within, rather than across, knowledge systems (Tengö *et al.* 2014).

Indigenous peoples make and use maps to: assert their rights to lands and waters; manage their territory; preserve knowledge of their own history, culture and environment; and communicate some of this knowledge to others (Tebtebba Foundation 2015).

Successful integration of traditional knowledge with modern science, technology and innovation can be seen in the example of a recent tech start-up called Indigital. This Aboriginal-owned and operated social enterprise, based in the Kakadu World Heritage Area in the Northern Territory of Australia, uses digital technology to showcase local sacred sites, knowledge and stories in augmented and virtual realities, contributing to the preservation of heritage while creating jobs in the digital economy (Cooper and Kruglikova 2018).

Table 25.4: Studies that combine traditional knowledge with Western scientific knowledge

Study	Summary
Genome-wide association study to identify the genetic base of smallholder farmer preferences of durum wheat traits (Biodiversity International 2017; Kidane *et al.* 2017)	The study demonstrates that researchers using modern scientific tools (i.e. genetic analysis), and holders of traditional knowledge using cultural practices in crop selection can work together to advance crop breeding to cope with the changing climate.
Arbediehtu pilot project on documentation and protection of Sami traditional knowledge (Porsanger and Guttorm eds. 2011)	The study highlights community work, legal questions and legislation, ethics of documentation, institutional relationships, history and identity, information technologies, transmission, management, and legitimacy.
Traditional knowledge and nutritive value of indigenous foods in the Oraon tribal community of Jharkhand: an exploratory cross-sectional study (Ghosh-Jerath *et al.* 2015)	The study identifies more than 130 varieties of indigenous foods, many of which are rich sources of micronutrients and medicinal properties, and explains how they can be leveraged to address malnutrition in tribal communities.
Cree traditional ecological knowledge and science: a case study of the sharp-tailed grouse (Tsuji 1996)	The study shows that Cree traditional knowledge is factual and often quantitative in nature. While limitations exist in the distinction between observations and interpretations, this knowledge can be added to databases to facilitate resource co-management.
Collaborative Partnership on Forests (2018)	The project shows that the integration of traditional knowledge into forest management practices is a prerequisite for achieving sustainable forest management because it strengthens the rights and participation of indigenous peoples and local communities, and clarifies land tenure.

Figure 25.10: Recognition of indigenous peoples in the 2030 Agenda for Sustainable Development

Source: United Nations (n.d.).

indigenous peoples. Aside from these six, many other SDGs address indigenous peoples indirectly, through the principles of human rights, equality, non-discrimination, sustainability, and participation by right-holders. Sustainability, as the banner of Agenda 2030, is underpinned by environmental health.

Having evolved after years of observation and experience from a holistic relationship between people and nature, traditional knowledge sustains life and landscapes. Meanwhile, current development strategies are largely based on Western science and technology, which are often detrimental to the environment and our well-being in many ways. Therefore, there is a need to harness scientific knowledge, technology, and traditional knowledge to solve many of the issues related to sustainable natural resource management and biodiversity conservation. Some examples of studies on the potential of traditional knowledge for sustainability are presented in **Table 25.5**.

Figure 25.11: Lands/territories of indigenous peoples are the base of their knowledge

© Tebtebba Foundation 2008.

The promotion and development of traditional knowledge is hampered by contemporary bias in existing power relations. In many places, holders of traditional knowledge continue to face exploitation when trying to defend their territories against further abuse. As a result, their lands – which constitute the basis of their knowledge systems – have been subjected to resource development projects such as mono-crop plantations for rubber, timber and palm oil, large hydroelectric dams, mineral extraction activities (Asia Indigenous Peoples Pact 2015), and conservation projects (Vidal 2016).

To address historical and continuing injustices adequately, indigenous peoples and local communities, their land, and the knowledge systems that hold these together must be included in the development process. As expressed in the United Nations

Table 25.5: Studies on the potential of traditional knowledge for sustainable development

Study	Summary
Local biodiversity outlooks: indigenous peoples' and local communities' contributions to the implementation of the strategic plan for biodiversity 2011-2020 (Forest Peoples Programme 2016).	This study presents snapshots of on-the-ground initiatives by indigenous peoples and local communities and demonstrates that they are making vital contributions to the implementation of the five strategic goals and the 20 Aichi biodiversity targets, although many challenges remain. It outlines the way forward, highlighting key potential actions to accelerate progress in the implementation of the strategic plan for biodiversity as it relates to indigenous peoples and local communities.
Sustaining and enhancing forests through traditional resource management (Enchaw and Njobdi 2014)	This study highlights women's roles in forest management and sustainable farming within the context of indigenous knowledge.
United Nations Educational Scientific and Cultural Organization (UNESCO) World Heritage: Agricultural Landscape (UNESCO 2013)	This study deals with communities that demonstrate rich cultural and landscape diversity and sustainable land-use systems. It also highlights how some people struggle for daily survival under extreme climatic and environmental conditions.

Declaration on the Rights of Indigenous Peoples (United Nations, General Assembly 2007), these peoples have the right to self-determination, as well as to their land, resources and the freedom to pursue their own way of life. These rights are necessary for peoples to sustain, innovate and develop traditional knowledge systems and customary practices. Granting and implementing these rights is crucial for ensuring the maintenance of a balance between the economic, social and environmental dimensions of sustainable development.

Traditional knowledge for a Healthy Planet, Healthy People

Traditional knowledge is an invaluable resource for sustaining a healthy population and planet. Indigenous territories constitute up to 22 per cent of the world's land surface and sustain 80 per cent of the planet's biodiversity (Food and Agriculture Organization of the United Nations [FAO] 2017). The vast majority of the world's genetic resources and a considerable part of global biodiversity survives within indigenous and community-conserved areas. This correlation is not coincidental but is due to the application of traditional knowledge and the customary sustainable use of biological resources over centuries (Independent Expert Advisory Group Secretariat 2014). Moreover, the indigenous ethics and values, plus culture and identity, related to land and wildlife stewardship, hold great promise for more effective resource management as well as for more effective risk reduction in human health (Houde 2007). **Figure 25.12** is just one illustration of the wealth and potential of traditional knowledge and practices for environmental management that need to be optimized in their contribution to improving the health of both land and people.

Further examples of benefit include the many advantages of maintaining a well-functioning food web – including the enhanced diversification of wild and cultivated food systems, enhanced nutrition, and a healthy environment (Kuhnlein *et al.* eds. 2013).

However, traditional knowledge remains underutilized in environmental assessment and management. While advances have been made, there are still challenges that need to be addressed. As cited by Genetic Resources Action International [GRAIN] and Kalpavriksh (2002), these include:

i. the continued loss of indigenous peoples' lands, making it challenging for indigenous communities to sustain their knowledge;
ii. the risk of misappropriating traditional knowledge or patenting life forms for commercial purposes without sharing the benefits with knowledge holders, as demonstrated by the case of turmeric in India;
iii. the proliferation and all-out promotion of so-called modern medicine and agriculture, which replace diverse plants and crops that would have been able to resist pests, diseases, and changing climatic and economic conditions (GRAIN and Kalpavriksh 2002);
iv. the co-production of knowledge processes that do not always guarantee fairness, equal standing, or power symmetries (Williams and Hardison 2013); and
v. the rapid erosion of linguistic diversity is accompanied by the loss of indigenous ways of knowing and understanding of the natural world (UNESCO 2017).

Other challenges relate to the full and effective implementation of free prior and informed consent (United Nations, General Assembly 2007). To address these challenges in an efficient way, policies need to be formulated and implemented on the basis of "further interdisciplinary action research that brings together indigenous knowledge holders and scientists, both natural and social, to build mutual understanding and reinforce dialogue" (Nakashima *et al.* 2012, p. 97).

25.3 Environmental monitoring for the future

25.3.1 Measuring what matters

There is a maxim that what gets measured gets done. Its origin is debatable, but the message is clear: measuring something gives us the information we need to make sure we actually achieve what we set out to do.

As noted in Section 3.3.1, the data requirements for the SDG indicators are almost as unprecedented as the SDGs themselves, and constitute a tremendous challenge to all countries. Unfortunately, much of the data required to monitor the SDGs are unavailable. Issues relating to quality, timeliness, human and financial capacity, and the lack of standardized methodologies all hamper our ability to comprehensively track this important agenda. As highlighted in the United Nations (2016) report on the SDGs, tracking their progress will require a shift in how data are collected, processed, analysed and disseminated, including a move to using data from new, diverse and innovative sources. More than ever, this demands that

Figure 25.12: Indigenous peoples as stewards of the environment

Source: FAO (2016).

we explore other data streams – citizen science, Big Data and traditional knowledge – to complement conventional, official statistics and Earth observations.

25.3.2 Translating local information into national data

A shift in data generation by electronic devices, data modelling, cloud computing and other technologies has produced unprecedented volumes of information. The availability of these data and information varies, however, between developed and less developed countries, between and within social groups – by gender, ethnicity, and social and income status – and markedly between the global and local levels. Data and information gathering at national and local levels is given lower priority, especially in developing countries. But to solve global problems, actions for solutions must emanate from the local and national level. Policy effectiveness should be measured according to its implementation and fulfilment at the local to national level (see Part B).

Over time, there has been an accumulation of knowledge on good and bad practices of environmental management, enabling some conclusions as to what needs to be done. However, there remains a need for making traditional knowledge more accessible and to integrate traditional knowledge with other sources of information.

On the global level, states and civil society organizations, academia, indigenous peoples and activists all agree that for the promotion of environmental protection and restoration, biological loss and climate change need to be addressed. Yet at the country level, positions about environmental protection vary, and often even contradict each other.

The SDGs aim to reduce environmental degradation while at the same time upholding basic human rights and promoting economic empowerment. This will require knowledge and technology to be shared between communities, businesses and governments – from local, through national, to global levels.

25.3.3 Open data and reproducible research

The open-data movement has gained significant traction in recent years and is expected to continue to grow. The concept of open data is that data resulting from publicly funded research should be freely available to all, for equity, transparency, and to catalyse the advancement of science. The principles of open data have been proposed in a number of different variations, including the Open Data Charter (2015), which stipulates that data should be:

i. open by default;
ii. timely and comprehensive;
iii. accessible and usable;
iv. comparable and interoperable;
v. for improved governance and citizen engagement; and
vi. for inclusive development and innovation.

There are clearly demonstrated benefits of open data, particularly in the health sector (Kostkova *et al.* 2016). The principles of open science, open innovation, open access and open source adopted by the malaria research community, for example, have allowed it to achieve more progress than would otherwise be possible (Wells *et al.* 2016). Additionally, many countries now have open government data portals that include environmental data.

There is a growing call for reproducible research alongside open data, and the two are often considered in tandem by the open-science movement. To be reproducible, research should be reported in a manner that allows it to be replicated precisely (Mesirov 2010). There are three aspects of research reproducibility – it should be:

i. empirical – based on scientific experiments and observations;
ii. computational – code, software, hardware and implementation details are made available; and
iii. statistical – based on statistical tests and model parameters (Stodden 2014).

Computational and statistical reproducibility are most pertinent to data practices. In theory, the publication of code and data together means that users can understand and critique the entire process of analysis and inference, including details of the techniques used and any assumptions made. Publication of code is now an essential prerequisite of many scientific journals, and this trend is increasing such that it is expected to become the norm in coming years.

Reproducible data analyses are especially important in an era of open data, since users of data will become increasingly detached from those who collect and curate the data. Open-access data increases the risk of data being misused or misinterpreted, but the publication of code circumvents this problem in that the treatment of data is transparent and can be scrutinized by readers. Reproducibility not only improves the quality of scientific output, but also increases trust in the results, and therefore uptake (Laine *et al.* 2007).

25.3.4 Coping with the changing data landscape

Strengthening the ability to gather, interpret and use data for effective planning, policymaking, management and evaluation is necessary for providing countries with a comprehensive view of environmental impacts, ranging from geo-political perspectives through to industrial operations, naturally occurring or anthropogenic environmental change or a combination of all these. The challenge escalates as the magnitude and types of generated data, both structured and unstructured, grow over time. Management information systems alone are insufficient to draw the full value from the exponential growth of potential data assets. It is imperative that governments and society learn to cope with the evolving data landscape, to shift from mere reporting and conventional data repository functions towards a predictive and prescriptive analysis for both modelling different environmental scenarios and creating appropriate policies to address these foreseen challenges.

The information landscape is changing as the technologies for harnessing data evolve from data-as-a-service (DaaS) or software-as-a-service (SaaS) to insights-as-a-service (IaaS), which uses prescriptive analytics **(Figure 25.13)**. With DaaS, data could reach several users beyond geographical limits and organizational segmentation, bringing together data in a central repository (Olson 2010). DaaS and SaaS have the capacity to present both historical and current states, reporting what happened, why it happened, and what is happening now with regard to both environment and society. From cleaning and consolidating Big Data, analytics-as-a-service (AaaS) shifts the demand from internal manual services to web-delivered technologies, outsourcing the needs from the Internet of things or IoT (Atos 2013) and providing virtual services. AaaS accompanies users throughout their experience, imparting knowledge through artificial intelligence to find expertise and support throughout the entire journey (Takahashi 2017). AaaS and knowledge-as-a-service (KaaS) apply predictive analytics and modelling to interpret Big Data from multiple sources and project the future. Tapping the potential of artificial intelligence could develop strategies not only to foresee future environmental and social challenges, but also to advance solutions by predicting the outcomes of the countries' efforts, responding to the question, how do we make it happen?

The exponential growth of Big Data, technological solutions, complex algorithms and open data sources propels and integrates artificial intelligence into our everyday lives, cities and world networks each year (Herweijer 2018). Artificial intelligence is a powerful tool for countries to navigate in managing environmental concerns and furthering the SDGs, but risks around privacy, biases, declining human intervention and autonomy have to be considered. A road map is needed for how artificial intelligence could transform traditional systems and add value in delivering services, combating the impacts of climate change, building sustainable and liveable cities, and protecting environmental and social welfare. As the scale of economic and health impacts broadens as a result of environmental degradation, strategic measures have to be developed to establish not only human-friendly but also Earth-friendly artificial intelligence.

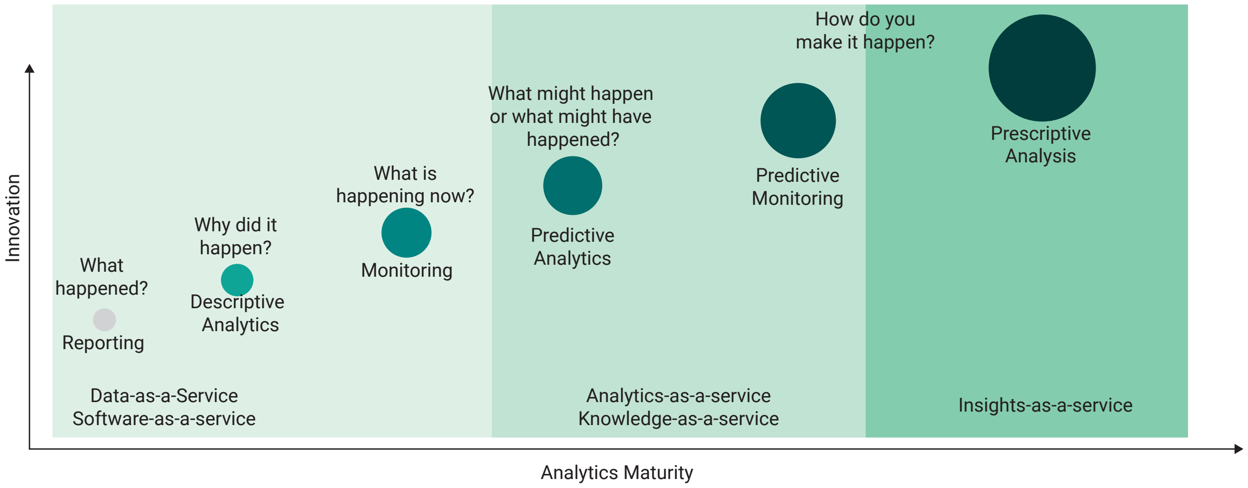
The new forms of data and knowledge, coupled with conventional tools, will dramatically influence the way solutions are created and delivered. Coping with the shift in the data landscape will need new information technology skills and a holistic approach to utilize emerging and existing data and knowledge tools – thus making data on healthy people and a healthy planet more accessible for environmental assessment and other purposes.

25.3.5 Crucial assets of technology

The collection of data in statistical operations follows well-founded methodological approaches, such as sample survey methods and designs. These include well-identified and defined sources, on which systematic methods are applied to transform the data into statistics that lend themselves to time series – the fundamental value of any statistic. New technologies such as remote sensing, transactional data, block chain and artificial intelligence algorithms have the potential to create a wealth of information that is useful for environmental purposes. However, there is a challenge to harness these new technologies to produce time series data, to achieve real-time monitoring and to bring this information into the scope of official statistics.

Only with time series do statistics transform into information, and thereby a fundamental knowledge system. The roles and responsibilities of those who count and those who are counted are also clear cut. The governance of official statistics is predicated on a set of fundamental principles laid down by the United Nations (2014). There are ten principles, including covering the fact that statistical practice should be impartial, should ensure the protection of the privacy of individuals, should be transparent, and should work to ensure the quality of the information produced. Technology has crucial assets which are founded in their development in compliance with standards. Beyond compliance is replication and scaling up. In a way, statistical operations and technological advancement can speed up optimization and give value to visibility, application and transformation of, and from, systems of data, statistics, and can give information to verifiable knowledge systems.

Figure 25.13: The evolution of the data landscape



Source: Chartered Property Causality Underliner [CPCU] (2015); ASI Business Solutions (2016).

But are these fundamental principles still relevant in the changed data environment? Maintaining the principles of official statistics involves many challenges in the new data landscape. These include definitional considerations such as concerning the selection of statistical units, data items and their associated spatio-temporal nature. The major paradigm leap for official statistics is how established methodologies of data collection and statistical time series can be adapted in the context of a deluge of unstructured, private (and potentially anonymized) data recorded on electronic devices. Phones, computers and other devices lend themselves readily to standardization and the possibility of replication and scale. For statistics to function optimally, though, standards are needed. Statisticians, technology gurus and data purveyors will need to navigate this space carefully as the hard and laborious slog of carefully designed statistical operations faces a real existential challenge of easy-to-use and fashionable observational tools. Statistics – and its essence, time series – can be greatly challenged in this new environment. The truth is, in the absence of time series, all data can be rendered useless. If a zealous adaptation of technology results in the loss of time series data, it will become impossible to track trends over time. The intersection of technology, data, statistics, knowledge, finance and governance needs to be found.

25.3.6 Data assurance and quality practices

With the increasing use of complementary data alongside traditional statistics to support environmental and sustainability policy, questions of how data quality, pedigree and provenance can be assured will need to be systematically answered to determine data that are credible and fit for purpose. Environmental data may come to include digital sources, incorporating Earth observations, citizen science, environmental monitoring, development data and statistics, administrative data sets, and population- and survey-derived data.

In addition to the fundamental principles of statistical practice mentioned above, references and standard practice documents are also emerging for complementary data sources. For example, metadata standards and practices now serve as a basis for data description, including methodological description and data quality. Metadata – the documentation of data – serves the purpose of making data discoverable, usable and understandable. Many discipline-specific or community-specific metadata standards have been developed to support systems of data management and data discovery, and to capture and convey information to users. Examples include directory interchange format (DIF), ecological metadata language (EML), sensor model language (SensorML), climate science modelling language (CSML), and netCDF markup language (NcML). Additionally, the International Organization for Standardization (ISO) has developed a series of standards to describe geographical information – ISO 19115 and ISO 11179.

Data-quality assurance for citizen-science data is in its formative stages, engaging a variety of digital-platform techniques for quality checking field observations contributed by citizens. Examples include the Local Environmental Observer (LEO) Network in the North American Arctic region, using a smartphone app that uploads observations for expert checks before being used in graphic displays on maps and in tabular data sets (LEO Network 2017). Recent workshops convened by the European Commission's Joint Research Centre are also beginning the process of establishing principles for mobile apps and platforms.

The Group on Earth Observations has propagated its Global Earth Observation System of Systems (GEOSS) data management principles, which are being widely adopted by Earth Observation entities (Group on Earth Observations 2014). Included in the coverage of the principles are: discoverability, accessibility, usability, preservation and curation.

As open-source analytics, community-sourced query codes and custom data-integration methods advance, there will be more community curation of data sets, exchange standards and application-programming interfaces. Code sets incorporating analytics and queries are now routinely community-curated on open collaboration platforms such as the GitHub development platform.

The Research Data Alliance sets registry standards for long-term curation and for defining the parameters of data sets in Earth science and other research domains. Complementary data sources will be judged to be credible as far as they conform with these registry standards.

The ultimate test for the data revolution of open and accessible digital data will be user satisfaction and integrated platform requirements for aligning with an array of recognized standards, practices and open source community-driven testing for methodology and data quality from across environmental, natural resource and development data.

25.4 Conclusion: Challenges, gaps and opportunities

The challenges, gaps and opportunities related to environmental data and statistics are presented below. Data and knowledge are valuable assets that need to be shared.

25.4.1 Data disaggregation

The SDGs call for a data revolution that leaves no one behind, incorporating disaggregated data and reporting at all levels of the 17 goals. As highlighted in Chapter 3 of this report, assessing the nexus between society and the environment can be done only if there is disaggregated information on different populations because not all people have the same level of dependence on the environment, nor the same impact on it. To tease out these differences, then, there is a need for information that can be disaggregated by income, gender, age, ethnicity, migratory status, disability, geographical location and other characteristics relevant in national contexts. Unfortunately, there is currently a dearth of environment-related information that can be disaggregated, and data from household surveys on access to water, energy and other natural resources is available only at the household level, which makes understanding differences at other levels difficult.

In addition to disaggregation by socioeconomic variables there is also a need for geospatial disaggregation of environmental information. Biological ecosystems do not follow national boundaries, so to understand both the state of particular ecosystems and the interactions between them and people and the economy requires spatially disaggregated information. Disaggregation will require bringing together Earth observation data, data from traditional censuses and surveys, as well as citizen science, traditional knowledge and Big Data into an integrated data ecosystem.

25.4.2 Open data access

Many developing countries will need access to unrestricted open data. In terms of data governance, the overall management of the availability, usability, integrity and security of the data should be made transparent. The continental leaders of vast, validated data sets – North America and Europe – are currently the data stewards making their valuable resources readily available to the rest of the world as public goods. What may be needed is a sound data-governance programme that includes a governing body or council and a defined set of procedures, with a plan for their execution. Governments should ensure that there are legal frameworks in place to promote good data governance. A documented, transparent data-governance policy will establish a set of easy-to-follow guidelines for ensuring the proper management of information, helping to answer questions revolving around sovereignty, security, data quality and privacy. In particular, protecting the privacy of individual people is an essential component of data governance. To help shepherd this governance programme, multilateral data stewards, responsible for the management and fitness of data elements, should be assigned to ensure that data exchanges are executed precisely and consistently between computer systems and between different collection methods (see 'Institutional arrangements' below).

25.4.3 Data and information governance

Governance of data involves managing and leveraging all data assets (dispersed among large external open data sources, central databases, and other existing and emerging data and knowledge sources from governments, institutions, research, and studies). Information technology alone is insufficient to draw full value from the exponential growth of potential data assets. Policy makers generally lack an understanding of the distinct issues that information technology has to deal with, and the need for active organizational involvement and institutional ownership. Addressing the challenges posed by Big Data means the integration of valuable information through a combined process of 'data semantics' and the intervention of a data-governance specialist.

Dealing with Big Data involves the storage of all data resources, sorting them out and identifying which data make sense beyond local use, and accessing the disparate data and finding their relationships. Having the right tools is not enough to sift through Big Data and find any relevant patterns and relationships. Like any valuable asset, information needs to be managed and secured through careful practices of designing, creating, developing and deploying data. With proper governance, the meaning, use and integrity of the data are preserved over time, increasing the value of data as an asset.

25.4.4 Public-private partnerships

The challenges of scarce resources and lack of financial capability in progressing the use of big-data analytics, particularly in developing countries, could be helped through public-private partnerships. The public sector could exercise the social implementation of big-data analytics as private entities finance and build on the technical know-how (Groff 2017). Public-private projects provide opportunities to strengthen statistical and information systems, develop innovative data-collection methods and initiate co-creation through people-first partnerships. Partnerships between the government and the private sector that involve civil society may result in the adoption of best practices in data collection, analytical assessment and monitoring for SDG-related information.

25.4.5 Institutional arrangements

The institutional arrangements between different authorities and agencies producing and holding environmental information are in their early stages (e.g. official statistics, map-based information on natural resources, hydrometeorological, geospatial monitoring, environmental and geospatial portals, open-data governmental portals, etc.). Any future measures should aim to improve institutional arrangements to facilitate the effective integration, sharing and reuse of data across barriers, including those of sovereignty, mandates, bureaucracy, knowledge gaps, standards and standardization, digitization and digitalization. These challenges need to be resolved. Open data and instruments that facilitate its use, such as the Statistical Data and Metadata eXchange (SDMX), hold promise for seamless data sharing. However, many countries face challenges in terms of implementing open-data instruments. National governments need to commit to common global goods, and intergovernmental arrangements are necessary to facilitate solutions to some of the political problems arising from government limitations.

25.4.6 Capacity-building

Sustainable capacity is achieved when the entire value chain works – where data from collation and analysis through to use are dealt with competently to deliver the value of the information, disseminate it through the most effective and efficient channels, and ensure that those for whom its value is intended are able to use it. To reach this stage of competency, environmental education should be implemented at all levels. Indigenous knowledge systems, including emerging opportunities offered by technological advances to democratize the participation of citizens in the exercise of science and scientific discovery should be facilitated.

25.4.7 Traditional knowledge

The challenge around the continued use and development of traditional knowledge is the lack of effective implementation of rights to indigenous knowledge, resulting in the loss of lands and the misappropriation of traditional knowledge. The inclusion of traditional knowledge does not guarantee the equality of benefit sharing.

The opportunities for traditional knowledge to contribute to sustainable development include the improved effort in documenting traditional knowledge and the increasing network building partnerships to develop innovative tools and methodologies for a better understanding of traditional knowledge to be used in strategic plans and interventions and to assess changes in ecosystems. The commitment of UN member states for sustainable development also matters when discussing traditional knowledge.

25.4.8 Integrated data systems

The physical and social systems, hand in hand, comprise the real state of the environment. Without both, the picture of where we are now and where we are heading would be incomplete. A combination of conventional methods with emerging forms of data and knowledge, and with the social data, could provide a holistic view of the environment, encompassing social, physical and economic perspectives.

Societies often notice change before measurable environmental change occurs – but, we also need to look into how data, information and knowledge are used and understood. In the era of information technology, the problem may not be the lack of knowledge or data on what is happening, but rather on the question of which available data should be acted on, what kind of information, based on which assumptions, is considered legitimate, and which reality counts (Davies 1994).

A shift in thought processes on environmental assessments would bring data, including from Big Data, citizen science and traditional knowledge to the forefront with Earth observation and official statistics. This would allow the environmental challenges to be identified, the progress in addressing the SDGs to be monitored, and the solutions to be drafted for sound and evidence-based policy-making. This would require not only turning the challenges of environmental data into opportunities, but also improving data integration, utilizing both existing and emerging tools for environmental assessment, and making data openly available for use.

References

Agrawal, A. (1995). Dismantling the divide between indigenous and scientific knowledge. *Development and Change* 26(3), 413-439. https://doi.org/10.1111/j.1467-7660.1995.tb00560.x

Anttila-Hughes, J., Dumas, M., Jones, L., Pestre, G., Qiu, Y., Levy, M. *et al.* (2015). *Big Data for Climate Change and Disaster Resilience: Realising the Benefits for Developing Countries.* Data-Pop Alliance. http://datapopalliance.org/wp-content/uploads/2015/11/Big-Data-for-Resilience-2015-Report.pdf.

ASI Business Solutions Inc (2016). *Evolution of analytics: Where does your company stand?* http://asi-solutions.com/2016/12/evolution-of-analytics-where-does-your-company-stand/ (Accessed: 06 February 2018).

Asia Indigenous Peoples Pact (2015). *Recognition of Indigenous Peoples' Customary Land Rights in Asia. Asia Indigenous Peoples Pact.* Chiang Mai: Asia Indigenous Peoples Pact. https://aippnet.org/wp-content/uploads/2015/07/iva.aippnet.org_wp-content_uploads_2015_07_CLR-AIPP-Corrected-2.pdf.

Atos (2013). *Data Analytics as a Service: Unleashing the Power of Cloud and Big Data.* Atos. https://atos.net/wp-content/uploads/2017/10/01032013-AscentWhitePaper-DataAnalyticsAsAService.pdf.

Baker, J., Rayner, A. and Wolowic, J. (2011). *Native Science: A Primer for Science Teachers.* https://ctabobandung.files.wordpress.com/2011/11/ns-primer.pdf.

Berkes, F., Colding, J. and Folke, C. (2000). Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications* 10(5), 1251-1262. https://doi.org/10.1890/1051-0761(2000)010[1251:ROTEKA]2.0.CO;2.

Bhunia, P. (2017). *Brief look at open government data in 6 ASEAN countries*. [OpenGov https://www.opengovasia.com/brief-look-at-open-government-data-in-6-asean-countries/ (Accessed: 02 January 2018).

Biodiversity International (2017). *Modern science meets traditional knowledge to improve crop breeding*. https://www.bioversityinternational.org/news/detail/modern-science-meets-traditional-knowledge-to-improve-crop-breeding/ (Accessed: 11 October 2018).

Blaney, R.J.P., Philippe, A.C.V., Pocock, M.J.O. and Jones, G.D. (2016). *Citizen science and Environmental Monitoring: Towards a Methodology for Evaluating Opportunities, Costs and Benefits.* Wiltshire: UK Environmental Observation Framework. http://www.ukeof.org.uk/resources/citizen-science-resources/Costbenefitcitizenscience.pdf.

Bowser, A., Brenton, P., Stevenson, R., Newman, G., Schade, S., Bastin, L. *et al.* (2017). *Citizen science Association Data and Metadata Working Group: Report from CSA 2017 and Future Outlook.* Washington, DC.: Woodrow Wilson International Center for Scholars. https://www.wilsoncenter.org/sites/default/files/wilson_171204_meta_data_f2.pdf.

Chartered Property Casualty Underliner (2015). *Big Data and Analytics Webinar.* 7 April. https://www.slideshare.net/PatriciaSaporito/cpcu-big-data-analytics-webinar-april-7-2015sedslideshare.

Collaborative Partnership on Forests (2018). *About the collaborative partnership on forests*. http://www.cpfweb.org/73947/en/ (Accessed: 11 October 2018).

Conrad, C.C. and Hilchey, K.G. (2011). A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environmental Monitoring and Assessment* 176(1-4), 273-291. https://doi.org/10.1007/s10661-010-1582-5.

Cooper, D. and Kruglikova, N. (2018). *Augmented Realities: The Digital Economy of Indigenous Knowledge.* International Labour Organization.

Data Pop Alliance (n.d.). *Vision and members*. Data Pop Alliance. http://datapopalliance.org/about/vision-and-members/ (Accessed: 02 July 2018).

Davies, S. (1994). Introduction: Information, Knowledge and Power. *IDS Bulletin* 25(2), 1-13. https://doi.org/10.1111/j.1759-5436.1994.mp25002001.x.

Dell EMC and International Data Corporation (2014). *IDC study: Digital universe in 2020*. https://www.emc.com/leadership/digital-universe/index.htm (Accessed: 05 May 2018).

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J. *et al.* (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology the Environment* 10(6), 291-297. https://doi.org/10.1890/110236.

Dickinson, J.L., Zuckerberg, B. and Bonter, D.N. (2010). Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics* 41, 149-172. https://doi.org/10.1146/annurev-ecolsys-102209-144636.

Enchaw, G.B. and Njobdi, I. (2014). *Sustaining and Enhancing Forests through Traditional Resource Management.* Tebtebba Foundation. http://www.tebtebba.org/index.php/content/276-sustaining-a-enhancing-forests-through-traditional-resource-management-volume-2.

Estonia Digital Society (n.d.). *x-road*. https://e-estonia.com/solutions/interoperability-services/x-road/ (Accessed: 20 November 2018).

Ferrari, M., de Jong, C. and Belohrad, V.S. (2015). Community-based monitoring and information systems (CBMIS) in the context of the Convention on Biological Diversity (CBD). *Biodiversity* 16(2-3), 57-67. https://doi.org/10.1080/14888386.2015.1074111.

Food and Agriculture Organization of the United Nations (2016). *Indigenous Peoples can feed the world*. Rome. http://www.fao.org/3/a-c0386e.pdf

Food and Agriculture Organization of the United Nations (2017). *6 ways indigenous peoples are helping the world achieve #ZeroHunger*. http://www.fao.org/indigenous-peoples/news-article/en/c/1029002 (Accessed: 11 October 2018).

Forest Peoples' Programme (2016). *Local Biodiversity Outlooks: Indigenous Peoples and Local Communities Contributions to the Implementation of the Strategic Plan for Biodiversity 2011-2010. A complement to the Fourth Edition of the Global Biodiversity Outlook.* Development Southern Africa. Moreton-in-Marsh. https://www.cbd.int/gbo/gbo4/publication/lbo-en.pdf.

Garcia-Soto, C., van der Meeren, G.I., Busch, J.A., Delany, J., Domegan, C., Dubsky, K. *et al.* (2017). *Advancing Citizen Science for Coastal and Ocean Research: Position Paper 23.* French, V., Kellett, P., Delany, J. and McDonough, N. (eds.). Ostend: European Marine Board. http://www.marineboard.eu/sites/marineboard.eu/files/public/publication/EMB_PP23__Citizen_Science_web_4.pdf.

Genetic Resources Action International and Kalpavriksh (2002). *Traditional Knowledge of Biodiversity in Asia-Pacific: Problems of Piracy and Protection.* New Delhi. https://www.grain.org/article/entries/81-traditional-knowledge-of-biodiversity-in-asis-pacific-problems-of-piracy-and-protection.pdf.

Ghosh-Jerath, S., Singh, A., Kamboj, P., Goldberg, G. and Magsumbol, M.S. (2015). Traditional knowledge and nutritive value of indigenous foods in the oraon tribal community of jharkhand: An exploratory cross-sectional study. *Ecology of Food and Nutrition* 54(5), 493-519. https://doi.org/10.1080/03670244.2015.1017758.

Global Learning and Observations to Benefit the Environment (2018). *About GLOBE program*. https://www.globe.gov/about/overview (Accessed: 20 July 2018).

Gonzales, A., Gilmer, A., Foley, R., Sweeney, J. and Fry, J. (2007). Developing and applying a user-friendly web-based GIS for participative environmental assessment. *Proceedings of the Geographical Information Science Research UK Conference*. Maynooth, 11-13 April 2007. http://eprints.maynoothuniversity.ie/1428/1/FoleyDevelopingGISRUK.pdf

Google (2018). *Google reCAPTCHA: The new way to stop bots*. https://www.google.com/recaptcha/intro/v3beta.html# (Accessed: 26 May 2018).

Groff, S.P. (2017). 'Better data is the key for developing Asia to successfully implement the SDGs by 2030'. https://blogs.adb.org/blog/data-matters-meet-sdgs (Accessed: 30 June 2017)

Group on Earth Observations (2014). *The GEOSS data sharing principles post-2015.* http://www.earthobservations.org/documents/dswg/10_GEOSS%20Data%20Sharing%20Principles%20post%202015.pdf (Accessed: 11 October 2018).

Haklay, M. (2013). Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing Geographic Knowledge.* Berlin: Springer. 105-122.

Harvard Business Review (2016). *The Explainer: Big Data and Analytics*. https://hbr.org/video/3633937151001/the-explainer-big-data-and-analytics (Accessed: 22 January 2018).

Herweijer, C. (2018). *8 Ways AI can help save the planet*. [World Economic Forum https://www.weforum.org/agenda/2018/01/8-ways-ai-can-help-save-the-planet/ (Accessed: 21 February 2018).

Hochachka, W.M., Fink, D., Hutchinson, R.A., Sheldon, D., Wong, W.-K. and Kelling, S. (2012). Data-intensive science applied to broad-scale citizen science. *Trends in Ecology Evolution* 27(2), 130-137. https://doi.org/10.1016/j.tree.2011.11.006.

Hoffmann, K. (2017). Using vessel data to study rescue patterns in the Mediterranean Sea. 20 September 2017 https://www.unglobalpulse.org/news/using-vessel-data-study-rescue-patterns-mediterranean-sea

Houde, N. (2007). The six faces of traditional ecological knowledge: challenges and opportunities for Canadian co-management arrangements. *Ecology Society* 12(2). https://doi.org/10.5751/ES-02270-120234.

Independent Expert Advisory Secretariat (2014). *A World that Counts Mobilising the Data Revolution for Sustainable Development.* New York, NY: United Nations. http://wedocs.unep.org/bitstream/handle/20.500.11822/20065/ieag_world.pdf?sequence=1&isAllowed=y.

International Business Machines (2017). *Extracting business value from the 4 V's of big data.* International Business Machines Corporation http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4Vs_Infographic_final.pdf

International Data Corporation (2012). *Worldwide Big Data Technology and Services 2012-2015 Forecast.* International Data Corporation. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=6242.

International Rivers (2012). Citizen science Supports a Healthy Mekong. *World Rivers Review* 4. https://www.internationalrivers.org/sites/default/files/attached-files/wwrdec_web.pdf.

Japan Science and Technology Agency (2011). *National bioscience database center*. https://biosciencedbc.jp/en/ (Accessed: 20 November 2018).

Jensen, M. (2017). *UNEP Sustainable Development Goals Initiatives Ontology.* http://ncgia.buffalo.edu/OntologyConference/Abstract/Jensen.pdf.

Kidane, Y.G., Mancini, C., Mengistu, D.K., Frascaroli, E., Fadda, C., Pè, M.E. *et al.* (2017). Genome wide association study to identify the genetic base of smallholder farmer preferences of durum wheat traits. *Frontiers in Plant Science* 8(1230). https://doi.org/10.3389/fpls.2017.01230.

Kim, S., Robson, C., Zimmerman, T., Pierce, J. and Haber, E.M. (2011). Creek watch: Pairing usefulness and usability for successful citizen science. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Vancouver, 7-12 May 2011. ACM 2125-2134 doi: ttps://doi.org/10.1145/1978942.1979251

Kirkpatrick, R. (2016). 'The importance of big data partnerships for sustainable development'. 31 May 2016 https://www.unglobalpulse.org/big-data-partnerships-for-sustainable-development

Kostkova, P., Brewer, H., de Lusignan, S., Fottrell, E., Goldacre, B., Hart, G. *et al.* (2016). Who owns the data? Open data for healthcare. *Frontiers in Public Health* 4(7). https://doi.org/10.3389/fpubh.2016.00007.

Kuhnlein, H.V., Erasmus, B., Spigelski, D. and Burlingame, B. (eds.) (2013). *Indigenous Peoples' Food Systems and Well-being: Interventions and Policies for Healthy Communities*. Rome: Food and Agriculture Organization of the United Nations. http://www.fao.org/docrep/018/i3144e/i3144e.pdf.

Laine, C., Goodman, S.N., Griswold, M.E. and Sox, H.C. (2007). Reproducible research: Moving toward research the public can really trust. *Annals of Internal Medicine* 146(6), 450-453. https://doi.org/10.7326/0003-4819-146-6-200703200-00154.

Local Environmental Observer Network (2017). *LEO reporter mobile app*. https://www.leonetwork.org/en/docs/about/mobile (Accessed: 11 October 2018).

Macauley M. and Brennan T. (2016). Data from drones: A new way to see the natural world. *Resources for the Future*, 192. 25 May 2016. http://www.rff.org/files/document/file/RFF-Resources-192_DataFromDrones.pdf.

Manley, L. (2015). 'Open data for business tool: Learning from initial pilots'. 24 August 2015 http://blogs.worldbank.org/ic4d/open-data-business-tool-learning-initial-pilots

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. *et al.* (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity.* McKinsey Global Institute. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx.

Mathieu, P., Borgeaud, M., Desnos, Y., Rast, M. and Brockmann, C. (2016). *ESAEarth Observation Open Science: Accelerating Data-Intensive Research in the Digital Age.* IEEE Magazine.

Mazumdar, S., Wrigley, S. and Ciravegna, F. (2017). Citizen science and crowdsourcing for earth observations: An analysis of stakeholder opinions on the present and future. *Remote Sensing* 9(1), 87. https://doi.org/10.3390/rs9010087.

McCallum, I., See, L., Sturn, T., Salk, C., Perger, C., Duerauer, M. *et al.* (2018). Engaging citizens in environmental monitoring via gaming. *International Journal of Spatial Data Infrastructures Research* 13, 15-23 https://doi.org/10.2902/1725-0463.2018.13.art3.

Mesirov, J.P. (2010). Accessible reproducible research. *Science* 327(5964), 415–416. https://doi.org/10.1126/science.1179653.

Monnappa, A. (2017). *Data science vs. big data vs. data analytics*. https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article (Accessed: 27 April 2018).

Nakashima, D., McLean, K.G., Thulstrup, H.D., Castillo, A.R. and Rubis, J.T. (2012). *Weathering Uncertainty: Traditional Knowledge for Climate Change Assessment and Adaptation.* Paris: United Nations Educational, Scientific, and Cultural Organization and United Nations University. http://unesdoc.unesco.org/images/0021/002166/216613e.pdf.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S. and Crowston, K. (2012). The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment* 10(6), 298-304. https://doi.org/10.1890/110294

Olson, J.A. (2010). Data as a service: Are we in the clouds? *Journal of Map and Geography Libraries* 6(1), 76–78. https://doi.org/10.1080/15420350903432739.

Open Data Charter (2015). *International Open Data Charter.* https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf.

Porsanger, J. and Guttorm, G. (2011). *Working with Traditional Knowledge: Communities, Institutions, Information Systems, Law and Ethics.* Trondheim: Sámi University College. https://brage.bibsys.no/xmlui/handle/11250/177065.

Pottinger, L. (2012). *Why our rivers need a citizen science movement*. https://www.internationalrivers.org/resources/why-our-rivers-need-a-citizen-science-movement-7764.

Protocol on Pollutant Release and Transfer Registers (2003). Signed 21 May 2003. https://www.unece.org/fileadmin/DAM/env/documents/2003/pp/ch_XXVII_13_ap.pdf.

Pulse Lab Kampala (2017). Using machine learning to analyse radio content in Uganda: Opportunities for sustainable development and humanitarian action. 11 September 2017 https://www.unglobalpulse.org/news/using-machine-learning-accelerate-sustainable-development-solutions-uganda-0 (Accessed: 6 November 2017).

Raygorodetsky, G. (2017). 'Braiding science together with indigenous knowledge'. *Observations*, 21 December 2017 https://blogs.scientificamerican.com/observations/braiding-science-together-with-indigenous-knowledge/

Roelfsema, C., Thurstan, R., Beger, M., Dudgeon, C., Loder, J., Kovacs, E. *et al.* (2016). A citizen science approach: A detailed ecological assessment of subtropical reefs at Point Lookout, Australia. *PLoS ONE* 11(10), e0163407. https://doi.org/10.1371/journal.pone.0163407.

Rominger, J. and Ikeda, S. (2015). The role of big data in solving environmental problems. *Gradient Trends Risk Science and Application*, Fall 2015, 64. https://gradientcorp.com/pdfs/newsletter/Trends64.pdf.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C. *et al.* (2012). Dynamic changes in motivation in collaborative citizen-science projects. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 217-226. https://doi.org/10.1145/2145204.2145238

Secretariat of the Convention on Biodiversity (n.d.). *Article 8(j) -Traditional Knowledge, Innovations and Practices*. https://www.cbd.int/traditional/ (Accessed: 14 March 2018).

See, L., Fritz, S., Dias, E., Hendriks, E., Mijling, B. and Snik, F. (2016). Supporting earth-observation calibration and validation: A new generation of tools for crowdsourcing and citizen science. *IEEE Geoscience and Remote Sensing Magazine* 4(3), 38-50. https://doi.org/10.1109/MGRS.2015.2498840.

Shirk, J., Ballard, H., Wilderman, C., Phillips, T., Wiggins, A., Jordan, R. *et al.* (2012). Public participation in scientific research: A framework for deliberate design. *Ecology and Society* 17(2), 29. https://doi.org/10.5751/ES-04705-170229.

Spitz, R., Pereira, C., Leite, L.C., Ferranti, M.P., Kogut, R., Oliveira, W. *et al.* (2017). Gamification, citizen science and civic engagement: In search of the common good. *Balance-Unbalance 2017*. Plymouth, 21-23 August 2017. http://balance-unbalance2017.org/events/gamification-citizen-science-and-civic-engagement-in-search-of-the-common-good/

Stodden, V. (2014). *What scientific idea is ready for retirement?* [Edge Foundation https://www.edge.org/responses/what-scientific-idea-is-ready-for-retirement (Accessed: 05 June 2018).

Storksdieck, M., Shirk, J.L., Cappadonna, J.L., Domroese, M., Göbel, C., Haklay, M. *et al.* (2016). Associations for citizen science: Regional knowledge, global collaboration. *Citizen science: Theory and Practice* 1(2), 1-10. https://doi.org/10.5334/cstp.55.

Takahashi, D. (2017). *Got It debuts knowledge-as-service that uses AI to help you find human experts*. [VentureBeat https://venturebeat.com/2017/06/06/got-it-debuts-knowledge-as-a-service-that-uses-ai-to-find-you-human-experts/ (Accessed: 22 February 2018).

Tebtebba Foundation (2015). Mapping our lands and waters: Protecting our future. *Global Conference on Community Participatory Mapping in Indigenous Peoples' Territories*. North Sumatra, 25-27 August 2013. Baguio City http://www.iapad.org/wp-content/uploads/2017/01/mapping-our-lands-waters-protecting-our-future.pdf

Tengö, M., Brondizio, E.S., Elmqvist, T., Malmer, P. and Spierenburg, M. (2014). Connecting diverse knowledge systems for enhanced ecosystem governance: The multiple evidence base approach. *Ambio* 43(5), 579-591. https://doi.org/10.1007/s13280-014-0501-3.

The world's most valuable resource is no longer oil, but data (2017). *The Economist,* 6 May. https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource.

Thomas, R. and McSharry, P. (2015). *Big Data Revolution: What Farmers, Doctors and Insurance Agents Teach Us About Discovering Big Data Patterns.* Wiley. https://www.wiley.com/en-us/Big+Data+Revolution%3A+What+farmers%2C+doctors+and+insurance+agents+teach+us+about+discovering+big+data+patterns-p-9781118943717.

Tsuji, L.J.S. (1996). Cree traditional ecological knowledge and science: A case study of the sharp-tailed grouse, *Tympanuchus Phasianllus Phasianllus*. *The Canadian Journal of Native Studies* XVI (1), 67-79. http://www3.brandonu.ca/cjns/16.1/tsuji.pdf.

Tsuji, L.J.S. and Ho, E. (2002). Traditional environmental knowledge and western science: In search of common ground. *Canadian Journal of Native Studies* 22(2), 327-360. http://www3.brandonu.ca/cjns/22.2/cjnsv.22no.2_pg327-360.pdf.

Uitto, J. (2016). Use of big data in environmental evaluation: A focus session on use of new technologies in M&E and implications for evaluation. *19th Meeting of the DAC Network on Development Evaluation*. Paris, 26-27 April 2016. https://www.gefieo.org/sites/default/files/ieo/ieo-documents/Presentation-big-data.pdf

United Nations (1998). *Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters*. https://treaties.un.org/pages/viewdetails.aspx?src=ind&mtdsg_no=xxvii-13&chapter=27&clang=_en (Accessed: 24 January 2018).

United Nations (2014). Resolution adopted by the General Assembly on 29 January 2014. Fundamental Principles of Official Statistics. A/RES/68/261 New York, NY https://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf

United Nations (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development.* New York, NY. https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf.

United Nations (2016). *The Sustainable Development Goals Report 2016.* New York, NY. https://unstats.un.org/sdgs/report/2016/The%20Sustainable%20Development%20Goals%20Report%202016.pdf.

United Nations (2018a). *About*. https://www.unglobalpulse.org/about-new (Accessed: 11 October 2018).

United Nations (2018b). *News*. https://www.unglobalpulse.org/blog (Accessed: 11 October 2018).

United Nations (2018c). *Big data: News*. https://unstats.un.org/bigdata (Accessed: 11 October 2018).

United Nations, General Assembly (2007). *United Nations Declaration on the Rights of Indigenous Peoples.* New York. https://www.un.org/esa/socdev/unpfii/documents/DRIPS_en.pdf.

United Nations (n.d.). *Indigenous Peoples: 2030 Agenda for Sustainable Development*. https://www.un.org/esa/socdev/unpfii/documents/2016/Docs-updates/Indigenous-Peoples-and-the-2030-Agenda-with-indicators.pdf

United Nations Convention to Combat Desertification (2016). *Report of the Conference of the Parties on its Twelfth Session, held in Ankara from 12 to 23 October 2015 Part two: Action Taken by the Conference of the Parties at its Twelfth Session - Addendum* 12-23 October. ICCD/COP (12)/20/Add.1. https://www.unccd.int/sites/default/files/sessions/documents/ICCD_COP12_20_Add.1/20add1eng.pdf.

United Nations Educational Scientific and Cultural Organization (2013). *Agricultural Landscapes.* October 2013 (69). United Nations Educational, Scientific and Cultural Organization http://es.calameo.com/read/00332997247675cccaf1e

United Nations Educational Scientific and Cultural Organization (2017). *Local Knowledge, Global Goals.* Paris. http://unesdoc.unesco.org/images/0025/002595/259599e.pdf.

United Nations Environment Programme (2012). *Global Environment Outlook-5: Environment for the Future We Want.* Nairobi: United Nations Environment Programme. http://wedocs.unep.org/bitstream/handle/20.500.11822/8021/GEO5_report_full_en.pdf?sequence=5&isAllowed=y.

United States Environmental Protection Agency (n.d.). *AirNow*. https://airnow.gov/ (Accessed: 20 November 2018).

United States National Aeronautics and Space Administration (2018). *Engaging citizen scientists with GPM*. https://pmm.nasa.gov/articles/engaging-citizen-scientists-gpm (Accessed: 20 July 2018).

University of the West of England, Science Communication Unit, (2013). *Science for Environment Policy In-depth Report: Environmental Citizen Science.* Report produced for the European Commission. http://ec.europa.eu/environment/integration/research/newsalert/pdf/IR9_en.pdf.

UrbanEmissions.info (2016). *Air quality forecasting for all India*. http://www.urbanemissions.info/ (Accessed: 11 October 2018).

Vacarelu, F. (2017). 'New paper from UN Global Pulse and UNHCR explores use of digital data for insights into forced displacement'. https://www.unglobalpulse.org/news/new-paper-un-global-pulse-and-unhcr-explores-use-digital-data-insights-forced-displacement (Accessed: 6 November 2017)

Van Vliet, K. and Moore, C. (2016). Citizen science initiatives: Engaging the public and demystifying science. *Journal of Microbiology and Biology Education* 17(1), 13-16. https://doi.org/10.1128/jmbe.v17i1.1019.

Vidal, J. (2016). The tribes paying the brutal price of conservation. *The Guardian.* https://www.theguardian.com/global-development/2016/aug/28/exiles-human-cost-of-conservation-indigenous-peoples-eco-tourism.

Webb, A. and Usher, D. (2017). 'New vulnerability monitoring platform to assist drought-affected populations in Indonesia'. https://www.unglobalpulse.org/news/new-vulnerability-monitoring-platform-assist-drought-affected-populations-indonesia

Wells, T.N.C., Willis, P., Burrows, J.N. and van Huijsduijnen, R.H. (2016). Open data in drug discovery and development: Lessons from malaria. *Nature Reviews Drug Discovery* 15(10), 661-662. https://doi.org/10.1038/nrd.2016.154.

Williams, T. and Hardison, P. (2013). Culture, law, risk and governance: Contexts of traditional knowledge in climate change adaptation. *Climate Change*, 531–544. https://doi.org/10.1007/s10584-013-0850-0.

World Bank (2016a). *The World Bank Annual Report 2016.* Washington, D.C. https://openknowledge.worldbank.org/bitstream/handle/10986/24985/9781464808524.pdf?sequence=3&isAllowed=y.

World Bank (2016b). *World Development Report 2016: Digital Dividends.* Washington, D.C: World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/23347/9781464806711.pdf.

World Intellectual Pr