Science & Technology Innovation Program

Wilson Center | Science and Technology Innovation Program

Authors
Mark Chandler
Alexandra Novak
Alison Parker
Ashley Schuett
Alex Long
Anne Bowser

Rethinking data quality: Considerations for low-cost (and open source) science tools

March 2022

This publication is part of The Science and Technology Innovation Program's THING Tank. The Science and Technology Innovation Program's work in low-cost and open hardware is supported by the Alfred P. Sloan Foundation.

THING tank
understanding low cost tools for science

ALFRED P. SLOAN FOUNDATION

Special thanks to reviewers Sara Wylie, Luigi Ceccaroni, Claire Narraway, and Daniel Rubenstein. Their insightful reviews helped the authors re-think our own assumptions about data quality, and broadened our perspective. However, the authors alone are responsible for the analysis and content included here. The authors would also like to thank the Science and Technology Innovation Program and to Elizabeth Newbury for feedback and comments throughout the editing process.

TABLE OF CONTENTS

Abstract

Low-cost tools, including open source hardware, are changing how research is done and by whom. With broadened participation, factors such as the amount, type, and frequency at which data is being collected are changing. This shift can pose challenges to interpreting the quality—and ultimately credibility—of low-cost tools and their data.

The concept of "fitness for use" can be defined as "the degree to which a dataset is suitable for a particular application or purpose, encompassing factors such as data quality, scale, interoperability, cost, and data format" (Holdren, 2015). In the context of low-cost tools, we aim to outline challenges and potential solutions related to data quality and fitness for use by analyzing the tools themselves, as well as their digital environment, their regulatory context, and social context. Building a deeper understanding of how intentionally adopting a fitness for use perspective will facilitate not only greater uptake of low-cost tools but potentially more equitable access to science.

Introduction

Opening access to full participation in scientific research has many societal benefits, yet participation in science is often limited to recognized experts with access to funding, dedicated training, specialized tools, and other institutional resources. Citizen science is broadening access and inclusion in who can do science; alongside citizen science, low-cost tools (including open source hardware) has potential to open up both the processes and products of scientific research.

Compared to tools traditionally used for scientific research, low-cost tools are more available and accessible—a factor relevant to those communities typically on the outside of most research activities. However, there is a perception among some practitioners, researchers, and the policy community that the decrease in cost corresponds to a decrease in data quality. As Lewis et al. (2018) observe related to low-cost air quality sensors:

> "Based on the scientific literature available up to the end of 2017, it is clear however that some trade-offs arise when LCSs are used in place of existing reference methods. Smaller and/or lower cost devices tend to be less sensitive, less precise and less chemically-specific to the compound or variable of interest. This is balanced by a potential increase in the spatial density of measurements that can be achieved by a network of sensors."

In addition, data users sometimes question whether the new types of stakeholders involved in low-cost tools have the requisite expertise to create instruments, conduct research, and interpret results.

Together, these concerns have resulted in some doubt in the research community on the value of low-cost tools for use for a wide range of purposes, including research and decision-making (e.g. regulatory purposes). Yet with new research fields opening up and with shifts in the use and users of research tools, traditional mechanisms and processes to assess, certify and communicate data quality may no longer be sufficient for a wide range of low-cost tools that are fit for different uses. In addition, the rapid change and customization of low-cost tools and the use of these tools means that traditional processes for evaluating tools and their data may not be applicable, are too rigid, or are cost prohibitive. Without a critical review of how new tools are adopted and valued, credentialized users of scientific data may (inadvertently) act as "gatekeepers"–in effect controlling the use of data for science from non-traditional tools and their users, in a way that limits their impact. For example, EPA regulates air quality on 24-hour averages, but many community science groups believe it is critical to study shorter temporal periods of higher exposure (Ottinger, 2010).

Here, we describe how a "fitness for use" perspective helps us assess, improve, and communicate the data quality associated with low-cost tools. We begin by describing how low-cost tools are used, and then unpack the concepts of data quality and "fitness for use" as they relate to low-cost tools. Then we describe the challenges that low-cost tools present for traditional assessments of data quality, and suggest specific solutions to address these challenges with a "fitness for use" perspective. Our goal is to identify specific policies and strategies that can help create a culture of confidence and reduce uncertainty around low-cost tools, to benefit tool designers and builders, tool users, and potential data users, including practitioners, researchers, and decision-makers.

Key Terminology

Low-Cost Tools

Low-cost tools are becoming more and more prevalent as alternatives to expensive and inaccessible equipment traditionally used for scientific applications. These tools span a range of disciplines, and include tools that are commercial or off-the-shelf, customizable, do-it-yourself (DIY), or obtained as kits to be assembled. Low-cost tools take many forms: environmental sensors, laboratory instruments, medical devices and biometric sensors, accelerometers, microscopes, CubeSats, robots, drones (Science Stack: Tools within Reach). Although some tools are incrementally lower cost, most are significantly so, to the extent that they significantly broaden who can access and use them (Parker & Novak, 2020).

Throughout the publication we use "tool" as a descriptor of physical devices; this term can be considered interchangeable with "hardware."

Open Source Hardware (OSH)

Open source hardware are physical tools whose design is "publicly available so that anyone can study, modify, distribute, make, and sell the design" (Definition (English), n.d.). Because open source hardware tends to be significantly cheaper than proprietary alternatives, we include open source in our general consideration of "low-cost" tools that accelerate research, broaden participation, and lead to other beneficial outcomes. To a greater extent than proprietary tools, open source hardware enables collaborative participation in the design of the tool, and allows users to replicate, repair, and customize the tool.

Background

FITNESS FOR USE

The concept of "fitness for use" can be defined as "the degree to which a dataset is suitable for a particular application or purpose, encompassing factors such as data quality, scale, interoperability, cost, and data format" (Holdren, 2015). A fitness for use perspective is particularly helpful when using low-cost tools for science, because different stakeholders may use any given tool or data set for different goals. Beginning with an intended use (see Figure 1), a data user can define the requirements needed for that intended use, and if all defined requirements are met, the data can be considered fit for use. For example, a water quality sensor used to enforce regulations will require high levels of precision and accuracy. However, when used for exploratory research or by communities interested in local monitoring, high precision and accuracy may not be essential, and accessibility, cost-effectiveness, and timeliness may be.

Figure 1

The spectrum of citizen science data use.

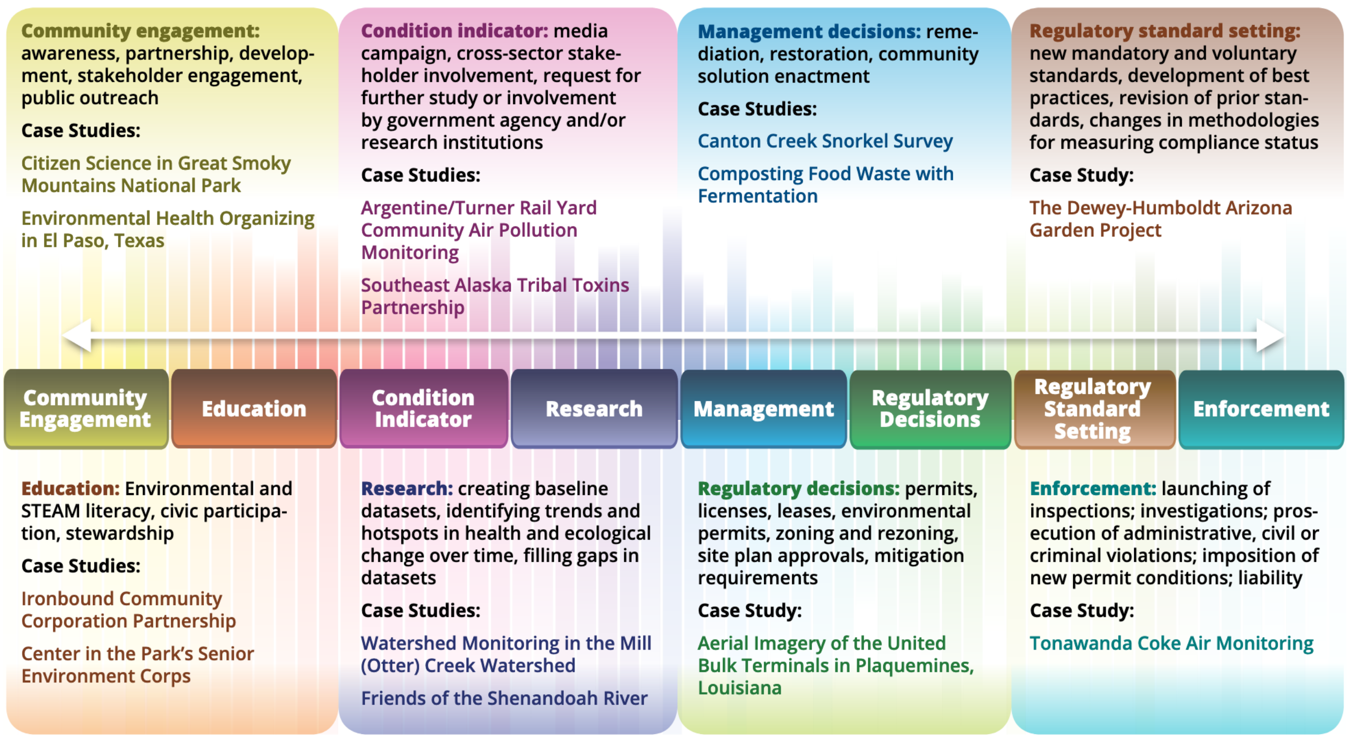


Image source: From Parker, A. & Dosemagen, S. (2016). Environmental Protection Belongs to the Public: A Vision for Citizen Science at EPA. National Advisory Council for Environmental Policy and Technology (NACEPT).

Fitness for use is becoming more recognized in research and policy communities. For example, the 2015 memorandum Addressing Societal and Scientific Challenges through Citizen Science and Crowdsourcing highlights fitness for use as a core principle for using citizen science and crowdsourcing to accomplish agency missions: "Recognizing that a 'one-size-fits-all' quality-assurance approach will not work for all projects, Federal agencies should apply the principle of 'fitness for use,' ensuring that data have the appropriate level of quality for the purposes of a particular project" (Holdren, 2015). The United States Environmental Protection Agency (EPA)'s National Advisory Council for Environmental Policy and Technology (NACEPT) charts (Figure 1) a spectrum of uses that citizen science data—often collected via the use of low-cost tools—are fit for, ranging from community engagement to regulation and enforcement (Parker & Dosemagen, 2016).

> The goal of using a fitness for use perspective is to move from a one-size-fits-all approach to data quality, to an approach that recognizes a wide range of potential uses, each associated with different data quality requirements.

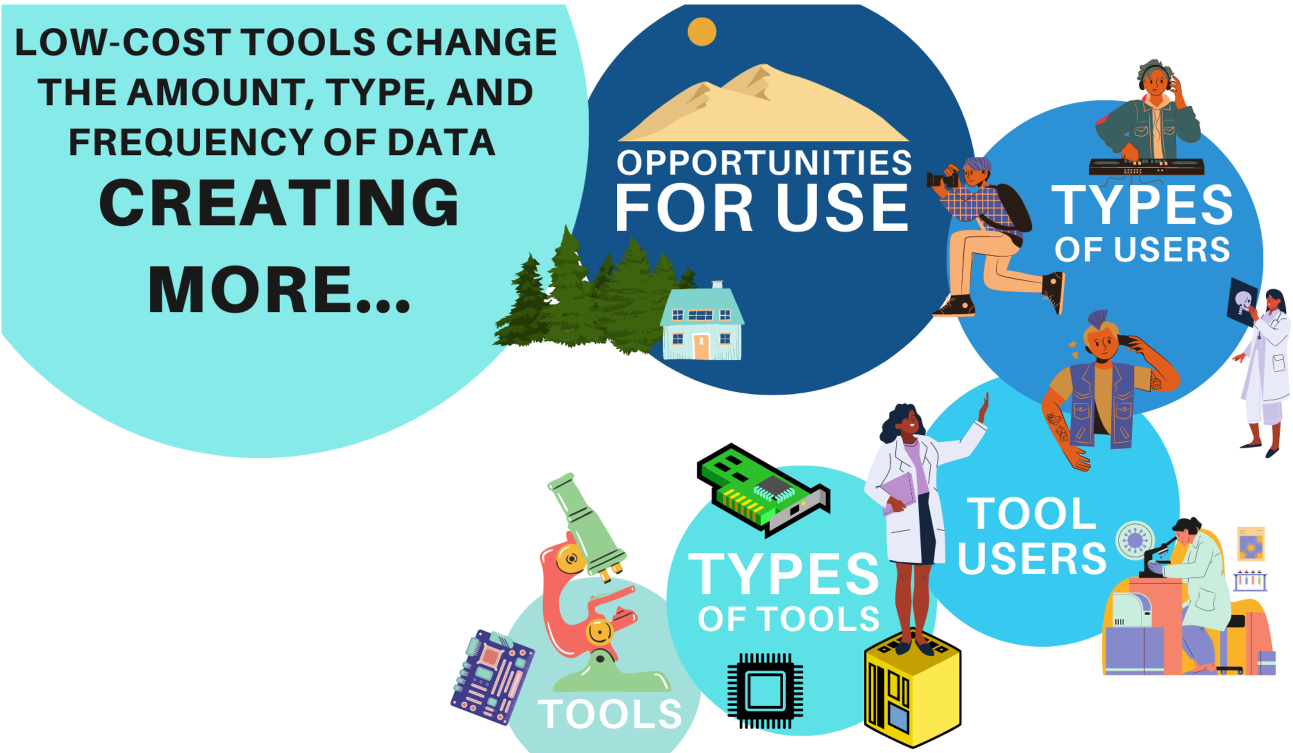
Factors beyond those traditionally used in data quality assessments are important here; implementation conditions, or the "enabling environment," can allow otherwise limited data to become more useful for a range of purposes. The enabling environment is a critical opportunity to expand the use of low-cost tools for accessible, relevant, and novel science.

How Low-Cost Tools are Used

Low-cost tools are improving the research process for professional scientists in traditional institutions by reducing costs to create greater efficiency and return on investment, improving reproducibility, and catalyzing new fields of scientific growth, including by enabling exploratory research (Maia Chagas, 2018; Pearce, 2017). However, the true power of low-cost tools might lie in their potential to broaden who participates in scientific research. By increasing availability and access for communities to research questions of interest to them, current gaps in monitoring and research (for example, localized air and water quality) might be addressed.

Figure 2

Low-cost tools change the amount, type, and frequency of data.



Low-cost tools are expanding where science happens; enabling research and innovation by individuals, by community based organizations, as well as in low to middle-income countries. In addition, low-cost tools for environmental monitoring specifically enable different approaches to data collection and analysis, such as community monitoring, and can ultimately help inform regulators of the lived experiences of community members in an area affected by pollution—instead of, or in addition to, traditional types of data (Gabrys et al., 2016). For example, Purple Air, a low-cost air quality sensor for particulate matter, is being used by citizen scientists to help enhance local air quality estimates from data derived from National Aeronautics and Space Administration (NASA) satellite observations (Doraiswamy, n.d.). Low-cost tools are used in crisis response at both global and local levels, including for COVID-19 (Bowser et al., 2021). Some low-cost tools, like PocketLab, are specifically designed for education, and support Science, Technology, Education, and Math (STEM) learning goals (PocketLab: Science Everywhere, n.d.).

Unpacking Data Quality and Fitness for Use

Data Quality

Defining and communicating data quality is complex, and often entails "a multifaceted evaluation of states such as completeness, validity, consistency, precision, and accuracy" (Wiggins et al., 2011). For example, at a workshop on low-cost air quality monitors, experts identified several dozen defined performance criteria that relate to data quality when assessing particulate matter pollution (Williams et al., 2019). The following common set of data quality attributes are found across many frameworks. These include:

- **Accuracy:** How well does the data reflect reality?
- **Precision:** Does a tool produce the same data each time under the same conditions?
- **Completeness (or comprehensiveness):** Does the data meet sufficient expectations of what's needed to answer a question?
- **Timeliness (or latency):** Are the data available when needed?
- **Validity (or conformity):** Are the data available in a format that aligns with established standards or meets other articulated needs?
- **Transparency:** Can communicating the initial goals of data use and the extent of quality assessments help ensure trust, responsibility, and impact?

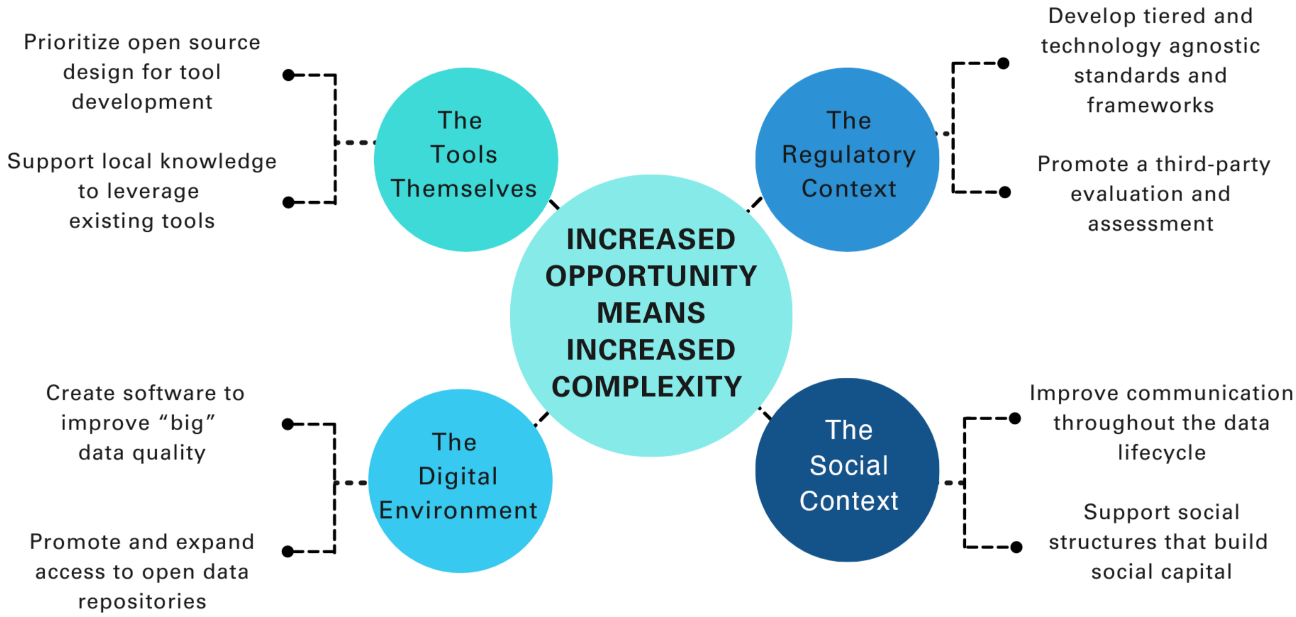
Assessing Data Quality of Low-cost Tools: Challenges and Solutions

With a fitness for use perspective, the key challenge is ensuring and documenting the appropriate dimensions of data for different uses. In this process, a number of factors need to be considered. These include challenges and solutions related to:

1. The **tools themselves**: the design, maintenance, and manufacturing of low-cost tools, including their interactions with the physical environment;
2. Their **digital environment**: the interaction (interface) between the physical (hardware) and digital (software) components of the tool;
3. The **regulatory context**: processes and protocols for assessing data quality and appropriate use; and,
4. The **social context**: including social structures for communicating data quality and uses.

Figure 3

The increased opportunity of more and different uses and users also means increased complexity in assessing and communicating data quality. These four challenges and their associated solutions help match data quality with the fitness for use.



The Tools Themselves

Challenges

A range of challenges relate to the design, maintenance, and manufacturing of physical tools, and their performance in their physical environment.

When the tools available do not meet the data quality requirements of the intended use(s), there is a "mismatch" between the need for a research tool and its availability. A mismatch can also occur when tools are too expensive or otherwise inaccessible for the intended use.

Technical considerations around the physical environment in which a tool is created and used influence how different component parts perform. In addition to sensitivity across different environmental conditions, like varying temperatures or humidity levels, successful performance of a particular tool can also depend on the effect of weather, its power requirements, size, and other factors. For example, the use of laser counters commonly used in low cost air quality sensors such as those manufactured by Plantower Technologies to estimate particulate matter are sensitive to both humidity and temperature, as well as ambient wind conditions. The results from lab conditions are not necessarily directly transferable to field conditions—necessitating multiple tests under different conditions to provide the full picture of how tools are reliable under different conditions.

Solutions

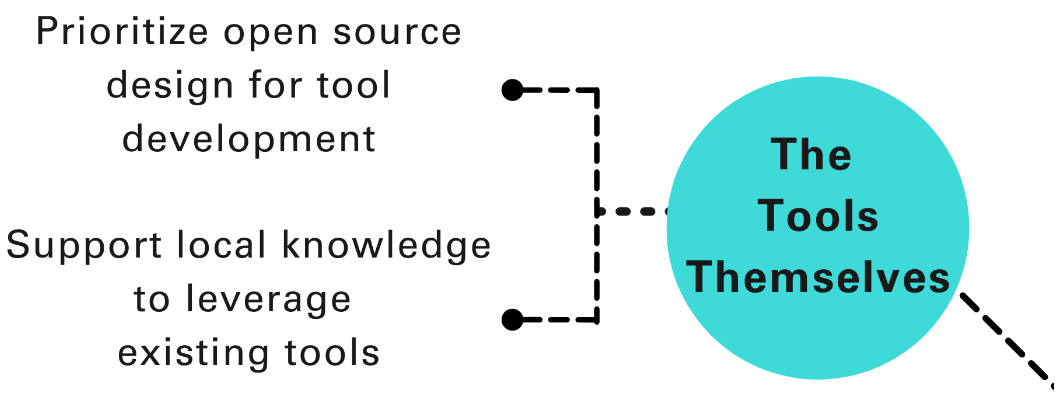


Figure 4

The Tools Themselves

Mismatches can be minimized by investing in solutions that help improve the design, production, and use of low-cost tools themselves, including prioritizing open source designs and evoking local knowledge and expertise.

Prioritize Open Source Design for Tool Development

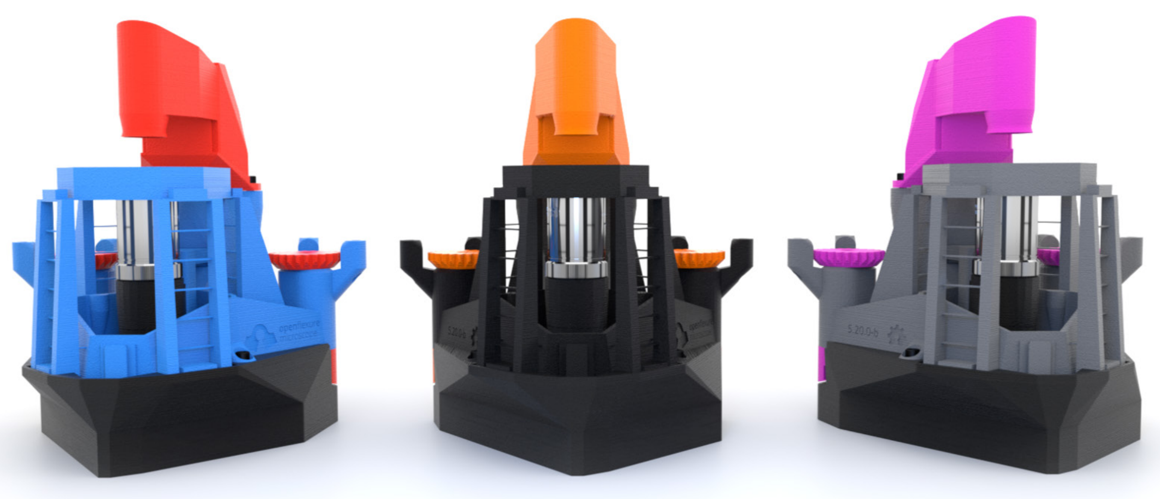
As open source hardware relies on collaborative participation in design, evaluation, and sharing, using open practices can address mismatch in cost by driving down both design and maintenance costs to help provide economic savings of 87% compared to equivalent proprietary tools (Pearce, 2020). Open source practices also encourage customization through the use of open licenses. This practice "enables the design [of a tool] to be modified for very specific uses" (Niezen et al. 2016). For example, someone using EnviroDIY's Mayfly Data Logger can customize the device to have a GPS add-on if their intended use requires location data (Stroud Water Research Center, n.d.). From a legal perspective, this is made possible by the use of the CERN Open Hardware License 1.2, which clearly spells out acceptable provisions for re-using the original blueprint to create derivative devices (CERN Open Hardware License, n.d.).

Open source is an opportunity for improved scientific reproducibility, and therefore data quality (Hill, 2021). Open source hardware, together with improved documentation, can make it easier for scientific protocols to be understood and replicated, increasing replicability, transparency, and trust. For example, the Oxford University's Penguin Watch program uses open source cameras to monitor penguin populations in Antarctica. Because they used open source tools, researchers were able to deploy more cameras and, as a result, improve the sample size and replicability of the research and increase confidence in the results (Hill, 2021).

Support Local Knowledge to Leverage Existing Tools

Local environmental, social, or economic conditions can significantly alter a tools' ability to generate data quality that is fit for purpose, often requiring specific tailoring or modification to the tool itself or the data generated to yield useful data. For example, philanthropically donated equipment from higher-resourced regions to lower-resourced regions (e.g. low- and middle-income countries) can be both expensive and ineffective; up to 50% of donated medical devices become unusable due to a lack of locally available maintenance and the inability to obtain spare parts (Garmendia et. al, 2020).

The use of low-cost tools, and in particular, open source hardware, benefits from local knowledge and resources to provide both necessary and long-term development and support. For example, during the COVID-19 pandemic, low-cost tools, appropriately modified by local know-how, were critical when responding to a pressing need; customized designs of medical devices were created locally using 3D printers and locally available materials. These tools allowed local experts to assess data quality and navigate local standards and regulations to ensure ethical, safe and legal use (Longhitano et al. 2020; Mueller et al., 2020), effectively reducing a serious shortage of



ventilators.

Figure 5

Image Source: "Microscope Blender Trio" by OpenFlexure is licensed under CC BY

Case Study 1: Open Source Microscopes & Medical Diagnostics

Low-cost microscopes come in many different forms, including foldable paper microscopes that cost $1 in parts, smartphone plugins that cost between $5 and $20, and automated, 3D printable microscopes available at a range of price points (Parker et al., 2021).

Although these microscopes are more accessible and affordable, these microscopes do not always meet sufficient data quality standards for all uses. Validation studies show that Foldscope produces efficient and high quality data for tick identification (Parada-Sánchez et al., 2018) but not medical diagnostics applications (Yong, 2019).

For medical diagnostics, such as screening blood smears for pathogens, images need high resolution and a large field of view. A major challenge to using low-cost microscopes for medical diagnostics is the inherent trade-off between resolution and field of view that typically requires expensive technical features to bridge (Switz et al., 2014).

To enable utilization for medical diagnostics, a team of researchers in the UK and Tanzania used open source design to create OpenFlexure, a low-cost, customizable microscope under trial for malaria diagnosis. The microscope was designed with a flexure mechanism for precise stage movement. This feature means that OpenFlexure can take automated images at a low field of view and high resolution and stitch them together to produce a larger field of view, successfully addressing the technical barriers that exist for other low-cost microscopes (Collins et al., 2020).

The microscope was co-designed by researchers in the UK and Tanzania so local knowledge was used to ensure that the microscope was fit for use in Tanzania (Stirling et al., 2020). This led to the design of an almost entirely 3D printed microscope, so that most parts can be produced locally. The use of open source design also enables users of the tool to customize technical features for their intended use, such as swapping out the lens for different magnifications.

With appropriate forethought in design and support, low-cost tools may provide more sustained useful data than higher cost alternatives. Low-cost, and in particular open source, tools can allow for access to local knowledge, and the transparency and openness required to understand and customize solutions.

The Digital Environment

Challenges

A range of challenges relate to the interaction between the physical (hardware) and digital (e.g. software) elements of a particular tool.

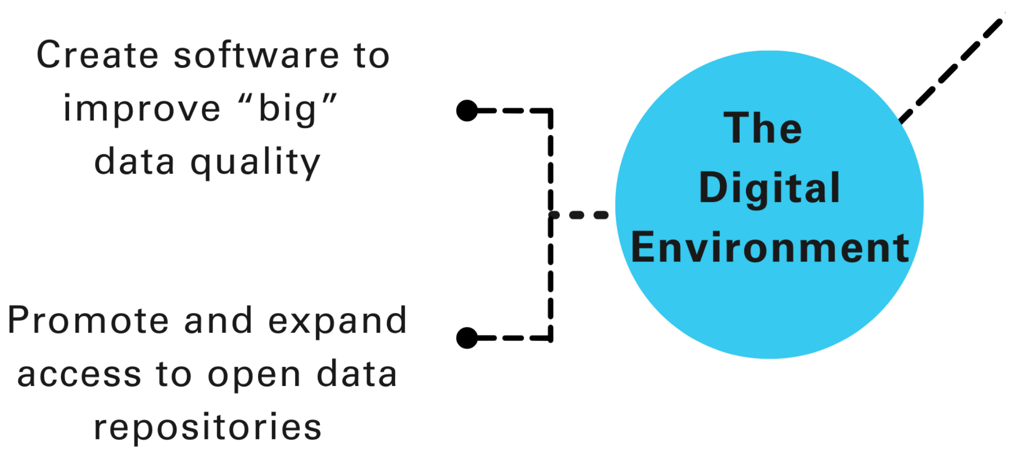
As low-cost tools become widely used by more people across different fields, there is an increase in the amount of the data being produced (e.g. terabytes, petabytes) (Sharma et al., 2015). The types of data collected are also expanding with more novel tools, diverse research applications, and different user groups. The frequency of data collection is also increasing, with more real time data monitoring enabled by the Internet of Things (IoT) (Zaslavsky et al., 2013).

As amount, type, and frequency increase, data can become more difficult to understand and access. Therefore, data produced by low-cost tools may require more complex protocols and softwares for Quality Assurance (QA) and Quality Control (QC) than higher-cost versions; without these, these data may be less useful or less credible. As another related challenge, scientists using low-cost tools may lack access to data infrastructure that follows FAIR (Findable, Accessible, Interoperable, Reusable) guidelines, or data analysis tools (FAIR Principles, 2021). Infrastructure, strategies, and tools to store, retrieve, extract, and analyze data are increasingly available (e.g. R for statistical analysis). However, these are not always accessible to groups like citizen and community scientists, inhibiting their ability to analyze and act on data produced from low-cost tools.

Solutions

Figure 6

The Digital Environment



Once tools are fully developed and available, the usefulness and credibility of the tool's data is facilitated by how the data it produces interacts with other relevant data. To address barriers related to the diverse data sets produced by low-cost tools, it is important to have open access to software tools to help improve data quality, and access to repositories for storing and retrieving data.

Software to Improve "Big" Data Quality

The availability of Artificial Intelligence (AI) and Machine Learning (ML) techniques allows data from low-cost tools to be transformed to improve data quality, increasing the potential reliability of these data and increasing their fitness to answer research questions. For example, the development of Next Generation Sequencing tools has dramatically reduced the cost of DNA and RNA sequencing (ThermoFisher Scientific, n.d.). These tools are expanding who can afford to do sequencing (including smaller labs or research institutes), and what research questions scientists can ask (including more population and exploratory-based questions). Taking advantage of the increased volume of sequences in repositories such as GenBank, an open database run by the National Institutes of Health (NIH), neural network software has improved the processing of the shorter and more error prone sequences generated by low-cost nanopore technology. This leads to improved accuracy at higher throughput (Bowden et al., 2019).

Promote and Expand Access to Open Data Repositories

Large, open, and accessible data repositories are making many kinds of data more useful, including data produced from low-cost tools. Such data repositories can be dedicated to individual kinds of tools such as specific kinds of air quality monitors (e.g. PurpleAir.com, Safecast) (Purple Air, n.d.; Safecast, n.d.). The repositories also can aggregate across different sources of data. Biodiversity occurrence records hosted by the Global Biodiversity Information Facility (GBIF) include data from low-cost sensors—camera traps—along with citizen science apps, professional field observations, and digitized museum specimens (Global Biodiversity Information Facility).

Open data repositories can be especially important for low-resourced contributors and users. First, they can provide a resource for storing data that others can access beyond the initial project database, elevating access and impact. Second, if similar research is included in the same repository, more modest data generating activities can be combined with the effort of others through data harmonization efforts or meta-reviews. This can increase the value of data produced by low-cost tools and thus the credibility, especially when the platforms contributing to a repository have established QA/QC mechanisms in place (see Case Study 2).

Associated metadata and licensing allow for interoperability and access for diverse users, as well as the legal grounding required for data reuse (Bowser et al., 2020). The promotion of citizen science data sharing practices by authorities such as NASA can help set norms for a range of open source science communities to follow (Amos et al., 2020). It is also important that these databases include well-documented information about the intended use context. The Findable Accessible Interoperable and Reusable (FAIR) data framework is one of the common sets of standards that is leveraging documentation to enhance data sharing and effective reuse for science, though—like open data ideals—realizing the FAIR principles can be difficult to achieve in practice.

Case Study 2: Big Biodiversity Data Verification and Integration

Simple biodiversity observations form the basis of much ecological research and monitoring efforts. With the wide availability of low-cost tools like camera traps, along with smartphones and digital cameras, there is significant increase in interest by the traditional research community as well as the general public in leveraging these technologies to generate data valuable to research and conservation outcomes.

Figure 7

Image Source: Burne, Clayton. Lion and Camera Trap. n.d. Retrieved shutterstock

To unlock the value of data collected through these devices, the use of "nature" images as scientific data requires accurate interpretation of the biodiversity in the images, often using a "crowd-sourcing approach" (Goodchild et al., 2012). AI algorithms are also being developed to assign identification. In both cases, an engaged community of biodiversity experts is required to oversee and verify the specific identification of the AI models, by labeling data for training purposes, as well as by providing direct identifications on images. Often, both the use of tools to collect images and the interpretation of images to generate data is performed by citizen scientists. The mobile application iNaturalist supports in-app crowdsourced verification through a multi-party review process that elevates an observation from "casual" to "research grade," and is the cornerstone of data quality for these tools (iNaturalist, 2015). Platforms like Zooniverse.org support a number of projects that involve volunteers in labeling, classifying, or annotating images of biodiversity from camera traps.

Once a certain quality threshold is reached, many citizen science projects share data with the Global Biodiversity Information Facility (GBIF) (Boone & Basille, 2019). The whole data production pipeline, from the users of "cameras" who take and upload still images or videos to the research community that uses the data for science, requires validation. In addition to identification and verification processes, training materials—like GBIF's guide to best practices for publishing camera trap data—provide informal standards or guidelines that researchers can adhere to (Cadman & González-Talaván, 2014).

Relying on crowdsourcing to add value requires platforms that not only provide digital tools for scientific workflows, but infrastructure for managing and supporting online communities. To build a robust and efficient data verification process, iNaturalist provides social structures such as moderated comment sections which enable conversations between those knowledgeable about species and those eager to learn. Data filters allow community members who want to specialize on specific taxa or locations to easily focus on images meeting their criteria (for example, "bees from Spain"). Zooniverse has also designed a number of opportunities for social interaction and community support, including a general "talk" discussion forum as well as specific "talk" components to each hosted project's page (Zooniverse Talk, n.d.). Building on these examples to replicate their success requires investment in both technical infrastructure and social structures.

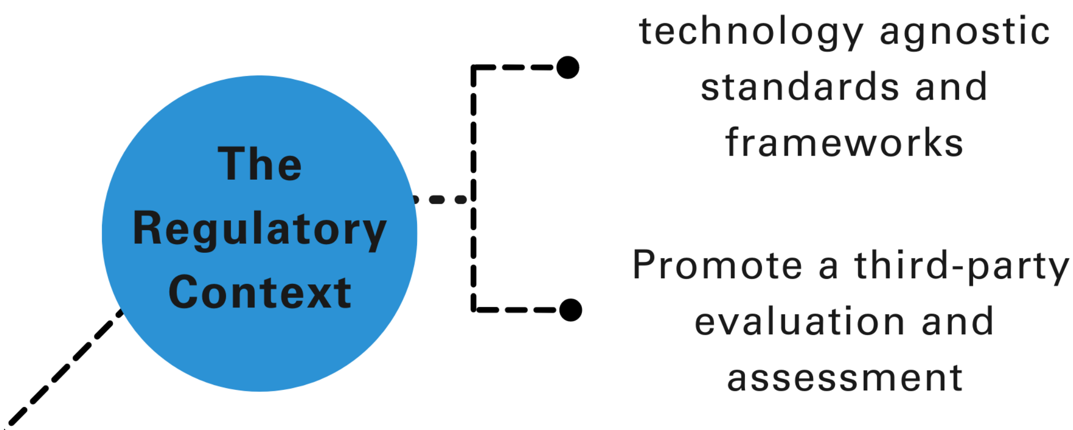
The Regulatory Context

Challenges

Standards developed by governmental agencies, or international professional organizations such as the International Standards Organization (ISO) or Institute of Electrical and Electronics Engineers (IEEE) (International Standards Organization, n.d.; Institute of Electrical and Electronics Engineers Standards Association, n.d.) often oversee and regulate the assessment of tool performance. Mandatory standards for data quality can sometimes be helpful by revealing a clear pathway toward use of tools (e.g. for regulatory purposes) or accelerating the development and adoption of new tools. At the same time, these same standards can become a boundary that can be unfit or inaccessible to the users of low-cost tools.

Solutions

Figure 8
The Regulatory Context



Solutions to these challenges could include focusing on technology agnostic standards, evaluating or assessing low-cost tools in a dynamic and open way, and brokering partnerships between tool designers on one hand, and tool producers on the other.

Develop Tiered and Technology Agnostic Standards and Frameworks

A tiered evaluation protocol for low-cost tools, similar to the EPA's four-tiered quality assurance criteria, could help assess data quality and fitness for use for different use cases. Tiered evaluation criteria offer tool creators and users an opportunity to articulate a particular use case, and document whether or how performance aligns with that particular use. Users can then select or customize tools to fit a similar use based on the understanding of contextual standards and requirements. For example, laboratory calibration may be necessary for an air quality sensor used in extreme weather conditions, but not for basic outdoor monitoring in a temperate location.

When developing standards and assessment methods based on fitness for use, it is important to continue to use a "technology agnostic philosophy" (Duvall et al., 2021). Creating standards that are technology agnostic requires defining data quality in terms of standards for an intended use, instead of focusing on the absolute performance of the technology itself. This would allow for the standards and methods to be applied to any type of similar technology (Duvall et al., 2021).

Figure 9

Source Image: "Making air quality sensing balloons at HackPittsburgh" by thelagged is licensed under CC BY-NC-SA 2.0

For example, instead of standards for a specific type of air quality sensor, like an electrochemical sensor, a standard could be linked to a particular use case, such as regulatory enforcement or exploratory research. Intermediaries can play important roles in working with regulatory bodies to make such standards accessible to tool manufacturers and users, as explored in depth below.

Third Party Evaluation and Assessment

Rigorous, transparent, and accessible testing and assessment of a tool under a range of conditions by a credible and neutral testing agency can build trust and confidence in the tool's data. Of particular value can be comparative testing of different tools under similar conditions, allowing users to compare and understand which tools might be of greatest value to their research. Government agencies and international organizations such as the International Standards Organization (ISO, see Case Study 1) and the International Electrical and Electronics Engineers (IEEE) can help by setting standards to assess data quality and tool performance. The State of California's South Coast Air Quality Management District has developed a lab dedicated to evaluating a large number of air quality tools and made the results of these analyses available for users to review (see Case Study 3). In addition to government agencies and other institutions, expert communities can be effective at assessing and evaluating the quality of data of low-cost tools through different mechanisms of peer review. Many of the same journals that publish information on tool designs also publish information on testing under different scenarios. For example, JOH is an open access, academic venue for publishing peer-reviewed articles including hardware metapapers. These papers provide detailed descriptions for tools used in areas as diverse as canine operant conditioning to innovations in 3D printing (Arce & Stevens, 2020; Delmans & Haseloff, 2018). Metapapers include discussions of quality control, provide links to external design files for additional transparency and assessment, and—like research articles—are peer-reviewed.

Social platforms such as WILDLABS are another model of informal peer review for low-cost tools; platform members can post information on the tool's intended use for audiences like educators or citizen science volunteers (Wildlabs.net, n.d.). Though often dubbed "informal," particularly compared to regulatory compliance reviews, the critique that happens in these venues is a legitimate form of community peer-review that should be regarded as trustworthy for many uses. Academic peer review and community-based review processes may be particularly helpful in cases where fitness for use does not align with existing regulatory standards.

Case Study 3: Monitoring Air Quality

Before the rise of low-cost air quality sensors, air quality monitoring of critical pollutants was limited to government operations for regulation and compliance. Although government monitoring provides absolute numerical value in relation to regulatory standards, it has location and cost limitations. For example, traditional air quality monitors for particulate matter (PM) could cost over $100,000, and, Federal Reference Methods or Federal Equivalent Methods could cost up to $10,000 (Levy Zamora et al., 2018). In addition, there are not many of these monitors, meaning that there are significant gaps in their coverage.

Figure 10

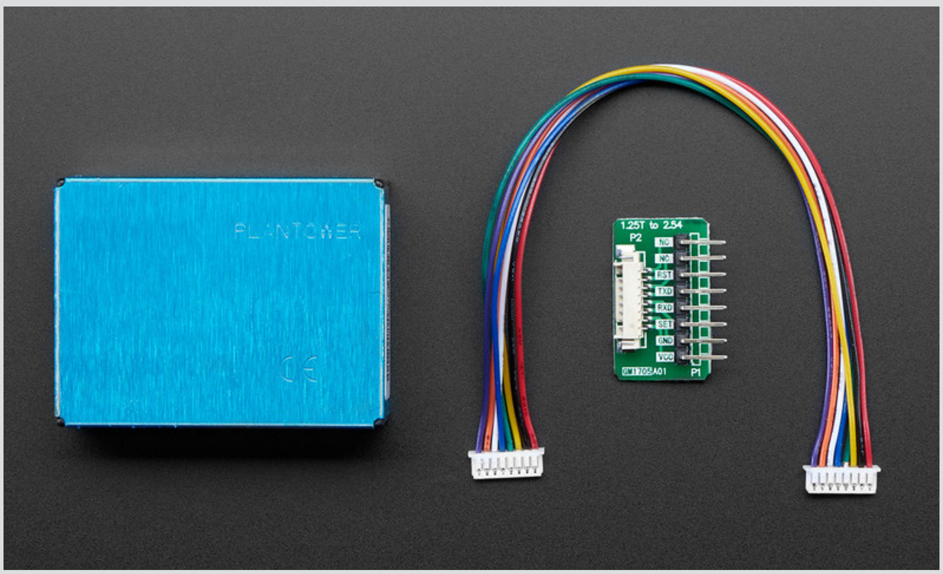


Image Source: "PM2.5 Air Quality Sensor and Breadboard Adapter Kit - PMS5003" by adafruit is licensed under CC BY-NC-SA 2.0

In recent years, the price of air quality sensors has gone down (Morawska et al., 2018). Although the EPA considers "low-cost" sensors to be less than $2500, some personal and community monitoring sensors are less than $100. Data can be collected in diverse locations by different types of people, including individuals with no formal training in environmental monitoring. Many low-cost sensors provide measurements which reliably indicate general trends and patterns but are less accurate than traditional monitoring's absolute numerical measurements; these "indicative" measurements can provide data of sufficient value on local patterns and changes in air quality data for many community uses (Gabrys et al., 2016). Additionally, simultaneous, large quantities of lower precision air quality monitoring can produce high resolution air quality maps, can ground-truth other datasets, and may close gaps in coverage.

In order to adapt to a paradigm shift towards expanded air quality monitoring, new processes and protocols for assessing data quality are needed. Traditional evaluation and assessment methods often require access to expensive benchmarking instruments (e.g. Federal Reference Methods), extensive laboratory studies, or long testing durations. These methods require a high level of expertise, along with significant monetary, physical, and professional resources. In fact, in one review that analyzed 57 studies on low-cost air quality sensor evaluation, only 5 studies were found to use existing EPA and European evaluation protocols (Morawska et al., 2018). Stakeholders also identified lack of performance targets for non-regulatory use cases as a barrier to successful comparison of sensors and, ultimately, use (Williams et al., 2019).

A number of solutions are emerging to address these barriers. For example, the State of California's South Coast AQMD Air Quality Sensor Performance Evaluation Center has developed protocols for using and evaluating low-cost air quality sensors in community monitoring as well as suggestions for testing outdoor community monitors. The South Coast AQMD has published results on evaluations of various air quality tools that can be publicly accessed by new types of data users and tool users, and EPA has created a website with evaluation results (Air Quality Sensor Performance Evaluation Center, n.d.; United States Environmental Protection Agency, n.d.). These efforts demonstrate the importance of expanding traditional testing to meet a range of low-cost tools use cases, and show how an intermediary—in this case, the South Coast AQMD—can be helpful in interfacing between community groups on one hand, and regulatory authorities like EPA on the other.

The Social Context

Challenges

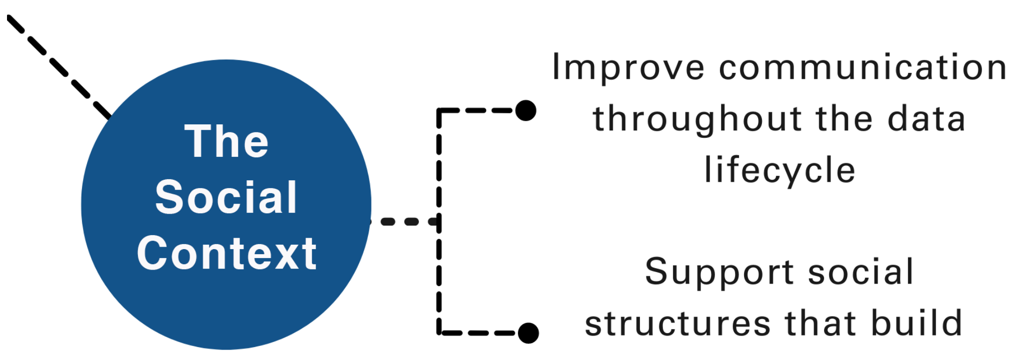
The social context of tool and data use can be very important to the resulting data's fitness for use. For example, there are many social structures in place for communicating data quality and appropriate use of different tools. Social structures can be formal—such as a user group dedicated to developing and sanctioning data quality standards, or informal—such as loose networks of users who interact via social media, conferences, or other shared activities. These social structures can facilitate tool distribution and acceptance, or prevent uptake.

In addition to questions about data quality from low-cost tools, professional scientific communities may doubt that users of low-cost tools are able to use them. Similar to perceptions of citizen or community science, some professional scientific communities question the ability of non-expert volunteers to conduct objective research in line with scientific standards (Kosmala et al., 2016). This can limit the impact of a tool or research project, and is often reinforced by the lack of social structures to bring different stakeholders together in discussions around data quality, fitness for purpose, and trust.

Solutions

Figure 11

The Social Context



To build trust and credibility, improved communication is needed across a range of stakeholders creating, using, and acting on research conducted with low-cost tools. Shared discussions should include topics such as data governance, data management strategies, and effective QA/QC.

Improve Communication throughout the Data Lifecycle

Tool developers, manufacturers, tool users (including both "traditional" tool users like professional scientists, and "new" tool users like environmental justice groups), and data users would better understand the capabilities of low-cost tools with improved communication on how data are managed for quality across the entire life cycle of data use—from creation to storage and use. Better cross-sector communication would allow each of these groups to develop an understanding and confidence in how data is being used, what performance metrics are important for manufacturers to provide information on, how the tools should be calibrated, and other factors that are important in developing credibility around low-cost tools.

Data stewards or other intermediaries can play an important role in facilitating communication between these groups and ensuring that guidelines are enforced, as well as recommending improvements to data governance processes. Public Lab is an example of an organization that plays a data steward role for environmental justice work (Public Lab, n.d.). For example, during the BP 2010 oil spill in the Gulf of Mexico, Public Lab brought together tool designers to help build low-cost, open source "community satellites" to map the coastline affected by the oil spill. The organization worked closely with grassroots partners to develop scientific and technological skills and produce high-resolution mapping of the coastline. Public Lab also engaged with environmental advocates and other data users to generate impact around the issue, including their data being featured in news sources such as BBC and New York Times. By facilitating communication and collaboration between tool developers, tool users, and data users, Public Lab ensured effective data governance and the production of high quality data for their intended use.

Figure 12

Source Image: "Public Lab Barataria Trip V boat" by eustatic is licensed under CC BY-NC-SA 2.0

Support Social Structures that Build Social Capital

Users of low-cost tools benefit from the support of social structures that build social capital and trust. Academic and government researchers typically leverage professional associations for knowledge sharing. New users of tools, especially those outside traditional research institutions, also benefit from social networks to build and access the knowledge around specific research areas or specific tools. In agricultural settings, for example, farmers can benefit from informal social knowledge networks, relying on known trusted relationships among users when choosing to adopt new technologies to monitor and manage tools to enhance crop production (Shang et al., 2021; Wang et al., 2020).

Some platforms evolve into "knowledge commons," where members can share resources on new low-cost tools, guidelines on best practices around fitness for different uses, recommendations on navigating regulatory processes, and information on funding opportunities. The impact of these platforms is illustrated by WILDLABS, a community platform and social network where conservation scientists, engineers, data scientists, and other related professions openly share information on how innovative technology and data can be applied to a range of conservation issues. This network helps increase the adoption of low-cost tools. For example, a community of users and builders that includes WILDLABS coalesced around the open source acoustic monitoring device AudioMoth, helping to create shared understanding and elevating its impact on conservation (AudioMoth, n.d.; Lahosz Monfort et al., 2019).

Regardless of type, effective social structures often include both "informal" communities, such as loosely connected networks of tool creators and users, and "formal" institutional actors, such as government and industry partners. Such a structure is demonstrated by the Federal COVID 3D Trust, led by federal authorities including the U.S. Food and Drug Administration (FDA), NIH, and U.S. Veterans Affairs (VA), along with America Makes. This group collaborated with grassroots communities as well as hospitals to share information and resources during shortages of Personal Protective Equipment (PPE) during the COVID-19 pandemic. Through various partnerships, the COVID 3D Trust was able to help assess whether or not different open source PPE devices were of sufficient quality to be used in clinical or community settings. In addition, the COVID 3D Trust helped maker communities distribute PPE (McCarthy et al., 2021; Bowser et al., 2021).

Many social structures for low-cost tools, including communities of practice, knowledge commons, and other structures, strive to be inclusive of all tool and data users, including those that lack institutional connections, resources to attend meetings or otherwise participate in community activities, or familiarity with social dynamics and groups. Significant infrastructure, including financial and in-kind support for technologies, community management tasks, and inclusion, is required to help these knowledge commons thrive.

In addition to the social support structures mentioned above—which often emerged out of low-cost tool communities—a number of already-existing support structures might expand to accommodate low-cost tool perspectives.

Case Study 4: The Gathering for Open Science Hardware (GOSH)

Individuals and groups around the world were developing and sharing open hardware designs for science, however, no official network existed until the first Gathering for Open Science Hardware (GOSH) in 2016. With attendees representing perspectives from academic research, education, community-based work, and entrepreneurial backgrounds, this initial gathering led to the development of a set of core values (the GOSH Manifesto) for open science hardware and the formalization of a community dedicated to making open science hardware ubiquitous by 2025 (GOSH Manifesto, n.d.).

Figure 13

Image Source: "GOSH 20181012-(Day3-GROW)-5.X 6.Factory reception -Social-Taken by Laura Olalde (144)" by GOSH Community is licensed under CC0 1.0

GOSH members use the GOSH forum and blog to discuss activities and opportunities, share skills, and discuss issues around key topics like licensing. The GOSH community developed a roadmap in 2017 and an action plan in 2018, both of which outline concrete strategies the community can take to increase open hardware's impact on science (Gathering for Open Science Hardware, 2018). The presence of the GOSH network has helped advance the use of open source hardware in many sectors, including community science and academia.
In addition to global gatherings, GOSH hosts targeted workshops to develop skill sets or address certain issues. Outcomes are shared online to seed further collaboration and solicit

feedback from members, as well as to amplify the reach of the materials and to build a member base.

In addition to sharing knowledge, the GOSH community helps open science hardware gain legitimacy in the eyes of scientific institutions (Arancio, 2020). For example, to gain trust from academia, GOSH develops research publications that demonstrate the quality performance of open science hardware. The community has also developed use cases for academic, community, and media audiences. GOSH now has members in universities, other research institutions, non-profits, and policy communities.

GOSH is a social structure that builds legitimacy, shares knowledge, and mobilizes resources to promote both the development and use of open science hardware. Supporting both general-purpose communities, such as GOSH, as well as platforms specific to a particular research discipline, can support data quality outcomes by providing spaces for discussion and targeted work around the tools themselves, their digital environment, and the regulatory landscape, among other topics.

Conclusion

Low-cost tools have the potential to accelerate scientific discovery, collect new types of data to answer diverse research questions, and broaden participation in science to include new stakeholder groups. However, the current misconception that low-cost tools produce lower quality data still serves as a barrier to their widespread uptake and adoption.

Low-cost tools typically reshape several dimensions of data such as amount, type, and accuracy. Drops in one dimension may be compensated for by gains in a different dimension. Thus low-cost tools do not necessarily produce lower quality data, instead they change data quality. Furthermore, low-cost tools can compare in performance to more expensive tools and have additional benefits, such as the collection of new types of data, participation of new communities in science, and improved science infrastructure. Misconceptions can be shifted by re-defining data quality as multidimensional and dynamic through fitness for use.

Consequently, key dimensions of data quality can be effectively communicated and assessed using a "fitness for use" perspective.

Using fitness for use as an overarching framework, we have identified opportunities for elevating data quality to ensure fitness for diverse uses that focus on: improving the tools themselves; building out a digital environment, including one that considers hardware-software dependencies; ensuring that standards and the regulatory context do not unduly limit the use of low-cost tools and data; and exploring how a social context facilitates knowledge sharing and trusted communication across users and uses.

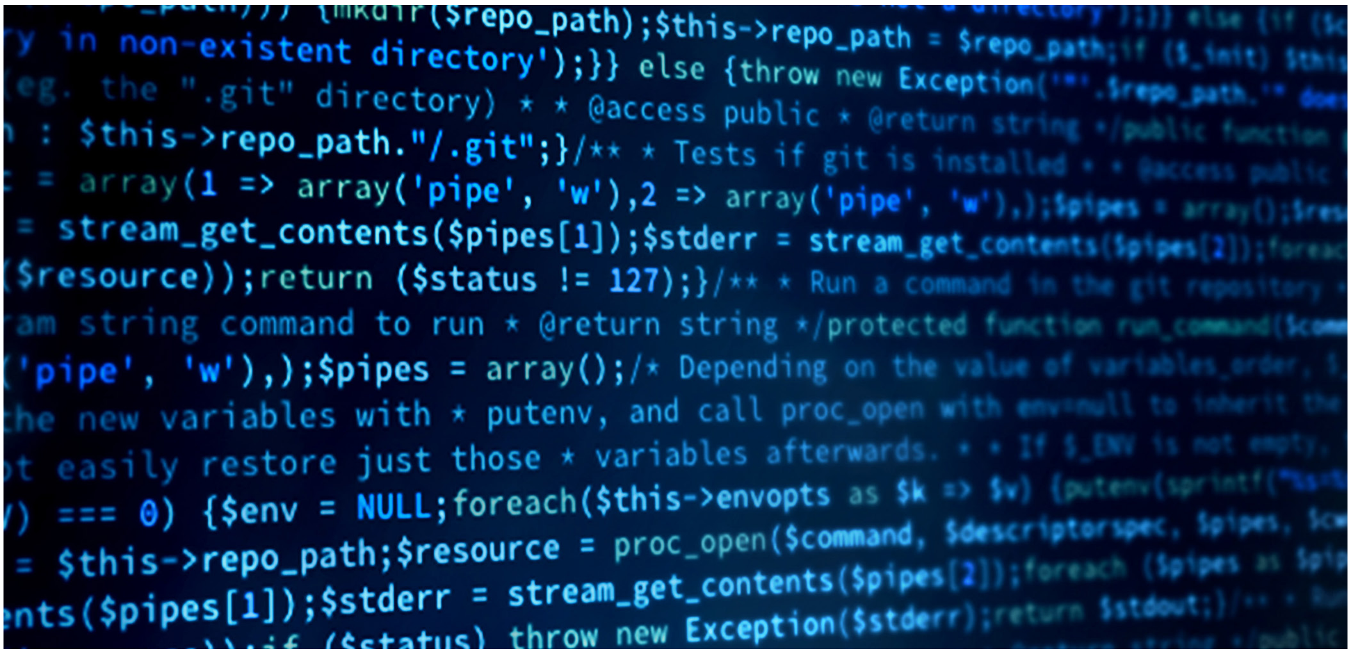
Recognizing these opportunities, there are a number of actions that government and policy audiences can take to support these data quality solutions and elevate the value of low-cost tools:

1) **Prioritize the use of open source hardware—tools open in their design and open to being changed and shared.** From the perspective of data quality and fitness for use, open source hardware has a number of benefits over low-cost, proprietary equivalents. Open practices make understanding, reviewing, and replicating tools and experimental protocols more easily accessible through publicly available design files. With more participation in and access to information on scientific tools, experiments can be better replicated, resulting in improved scientific reproducibility. In addition, open source hardware is licensed so that tools can be modified; this can improve mismatch in tool availability and data quality as open source hardware can be customized to fit a particular intended use.

2) **Invest in accessible infrastructure to improve the digital environment of low-cost tools.** A tool's particular digital environment is often provided by increasingly sophisticated software models and accessible data repositories to support data quality assurance. Making different software models more accessible, such as publishing them as open source, is one opportunity for elevating the value of low-cost and open source tools across a range of scientific research domains. Existing data aggregation facilities or data centers could also clarify opportunities for (re)publishing open and FAIR data from low-cost tools. In particular, promoting the use of FAIR principles—including a more complete documentation along the data life cycle--can build trust across a range of stakeholders. It may also be helpful for existing groups providing data management

guidelines to citizen and community based groups, such as the NASA community working on citizen science data management and documentation, to explicitly consider relevance to stakeholders using low-cost tools. U.S. federal agencies can also partner with non-governmental groups to help develop infrastructures for generating and elevating the value of data from low-cost tools. For example, Fundación Española para la Ciencia y Tecnología (Spanish Foundation for Science and Technology) helped improve iNaturalist's open source software.

Figure 14



Source Image: Rigon, Dario. Software background. Website codes on computer monitor. Software development. Programmer developer screen. Machine learning code. Java code. Php code. Linux code. n.d. shutterstock.com

3) Create flexible assessment processes that take the burden off tool users and recognize a range of desired use cases. The development of tiered and/or technology agnostic standards and regulatory frameworks and guidelines can make assessing fitness for use of low-cost tools more accessible to different intended use cases. Of importance is the inclusion of the full range of user communities in the development of standards. International standards organizations such as ISO and IEEE could play a role in spearheading some initiatives. Although it is important that these evaluations and assessments are accessible to tool users, supporting third party assessment of "formal" communities, in government agencies and "informal" community groups, can take the burden of assessment off tool users who may not have the resources or expertise to assess data quality. In addition, beyond formal assessment processes such as regulatory standards, the scientific and technical community should recognize the value of academic and community-based peer review processes for evaluating data quality and fitness for use.

4) **Support social structures that ensure access to low-cost tools for all those interested in doing research, including those with limited resources.** The uptake of a tool depends on social networks—informal and formal—of users, supporters, promoters, and regulators among others. With adequate governance and support, these social networks can accelerate the development of trust around these tools. Policy communities and regulators who are data users can engage with low-cost tool data stewards and intermediaries to communicate data quality requirements for different uses. Supporting existing social structures that strive to be inclusive of limited resource groups, such as professional networks and online platforms can foster trust for a range of low-cost tool users and data users. Leveraging other existing social support structures, including extension programs and Digital.gov communities of practice, can also help disciplinary and government communities discuss key aspects of low-cost tools that are relevant to their work.

The use of low-cost tools in scientific research is contingent upon trust in the quality of data they produce. By developing new, flexible ways to define, analyze, communicate, and build community around data quality and providing the resources and infrastructure to do so, the challenges of using low-cost tools in scientific research can be addressed. With these solutions in place, low-cost tools can facilitate more equitable and innovative ways of doing science.

About the Authors

Mark Chandler, PhD is a scientist, program director and educator with 25 years developing and sustaining social and environmental sustainability programs. For over 20 years, Mark designed and implemented community science programs at Earthwatch Institute resulting in high impact scientific publications, conservation outcomes and the engagement of the public in field research. Most recently, he has worked on community-supported sensor networks to assess air temperature, ozone and particulate matter in southern California, Massachusetts and India. His work seeks innovative solutions to environmental challenges at scale with a focus on improving the health of communities and environment. Mark holds a Ph.D. from McGill University and a B.Sc. from University of Guelph.

Alison Parker, PhD serves as a Senior Program Associate with the Science and Technology Innovation Program (STIP) at the Woodrow Wilson International Center for Scholars. With STIP, Alison evaluates and amplifies emerging approaches to science and technology, including low-cost and open source hardware and citizen science. Previously, Alison served on the Board of Directors for the Citizen Science Association, including as Chair. Prior to joining the Wilson Center, Alison held a fellowship in the Office of Research and Development at the US Environmental Protection Agency. Alison received her B.Sc. in Biological Sciences from the George Washington University and PhD in Ecology and Evolutionary Biology from the University of Toronto.

Alexandra Novak is a Staff Research Intern with the Science and Technology Innovation Program (STIP) at the Woodrow Wilson International Center for Scholars. Alexandra conducts research on low-cost and open hardware for the Thing Tank Initiative. Previously, she worked in microscopy and held a Fulbright research fellowship in Paraguay. Alexandra received her B.Sc. in Chemistry from Union College and is currently a NSF Graduate Research Fellow in environmental engineering at Columbia University.

Ashley Schuett is a Staff Assistant Intern with the Science and Technology Innovation Program (STIP) at the Wilson Center. Ashley is conducting research on open science and its intersection with cybersecurity with her past work as a software engineer in open-source software informing her perspective. Her interests lie in privacy and equitability in emerging technology.

Alex Long is a Program Associate for the Science and Technology Innovation Program (STIP) at the Wilson Center. Alex's work centers around the interplay between science policy and society. His primary focuses are global health and open science.

Anne Bowser, PhD is a Global Fellow at the Wilson Center, the former Deputy Director and Director of Innovation with the Science and Technology Innovation Program (STIP). Her work investigates the intersections between science, technology and democracy.

References

Air Quality Sensor Performance Evaluation Center. (n.d.). South Coast Air Quality Monitoring District. https://www.aqmd.gov/aq-spec/home

Amos, H.M., Andersen, T., Arendt, A., Byrnes, J., Clark, M., Dallas, L., … & Wilson, B. (2020). NASA ESDS Citizen Science Data Working Group White Paper. *Earth Science Data Systems*. https://cdn.earthdata.nasa.gov/conduit/upload/14273/CSDWG-White-Paper.pdf

Arancio, J.C. (2020). Opening up the tools for doing science: The case of the Global Open Science Hardware Movement. PrePrint. Submitted to *International Journal of Engineering, Social Justice & Peace*, Review in progress. https://osf.io/preprints/socarxiv/46keb/

Arce, W., & Stevens, J. R. (2020). Developing a Computer-Controlled Treat Dispenser for Canine Operant Conditioning. *Journal of Open Hardware, 4(1).* http://doi.org/10.5334/joh.27

AudioMoth. (n.d.). Open Acoustic Devices. https://www.openacousticdevices.info/audiomoth

Boone, M. E., & Basille, M. (2019). Using iNaturalist to contribute your nature observations to science. EDIS, 2019(4), 5-5. https://doi.org/10.32473/edis-uw458-2019

Bowden, R., Davies, R. W., Heger, A., Pagnamenta, A. T., de Cesare, M., Oikkonen, L. E., ... & Donnelly, P. (2019). Sequencing of human genomes with nanopore technology. *Nature communications, 10(1), 1-9.* https://doi.org/10.1038/s41467-019-09637-5

Bowser, A., Cooper, C., De Sherbinin, A., Wiggins, A., Brenton, P., Chuang, T. R., ... & Meloche, M. (2020). Still in need of norms: the state of the data in citizen science. *Citizen Science: Theory and Practice*, 5(1). https://doi.org/10.5334/CSTP.303

Bowser, A., Long, A., Novak, A., Parker, A., & Weinberg, M. (2021). Stitching Together a Solution: Lessons from the Open Source Hardware Response to COVID-19. T*he Wilson Center.* https://www.wilsoncenter.org/publication/stitching-together-solution-lessons-open-source-hardware-response-covid-19

Cadman, M. & González-Talaván, A (eds.). (2014). Publishing Camera Trap Data: A Best Practice Guide. *Global Biodiversity Information Facility.* https://www.gbif.org/document/1o6HNHuCxKaiAC8yG86gQq/publishing-camera-trap-data-a-best-practice-guide

Collins, J. T., Knapper, J., Stirling, J., Mduda, J., Mkindi, C., Mayagaya, V., ... & Bowman, R. (2020). Robotic microscopy for everyone: the OpenFlexure microscope. *Biomedical Optics Express, 11*(5), 2447-2460. https://doi.org/10.1364/BOE.385729

Definition (English). (n.d.). In Open Source Hardware Association. Retrieved July 20, 2021, from https://www.oshwa.org/definition

Delmans, M., & Haseloff, J. (2018). µCube: a framework for 3D printable optomechanics. *Journal of Open Hardware, 2*(1). http://doi.org/10.5334/joh.8

Doraiswamy, P. (n.d.). *Can Citizen Science and Low-Cost Sensors Help Improve Earth System Data?.* NASA EarthData. https://earthdata.nasa.gov/esds/competitive-programs/csesp/improve-earth-system-data

Duvall, R. M., Hagler, G. S. W., Clements, A. L., Benedict, K., Barkjohn, K., Kilaru, V., ... & Snyder, J. L. (2021). Deliberating Performance Targets: Follow-on workshop discussing PM10, NO2, CO, and SO2 air sensor targets. *Atmospheric Environment,* 246, 118099. https://doi.org/10.1016/j.atmosenv.2020.118099

United States Environmental Protection Agency. (n.d.). *Evaluation of Emerging Air Sensor Performance.* Air Sensor Toolbox. https://www.epa.gov/air-sensor-toolbox/evaluation-emerging-air-sensor-performance

FAIR Principles. (n.d.). GoFAIR. Retrieved July 21, 2021 from https://www.go-fair.org/fair-principles/

Gabrys, J., Pritchard, H., & Barratt, B. (2016). Just good enough data: Figuring data citizenships through air pollution sensing and data stories. *Big Data & Society, 3*(2), 2053951716679677. https://doi.org/10.1177%2F2053951716679677

Gathering for Open Science Hardware. (2018). Global Open Science Hardware Roadmap. *Gathering for Open Science Hardware.* https://openhardware.science/global-open-science-hardware-roadmap/

Garmendia, O., Rodríguez-Lazaro, M. A., Otero, J., Phan, P., Stoyanova, A., Dinh-Xuan, A. T., ... & Farré, R. (2020). Low-cost, easy-to-build noninvasive pressure support ventilator for under-resourced regions: open source hardware description, performance and feasibility testing. *European Respiratory Journal, 55*(6). https://doi.org/10.1183/13993003.01271-2020

Global Biodiversity Information Facility. (n.d.). *Free and open access to biodiversity data.* Global Biodiversity Information Facility. https://www.gbif.org/

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial statistics, 1*, 110-120. https://doi.org/10.1016/j.spasta.2012.03.002

GOSH Manifesto. (n.d.). Gathering for Open Science Hardware. Retrieved July 22, 2021 from https://openhardware.science/gosh-manifesto/

Hill, A. (2021, January 15). *Open hardware provides a better chance for scientific reproducibility.* Journal of Open HW Medium. https://journalopenhw.medium.com/open-hardware-provides-a-better-chance-for-scientific-reproducibility-829ed23bb71d

iNaturalist. (2015, August 14). *Changes to Quality Grade.* Tumblr. https://inaturalist.tumblr.com/post/126691814973/changes-to-quality-grade

International Standards Organization. (n.d.). International Standards Organization. Retrieved July 21, 2021 from https://www.iso.org/home.html

Institute of Electrical and Electronics Engineers Standards Association. (n.d.). Institute of Electrical and Electronics Engineers Standards Association. https://standards.ieee.org/

Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment, 14*(10), 551-560. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/fee.1436

Levy Zamora, M., Xiong, F., Gentner, D., Kerkez, B., Kohrman-Glaser, J., & Koehler, K. (2018). Field and laboratory evaluations of the low-cost plantower particulate matter sensor. *Environmental science & technology,* 53(2), 838-849. https://doi.org/10.1021/acs.est.8b05174

Longhitano, G. A., Nunes, G. B., Candido, G., & da Silva, J. V. L. (2020). The role of 3D printing during COVID-19 pandemic: a review. *Progress in Additive Manufacturing*, 1-19. https://doi.org/10.1007/s40964-020-00159-x

Maia Chagas, A. (2018). Haves and have nots must find a better way: The case for open scientific hardware. *PLoS biology, 16*(9), e3000014. https://doi.org/10.1371/journal.pbio.3000014

McCarthy, M., Di Prima, M., Cruz, P., Ribic, B., Wilczynski, J., Ripley, B., & Coburn, J. (2021, November 10). Trust in the time of covid-19: 3D Printing and Additive Manufacturing (3DP/AM) as a solution to supply chain gaps. NEJM Catalyst Innovations in Care Delivery. Retrieved December 1, 2021, from https://catalyst.nejm.org/doi/full/10.1056/CAT.21.0321.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., ... & Williams, R. (2018). Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?. *Environment international, 116,* 286-299. https://dx.doi.org/10.1016%2Fj.envint.2018.04.018

Mueller, T., Elkaseer, A., Charles, A., Fauth, J., Rabsch, D., Scholz, A., ... & Scholz, S. G. (2020). Eight weeks later—the unprecedented rise of 3D printing during the COVID-19 pandemic—a case study, lessons learned, and implications on the future of global decentralized manufacturing. *Applied Sciences, 10*(12), 4135. https://doi.org/10.3390/app10124135

Niezen, G., Eslambolchilar, P., & Thimbleby, H. (2016). Open-source hardware for medical devices. *BMJ innovations, 2*(2). http://dx.doi.org/10.1136/bmjinnov-2015-000080

Ottinger, G. (2010). Buckets of resistance: Standards and the effectiveness of citizen science. *Science, Technology, & Human Values, 35*(2), 244-270. https://doi.org/10.1177%2F0162243909337121

Parada-Sánchez, S.G., Meléndez-Salcido, C.G., Hernández-Castaños, M.R., Prado-Ávila, S.R., & Adame-Gallegos, J.R. (2018. Evaluation of Foldscope, a paper-based origami microscope useful for taxonomic identification of Rhipicephalus sanguineus ticks. *Acta universitaria 28*(4), 19-24. https://doi.org/10.15174/au.2018.2134

Parker, A. & Dosemagen, S. (2016). Environmental Protection Belongs to the Public: A Vision for Citizen Science at EPA. *National Advisory Council for Environmental Policy and Technology (NACEPT).* https://www.epa.gov/sites/production/files/2020-04/documents/nacept_cs_report_final_508.pdf

Parker, A., Dosemagen, S., Molloy, J., Bowser, A. & Novak, A. (2021). Open Hardware: An Opportunity to Build Better Science. *The Wilson Center.* https://www.wilsoncenter.org/publication/open-hardware-opportunity-build-better-science

Parker, A. & Novak, A. (2020). Building Blocks for Better Science: Case Studies in Low-Cost and Open Tools for Science. *The Wilson Center.* https://www.wilsoncenter.org/sites/default/files/media/uploads/documents/STIP%20THING%20Tank%20Building%20Blocks%20for%20Better%20Science.pdf

Pearce, J.M. (2017). Impacts of open source hardware in science and engineering. *The Bridge.* https://hal.archives-ouvertes.fr/hal-02111398

Pearce, J.M. (2020). Economic savings for scientific free and open source technology: A review. *HardwareX, 8*, e00139–e00139. https://doi.org/10.1016/j.ohx.2020.e00139

PocketLab: Science Everywhere. (n.d.). The PocketLab. Retrieved July 20, 2021, from https://www.thepocketlab.com/

Public Lab. (n.d.). *About Public Lab.* Public Lab. https://publiclab.org/about

Purple Air. (n.d.). *PurpleAir: Real-time Air Quality Monitoring.* PurpleAir. https://www2.purpleair.com/

Resources. (n.d.). Wildlabs.net. Retrieved July 21, 2021 from https://wildlabs.net/resources

Safecast. (n.d.). Safecast. Retrieved July 21, 2021 from https://safecast.org/

Science Stack: Tools within Reach. (n.d.). The Wilson Center. Retrieved July 20, 2021 from https://www.wilsoncenter.org/sciencestack

Shang, L., Heckelei, T., Gerullis, M. K., Börner, J., & Rasch, S. (2021). Adoption and diffusion of digital farming technologies-integrating farm-level evidence and system interaction. *Agricultural Systems, 190*, 103074. https://doi.org/10.1016/j.agsy.2021.103074

Sharma, S., & Mangat, V. (2015, February). Technology and trends to handle big data: Survey. In *2015 Fifth International Conference on Advanced Computing & Communication Technologies* (pp. 266-271). IEEE. https://doi.org/10.1109/ACCT.2015.121

Stirling, J., Sanga, V. L., Nyakyi, P. T., Mwakajinga, G. A., Collins, J. T., Bumke, K., ... & Bowman, R. (2020). The OpenFlexure Project. The technical challenges of Co-Developing a microscope in the UK and Tanzania.

In *2020 IEEE Global Humanitarian Technology Conference (GHTC)* (pp. 1-4). IEEE. https://doi.org/10.1109/GHTC46280.2020.9342860

Stroud Water Research Center. (n.d.). *Getting Started With the Mayfly Data Logger*. EnviroDIY. https://www.envirodiy.org/mayfly/

Switz, N. A., D›Ambrosio, M. V., & Fletcher, D. A. (2014). Low-cost mobile phone microscopy with a reversed mobile phone camera lens. *PloS one, 9*(5), e95330. https://doi.org/10.1371/journal.pone.0095330

ThermoFisher Scientific. (n.d.). *What is Next-Generation Sequencing (NGS)?*. ThermoFisher Scientific. https://www.thermofisher.com/us/en/home/life-science/sequencing/sequencing-learning-center/next-generation-sequencing-information/ngs-basics/what-is-next-generation-sequencing.html

Wang, G., Lu, Q., & Capareda, S. C. (2020). Social network and extension service in farmers' agricultural technology adoption efficiency. Plos one, 15(7), e0235927. https://doi.org/10.1371/journal.pone.0235927

Wiggins, A., Newman, G., Stevenson, R. D., & Crowston, K. (2011, December). Mechanisms for data quality and validation in citizen science. In *2011 IEEE seventh international conference on e-Science Workshops* (pp. 14-19). IEEE. https://doi.org/10.1109/eScienceW.2011.27

Williams, R., Duvall, R., Kilaru, V., Hagler, G., Hassinger, L., Benedict, K., ... & Ning, Z. (2019). Deliberating performance targets workshop: Potential paths for emerging PM2. 5 and O3 air sensor progress. *Atmospheric Environment: X, 2,* 100031. https://doi.org/10.1016/j.atmosenv.2020.118099

Yong, E. (2019, August 22). *An Ingenious Microscope Could Change How Quickly Disease Is Detected*. The Atlantic. https://www.theatlantic.com/science/archive/2019/08/cheap-automatic-microscope-could-change-how-diseases-are-detected/596440/

Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2013). Sensing as a service and big data. *arXiv preprint arXiv:* 1301.0159. https://arxiv.org/abs/1301.0159

Zooniverse Talk. (n.d.). Zooniverse. Retrieved July 21, 2021 from https://www.zooniverse.org/talk

List of Images:

Figure 5. OpenFlexure is an open source microscope designed so it can be 3D printed and customized around the world.: https://openflexure.org/assets/MicroscopeBlenderTrio.png

Figure 6. TheDigitalEnvironment.png

Figure 7. https://www.shutterstock.com/image-photo/lion-camera-trap-314483288

Figure 8. TheRegulatoryContext.png

Figure 9. Air quality sensing balloons developed at a hackathon.: https://www.flickr.com/photos/25648105@N03/5603187711

Figure 10. Plantower is a building block of low-cost air quality sensors.: https://www.flickr.com/photos/35434449@N08/27678963859

Figure 11. TheSocialContext.png

Figure 12. Image of Public Lab's community satellites used to map the BP oil spill. https://www.flickr.com/photos/eustatic/14039110521/in/photostream/

Figure 13. Members of the GOSH Community gather at a global event.: https://flickr.com/photos/goshcommunity/30631136007/

Figure 14. Supporting the development of machine learning, artificial intelligence and other software techniques can help enhance data quality of low-cost tools.: https://www.shutterstock.com/image-photo/software-background-website-codes-on-computer-507234217

WOODROW WILSON INTERNATIONAL CENTER FOR SCHOLARS

The Woodrow Wilson International Center for Scholars, established by Congress in 1968 and headquartered in Washington, D.C., is a living national memorial to President Wilson. The Center's mission is to commemorate the ideals and concerns of Woodrow Wilson by providing a link between the worlds of ideas and policy, while fostering research, study, discussion, and collaboration among a broad spectrum of individuals concerned with policy and scholarship in national and international affairs. Supported by public and private funds, the Center is a nonpartisan institution engaged in the study of national and world affairs. It establishes and maintains a neutral forum for free, open, and informed dialogue. Conclusions or opinions expressed in Center publications and programs are those of the authors and speakers and do not necessarily reflect the views of the Center staff, fellows, trustees, advisory groups, or any individuals or organizations that provide financial support to the Center.

THE SCIENCE AND TECHNOLOGY INNOVATION PROGRAM (STIP)

The Science and Technology Innovation Program (STIP) brings foresight to the frontier. Our experts explore emerging technologies through vital conversations, making science policy accessible to everyone.

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027

The Wilson Center

www.wilsoncenter.org
wwics@wilsoncenter.org
facebook.com/woodrowwilsoncenter
@thewilsoncenter
202.691.4000

STIP

www.wilsoncenter.org/program/science-and-technology-innovation-program
stip@wilsoncenter.org
@WilsonSTIP
202.691.4321

Wilson Center

Science and Technology Innovation Program