Author/s: Williamson, K., Kennan, M.A., Johansen, G., Weckert, J.

Title: Data sharing for the advancement of science: overcoming barriers for citizen scientists

Journal: Journal of the Association of Information Science and Technology

ISSN: 2330-1635

Year: 2016

Volume: 67

Issue: 10

Pages: 2392-2403

Abstract: Systematic study of data sharing by citizen scientists will make a significant contribution to science because of the growing importance of aggregated data in data-intensive science. This article expands on the data sharing component of a paper presented at the 2013 ASIST conference. A three-phase project is reported. Conducted between 2011 and 2013 within an environmental voluntary group, the Australian Plants Society Victoria (APSV), the interviews of the first phase are the major data source. Because the project revealed the importance of data sharing with professional scientists, their views are included in the literature review where four themes are explored: lack of shared disciplinary culture, trust, responsibility and controlled access to data, and describing data to enable reuse. The findings, presented under these themes, revealed that, within APSV, sharing among members is mostly generous and uninhibited. Beyond APSV, when online repositories were involved, barriers came very strongly into play. Trust was weaker and barriers also included issues of data quality, data description, and ownership and control. The conclusion is that further investigation of these barriers, including the attitudes of professional scientists to using data contributed by citizen scientists, would indicate how more extensive and useful data sharing could be achieved.

DOI: http://dx.doi.org/10.1002/asi.23564

Authors' final proof: Williamson, K., Kennan , M.A., Johanson, G. & Weckert, J. (2016) Data sharing for the advancement of science: Overcoming barriers for citizen scientists JASIST: Journal of the Association of Information Science and Technology. 67(10):2392–2403.

Data sharing for the advancement of science: Overcoming barriers for citizen scientists

Kirsty Williamson
CSU & Monash University
P.O. Box 197, Caulfield East,
VIC 3145, AUSTRALIA
Telephone: 61 3 9903 1083
Fax: 61 3 9903 1077
Email: kirsty.williamson@monash.edu

Mary Anne Kennan (Corresponding author)
Charles Sturt University (CSU)
Locked Bag 450, Silverwater
NSW 2128 AUSTRALIA
Telephone: 612 6933 4893
Fax: 612 6933 2733
Email: mkennan@csu.edu.au

Graeme Johanson
Monash University
P.O. Box 197, Caulfield East,
VIC 3145, AUSTRALIA
Telephone: 61 4 5760 4701.
Fax: 61 3 9903 1077.
Email: graeme.johanson@monash.edu

John Weckert
Centre for Applied Philosophy and Public Ethics
Charles Sturt University
PO Box 588, Wagga Wagga
NSW 2654 AUSTRALIA
Telephone and Fax: 61 2 6933 2372
jweckert@csu.edu.au

Data sharing for the advancement of science: Overcoming barriers for citizen scientists

Abstract

Extending data sharing by citizen scientists will make a significant contribution to science because of the growing importance of aggregated data in data-intensive science. This article expands upon the data sharing component of a paper presented at last year's ASIST conference. A three-phase project is reported. Conducted between 2011 and 2013 within a large environmental voluntary group, the Australian Plants' Society Victoria (APSV), the interviews of the first phase are the major data source. Because the project revealed the importance of data sharing with professional scientists, their views are included in the literature review where four themes are explored: lack of shared disciplinary culture; trust; responsibility and controlled access to data; and describing data to enable reuse. The findings, presented under these themes, revealed that, within APSV, sharing amongst members is mostly generous and uninhibited. Beyond APSV, when online repositories were involved, barriers came very strongly into play. Trust was weaker and barriers also included issues of data quality, data description, and ownership and control. The conclusion is that further investigation of these barriers, including the attitudes of professional scientists to using data contributed by citizen scientists, would indicate how more extensive and useful data sharing can be achieved.

Introduction

Increasingly governments and other organisations are funding repositories and data aggregations for the storage and sharing of research data and information, collected and created by both professional and citizen scientists. An assumption underpinning these developments and this article is that improving data sharing between these two groups of scientists will benefit science. This is because the questions which are now possible to ask, with the advent of big science and data-intensive science, require data at a huge scale to provide the answers (Hey, Tansley, & Tolle, 2009). As far as can be ascertained, there is no major research focussing on independent citizen scientists' data contributions to online repositories and the use of those data especially by professional scientists; yet such research

is clearly needed as more citizens contribute, more repositories are funded, and more data are collected.

This article reports a study which investigated barriers to data sharing by independent citizen scientists, the term referring to people who collect data independently of professional scientists, principally for their own purposes and interests, although most often as part of a community-based group. The project did not include citizen scientists who collect data in teams in which professional scientists are involved. Although Bonney et al. (2009) reported three models for citizen science which they label contributory, collaborative and co-created, all involving professional scientists, other studies (e.g., Grove-White et al., 2007 and Kennan, Williamson, & Johanson, 2013 – the present study) identify a further group, viz., independent citizen scientists.

The study of three phases, was conducted between 2011 and 2013, and involved members of a large environmental voluntary group, the Australian Plants Society Victoria (APSV). The first phase, which involved intensive interviews with fifteen APSV members, took place in 2011. The focus was data collection, management and storage, given that members of APSV management were aware that data practices amongst many members were far from ideal and were concerned about loss of valuable data. At this time the researchers were also researching and teaching in the field of research data management, resulting in a confluence of interest for them. The second phase was a simple survey of APSV members at the beginning of 2013, designed to ascertain if there had been changes in data collection, management, storage, and sharing practices following the awareness raised of these issues during phase one of the study. Sharing was specifically included at this point, since it had emerged as very important during the first phase. The third phase, taking place throughout 2013, was aimed at improving data-sharing practices amongst ASPV members, particularly providing training through an action research project to encourage the effective contribution of data to online repositories. Australian examples of repositories for environmental data, such as those produced by members of APSV, include NatureShare (http://natureshare.org.au/), BowerBird (http://bowerbird.org.au/); and the Atlas of Living Australia (ALA) (http://www.ala.org.au/about-the-atlas/). NatureShare grew out of the enthusiasm of one dedicated APSV member, though it is not an 'official'APSV site. Bowerbird was developed by Museum Victoria with assistance from the ALA. Data from both BowerBird and NatureShare are fed into the ALA, which is Australia's national biodiversity repository.

The findings reported in this article are principally from the first phase interviews (2011). As the interviews progressed, many comments about data sharing were made by interviewees meaning that the researchers became aware of the strong implications of the study for data sharing, which had not been an original research focus. During the course of the project, 'data sharing' had become a central topic because of the advent, and increasing use, of major repositories which enable such sharing of data with the wider scientific community to take place. The article builds on a paper delivered at the 2013 ASIST Conference (Kennan, Williamson, & Johanson 2013). The ASIST paper covered all aspects of the project as did two other publications by the authors (Kennan, Williamson, & Johanson, 2012; Johanson, Williamson, & Kennan, 2013). The present article further develops the information-sharing component. The research questions, central to this article, are:

(1) What are the barriers to data sharing by members of a large environmental voluntary group, APSV?
(2) What are the implications of these barriers for data sharing with the wider scientific community?
(3) What suggestions for overcoming barriers can be made?
(4) What further research is required to clarify further the barriers to data sharing between citizen and professional scientists and to identify strategies for overcoming them?

APSV is similar to other environmental voluntary groups which collect environmental and biodiversity data in Victoria and Australia (Kennan, Williamson & Johanson, 2013), and internationally (Grove-White et al., 2007). Indeed there is an overlap of membership between APSV and other groups such as the Field Naturalists' Club of Victoria and Birdlife Australia. Thus the findings of the project have implications which extend beyond APSV to other community-based groups of independent citizen scientists, but do not necessarily extend to citizen scientists in teams led by professional scientists.

Significance of the study

The significance of this article, and for the research it reports, is that data sharing and aggregation are increasingly vital components of big science and data-intensive science (e.g., Grey, 2009). The value of data is seen to increase as they are interconnected, networked, shared and used (Borgman, 2007). Reasons include that: (i) broader conclusions can be deduced from interconnected data (Lee, 2010); (ii) significant gaps, as well as strengths, in the data become

obvious (Funk, Richardson, & Ferrier, 2005); (iii) unexpected relationships between data may be discovered (Reshef et al., 2011) and (iv) large repositories of data provide the means for later analysis and retrospective findings (Tenopir et al., 2011).

Borgman (2012, p. 1060) defined the term 'data sharing' in the broad sense of releasing data for use by others. Large-scale biodiversity research, with major economic, environmental and social contributions to make, needs large numbers of data collectors over wide areas and extended periods of time. It is well recognised that there are insufficient resources to pay all such collectors, thus there is increasing focus on volunteers such as citizen scientists (Bonney et al., 2009; Miller-Rushing et al., 2012; Roy et al., 2012). Data sharing is essential to enable the data collected by citizen scientists, sometimes working alone, sometimes in teams, to fill this breach. Data of appropriate quality collected by volunteers can aid scientists in their research on topics, such as the effects of population growth or climate change on biodiversity, and can contribute to conservation policy (Grove-White et al., 2007; Atlas of Living Australia, n.d.). Thus data collected by citizen scientists are now vital in the context of what constitutes scientific knowledge. The distributed knowledge of the single citizen scientist is strengthened by sharing with a wider community of other citizen scientists and professional scientists (Fuller, 2015). Collectivity increases the quality and value of knowledge (Suroweicki, 2004).

The significance of the article is enhanced through consideration of professional scientists as recipients of citizen science data. The importance of professional scientists in the context of data sharing emerged as the study progressed and will be the focus of a future study by the researchers. Professional scientists' perceptions of the quality of citizen science data, and their preparedness to use those data, are crucial to the effectiveness of citizen scientists in assisting scientific progress. Professional scientists tend to be sanctioned as the authority to filter all sources of data and to draw objective conclusions from the best sources (Fuller, 2015). They have training in research and science, formally through academic studies and also through exposure to the culture and traditions of their discipline (Becher & Trowler, 2001). The data collected as a part of a professional science project are subjected to a set process of project development and rigorous scientific method, (Catlin-Groves, 2012; Flanagan & Metzger, 2008). In contrast, independent citizen scientists collect data on a small scale or through local projects based on their own interests and curiosity. Their data collection is usually undertaken without an overt research question or hypothesis. Yet the

data collected in these projects may be of use for professional science when aggregated with other data (Primack & Miller-Rushing, 2012).

Thus the following literature review includes not only the barriers to data sharing by citizen scientists, as underpinning the empirical findings for the APSV study reported here, but also considers the barriers to acceptance of citizen scientist data by professional scientists

Literature Review

Following a general introduction to the literature, the review highlights the following barriers which are important to the findings of the project, as well as to data sharing between citizen and professional scientists. The first is 'lack of shared disciplinary background' which is related to the second barrier, 'trust', the lack of which is discussed in relation to perceptions of data quality and the particular issues of online environments. The third barrier, responsibility, is associated with trust and includes the controlling of access to data for reasons of behavioural responsibility, but also in response to issues of intellectual property and attribution. Then follows discussions about the barriers related to effective description, storage, management and preservation of data, along with socio-technical issues which need to be addressed in order to facilitate data sharing. International research, focussed on barriers to data sharing, includes Socientize Consortium (2013); Catlin-Groves (2012); and Flanagin & Metzger (2008).

Much has been written about the differences between 'data', 'information' and 'knowledge' (e.g., Fricke, 2009). In certain situations, these distinctions may be necessary but, with regard to sharing, the issues are often much the same. For example, the exploration of information sharing (e.g., Talja, 2002; Pilerot, 2012; 2013) in the library and information science (LIS) literature raises issues such as the importance of shared culture and trust, as do explorations of data sharing (e.g., Borgman, 2006; 2012; Pryor, 2009). Pilerot and Limberg (2011) saw information as including "data and documents that can be regarded as 'informative' and therefore as information" (p.314) when shared. Similarly information such as texts and images can become research data when they are used as evidence to answer a research question or solve a research problem. "The recognition that an observation, artefact or record constitutes data is itself a scholarly act" (Borgman, 2012, p. 1061).

Another point, emphasised in the LIS literature, is reciprocity (Savolainen, 2007; Wilson, 2010). The concept of reciprocity implies that an action of sharing is met with a similar

sharing action in return. Yet, with some of the data sharing discussed in this paper, individual independent citizen scientists have no control over who uses their data or how they are used, even if there is attribution for their original contribution, and may not benefit from the experience of a similar sharing action in return. This is the case, for example, when individual observations or small datasets such as site surveys are incorporated into, or partially or wholly merged with, larger ones. If shared data are used as a part of a larger data set by professional scientists, or for policy purposes, it is likely that the original citizen scientist contributor will not be a part of the community that uses or re-uses the data. These issues constitute barriers to sharing, especially for citizen scientists. There are few explicit and tangible rewards for sharing data and even professional researchers report it as low on the list of their priorities (e.g., Tenopir et al., 2011). Further barriers are discussed below under individual topic headings.

Lack of shared disciplinary culture

Researchers who share data and collaborate with each other tend to have shared disciplinary backgrounds and analytical skills (Borgman, 2006). Sharing data across disciplines or between researchers with different motivations and organisational cultures, for example, via open data repositories, is likely to be more difficult. Disciplinary practice, and increasingly policy, guide data sharing (Cragin, Palmer, Carlson, & Witt, 2010). Bowker (2000) alluded to the difficulties of sharing beyond small informal groups through formal systems and repositories when the information to be shared comes from different contexts, where the naming systems or classifications used by participants are not shared, and where different information needs to be integrated.

Trust

An absence of shared disciplinary culture is seen to make trust more difficult. Trust is very important to information sharing (e.g., Pilerot, 2012) and is a widely discussed concept in a range of different fields, including information science (e.g., Cronin, 2003; Wilson, 2010), as well in sociological theory (e.g., Giddens, 1991); yet it remains an elusive concept (Lane & Bachman, 1998). Weckert (2005), a computer ethicist, provides an in-depth philosophical discussion of what trust is and Bolman and Deal (2013) discuss how trust can be institutionalised or structured, especially the role of leadership.

Trust in online environments

Much has been written about information sharing in online communities, particularly in comparison to face-to-face communities (Hersberger, Murray & Rioux, 2007). In the online context, Fukuyama (1995) emphasised that trust arises where there are commonly shared norms by members of a community; but, as Weckert (2005, p. 16) pointed out, "the Internet is not a localised community with commonly shared norms." Even where there are two major groups with parallel environmental interests contributing to online repositories (professional scientists and citizen scientists), these groups do not necessarily have shared norms. It has been suggested that, on the one hand, citizen scientists are motivated by both altruistic and egotistic reasons, such as curiosity, perceived community needs, and a commitment to conservation; while, on the other hand, professional scientists are more likely to be motivated by the formal processes and structure of the scientific method and to further their own professional careers (Rotman, Procita, Hansen, Sims Parr & Preece, 2012; Latour & Woolgar, 1979). Here may be a major obstacle to trust: lack of shared norms and different motivations when depositors and future users of shared data are different from, or unknown to, each other.

In digital environments, interactions that could lead to community formation can be depersonalised by technology (Rowley & Johnson, 2013). This point was earlier made by Nissenbaum (2001) who saw a key obstacle to trust in online environments as 'missing identities' and 'missing personal characteristics' (p. 647), meaning that personal characteristics of the information or data provider are unknown and cues to assess trustworthiness are missing. This is a problem which Weckert (2005) called 'disembodiment': "On-line it is not clear with whom we are communicating, so it is difficult to know whether it is someone trustworthy or not. And because they are anonymous, … there is less reason to believe that they will act responsibly" (p.17). The issue of responsibility is discussed below.

Trust in data quality

Existing limited research points to professional scientists' lack of trust regarding the quality and accuracy of data collected by volunteers, particularly those outside of scientist-led citizen science projects (Catlin-Groves 2012; Flanagin & Metzger, 2008; Socientize, 2013). Professional scientists can also be sceptical about the quality of data collected by the citizen scientists who assist them in their teams (Gommerman & Monroe, 2012). Indeed, there is some emphasis in the literature on data quality in citizen science projects led by professional scientists (e.g., Vianna, Meekan, Bornovski, & Meeuwig, 2014), the difference being that

mechanisms for ensuring data quality can be put in place where there is professional leadership (Wiggins, Newman, Stevenson & Crowston, 2011; Gommerman & Monroe, 2012). As in the project reported in this article, not all citizen scientists, who collect or generate valuable data, work in professionally-led teams or in teams where there is involvement of professional scientists. Many citizen scientists work independently.

Responsibility and controlled data access

In their book on professionalism in the information and communications technology industry, Weckert and Lucas (2013) discussed the 'responsibility' of a professional. They defined a professional as being "an expert, at least to some extent relative to the population at large". Professionals have "some special, and socially useful, knowledge and skills" which place them "in a position of power relative to those who lack that knowledge but who have need of it" (p.6). Being in this position means that professionals need to be morally, socially and possibly legally accountable for what they do, including professional scientists in relation to the data which they collect.

While citizen scientists' views about responsibility for their data do not appear to be explicitly addressed in the literature, there is a sense in which responsibility can be exercised through controlling access to data, which may apply to them as well as other scientists. A study by Eschenfelder & Johnson (2011) found many reasons for researchers requesting that repositories control access to data. The most frequently listed was 'avoiding misuse' of data, closely followed by protection of sensitive information, privacy and intellectual property concerns. In the environmental area, preventing misuse of data and protecting sensitive information could include not disclosing locations of endangered species from collectors. PARSE Insight (2009), which describes the results of surveys aimed at gaining insight into research in Europe, also found that "major barriers for sharing research data are *the fear of researchers regarding legal issues* and *the misuse of their data"*.

A study of sharing practices by scientists (Tenopir et al., 2011) found emphasis on other reasons for controlling data access, including getting credit through citation and notification of use, provision of secure but flexible infrastructure, and assistance with description and deposit. Intellectual property and ownership issues undoubtedly have an effect on data sharing (Borgman, 2006; Pryor, 2009). When individuals contribute open data to repositories,

they lose control of what happens to their data. Fear of this occurrence can result in both citizen and professional scientists restricting access to their data.

Describing data to enable reuse

Data collected for individual or voluntary purposes may or may not be suitable for re-use and sharing for scientific purposes (Palmer, Weber, & Cragin, 2011). One of the issues for the potential reuse of citizen science data is whether appropriate metadata are available. The question is whether there is enough information about the data for professional scientists to know that they can trust and use them. Data need to be described, stored, managed and preserved effectively if they are to be shared; and socio-technical obstacles to data sharing need to be overcome. For data to be re-usable, they require description, usually in the form of metadata, which are often manually created and documented. Metadata, sometimes known as 'data about data' or 'information resource description', are required to describe data and datasets, and to aid their discovery, management and re-use (Hider, 2012, pp.4-5). There are also quality requirements for wholeness, consistency and correctness of both data and metadata (Kowalczyk & Shankar, 2011), all deceptively simple terms which conceal thick layers of complexity (Borgman 2012).

Shared data vary from raw data to data which are richly described and curated (Borgman, 2012). As research data are heterogeneous, different metadata schema and/or standards have been developed to describe them. The Darwin Core metadata schema were developed to describe biological and biodiversity data (Hider, 2012, pp.140-141). In the case of the online repositories under discussion in this article, the required descriptive metadata for an observation or field survey are derived from the Darwin Core schema. For each online repository, the required metadata are slightly different depending on the nature of the service. Australian examples, relevant to the present article, include firstly NatureShare (limited to the State of Victoria). As this repository evolved from a community where there is trust in the expertise of members, the metadata required for an observation or survey are only species name, location, and date/time. A second Australian online repository, Bowerbird, requires date/time, location, and evidence, for example a photograph of a plant, or a recording of a frog or bird call as the social network itself can provide the species name from the evidence. The ALA, which aggregates data from a number of sources, including NatureShare and Bowerbird, requires species name, date/time, location, and evidence. ALA data are used in science and policy.

The literature discussed above relates first concerns of the people who are involved in sharing (issues of trust, responsibility and controlling access to data), and second to the data to be shared (its quality and requirements for reuse). Both aspects are relevant to the findings discussed below. In the next section, the research philosophy and method for the project are discussed.

Research Approach

As mentioned above, there are three phases to this project at present, implemented between 2011 and 2013 with members of APSV. More information about the Society is supplied at the beginning of the 'findings' section, below. The first phase of the project is most relevant to this article because it explored, in depth, issues of data management and sharing although the initial focus was not intended to be on 'data sharing'. The follow-up survey confirmed that data management and sharing practices had not changed significantly in the two years following the first phase. The third phase addressed issues raised in the first and second phases by introducing an action research cycle designed to investigate whether formal training can assist in overcoming the barriers to data deposit, thus resulting in sharing via online repositories. In the process, this phase addressed concerns raised in phases one and two about long terms storage, preservation and sharing of data.

For the first phase of the project, the researchers adopted an interpretivist/constructivist approach to exploring the issues around data and information, emphasising the natural setting of participants and seeking to gain deep understanding of the meanings of the actors involved in the social phenomenon under study (Williamson, 2013a). The researchers' approach was to look for the shared meanings of participants but, at the same time, to take note of the individual meanings emerging in different attitudes and approaches. In other words, the researchers looked for both consensus and dissonance. Within this framework, the method was broadly ethnographic in style, seeking to describe the various behaviours and approaches of participants regarding a range of issues concerned with data collection, processing, and storage. Although initially specific questions about data sharing were not included in the interviews, attitudes to sharing emerged strongly and came to be seen as very important to the participants and to project outcomes.

The APSV members involved were all active collectors of data with different approaches to the issues of interest. Fifteen members of the APSV were interviewed during 2011.

Interviewees were purposively selected to include those who were particularly active in data collection in the field, or who had been involved in data/information management and storage in the past. Table 1 presents the demographic profile of the interviewees. Note that participants in each category are broadly representative of the numbers in those categories in the overall membership. Members were mostly from older age groups.

Gender	No.	Age	No.	Length of APSV Membership	No.
Male	7	40-49	1	1-5 years	1
Female	8	50-59	4	6-10 years	3
		60-69	5	11-15 years	3
		70-79	3	21-25 years	3
		80-89	2	30+ years	5

Table 1: Profile of interviewees

Semi-structured individual interviews, lasting about one hour, were undertaken. The interviews were recorded with the permission of the interviewees and the audio-recordings were transcribed by a trained transcript typist. Although the research method was not 'grounded theory' in itself, the analysis was influenced by the 'constructivist grounded theory' approach of Charmaz (2003). Constructivist grounded theory recognizes that, despite all efforts by researchers to present the views of participants, "the viewer creates the data and ensuing analysis through interaction with the viewed" and therefore the data do not provide a window on an objective reality (p. 273).

The analysis involved initially identifying categories/codes from the interview data and then merging these into key themes (Morse, 2008). Quotations from the interviews which illustrate the themes were recorded, resulting in the construction of a 'voice sheet' on each theme, so named because the quotations represent the voices of the participants. As each voice sheet was completed, an overview or summary of the data in that voice sheet was constructed. A range of themes focussed on storage, management, and sharing of data were identified. The theme of most relevance to this article is 'data sharing'.

The second phase of the project involved a follow-up survey of 101 APSV members (about one-tenth of the membership), undertaken early in 2013. Only frequency counts of the simple survey questions were undertaken. The results indicated that the qualitative data, collected in

2011, still broadly reflected the situation regarding storage, management, and sharing of data within APSV. A mention of the survey is made at a relevant point of the 'findings' section.

The third and final phase of the project, from July to December 2013, consisted of a small action research (Williamson, 2013b) project. The aim was to investigate whether training in technical and evidence requirements for submitting data to an online repository could address concerns raised in phases one and two about long-term storage, preservation and sharing of data. Prior to the intervention short interviews with the ten participants were conducted to ascertain their familiarity with the technical and metadata requirements of uploading data to online repositories (NatureShare and Bowerbird). The interviews were followed by training, provided by the developers of the repositories, which focussed particularly on the uploading of data and metadata requirements. Volunteers were then given a few months to collect data and attempt to upload it to the repositories. The final stage of the action research consisted of two focus groups, conducted by the researchers, which were used to explore participants' views about the training and their problems in using the repositories. The participants were volunteers who agreed to participate in the training and afterwards use what they had learned to train other APSV members in data storage and sharing through repositories.

During the focus groups of the action research phase, a small number of comments, relevant to the present article, were made by participants. These comments, distinguished from Phase 1 interviewee comments by the designation (AR), are included in the findings where appropriate.

Findings

A key finding of the project is that data sharing was much less inhibited within APSV than it was beyond the Society, although the overlapping membership of APSV with similar groups, such as the Field Naturalists' Club of Victoria and Birdlife Australia, resulted in broader sharing than might otherwise have occurred. A key point is that attempted sharing beyond APSV and especially with the wider scientific community, through contributions to online repositories, was found to be quite limited. The barriers that were in play are discussed below, largely under the headings used in the literature review.

Understanding the findings in context requires background to APSV. APSV is a state branch of Australian Native Plants' Society Australia (ANPSA) and has approximately 1,700 members according to the most recent source available (Walter, 2007). The APSV website

(http://www.apsvic.org.au/) lists multiple objectives, a key one being to "encourage and facilitate the conservation and study of Australian plants and their habitats". APSV activities for members include monthly meetings where experts (often APSV members) share information about specific species and habitats, garden visits, and bush walks. Activities specifically related to data collection include surveys of specific habitats at a particular time, or over a period of time, and individual observations of a species in particular locations or times. Other than meetings, traditional means of sharing of APSV data and information are newsletters and magazines and often books, currently in printed form. These forms of sharing clearly limit access to non-members such as interested individuals, and groups and scientists outside of APSV. Most important in the APSV structure are the district groups to which there is strong local loyalty. At the time of the initial data collection for the research reported here (2011), some district groups had their own websites to which members contributed data. A few individuals had their own websites and databases but few were contributing their data to the online repository, NatureShare, which was the initiative of an APSV member and relatively new at the time. BowerBird, to which some members are now contributing, was yet to go online.

The APSV study revealed that participants collected a wide range of data, mainly individual plant species observations and site surveys. Photographs were the major data type, resulting in the digital camera being the most common data collection tool which was used by all participants. Nevertheless, surveys and personal notes on observations, such as location, season, time of flowering, habitat, pollinators, growth patterns - often collected over time - were also important, mostly collected with pen and paper and sometimes put into a spreadsheet or database.

The sharing culture of APSV

There was a pervasive sharing culture *within* APSV, with few obstacles perceived by interviewees. While this culture was not based on formal disciplinary background, members shared a common interest in conservation and Australian plants about which many of them were passionate. Society members thus indicated some comparability with researchers who are willing to share data and collaborate with each other due to shared disciplinary backgrounds and analytical skills (Borgman, 2006). The following comments were typical of the views expressed. The second quotation is an interesting comment about academic

behaviour, perhaps motivated by perceptions about the need for academics to further their own professional careers as identified by Rotman et al. (2012).

> "I've rarely come across anyone within the plant society not interested to share." (Interviewee 7)

> "Coming from a university … I was not used to people sharing information. … People always keep their information to themselves, and protect it and share it on the basis of them getting something out of it. ... All these [APSV] people are so willing to share information." (Interviewee 3)

Face-to-face communication was particularly prized. Members shared an interest and enjoyed pursuing it with like-minded people at regular get-togethers. Thus theirs was, at least to some extent, a shared culture where tribal customs (Cronin, 2003) were sufficiently similar. Many members revealed a strong interest in increasing their own knowledge and understanding in interaction with others, as well as contributing to society more generally, commitments often shared with professional scientists (Bell et al., 2008).

Digital photography and computers have made sharing easier. Having photos and other data on computer is a first step towards being able to share with others. Indeed, the few participants with their own websites found that they were sharing their data.

> "I often get requests through my website for people to use photos, in publications …, brochures and things like that." (Interviewee 11)

The district groups had also developed websites which enabled the sharing of data and information, at least within APSV, though they were not always easy to maintain.

> "We're just getting [our district's website] back on track again." (Interviewee 12)

The above quotation hints at an obstacle to sharing within APSV: the age profile of the membership impeded interest and participation of members in new technology and thus the level of computer and information technology knowledge and use which would facilitate optimal sharing.

> "The age group of the members of APSV is generally from 55 or 60 onwards ... so the people there are not really interested in technology." (Interviewee 9).

Lack of IT skills had a significant effect.

> "A lot of … the older ones, who have got fabulous information because they've been growing and out there in the bush for so long, and theirs is all stuck up here [points to

> the head] too; fabulous information but they are just not computer literate." (Interviewee 7)

In general, practices of data collection, management and storage did not support and encourage data sharing. Hand-written notes and information carried in the head and not written down, do not facilitate sharing. The resulting data, possibly useful to science, tend to be lost.

> "We realised that [we had lost amazing information] at Maroondah group a few years ago when we had a succession of our older members pass away." (Interviewee 8)

Some APSV communication practices also had limitations. The age profile of APSV membership was likely to have been influential in the fact that most interviewees were still strongly in favour of print forms of communication (books, magazines and newsletters). Nevertheless, the problem of information becoming dated in books was well recognized, especially given the ongoing changes of botanical names.

> "They're out of date almost instantly; it's changing all the time ... all the family names, everything." (Interviewee 5)

While interviewees enjoyed receiving their magazines and newsletters, the longer-term problems of retrieving information from them was well known.

> "The more recent newsletters are on the website … but [for some information] … there may well be an article in one of our newsletters from 20 years ago." (Interviewee 2)

Even a popular electronic form, the CD-ROM, on which one district group was compiling data at the time of the interviews, has limited communication value and is likely to be affected in time by the 'obsolete technology' syndrome.

Despite the above barriers to efficient sharing, member attitudes within APSV for the most part favoured collaboration with others in the Society. The story was quite different with regard to sharing beyond APSV, thus indicating a clear distinction between sharing with 'insiders' (i.e., other APSV members or close associates) and 'outsiders' (especially 'unknowns' through online repositories). Of the 15 interviewees in 2011, only the developer of the repository, NatureShare, (also an APSV member) had contributed to either it or to the national repository, the ALA. Although this was understandable given that both repositories were new at the time, the survey conducted

two years later, at the beginning of 2013 before the training action research was implemented, revealed that only seven of the 101 APSV members who responded, were contributing to either of these repositories – the two options available to them at the time. Given that those contributing were more likely to respond to the survey, these results revealed that the situation had not changed very much in the two intervening years. The main barriers, particularly, with regard to contributions to online repositories, are discussed below.

Trust

Obvious reasons for the generous attitude to sharing within APSV are that members know each other personally, making it easy to trust one other, a key ingredient to sharing as revealed in the literature (e.g., Cronin, 2003; Pilerot, 2013). As Weckert (2005) pointed out, there is often a lack of trust with Internet sharing because "the Internet is not a localised community with commonly shared norms" (p.16) and because the problem of "missing identities" and "missing personal characteristics" (Nissenbaum, 2001, p. 647) on the Internet means that personal characteristics of the information or data provider are unknown and cues needed to assess trustworthiness are missing.

Making trust in online repositories even more difficult for APSV members was participants' earlier experiences in contributing to an online flora database, the Victorian State Government repository (Flora Information System - FIS), which had ended badly when the database had been sold into private hands. Interviewees had refused to contribute further when they found that they had to pay for access to datasets to which they had originally contributed their own data voluntarily. Moreover, there was resentment that data contributed to FIS seemed to disappear, unacknowledged and unattributed.

> "Well at one stage … I saw myself sending all my observations [into FIS] … so it could be of use. And it was just so difficult. ... You're not encouraged to do it because you get no feedback and ... you're even wondering ... what's happened." (Interviewee 5)

Future trust that their data would be respected had thus been undermined.

The issue of data quality is complex (discussed further below) but there is an element of trust that is important. As the literature review noted, professional scientists can be critical of citizen science data, even where citizen scientists assist them in their teams (Gommerman & Monroe, 2012). It was therefore interesting to note a lack of trust on the part of some study

participants who were sceptical that 'quality data' would be contributed to online repositories (by other citizen scientists) and this deterred them from wanting to add data to a shared repository or website. In other words, they did not want to be associated with possibly inaccurate data in repositories.

> "It depends who has provided it [the data/ information]; some are just people out in the bush, you know ..., and they get it wrong, and they don't know, and so, errors creep in." (Interviewee 7)

> "It's too easy to get wrong information online … I look at stuff and I think, that just can't possibly be there, it's got to be a mistake; it only grows on the Murray River. … So, that's one of the big problems of freely being able to put information on, because I think it then becomes useless." (Interviewee 7)

Concerns about data and description quality are common, and shared with other volunteer scientists (Grove-White et al., 2007) and professional scientists (Borgman, 2007).

Responsibility and controlled data access

The literature review discusses the 'responsibility' of professionals (Weckert & Lucas, 2013) and their moral, social and legal accountabilities, prompting the researchers to speculate that some responsibility for their data might also be expressed by citizen scientists through efforts to control access to data as other researchers do. Indeed the APSV study confirmed the findings of Eschenfelder & Johnson (2011) and PARSE (2009) in that several comments were made about controlling access to data to avoid misuse. The first example is connected with the perceived need to protect the environment, a sentiment expressed by several interviewees, for example in relation to the plants which are occasionally discovered.

> "[I'm] happy to share data and information ... with other interested people ... but there's a certain protectiveness about it. ... You don't want to generate a lot of interest from people who are not used to conserving the bushland." (Interviewee 1)

The third, action research phase of the project, involving ASPV members who were prepared to undertake the training, also provided a confirmation of concern about information getting into the hands of those without respect, for example, for a rare orchid, and the need to protect the environment.

> "There is a hide location facility on Bowerbird [i.e., of location]. I had a question about that. Who does it actually hide the location from? ... It's a concern whether it is best to use that hide location facility or just not put the location in at all." (AR)

In another example, there were fears of loss of control and lack of recognition on the part of a participant with his own website, as well as concerns about data quality.

> "If you put all of your effort into that [an online repository] then you lose that recognition. ... plus I can't control it. You know, if it goes up and the data's wrong … then it's my stuff is then implicated as potentially being wrong as well; whereas if it's my site I'm in control of it – if it's wrong then it's my fault." (Interviewee 15)

All these reasons for controlling access to data expressed by the volunteers were also expressed by professional researcher participants in Eschenfelder & Johnson's (2011) study. The importance of intellectual property and ownership are emphasised by Borgman (2006) and Pryor (2009).

Describing data to enable reuse

As the literature review pointed out, data collected for individual or voluntary purposes may or may not be suitable for re-use and sharing for scientific purposes (Palmer, Weber & Cragin, 2011). Issues here also relate to data quality. One of the issues emerging from the APSV project was the problem of inconsistent naming systems. In both the first and third phase of the project, several interviewees mentioned the frequency with which name changes of plants occurred. Online repositories need to be well designed to overcome naming issues (Bowker, 2000) by automatically updating and referencing taxonomy changes and should also assist with identification. Comments revealed on the one hand that naming issues had not been entirely overcome in the new repositories.

> " ... some of the clan names are not entirely up to date ... [For example, the leek orchid] … now has a name of its own. When I uploaded it, it wasn't in the database, I had to try and find the old name and in comments, put [that the] name has been changed." (AR)

On the other hand, assistance with identification seems to have worked from the social network encouraged by both repositories.

> "... anything I didn't know in the way of the little critters, I would quickly put on BowerBird because within a few hours you would have the answer of what it was." (AR)

In terms of accurate data description, storage, management and preservation, to enable effective sharing, not only was there a skills deficit among many members, particularly because of the age profile, but participants often referred to issues of 'time'.

> "They [APSV members] could provide information to be put online. And that relies on someone having a lot of time … [With volunteers] that can very soon get out of hand … and how are you going to double check that it's correct?" (Interviewee 7)

Other barriers arose from members' loyalties to their district groups and their preference to contribute their data to those groups using their traditional methods of meeting presentations, newsletter articles and the district websites. Lack of time was particularly mentioned as an impediment to sharing on a wider basis. The implications of the findings and the conclusions follow.

Discussion and conclusion

The first question, addressed by this article, concerned identifying barriers to data sharing by members of a large environmental voluntary group, APSV. The project found that, while data sharing within APSV occurred very frequently, although often restricted to oral networks and with considerable emphasis on print sources, there was very little data sharing beyond the APSV membership and associated groups through online repositories. The barriers included members' lack of trust in how their data might be used, as well as scepticism about data quality in online repositories, including data of fellow members. Past negative experiences of sharing through one particular online repository (which had become commercialised) had also undermined trust. Desire to protect the environment created another reason for controlling data access, along with the desire for acknowledgement and attribution of their data. Members' concern about protecting the environment and data quality, i.e., not wanting to be associated with inaccurate data in repositories, indicates a level of responsibility at least in the case of some citizen scientists. These concerns are not dissimilar to those of professional scientists in the academy and government (Douglass et al., 2014; Tenopir et al., 2011). For example, both communities (citizen and professional scientists) report concerns about quality, which indicate that responsibility in this regard is felt by professionals and non-professionals alike. Finding the time and having the skills to upload data into repositories, along with some inefficiencies in the repositories themselves, were also key barriers for many APSV participants. The third phase of the project provided a small step towards improving the skills of members, but more remains to be done.

The second question concerns the implications of these barriers for data sharing with the wider scientific community. The ramifications of present attitudes are stark. At the moment a

considerable amount of valuable citizen science data from these environmental groups, who work independently from professional scientists, is not widely available. This situation is unhelpful in the era of big science and data-intensive science, although there is a flip side here: citizen science data needs to be trusted sufficiently by professional scientists for use to be made of them. There are advantages not only for science but also for independent citizen scientists if they contribute their valuable field data to repositories, given the ability of repositories to automatically update and reference taxonomic changes, to manage data and metadata quality, and offer data curation and preservation for long-term sustainability.

Suggestions for overcoming barriers (the third question) at this stage include developing a number of repository practices to encourage sharing behaviours. Some of the recommended practices are relatively simple to implement, such as over-coming IT and time constraints by making upload of data as simple and easy as possible, perhaps even providing instruction. Encouraging organisational change, such as delegating a role of data sharing to specific APSV volunteer members, who are IT savvy and convinced of the importance of collaboration, may provide some improvement. The developers of both BowerBird and NatureShare, the two online repositories on which volunteers were trained, have both tried to assist users to upload their data as easily as possible by providing guidance, online submission templates and support with data quality, metadata and technology requirements, addressing the issues identified by Palmer, Weber, & Cragin (2011) of appropriate description, storage, management and preservation.

Other changes in attitudes to data sharing are more difficult to implement, requiring individual, social and cultural change. These include professional scientists recognising their shared concerns with citizen scientists, and acknowledging that citizen scientists also have concerns about quality. Most project interviewees did not want to be seen to be contributing data to a repository where there may be concerns about quality and were aware that not all APSV members were as concerned about data quality as they were. Nevertheless, previous publications by the authors confirmed that independent citizen scientists have useful data to contribute to many aspects of professional science (Johanson et al., 2013; Kennan et al., 2012, 2013). Another issue not easy to resolve is that of trust which is fundamental to the activities of both citizen and professional scientists. Trust between citizen scientists and professional scientists needs to be improved before data sharing can reach its potential in the scientific endeavour.

While this article indicates barriers for citizen scientists to sharing their data, there is much further scope for research. The attitudes of professional scientists with regard to the use of data collected and contributed by independent citizen scientists also deserves further exploration. The researchers intend to undertake a further project, this time using samples of both citizen and professional scientists and a mixed methods approach (surveys and interviews with both groups). There will be value in obtaining a broad overview of practices and attitudes, as well as in providing deep exploration through intensive interviews. The intention is to focus the research on specific online repositories. Interestingly, Pilerot's (2014) article provides a case for making information sharing visible through material objects, which in this case will be the selected online repositories. If the research reported in this article is an indication, there may be fewer differences in values and attitudes between the two groups than their separate members currently recognise. As Borgman (2012) suggested, for improved data sharing to occur, more needs to be done to understand and respond to the interests of the data communities themselves. This can only be achieved by continuing close engagement and study.

Acknowledgements

The authors acknowledge, with thanks, the support of the APSV, especially the assistance of Cathy Powers, President; Russell Best, Research Officer and developer of NatureShare; and the participants who gave us their time and their views. Also thanks to Ken Walker of Museum Victoria and developer of BowerBird. We also acknowledge with thanks the funding from the Wettenhall Foundation for the action research component and the initial Small Grant funding received by Kennan and Williamson from the Faculty of Education, Charles Sturt University.

References

Atlas of Living Australia (n.d.) *Atlas background* [website]

Bell, S., Marzano, M., Cent, J., Kobierska, H., Podjed, D., Vandzinskaite, D., ... & Muršič, R. (2008). What counts? Volunteers and their organisations in the recording and monitoring of biodiversity. *Biodiversity and Conservation*, **17**(14), 3443-3454.

Becher, T. & Trowler, P.R. (2001) *Academic Tribes and Territories*, 2nd edition, Buckingham, the Society for Research into Higher Education and Open University Press.

Bolman, L. G. & Deal, T. E. (2013). *Reframing organizations: Artistry, choice, and leadership*. New York, Jossey-Bass.

Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J., & Wilderman, C. (2009). Public participation in scientific research: Defining the field and assessing its potential for informal science education. *A CAISE Inquiry Group Report*, Washington DC, Center for the Advancement of Informal Education.

Borgman, C. L. (2006). What can studies of e-learning teach us about collaboration in e-esearch? Some findings from digital library studies. *Computer Supported Cooperative Work: The Journal of Collaborative Computing,* **15**(4), 359-383.

Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet* Cambridge: The MIT Press.

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, **63**(6), 1059-1078

Bowker, G. C. (2000). Biodiversity datadiversity. *Social Studies of Science,* **30**(5), 643-683.

Catlin-Groves, C. L. (2012). The citizen science landscape: From volunteers to citizen sensors and beyond. *International Journal of Zoology* **2012**. doi:10.1155/2012/349630.

Charmaz, K. (2003). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. Lincoln (Eds.), *Strategies of qualitative inquiry,* (2nd ed., pp. 249-291). Thousand Oaks, CA: Sage.

Cragin, M.H., Palmer, C.L., Carlson, J.R., Witt, M. (2010). Data sharing, small science, and institutional repositories. *Philosophical Transactions of the Royal Society A,* **368**(1926), 4023-4038.

Cronin, B. (2003). Scholarly communication and epistemic cultures. *New Review of Academic Librarianship*, **9**(1), 1-24.

Douglass, K., Allard, S., Tenopir, C., Wu, L., & Frame, M. (2014). Managing scientific data as public assets: Data sharing practices and policies among full-time government employees. *Journal of the Association for Information Science and Technology*, **65**(2), 251-262

Eschenfelder, K. & Johnson, A. (2011). Limits of sharing: Controlled data collections *ASIST Annual Meeting 2011*, October 9-13, New Orleans.

Flanagin, A. J. & Metzger, M.J. (2008). The credibility of volunteered geographic information. *GeoJournal* **72**(3-4),137-148.

Fricke, M. (2009). The knowledge pyramid: a critique of the DIKW hierarchy. *Journal of Information Science,* **35**(2), 131-142.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*, London: Penguin Books.

Fuller, S. (2015). *Knowledge; the philosophical quest in history*. London: Routledge, pp 3, 173.

Funk, V.A., Richardson, K.S. & Ferrier, S. (2005). Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnean Society*, **85**, 549–567

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age.* Cambridge, UK: Polity Press.

Gommerman, L. & Monroe, M. C. (2012). *Lessons learned from evaluation of citizen science projects.* Gainesville FL, School of Forest Resources and Conservation Department, University of Florida, Institute of Food and Agricultural Science.

Grey, J. 2009. Jim Grey on e-science: A transformed scientific method. In *The fourth paradigm: Data intensive scientific discovery*, edited by A.J.G. Hey, S. Tansley and K. M. Tolle, 177-184. Redmond, Washington: Microsoft Research.

Grove-White, R., Waterton, C., Ellis, R., Vogel, J., Stevens, G., & Peacock, B. (2007). *Amateurs as experts: Harnessing new networks for biodiversity*. Lancaster: Lancaster University, Institute for Environment, Philosophy, and Public Policy.

Hersberger, J.A , Murray, A.L. & Rioux, K. S. (2007). Examining information exchange and virtual communities: an emergent framework. *Online Information Review,* **31**(2), 135-147.

Hey, A. J. G., Tansley, S., & Tolle, K. M. 2009. *The fourth paradigm: Data-intensive scientific discovery.* Redmond, Washington : Microsoft Research.

Hider, P. (2012). *Information resource description: Creating and managing metadata.* London: Facet.

Johanson, G., Williamson, K. & M.A. Kennan (2013). Multiple communities: botanical data from citizen scientists for digital repositories. In *Nexus, Confluence, and Difference: Community Archives meets Community Informatics*; Prato CIRN Conference Oct 28-30, 2013, Monash Centre, Prato Italy. Centre for Community Networking Research, Centre for Social Informatics, Monash University.

Kennan, M.A., Williamson, K. & Johanson, G. (2012). Wild data: collaborative e-research and university libraries. *Australian Academic & Research Libraries*, **43**(1) pp.55-78.

Kennan, M.A., Williamson, K & Johanson, G. (2013). Environmental Voluntary Groups: Towards Curating Data for Sharing, Access and Preservation, 76th Annual ASIST Annual Meeting, *Beyond the Cloud: Rethinking information boundaries*, Montreal Canada. November 1-5 2013. Proceedings of the 76th ASIS&T Annual Meeting, Volume 50.

Kowalczyk, S., & Shankar, K. (2011). Data sharing in the sciences. *Annual Review of Information Science and Technology,* **45**, 247-294.

Lane, C. & Bachmann, R. (1998). Trust within and between organizations: Conceptual issues and empirical applications. Oxford: Oxford University Press.

Latour, B. & Woolgar, S. (1979). *Laboratory Life: The Construction of Scientific Facts.* Princeton Sage, London.

Lee, R. (2010). *Knowledge matters; the structures of knowledge and the crisis of the modern world system*. Brisbane: University of Queensland Press.

Miller-Rushing, A., Primack, R. & Bonney, R. 2012. The history of public participation in ecological research. *Frontiers in Ecology and the Environment* **10**(6), 285-290.

Morse, J.W. (2008). Confusing categories and themes *Qualitative Health Research,* **18**(6), 727-728.

Nissenbaum, H. (2001). Securing trust online: Wisdom or oxymoron. *Boston University Law Review*, **81**(3), 635-654.

Palmer, C., Weber, N. & Cragin, M. (2011) The Analytic potential of scientific data: Understanding re-use value. *ASIST Annual Meeting 2011*, October 9-13, New Orleans.

PARSE Insight (2009). *Insight into digital preservation of research output in Europe: Survey report*. Didcot, UK, Permanent Access to the Records of Science in Europe (PARSE) project.

Pilerot, O. (2012). LIS research on information sharing activities–people, places, or information. *Journal of Documentation,* **68**(4), 559-581.

Pilerot , O. (2013). A practice theoretical exploration of information sharing and trust in a dispersed community of design scholars. *Information Research*, **18**(4).

Pilerot, O., & Limberg, L. (2011). Information sharing as a means to reach collective understanding: A study of design scholars' information practices. *Journal of Documentation,* **67**(2), 312-333.

Pilerot, O. (2014). Making design researchers' information sharing visible through material objects. *Journal of the Association for Information Science and Technology*, **65**(10), 2006-2016.

Primak, R.B. & Miller-Rushing, A.J. (2012) Uncovering, collecting, and analysing records to investigate the ecological impacts of climate change: a template from Thoreau's Concord. *BioScience* **62**(2), 170-181.

Pryor, G. (2009). Multi-scale data sharing in the life sciences: some lessons for policy makers. *International Journal of Digital Curation*, 4**(**3), 72-82.

Putnam, Robert D. (1994). *Making democracy work: Civic traditions in modern Italy*. Princeton: Princeton University Press.

Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M. & Sabeti, P.C. (2011). Detecting novel associations in large data sets. Science, **334**, 1518-1524.

Rotman, D., Procita, K. Hansen, Sims Parr, C. & Preece, J. (2012). Supporting content curation communities: The case of the Encyclopedia of Life. *Journal of the American Society for Information Science and Technology*, **63(**6),1092-1107.

Rowley, J. & Johnson, F. (2013). Understanding trust formation in digital information sources: The case of Wikipedia. *Journal of Information Science,* **39**(4), 494–508.

Roy, H., Pocock, M., Preston, C., Roy, D., Savage, J., Tweddle, J., & Robinson, L. (2012). *Understanding citizen science and environmental monitoring: final report on behalf of UK Environmental Observation Framework*. London, Natural History Museum..

Savolainen, R. (2007). Motives for giving information in non-work contexts and the expectations of reciprocity. The case of environmental activists. *Proceedings of the American Society for Information Science and Technology*, **44**(1), 1-13.

Socientize Consortium (2013). *Green Paper on citizen science, citizen science for Europe: Towards a better society of empowered citizens and enhanced research*. Aragon, Spain, The Socientize Consortium of the European Commission.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations.* New York: Doubleday, Anchor.

Talja, S. (2002). Information sharing in academic communities: types and levels of collaboration in information seeking and use. *New Review of Information Behavior Research,* **3**(1), 143-159.

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read, E., Manoff, M. & Frame, M. (2011) Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* **6**(6): e21101.

Vianna, G.M.S., Meekan, M.G., Bornovski, T.H., & Meeuwig, J.J. (2014). Acoustic telemetry validates a citizen science approach for monitoring sharks on coral reefs. PloS ONE, **9**(4) doi:10.1371/journal.pone.0095565.

Walter, J. (2007). The Story of Arthur Swaby and the Society for Growing Australian Plants: Melbourne, Vic.: Australian Plants Society Inc.

Weckert, J. (2005). On-line trust. In Cavalier, R. (Ed). *The Impact of the Internet on Our Moral Lives*. Albany, NY: SUNY Press, 95-117

Weckert, J. & Lucas, R. (2013). *Professionalism in the information and communication technology industry.* Canberra, ACT, ANU E-Press.

Wiggins, A., Newman, G., Stevenson, R. D., & Crowston, K. (2011, December). Mechanisms for data quality and validation in citizen science. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on* (pp. 14-19). IEEE.

Williamson, K. (2013a). Research concepts. In K. Williamson & G. Johanson (Eds.) *Research methods: Information, systems and contexts* (pp. 3-23)**.** Prahran, VIC., Tilde University Press.

Williamson, K. (2013b). Action research. In K. Williamson & G. Johanson (Eds.) *Research methods: Information, systems and contexts* (pp.188-202). Prahran, VIC., Tilde University Press.

Wilson, T. (2010). Information sharing: an exploration of the literature and some propositions. *Information Research*, **15**(4).