5
Technology infrastructure for citizen science

Peter Brenton[1], Stephanie von Gavel[1], Ella Vogel[2] and Marie-Elise Lecoq[3]

[1] *Atlas of Living Australia, CSIRO, Canberra, Australia*
[2] *National Biodiversity Network Trust (NBN), Nottingham, UK*
[3] *GBIF France, MNHN Géologie, Paris, France*

corresponding author email: Peter.Brenton@csiro.au

In: Hecker, S., Haklay, M., Bowser, A., Makuch, Z., Vogel, J. & Bonn, A. 2018. *Citizen Science: Innovation in Open Science, Society and Policy*. UCL Press, London. https://doi.org/10.14324/111.9781787352339

Highlights

- Information technology (IT) infrastructure is a vital enabler of successful citizen science projects.
- There are numerous IT tools available to citizen science projects and navigating them can be confusing. When choosing tools, it is important to consider their compliance with applicable process and data standards, their ability to connect with the information supply chain and their fitness for the required use.
- The information and data generated by citizen science projects is likely to be their most enduring and impactful legacy if they are made publicly accessible in a timely manner and in a form which is suitable for multiple downstream uses. To do this, they need to conform as much as possible to existing data and process standards.

Introduction

The chapter considers what infrastructure means in a citizen science context and characterises the types of technology-based infrastructure being used by the global citizen science community, with a focus on the environmental domain. Some issues emerging around the application of different infrastructure solutions in current use are also raised and, using some examples and case studies, existing infrastructure solutions are discussed in an 'information supply chain' framework. An information supply chain refers to the process flow or movement of a piece of information (data) from being acquired or collected, to being used in one or more transformative actions such as policy settings, physical management and/or educational or behavioural change campaigns. Invariably, this will also involve intervening processes on the data, potentially by parties other than the collectors, such as data curation, management, aggregation and analysis.

This chapter draws on the authors' experience and expertise in citizen science infrastructure in Australia and primarily in the environmental domain.

The notion of 'best practice' in the context of citizen science infrastructure is also considered, concluding that 'best practice' is relative to available solutions and practices at a given time and that it will inevitably change over time.

What is citizen science infrastructure?

The online version of the Merriam-Webster Dictionary defines 'infrastructure' as: 'the underlying foundation or basic framework (as of a system or organisation); . . . *and*: the resources (as personnel, buildings, or equipment) required for an activity'. Thus infrastructures are the physical structures, equipment and tools, processes, services, human capital and social networks which enable systems and enterprises to function effectively. In a citizen science context, this includes:

a. Physical kit – buildings, vehicles, telescopes, microscopes and binoculars, measuring instruments, cameras, scanners, sensors, drones and various other equipment;
b. Social assets – the organisers of projects, events and collaboration services, sponsors and funding bodies, the public participants in projects and events, and the social networks of connected individuals; and
c. Technology assets – the information technology–based platforms/tools and services used to collect, store, manage and process, share, visualise and analyse information (data and metadata) which is produced by citizen science endeavours, as well as those used to organise and manage citizen science projects and events.

This chapter deals only with the information technology–based infrastructures which support data produced by citizen science endeavours, not those used for stakeholder and event management (see also Wyler & Haklay in this volume on the infrastructure provided by universities).

Historically, data generated by citizen science projects – and indeed many non-citizen science projects, too – was often only used within the context of the project for which it was collected. However, aggregated data from multiple sources is becoming increasingly important as features of research work and as inputs to policy and management actions. It is therefore also useful to consider IT infrastructures which support citizen science in a broader context; that is, the role they play in the information supply chain. This helps us to understand the relevance and role of individual projects in contributing to new knowledge and improved management outcomes, and hence the significant role of public participation in this larger context.

Figure 5.1 shows a conceptual information supply chain model in which citizen science projects are involved in data acquisition and analysis processes. People use all sorts of tools and infrastructure to collect raw data which gets stored somewhere, usually in local databases or cloud services. However, raw data by itself has little intrinsic value or usefulness – raw data only has value and meaning when it is interpreted in conjunction with the context in, and by which, it was collected (see also Williams et al. in this volume).

The reasons for collecting raw data are many and varied, and include:

- Answering specific research questions or modelling and understanding real-world processes;
- Support social, political, environmental or economic objectives;
- Gaining personal satisfaction and fulfilment;
- Enhancing social opportunities;
- Connecting with nature; and much more.

Data aggregators procure and combine raw data from those who collect and produce it, and provide repositories in which data producers can proactively lodge their data. Aggregators typically transform inbound data to fit into a standardised data structure and add value to the raw data by providing a range of products and services to data producers and consumers.

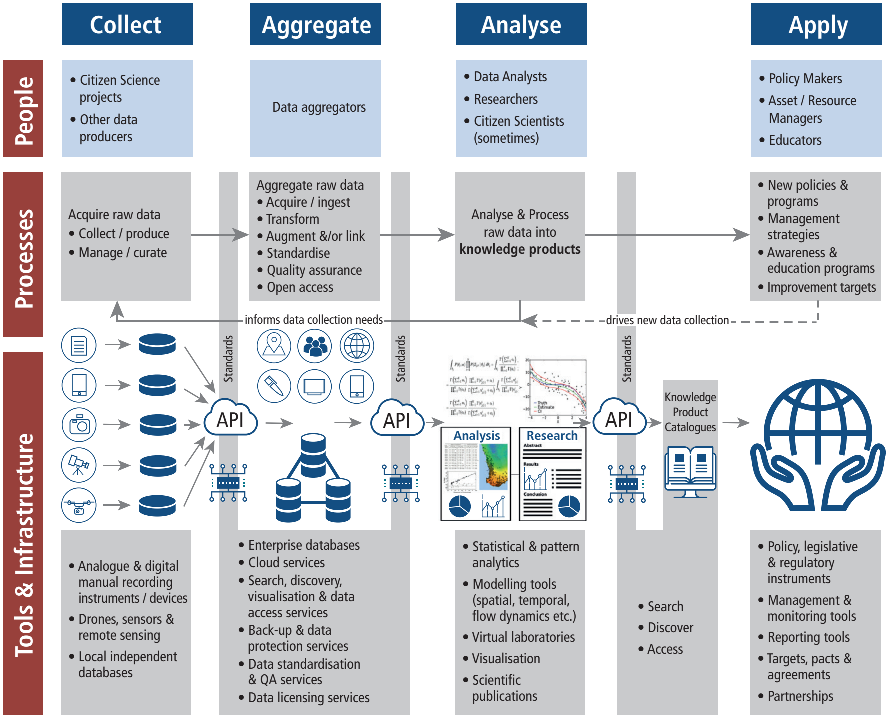

Fig. 5.1 A conceptual model for a digital information supply chain. (Source: Icon made from http://www.onlinewebfonts.com/icon fonts is licensed under CC BY 3.0)

Aggregated data is then accessed by data analysts and researchers who use tools and expertise to gain meaning and knowledge. This results in knowledge products, which can then be used to inform policy, planning and management decisions, facilitate assessment against national and international benchmarks and target measures, and many other applications. Information technology infrastructure is also used to make knowledge products more discoverable and accessible.

Policies and management actions invariably have impacts which require measurement and monitoring, which in turn drives further raw data collection. Outputs from analysis can also identify gaps in information and stimulate further focused raw data collection.

Data and procedural standards provide a common language which allows similar information from disparate sources to be efficiently aggregated and exchanged, thus giving raw data potential value, utility and impact beyond the purpose for which it was originally collected. Application interfaces (APIs) provide a simple mechanism for exchanging data between different electronic systems, facilitated by growing access to high-speed internet technologies. These are becoming increasingly important enablers by supporting 'linked data' and 'big data' approaches to understanding the complexity of the world and informing policy and management responses to complex global challenges (Ceccaroni, Bowser & Brenton 2017; Ottinger 2010).

Information technology infrastructure plays a significant and important role in the information supply chain by supporting human interactions with data, as well as enforcing standards, automating processes and performing computational functions (Wiggins et al. 2011; Newman et al. 2012), such as:

- Connecting and linking system components;
- Standardising data definitions so that there is a shared data language;
- Mobilising data from analogue (non-digital) and siloed digital systems into standardised digital formats which can be transported through and used by all tiers of the information supply chain;
- Supporting data curation and data quality improvement;
- Improving data flow and processing efficiency; and
- Much more.

Citizen science and IT infrastructure – a natural partnership

Internet and wireless technologies are enabling unprecedented access to scientific materials and facilitating mass public participation in science (Couvet et al. 2008; Hochachka et al. 2012; Newman et al. 2010).

Information technology platforms can codify and enforce rules and processes which help to improve the quality and hence the reliability, reusability and scientific trustworthiness of information generated through non-traditional scientific channels. Technology infrastructures are therefore an important enabler of citizen science and are arguably the single biggest factor driving the recent rise of citizen science and the democratisation of science generally (Nov, Arazy & Anderson 2011a).

However, it is difficult to keep pace with the constant and rapid changes in technology. Such changes generally bring improvements in usability, functionality, performance, reliability, accessibility, accuracy and precision, as well as new beneficial features. At the same time, they introduce more and potentially confusing options, creating a potentially bewildering technology landscape for citizen science project co-ordinators. The cost of hardware and sometimes software can also impede a project's uptake and benefits. Later, this chapter looks at some things to consider when choosing an IT solution for a project.

As evidenced elsewhere in this book and in many other published works, citizen science is a significant public good endeavour which provides numerous social, environmental and economic benefits in addition to enhancing science engagement and literacy. The important role of IT infrastructure in supporting citizen science makes it reasonable to consider issues such as:

- The role governments, non-governmental organisations (NGOs) and philanthropic organisations should play in facilitating access to, and reducing the cost of, technology infrastructures for citizen science;
- How citizen scientists can access and make the most effective use of technology to improve the efficiency, effectiveness, accuracy and impacts of their contributions to scientific endeavour; and,
- How technology can be used to demonstrate the impact of citizen science contributions on social, policy and management outcomes, and thus empower and enhance the engagement of the public in these areas, as well as to improve recognition of citizen science contributions in traditional science and policy circles.

Such questions are being addressed in numerous studies around the world such as Bonter & Cooper (2012); Couvet et al. (2008); Nov, Arazy & Anderson (2011a); Sequeira et al. (2014); Kaartinen et al. (2013); and many others.

The recent worldwide explosion in the number and scope of citizen science projects has seen a growing need to develop effective mechanisms to assist the public in finding, discovering and connecting with citizen science projects; and for project owners to promote and connect their projects with citizen scientists. This has resulted in the emergence of several independently developed 'Project Finders' – searchable project catalogues. Some of the open public facilities have become channels for citizen science projects to promote themselves, but with worldwide access and broadly similar functionality, they are sometimes perceived as competing with each other, which has led to some community confusion as to where they should register their projects. In an ideal world, any citizen science project registered in any catalogue system should be discoverable and accessible via any project finder – thus giving people the most comprehensive and current information possible about projects at their location, and allowing them to directly connect to projects of interest. To achieve this however, information needs to be shared between systems using common standards and protocols as described by Ceccaroni, Bowser & Brenton (2017) (and see also Williams et al. in this volume). To this end some key public catalogue managers are collaborating to develop a standard core set of data attributes for citizen science projects, as well as a standard data schema and data exchange protocols, known as the PPSR-Core project.

Information technology platforms, both desktop and mobile, facilitate vast networks of human observers, stationary autonomous sensors (e.g., camera traps, weather and environmental sampling stations, etc.), and mobile remote platform sensors (e.g., drones and satellites) to collect reasonably consistent quality spatial and temporal data. This enables large-scale spatial and temporal analyses of patterns and distributions which would otherwise be impossible using traditional scientific data collection methods (Sullivan et al. 2009). Some successful early examples of these in citizen science, such as eBird (http://ebird.org/) and Galaxy Zoo (https://www.galaxyzoo.org/), have become benchmarks for large-scale global citizen science programmes.

There are many tools currently available (box 5.1) and many more are likely to emerge in the future. This chapter does not endorse particular tools, but instead aims to illustrate the complex array of tools available. All tools have strengths and weaknesses and differ in their suitability for different projects and situations. In addition, significant gaps remain where infrastructure is not yet fully servicing the scope of requirements for technology support in the citizen science domain – for example, species identification in the biodiversity domain and portals focused on communities of interest more generally.

When choosing a tool, there are important factors to consider:

i. Is there an existing tool available at an acceptable cost? Why build a new tool when something suitable already exists or can be adapted to fit?
ii. Is the tool already connected or designed to connect and share with open data infrastructures? Most tools do not do this, but it is critical for data sharing.
iii. Are the data capture and storage structures compliant with domain-relevant standards? Most are not, and this is also critical for data sharing.

Box 5.1. Citizen science infrastructure tools

The citizen science sector has produced an impressive array of tools operating at varying spatial scales, as well as with different temporal scopes and topics of interest. These can be broadly categorised as follows:

1. **Project catalogues/finders** provide a central point of discovery and connection to citizen science projects. Examples include: CitSci.org (www.citsci.org); SciStarter (www.scistarter.com); Federal Crowdsourcing and Citizen Science Catalog (https://ccsinventory.wilsoncenter.org/); Zooniverse (https://www.zooniverse.org/projects?status=live); EU BON (http://biodiversity.eubon.eu/zh/web/citizen-science/view-all); and BioCollect (https://biocollect.ala.org.au/acsa). These facilities also support community engagement and, in some cases, data collection services. There are also commercial providers serving the citizen science community with data recording capabilities and small project catalogues.

   In addition, organisations which fund/sponsor projects often monitor their progress and have their own project catalogues, examples include: the European Commission's 'CORDIS' system (http://cordis.europa.eu/project/rcn/51266_en.html); the Alfred P. Sloane Foundation (https://sloan.org/search?q=citizen+science); the Myer Foundation (http://myerfoundation.org.au/grants/grant-finder/); the National Geographic Society (https://www.nationalgeographic.org/idea/citizen-science-projects/); and many other government, NGO and philanthropic organisations.
2. **Generic domain-agnostic tools** provide general data collection/capture capabilities for any type of science project – for example CitSci.org (http://citsci.org/cwis438/websites/citsci/home.php?WebSiteID=7); Zooniverse (https://www.zooniverse.org/); CyberTracker (http://www.cybertracker.org/); Fulcrum (http://www.fulcrumapp.com/?gclid=CMzT5IDyidICFQybvAodlWMCuQ); BioCollect (http://www.ala.org.au/biocollect/); and others.
3. **Generic domain-specific tools** provide general data collection/capture capabilities for projects within a specific area of science. There are many variations available (and a great

deal of non-compliance with standards), but typically these systems are based on a core domain-relevant data standard and/or schema such as Darwin Core (http://rs.tdwg.org/dwc/) in the biodiversity domain for species observational and collections data. Examples include: iNaturalist (http://www.inaturalist.org/); iSpot (http://www.ispotnature.org/communities/global); Indicia (https://nbn.org.uk/news/instant-indicia/); Natusfera (http://natusfera.gbif.es/); and NatureMapr (http://naturemapr.com/). Some of these have also established large communities of users and include a range of different community-based mechanisms for verifying the accuracy and identifications of contributed records.

4. **Bespoke project-specific tools** are developed specifically for a particular project as either desktop or mobile apps, or a combination of both. Examples include: CrowdMag (https://www.ngdc.noaa.gov/geomag/crowdmag.shtml); Project Noah (http://www.projectnoah.org/); QuestaGame (https://questagame.com/); OPAL Water Survey (https://www.opalexplorenature.org/WaterSurvey); and hundreds of others.
5. **Data transcription tools** are open platforms which facilitate crowd-sourced data transcription, enabling large amounts of data locked in analogue records to be mobilised as digital information and used in previously impossible ways. Such tools include DigiVol (https://volunteer.ala.org.au/); Trove (http://trove.nla.gov.au/); Notes from Nature (http://www.notesfromnature.org/); Ancient Lives (http://ancientlives.org/); Old Weather (http://www.oldweather.org/); the Smithsonian Transcription Centre (https://transcription.si.edu/); and others.
6. **Education, engagement and support tools** provide mainly look-up and read-only support information for specific domain areas. Examples include field guides and identification support apps such as versions of Australian museum–sponsored field guides to Australian fauna apps; various thematic versions of the Gaia Guide apps; the Waterbug App; various thematic Lucid key apps); etc. All of these are available in the Google Play and Apple iTunes app stores.

iv. How will the tool support the project and the community using it? Does it have all of the functionality and features required for the project? Can the project live with any deficiencies? Is it already used by similar communities elsewhere?
v. Is customisation required and how customisable is it?
vi. Does the tool have:
   a. A long-term future – is it sustained/maintained by an active community or vendor;
   b. A technology upgrade pathway; and
   c. User and/or technical support?

Best practice solutions

The Business Dictionary defines 'best practice' as: 'A method or technique that has consistently shown results superior to those achieved with other means, and that is used as a benchmark'. This assumes a static, or at least slow-moving environment, but technology is changing at a dizzying rate – therefore, this concept needs to be considered in the context of continuous improvement when it is applied to information technology.

Technology, like most things, does not stand still, it will always have innovators leading and pushing the boundaries of what is possible in both hardware and software, as well as early-adopter consumers with needs to be met that current solutions do not satisfy. It is both an enabler and supporter of current needs as well as a driver of new needs, because as new technologies fulfil current needs it is possible to see opportunities and applications for even newer innovations and technologies. In a nutshell: Innovators envision needs beyond the horizon and push the boundaries of the present; early-adopters consume innovations and through demand, fuel even more innovation; while old innovations become the new normal for the masses and old norms are displaced. This is how progress is made.

A multiplicity of different solutions is currently being independently developed to meet similar needs at different times and places, and the whole scene is constantly evolving. Therefore, the concept of 'best practice' solutions are only ever relative to a given point in time, essentially reflecting the solution available at a given time which best meets the requirements and needs of a demographic of consumers/users at that time.

There is unfortunately a long way to go to realise the goal of a fully connected and functioning information supply chain, but progress is being made by many dedicated people around the world. There is also a growing enthusiasm and commitment amongst many of the major global infrastructure providers to collaborate more effectively to deliver more unified (interoperating) and integrated technology platforms, as well as to build a global community of practice to maintain and enhance the platforms in the most cost-efficient and impactful ways possible (see also Williams et al. in this volume). For example, Australia's national biodiversity data aggregator, the Atlas of Living Australia, with the support of the Australian government, has developed a suite of current best-practice tools and made them freely available worldwide under open source licences.

Case studies below highlight how this 'Living Atlas' software platform (box 5.2) is now being adopted by other countries (boxes 5.3 and 5.4) and is facilitating major improvements in data quality; data mobilisation and processing efficiency; and data accessibility and reuse; as well

Box 5.2. Case study – The Atlas of Living Australia (ALA – www.ala.org.au)

Stephanie von Gavel, ALA Business Development

The Atlas of Living Australia was established by Australia's premier research body the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and many partner organisations including museums, biological collections, research organisations and

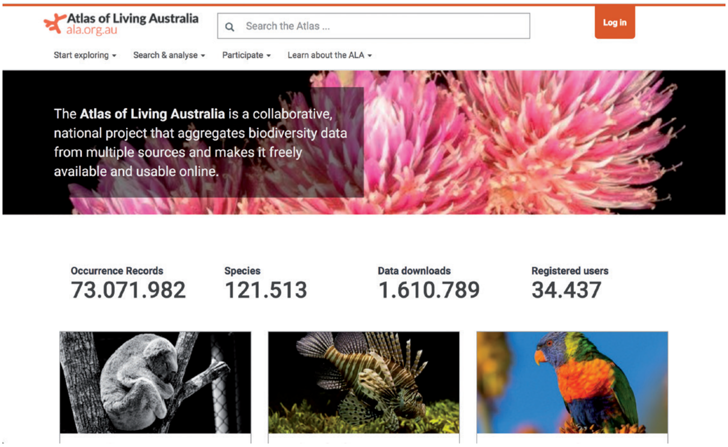

Fig. 5.2 Website Atlas of Living Australia

(continued)

government (state and federal) to provide a consistent comprehensive single point of access for Australia's biodiversity data and species information. It is funded by the Australian government via its National Collaborative Research Infrastructure Strategy (NCRIS) and is hosted by the CSIRO.

This web-based infrastructure comprises a modular suite of inter-connected databases, web applications (tools), APIs, and mobile apps. Data, which is not owned by the ALA, is also part of the infrastructure. The tools support the capture, aggregation, management, discovery, visualisation and analysis of all classes of biodiversity information. They are used for a wide range of purposes, including research, biodiversity discovery and documentation, environmental monitoring and reporting, conservation planning, biosecurity activities, education and citizen science. In addition, external enterprises and organisations are using the open infrastructure to create and enhance their own products and services. For more information on the Living Atlas platform, see http://living-atlases.gbif.org/.

Prior to the ALA, a major barrier to Australia's biodiversity research and management efforts was fragmentation and inaccessibility of data. Information was generated and siloed, housed in museums, herbaria and other collections; universities; research organisations; and government agencies, as well as with individual citizen scientists and researchers.

The ALA brings together biodiversity data and associated information from a wide variety of sources, processing and linking it together, and making it accessible from a single place in a standard format via a set of purpose-built tools and services. Accessing biodiversity data is now free and more efficient than ever before, as the ALA has already addressed a wide range of data access issues for all consumers which would otherwise have to be negotiated individually by each data consumer.

The ALA is the Australian node of the international open-data infrastructure the Global Biodiversity Information Facility (GBIF). The ALA has also 'open sourced' its software as the Living Atlas Platform to encourage the development of a collaborative community of practice around the infrastructure, and to facilitate interoperability and cost savings to the global biodiversity community. Accessible and affordable technology platforms empower and enable people to participate more actively in biodiversity knowledge-building activities. This democratises biodiversity science and develops fascination and enquiry among the next generation of scientists.

As an exemplar for open infrastructure, open data and data reuse, the Living Atlas platform is being adopted and used by an ever-increasing number of organisations both domestically and internationally. The data available via the ALA are also being used for a multitude of purposes. Atlas of Living Australia tools provide capability in many areas across the spectrum of the information supply chain.

The Living Atlas platform supports many different systems – whether they be separate instances of the software suite, or hubs (different thematic interfaces over one common instance of the platform). Open APIs also allow others to independently access data and some data processing services.

The ALA is a strong supporter of citizen science and has partnered with the Australian Citizen Science Association (ACSA) to provide the national citizen science projects catalogue. The Atlas of Living Australia also directly supports numerous projects collecting data through the BioCollect tool (http://biocollect.ala.org.au/acsa). The project finder exchanges project information with the SciStarter (www.scistarter.com) system in the United States and, through the PPSR-Core initiative, various catalogues of citizen science projects are being progressively connected to enable fast and simple discovery and access to projects of interest from a comprehensive list of projects from around the world.

Box 5.3. Case study — GBIF France

Marie-Elise Lecoq, GBIF France, Systems Development and Support

The Global Biodiversity Information Facility (GBIF) is an international open-data infrastructure for biodiversity data. GBIF encourages and helps participant countries to publish and share biodiversity data to support international biodiversity research, inform pan-national policy and improve management outcomes

(continued)

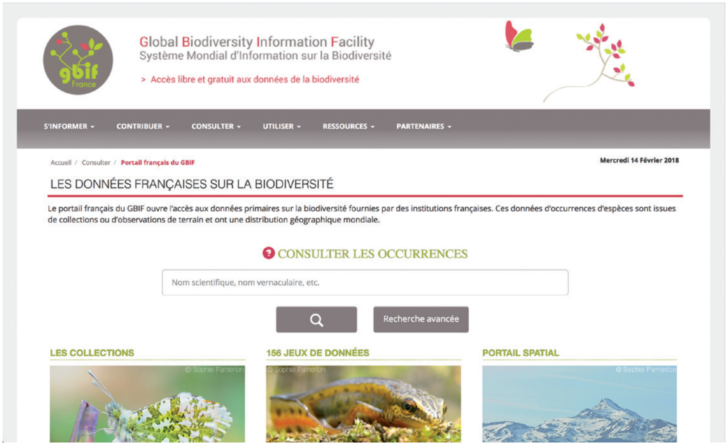

Fig. 5.3 Website Global Biodiversity Information Facility France

for biodiversity, in other words, better decisions to conserve and sustainably use the biological resources of the planet.

The Global Biodiversity Information Facility operates through a network of collaborating nodes which share skills, experiences and technical capacity. Its vision is, 'A world in which biodiversity information is freely and universally available for science, society and a sustainable future'. To achieve this, GBIF provides and endorses tools which help publishers share their own data using standards such as DwC (Darwin Core). Experiences and developments made by GBIF nodes increased the list of GBIF tools with a set of reusable ones. GBIF France decided to work with software developed by the community, especially the ALA platform (see box 5.2).

This platform was chosen because it is a powerful infrastructure that has already addressed a lot of GBIF France requirements, meaning that work was only needed to install the system and add national specificities (language, data, design, etc.). As a result, the French portal of GBIF France was established within a year and was later enhanced with the addition of the ALA spatial portal. GBIF France developed and optimised their performance to produce an attractive feature-rich portal within two years. Due to the efficiency of the development, GBIF France decided to participate more in growing and supporting the community around ALA modules.

Table 5.1 Countries currently using the Living Atlas platform

Country	Link to Platform
Australia	http://www.ala.org.au/
Argentina	http://datos.sndb.mincyt.gob.ar/
Brazil	https://portaldabiodiversidade.icmbio.gov.br
Costa Rica	http://www.crbio.cr/crbio/
France	http://portail.gbif.fr
Spain	http://datos.gbif.es/
Portugal	http://dados.gbif.pt/
UK	http://www.als.scot/ and https://nbnatlas.org/

The commitment of ALA, GBIF France, GBIF Portugal, GBIF Spain and several others to this community has multiple forms. Since 2013, a technical workshop held at least once a year has presented ALA modules to new users, to improve existing data portals and to learn from others' successes and achievements. For instance, GBIF France was the first outside the ALA team to install the spatial portal and gave feedback on this experience to the growing 'Living Atlas' community at the workshops. The meetings are motivating for new users because they can see that they can gain a powerful tool for themselves and for other participants with relatively little time and investment. Indeed, during training, technical teams get ideas from other projects and can also complete significant work on their own project. Community members have also shown the result of this collaborative work through presentations and posters at international conferences around the world. Finally, the international community around ALA have helped other institutions who do not have the technical competencies to implement their own data portal, especially in Africa.

Thanks to these engagements, seven data portals using the ALA platform were released between 2014 and 2016 (table 5.1), with several others currently in development and more investigating its use.

This ALA community is therefore helpful for organisations or associations who want to install a data portal but do not have the technical competence or staff to do so.

Box 5.4. Case Study – NBN, United Kingdom

Ella Vogel, NBN UK, Programme Development and Support

The UK National Biodiversity Network (NBN) has a long history of activity in biological recording and citizen science. In 2015 it undertook a review of its online data-sharing infrastructure and concluded that the current system was no longer fit to serve the growing needs of the Network. Three options were considered: (1) Develop a new platform from scratch; (2) re-engineer and enhance the existing platform to accommodate required functionality; and (3) adopt an existing platform to replace the old system. When the ALA open source platform (see box 5.2) was presented to the NBN Secretariat, it was clear that the most time and cost-efficient way to move forward was to adopt this infrastructure in the UK.

The pilot, NBN Atlas Scotland, was launched in 2016 as the precursor to the new core NBN Atlas. Implementation of the Living Atlas platform has enabled the UK to shift its attitude to data accessibility to being more open with improved data sharing both within the UK and globally. Previously, record sharing via the NBN Gateway was done under a bespoke NBN Data Exchange Format. Within the UK this worked well, but with a more global outlook it is important that common and interoperable formats are used. Data can now be shared both within the UK and internationally using com-

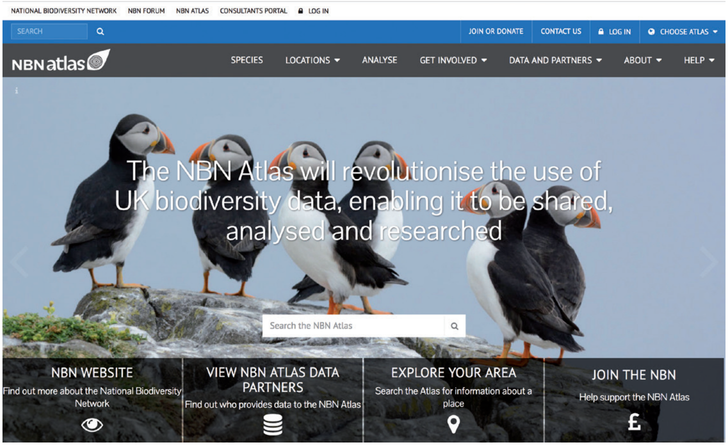

Fig. 5.4 Website National Biodiversity Network, UK

mon Darwin Core–based standards. The new system has also encouraged the use of creative commons licences, allowing datasets to be more easily used by others domestically and internationally in research, policy and planning at any scale.

Over the years, many questions have been raised about how to mobilise historic datasets; how to empower citizen scientists to collect biological records in a transparent, consistent and peer reviewed way so that their efforts are seen as equal alongside the work of professionals; how to provide access to biological records by network members; and how to combine datasets and data layers to undertake detailed analysis, without having to each have access to separate tools and different systems to perform each step. The Living Atlas infrastructure has provided solutions to these and many other issues and has given the NBN a clear direction for future development. With a global developer base to contribute to and learn from, there is stability in the future of the Atlas platform and endless opportunities for growth and development.

as facilitating change in the way that people think about the whole information supply chain and the value of their data beyond the project that they used to collect it.

Conclusion

It is not the aim of this chapter to pick 'winners' among the large pool of current technology solutions serving the citizen science community. Instead, it aims to highlight that 'best practice' in technology is a rapidly moving target and that at any given time there will always be a range of old and new technologies, features, capabilities and costing models among the wide array of tools available. However, within this environment, there are some fundamental considerations for citizen science projects when choosing appropriate infrastructure solutions to support their needs. These choices can determine the real value of a project's outputs to downstream scientific endeavours and supply chain outcomes.

Arguably, notwithstanding the direct and sometimes profound personal, social and environmental benefits of public participation in scientific activities, the most enduring element – where public contributions to science will likely have their greatest impact – is the information which they generate. However, to be of real value, this information must be accessible to the information supply chain in a timely manner and in a form which is suitable for use throughout.

Therefore, the application of standards in data collection, data transmission, and the descriptions of datasets and collection methods are critical to scientists and policymakers accepting and giving proper value and respect to, citizen science data and the enormous volunteer commitment made by citizen science participants worldwide. Well-designed IT infrastructures, which include in-built processes and rules to enforce standards and data quality, as well as mechanisms for standards compliant data sharing, can fulfil such requirements with minimal impact on users. Solutions that include such features should therefore be chosen over those that do not. Such market-based demand-driven choices will encourage all infrastructure providers to engage with the standards framework, which is critical to a functioning information ecosystem.

Acknowledgements

I acknowledge and thank my colleagues at the Atlas of Living Australia, Hannah Scott and Paul Box (CSIRO), for providing editorial comments during the drafting of this chapter.