

## CHAPTER 6

# Data quality in crowdsourcing for biodiversity research: issues and examples

Clemens Jacobs

GIScience Research Group, Institute of Geography, Heidelberg University,  
clemens.jacobs@geog.uni-heidelberg.de

### Abstract

The last few years have seen the emergence of a large number of worldwide web portals where volunteers report and collect observations of plants and animals, share these reports with other users, and provide data for scientific research purposes along the way. Activities engaging citizens in the collection of scientific data or in solving scientific problems are collectively called citizen science. Data quality is a vital issue in this field. Currently, reports of species observations from citizen scientists are often validated manually by experts as a means of quality control. Experts evaluate the plausibility of a report based on their own expertise and experience. However, a rapid growth in the quantity of reports to be processed makes this approach increasingly less feasible, creating a need for methods supporting (semi)automatic validation of observation data. This aim is achieved primarily by analysing the spatial and temporal context of the data. Relevant context information can be provided by existing observation data, as well as by spatial data of environmental factors, or other spatio-temporal factors impacting the distribution of species, or the process of observation and contribution itself. It is very important that the

---

#### How to cite this book chapter:

Jacobs, C. 2016. Data quality in crowdsourcing for biodiversity research: issues and examples. In: Capineri, C, Haklay, M, Huang, H, Antoniou, V, Kettunen, J, Ostermann, F and Purves, R. (eds.) *European Handbook of Crowdsourced Geographic Information*, Pp. 75–86. London: Ubiquity Press. DOI: <http://dx.doi.org/10.5334/bax.f>. License: CC-BY 4.0.

specific properties of data emerging from citizen science origins are taken into account. These data are often not produced in a systematic way, resulting in (for instance) spatial and temporal incompleteness. Also, the data structure is not only determined by the natural spatio-temporal patterns of species distribution, but by other factors such as the behaviour of contributors or the design of the citizen science project that produced the data as well.

### Keywords

crowdsourcing, citizen science, data quality, data validation, biodiversity

### Introduction

Learning more about biodiversity on our planet has become an important challenge as we face climate change and species extinction. Any conservation efforts need to be based on adequate knowledge about distribution, behaviour, and ecology of species. However, the long-term data covering broad geographic regions, which are necessary to gain said knowledge (Dickinson, Zuckerberg & Bonter 2010), cannot be collected using professional data collectors alone. High costs associated with professional data gathering pose another inhibiting challenge. One way of solving this problem is data collection by volunteers. Activities involving citizens in the collection of scientific data or, more general, in scientific research endeavours, are called citizen science. While citizen science itself is not a new phenomenon, we see a growing number of such projects being organized in web portals, revolutionising the way biodiversity data are collected and made available. Recent years have seen a growing number of projects using the possibilities offered by web 2.0 technologies (Dickinson, Zuckerberg & Bonter 2010; Miller-Rushing, Primack & Bonney 2012), where volunteers can upload, manage, and share their own observations of plants and animals, and make them available for scientific research. Opportunities for biodiversity monitoring and ecological research provided by this phenomenon, but also implications for project organisation and management, are extensively discussed in a book by Dickinson and Bonney (2012), and in numerous other publications (e.g., Connors, Lei & Kelly 2012; Chandler et al. 2012; Cosquer, Raymond & Prevot-Julliard 2012; Sullivan et al. 2014). Motivations of initiatives in this field range from furthering public interest in conservation issues and concerns (with data collection as a mere by-product), to systematic generation of such data for specific uses in scientific research, planning or public administration (e.g. monitoring of certain species or groups of species in certain areas or regions). Other projects aim at collection of data about the distribution of species without a predefined, specific goal.

This way of collecting data using the www and the general public is a specific form of crowdsourcing (Howe 2006), i.e. employing the general public to produce web content or to carry out certain labour-intensive tasks (especially tasks that cannot be easily automated using methods of data processing). Other terms are used in biodiversity citizen science depending on data collection procedures employed or goals pursued, such as community-based monitoring or CBM (Conrad & Hilchey 2011). As the data collected always have a geographic reference, they represent a specific type of Volunteered Geographic Information (VGI) (Goodchild 2007; Haklay 2013).

One of the most important concerns with these data is data quality. Assuring data quality is important because a general lack of trust will decrease their use for science or administration (Conrad & Hilchey 2011). A recent study by Theobald et al. (2015) showed that so far only a small portion of biodiversity-related citizen science projects contributed data to peer-reviewed scientific articles. The quality of the output of scientific research depends directly on the quality of the data used (Dickinson, Zuckerberg & Bonter 2010), as does the quality of administrative and planning decisions. On the other hand, citizen science approaches introduce great advantages, considering their ability to provide large amounts of data over broad geographic regions as well as long periods of time, often at relatively low cost (Dickinson, Zuckerberg & Bonter 2010). At the same time, this poses a challenge for data quality assurance: many projects acquire large amounts of observations - often hundreds of observations per day, or even more. Many projects employ manual validation procedures that do not scale well, making (semi)automatic validation methods necessary.

This chapter presents an overview of important issues related to quality of citizen science biodiversity data. Using examples from citizen science projects in the domain of biodiversity, it discusses specific problems and possible avenues to solutions concerning quality assurance for this specific kind of VGI. While there are many commonalities with VGI from other domains, allowing for the adoption of quality assurance approaches and strategies that are also used in other fields of VGI, there are also notable differences or features shared only with few other VGI domains, making adjustments of common approaches and strategies necessary. Most important among these differences are the diversity concerning project design and organisation (from strict monitoring schemes to rather open, opportunistic data collection, resulting in data properties and quality assurance needs varying between projects), and the nature of the information mapped (identification of species requiring some degree of expert knowledge, thereby raising issues of credibility).

### Quality of citizen science biodiversity data

When we examine the quality of citizen science data from the biodiversity domain, we need to look at how data quality can be defined, and how it is used

and handled in the relevant scientific practice. We approach the term data quality from two different perspectives:

- Data quality in terms of the sum of the data's properties, and
- data quality in terms of the data's fitness for use.

### *Data quality as the sum of the data's properties*

Observations of occurrences of species are geographic data. Therefore, their quality in terms of characteristic properties can be described using the quality features introduced by ISO standard ISO 19113 (ISO 2002). These include the following: completeness, logical consistency, positional accuracy, temporal accuracy, and thematic accuracy. Properties of the data regarding these attributes are determined mostly by the design of the project collecting the data (especially rules and guidelines concerning data collection). Therefore, they are diverse. Considering the aspect of **completeness**, we often find a pronounced spatial and temporal heterogeneity in citizen science biodiversity data. This is especially the case for data collected without using structured monitoring schemes or strict rules - so called casual data collected in an opportunistic way (Chapman 2005). There are many reasons for this heterogeneity, like contributor preference for certain species or groups of species, variable observation effort caused by different (seasonal) weather conditions, or differences in spatial density of observations associated with differences in population density, among many others. Bird et al. (2014) describe approaches to account for the variability in the resulting data caused by such factors. They use several statistical tools to demonstrate effects of certain types of error and bias in citizen science data on modelling results in biology, and describe how to address these issues. Van Strien, van Swaay and Termaat (2013) present a methodology to remedy several types of bias in the data when using them for occupancy models (modelling the distribution of species in space and time).

The **positional accuracy** of citizen science species distribution data depends primarily on the type of location information, e.g. exact point, assignment to an (arbitrary) area or to a map quadrant. The data of many relevant projects are heterogeneous in this respect. The positional accuracy of point data depends (among other factors) on the way the coordinates of an observation are determined, e.g. using a GPS device on site, placing the observation's location on a map or aerial photograph (in a map viewer), deriving the location from a specimen description, etc.

**Thematic accuracy** refers to the correctness of the classification of objects or of their non-quantitative attributes (Kresse & Fadaie 2004). An important issue regarding thematic accuracy of observational data of animals and plants from citizen science projects is the participants' lack of scientific training and its effect on the reliability or credibility of species identification (Conrad & Hilchey 2011).

The **temporal accuracy** of observational data of animals and plants from citizen science projects is determined by how accurately it can be determined at data collection. The day of observation is mandatory information provided in most cases. Sometimes, the time of day can be specified as well, or is recorded automatically if an observation is reported using a mobile device. Currentness, i.e. the correctness of data in relation to the state of the environment changing over time, is another important aspect of temporal accuracy.

**Logical consistency**, including aspects like consistency of data structure or compliance with certain rules (Kresse & Fadaie 2004) is usually ensured by adequate design of the reporting tools and data base.

### *Data quality in terms of fitness for use*

Data quality in terms of ‘usefulness’ can only be assessed for a certain intended use of the data (Devillers et al. 2007). Whether data quality is ‘good enough’ for a specific use depends on whether the data’s properties allow for the question(s) at hand to be answered (Devictor, Whittaker & Beltrame 2010). For example, a precise location in observation data of plants or animals is not important if the data are used for deriving seasonal occurrence for larger regions, but would be important for analysing fine-grained spatial distribution patterns. Bordogna et al. (2014) point to the need for all VGI to assess and improve data quality with respect to the data’s intended use and the data user’s expectations. They propose a framework to match users’ needs and data properties.

### *Principles of quality assurance for user-generated data*

Data quality assurance aims at identifying, correcting and eliminating errors. Chapman (2005) also uses the term ‘data cleaning’. On the one hand, this process includes the identification of formal errors, i.e. missing values, typing errors, etc. On the other hand, the suitability of a (formally correct) data set for a particular purpose depends, as we have already seen, on whether the data’s characteristics (e.g. position accuracy) are sufficient for this purpose. Such uses can be very diverse and are often not fully foreseen prior to data collection (Dickinson, Zuckerberg & Bonter 2010).

Goodchild and Li (2012) identify three basic approaches to quality assurance for VGI, which are also applicable to citizen science observation data in the field of biodiversity.

The ‘**crowd-sourcing approach**’ builds on the assumption that an error cannot persist if many users work on the same data. Hardisty and Roberts (2013) consider this the best method to identify errors in biodiversity data. Goodchild and Li (2012), however, present a good example where this assumption failed, with a wrong name of a golf course in California persisting for years in

Wikimapia (an online map project collecting information on locations from users). They also conclude that what they call 'obscure' objects (e.g. objects that exist only for short periods of time) may be more susceptible to such errors than others. Observations, especially of more mobile animal species, may well be counted among these.

Another principle, termed the '**social approach**' by Goodchild and Li (2012), uses privileged users as controllers validating the data collected in the project. This approach is widely used in citizen science projects in the biodiversity domain (Wiggins et al. 2011). Data validators are often regional experts for a certain species group (Sullivan et al. 2009), responsible for data validation in a certain area that they know well. The validation process sometimes involves communication between data reviewers and observers, when a reviewer requests more specific information about an unusual report (Bonter & Cooper 2012) that may help to validate it.

In the '**geographic approach**', Goodchild and Li (2012) summarize all methods using rules formalising geographic context. As Elwood, Goodchild and Sui (2012: 580) conclude, '... the richness of geographic context (...) makes it comparatively difficult to falsify VGI, either accidentally or deliberately.' Methods based on this principle allow for automatic verification of data. The necessary geographic context can be gained from observation data already existing in the project in question. This approach requires large amounts of existing data with a relatively high spatial density (Conrad & Hilchey 2011), often not (or not yet) available in citizen science data sets in the biodiversity domain. Consequently there is a need for methods relying on other context sources. Using external context data may provide a solution to this challenge (Elwood, Goodchild & Sui 2012), adding the question of data quality of these context data to the picture. Goodchild and Li (2012) conclude that there is a need for the formalization of relevant geographic context and the rules for describing it.

Using geographic context with distribution data of organisms shows certain methodological similarities with niche or habitat modelling, using known occurrences or absences of a species or of species communities in order to find correlations between these occurrences and a number of environmental factors, with the goal of predicting occurrences (or, at least, finding suitable habitats) in regions without available occurrence data (Engler et al. 2004). Many niche modelling methods need absence data (that is, data about locations where the species in question is definitely not present) to work (Engler, Guisan & Rechsteiner 2004). However, the inability to provide absence data is a notorious weakness of citizen science data in the biodiversity domain, especially if collected as casual data in an opportunistic way (Chapman 2005). This disadvantage can be overcome (or at least mitigated) by using an appropriate project design concerning the protocols and procedures to be followed at observation data collection. A well-established approach is the use of species checklists, allowing to differentiate between species that were observed at a certain place and time and species that were not (for example, the project eBird

or the German ornitho.de platform use this method). Certain issues like the detectability of species still need to be taken into account when working with this approach.

### **Quality assurance for user-generated data from citizen science projects: research and practice, shortcoming and possible solutions**

Wiggins et al. (2011) conducted a study analysing the quality assurance mechanisms used in citizen science projects. They found that many projects assure the quality of the data produced by implementing suitable measures before data collection (e.g. project design, training of participants, etc.), while manual validation of observation data by experts is the dominant approach for ex post verification of data. The assessment of correctness (or 'truth') of an observation is based on the plausibility of that observation in the light of the information provided with the observation. The expert's knowledge about the species and the region the observation comes from serve as reference information for the assessment. Also, photographs are often used as evidence.

Some projects employ automatic assessments of the plausibility of observations. For instance, the project eBird, considered as a 'gold standard' source for bird observations from citizen scientists for use in scientific research, checks the numbers of individuals of species specified by the observer for plausibility, taking into account the location and the season (Sullivan et al. 2009). If the numbers are considered implausible, the observer gets feedback right away. If he or she insists, the observation is passed on to a regional expert for validation. This is also the case for observations that contain species not listed in the species checklist provided to the observer for the location and season (observers can manually add species to the list). eBird now also uses the large amount of data already accumulated in the project to determine parameters for its filter mechanisms, improving filtering results concerning unusual observations (Sullivan et al. 2014). In the German portal 'naturgucker', observers get hints from the system if an observation has certain properties making it implausible. For example, the system checks whether the reported species usually occurs in the region and at the time the observation was made. Another filter checks whether the species has been reported from that region before. Reports of uncommonly rare species will also lead to appropriate feedback to the observer. This project does not flag reports or pass them on to experts for verification, leaving further data quality control entirely up to the crowd. Project Feeder Watch, a North American bird monitoring program, has automatic filters very similar to those of the project eBird, as well using species check lists for regions and seasons, and numbers of individuals observed. Bonter and Cooper (2012) point to the inability of such filters to detect plausible but false reports, and see a need for more research in this area. They expect advances

through combining different approaches for plausibility assessment, including assessment of the observers' expertise or experience. Concerning contributors and their properties, Schlieder and Yanenko (2010) explored approaches using social distance between contributors as a confirming factor for the reliability of VGI contributions closely related in space and time. However, such concepts are hardly applicable for citizen science data from the biodiversity domain, as suitable information about contributors to measure their social distance is very rarely available. For an overview of the data quality assurance strategies in projects mentioned in this section, see Table 1.

Many citizen science projects in the field of biodiversity collect observations of plants, animals and fungi in an opportunistic way, producing so called casual data without imposing strict rules or protocols on the contributors. Volunteers contributing to such projects are free to collect and submit observations of a large number of different species at any time and from any place (examples are the Swedish Artportalen project and iNaturalist, an American project with a world-wide scope; see Table 1 for an overview of their respective data quality assurance strategies). This approach has the potential of producing large amounts of data, as the effort required from volunteers is relatively low,

Project	Data quality assurance strategies and options, in terms used by	
	Goodchild and Li (2012)	Wiggins et al. (2011)
<b>eBird</b> ( <a href="http://ebird.org">http://ebird.org</a> )	Social approach	Filtering of unusual reports, contacting participants about unusual reports, expert review
<b>Project Feeder Watch</b> ( <a href="http://feederwatch.org">http://feederwatch.org</a> )	Social approach	Filtering of unusual reports, contacting participants about unusual reports, expert review
<b>Ornitho.de</b> ( <a href="http://www.ornitho.de">http://www.ornitho.de</a> )	Social approach	Contacting participants about unusual reports, expert review
<b>naturgucker</b> ( <a href="http://www.naturgucker.de">http://www.naturgucker.de</a> )	Crowd-sourcing approach	Filtering of unusual reports
<b>Artportalen</b> ( <a href="http://www.artportalen.se">http://www.artportalen.se</a> )	Social approach	Expert review
<b>iNaturalist</b> ( <a href="http://www.inaturalist.org">http://www.inaturalist.org</a> )	Crowd-sourcing approach	Filtering of unusual reports

**Table 1:** Data quality assurance strategies and options employed by the citizen science projects cited in this chapter, in terms used by Goodchild and Li (2012) and Wiggins et al. (2011), respectively. Information about the projects' data quality assurance strategies can be found on their web sites (see table).



encouraging participation and thus furthering high numbers of participants. However, this kind of data has increased needs for ex post quality assurance and suitable data quality parameters, because the usefulness of such projects and their data for science, administration, and planning is often questioned due to a lack of ex ante quality assurance measures (e.g. training of volunteers, implementation of monitoring schemes, etc.).

Most observations consist of at least the species, location, time, and observer, sometimes supplemented with more (project-specific) information. Therefore, methods for quality assurance or plausibility assessment needing only the four basic aspects of an observation have the potential to be useful for many different projects and data sets, but data properties have to be carefully examined in any case. For example, a seemingly exact location in the form of coordinates can have a wide range of accuracy, or even represent different types of locations (i.e. an exact location vs. the centre of a map quadrant).

## Conclusion

The scientific studies cited in this chapter, as well as the examples given, provide an overview of the most important aspects of quality of citizen science data from the biodiversity domain and its assurance. They show that manual validation of observations of species by experts based on an assessment of their plausibility in the light of available context information is the dominant approach in citizen science projects in the biodiversity domain. The use of automatic (or semi-automatic) approaches for plausibility assessment is increasing, yet they have important shortcomings as described in section 3. Employing the geographic context for plausibility assessment of crowd-sourced geographic data has high potential for assessing the plausibility of species observations in a (semi)automatic way, despite being rarely used so far. There is a great need for further research on methods to assess the plausibility of citizen science data in the biodiversity domain taking their specific properties into account.

## References

- Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J., Pecl, G. T., Barrett, N., & Frusher, S. 2014. Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation*, 173: 144–154. DOI: <http://dx.doi.org/10.1016/j.biocon.2013.07.037>
- Bonter, D. N., & Cooper, C. B. 2012. Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment*, 10(6): 305–307. DOI: <http://dx.doi.org/10.1890/110273>
- Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., & Rampini, A. 2014. On predicting and improving the quality of Volunteer Geographic Information

- projects. *International Journal of Digital Earth*. DOI: <http://dx.doi.org/10.1080/17538947.2014.976774>
- Chandler, M., Bebbler, D. P., Castro, S., Lowman, M. D., Muoria, P., Ogue, N., & Rubenstein, D. I. 2012. International citizen science: making the local global. *Frontiers in Ecology and the Environment*, 10(6): 328–331. DOI: <http://dx.doi.org/10.1890/110283>
- Chapman, A. D. 2005. *Principles and Methods of Data Cleaning – Primary Species and Species-Occurrence Data, version 1.0. Report for the Global Biodiversity Information Facility*. Copenhagen, Denmark: GBIF.
- Connors, J. P., Lei, S., & Kelly, M. 2012. Citizen Science in the Age of Neogeography: Utilizing Volunteered Geographic Information for Environmental Monitoring. *Annals of the Association of American Geographers*, 102(6): 1267–1289. DOI: <http://dx.doi.org/10.1080/00045608.2011.627058>
- Conrad, C., & Hilchey, K. 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental Monitoring and Assessment*, 176(1): 273–291. DOI: <http://dx.doi.org/10.1007/s10661-010-1582-5>
- Cosquer, A., Raymond, R., & Prevot-Julliard, A. 2012. Observations of Everyday Biodiversity: a New Perspective for Conservation? *Ecology and Society*, 17(4): 2–16. DOI: <http://dx.doi.org/10535/8668>
- Devictor, V., Whittaker, R. J., & Beltrame, C. 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions*, 16(3): 354–362. DOI: <http://dx.doi.org/10.1111/j.1472-4642.2009.00615.x>
- Devillers, R., Bédard, Y., Jeansoulin, R., & Moulin, B. 2007. Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International journal of Geographic Information Science*, 21(3): 261–282. DOI: <http://dx.doi.org/10.1080/13658810600911879>
- Dickinson, J. L., Zuckerberg, B., & Bonter, D. N. 2010. Citizen Science as an Ecological Research Tool: Challenges & Benefits. *The Annual Review of Ecology, Evolution, & Systematics*, 41: 149–172. DOI: <http://dx.doi.org/10.1146/annurev-ecolsys-102209-144636>
- Dickinson, J. L., & Bonney, R. (ed.). 2012. *Citizen Science: Public Participation in Environmental Research*. Ithaca and London, Comstock Pub. Associates: 300 pp.
- Elwood, S., Goodchild, M. F., & Sui, D. 2012. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers*, 102(3): 571–590. DOI: <http://dx.doi.org/10.1080/00045608.2011.595657>
- Engler, R., Guisan, A., & Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2): 263–274. DOI: <http://dx.doi.org/10.1111/j.0021-8901.2004.00881.x>

- Goodchild, M. F. 2007. Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4): 211–221. DOI: <http://dx.doi.org/10.1007/s10708-007-9111-y>
- Goodchild, M.F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1: 110–120. DOI: <http://dx.doi.org/10.1016/j.spasta.2012.03.002>
- Haklay, M. (2013). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: Sui, D., Elwood, S., & Goodchild, M. F. (Eds.), *Crowdsourcing geographic knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. New York: Springer, pp. 105–122. DOI: [http://dx.doi.org/10.1007/978-94-007-4587-2\\_7](http://dx.doi.org/10.1007/978-94-007-4587-2_7)
- Hardisty, A., & Roberts, D. 2013. A decadal view of biodiversity informatics: challenges and priorities. *BMC Ecology*, 13(16). DOI: <http://dx.doi.org/10.1186/1472-6785-13-16>
- Howe, J. 2006. The rise of crowdsourcing. *Wired*, 14(6).
- International Organisation for Standards. 2002. *ISO 19113 Geographic information – Quality principles*. Geneva, Switzerland: ISO.
- Kresse, W., & Fadaie, K. 2004. *ISO Standards for Geographic Information*. Heidelberg: Springer.
- 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. DOI: <http://dx.doi.org/10.1890/110278>
- Schlieder, C., & Yanenko, O. 2010. Spatio-temporal proximity and social distance: a confirmation framework for social reporting. In: Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks (LBSN 2010), San Jose, CA, USA, November 03–05, 2010: pp. 60–67.
- van Strien, A. J., van Swaay, C. A. M., & Termaat, T. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, 50(6): 1450–1458. DOI: <http://dx.doi.org/10.1111/1365-2664.12158>
- Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. 2009. eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10): 2282–2292. DOI: <http://dx.doi.org/10.1016/j.biocon.2009.05.006>
- Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A. A., Dieterich, T., Farnsworth, A., Fink, D., Fitzpatrick, J. W., Fredericks, T., Gerbracht, J., Gomes, C., Hochachka, W. M., Iliff, M. J., Lagoze, C., La Sorte, F. A., Merrifield, M., Morris, W., Phillips, T. B., Reynolds, M., Rodewald, A. D., Rosenberg, K. V., Trautmann, N. M., Wiggins, A., Winkler, D. W., Wong, W., Wood, C. L., Yu, J., & Kelling, S. 2014. The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, 169: 31–40. DOI: <http://dx.doi.org/10.1016/j.biocon.2013.11.003>

- Theobald, E. J., Ettinger, A. K., Burgess, H. K., DeBey, L. B., Schmidt, N. R., Froehlich, H. E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch M. A., & Parrish, J. K. (2015). Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, 181: 236–244. DOI: <http://dx.doi.org/10.1016/j.biocon.2013.11.003>
- Wiggins, A., Newman, G., Stevenson, R. D., & Crowston, K. 2011. Mechanisms for Data Quality and Validation in Citizen Science. In: *Proceedings of the IEEE Seventh International Conference on e-Science*. Stockholm, Sweden, December 05–08, 2011: pp. 14–19.