Report from Dagstuhl Seminar 17272

Citizen Science: Design and Engagement

Edited by

Irene Celino[1], Oscar Corcho[2], Franz Hölker[3], and Elena Simperl[4]

1 **CEFRIEL - Milan, IT,** irene.celino@cefriel.com
2 **Polytechnic University of Madrid, ES,** ocorcho@fi.upm.es
3 **IGB - Berlin, DE,** hoelker@igb-berlin.de
4 **University of Southampton, GB,** e.simperl@soton.ac.uk

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 17272 "Citizen Science: Design and Engagement". In this report, we briefly summarise the content of three invited keynote talks and two invited tutorials. We further outline the findings of five parallel working groups, which met on the first and third day of the workshop, in the areas of: sustainability, measuring success, community engagement, linking and quality.

Seminar July 2–5, 2017 – http://www.dagstuhl.de/17272
1998 ACM Subject Classification I.2 Artificial Intelligence, I.2.9 Robotics, Society, Human-computer Interaction
Keywords and phrases Citizen science, Crowdsourcing, Data Analytics, Gamification, Human Computation, Incentives Engineering, Online Community, Open Science
Digital Object Identifier 10.4230/DagRep.7.7.22
Edited in cooperation with Neal Reeves (University of Southampton)

1 Executive summary

Irene Celino
Oscar Corcho
Franz Hölker
Elena Simperl

Citizen science is an approach to science that is enlisting the help of millions of volunteers across a range of academic disciplines to complete tasks that would have otherwise been unfeasible to tackle using expert time or computational methods [2]. While it is a popular and effective way to solve various problems, with many examples of incredible success [3, 1], there remains a number of ongoing challenges that must be addressed in order to ensure the validity of citizen science as a widespread approach to research. The aim of this workshop – organised in partnership with the SOCIAM[1] and Stars4All[2] projects – was to discuss and explore aspects of the future of citizen science, focusing on design factors and engagement strategies, although this naturally required a holistic assessment of citizen science projects, platforms and applications as a whole.

[1] https://sociam.org/about
[2] http://stars4all.eu/

References

1 Jeehung Lee, Wipapat Kladwang, Minjae Lee, Daniel Cantu, Martin Azizyan, Hanjoo Kim, Alex Limpaecher, Snehal Gaikwad, Sungroh Yoon, Adrien Treuille et al. *RNA Design Rules from a Massive Open Laboratory*. Proceedings of the National Academy of Sciences, National Academy of Sciences, 111; 6, 2122–2127, 2014.

2 Chris J Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Bamford Steven, Daniel Thomas, Jordan M Raddick, Robert C Nichol, Alex Szalay and Dan Andreescu et al. *Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey*. Monthly Notices of the Royal Astronomical Society, Blackwell Publishing Ltd Oxford, UK, 389; 3, 1179–1189, 2008.

3 Chris J Lintott, Kevin Schawinski, William Keel, Hanny Van Arkel, Nicola Bennert, Edward Edmondson, Daniel Thomas, Daniel JB Smith, Peter D Herbert, Matt J Jarvis et al. *Galaxy Zoo:'Hanny's Voorwerp', a quasar light echo?*. Monthly Notices of the Royal Astronomical Society, Blackwell Publishing Ltd Oxford, UK, 399; 1, 129–140,2009.

2 Table of Contents

3 Introduction

While amateur involvement in science began long before the establishment of modern academic institutions, the web and digital technologies have fundamentally revitalized and expanded the ways and scale in which untrained citizens can participate in scholarly research. These 'Citizen Science' projects have thus far enlisted the help of millions of volunteers in a wide array of scientific inquiries, ranging from the taxonomic classification of galaxies and the creation of an online encyclopedia of all living species on Earth, to the derivation of solutions to protein folding problems and the tracking and measuring the population and migratory patterns of animals in the Serengeti national park [2]. This new, more inclusive way of pursuing science is proving successful in many ways: it gives scientists around the world an effective, affordable way to collect and analyze large amounts of data in a short period of time, popularizes scientific topics to wider audiences, and encourages the formation of amateur scientific communities, which initiate their own projects and deliver notable results.

The seminar was arranged as a platform to discuss and explore the aspects of and to outline the future of the citizen science research, platforms, and applications. The seminar was organized around to the following three perspectives:

1. A crowdsourcing perspective that views citizen science as a large-scale volunteer-driven human computation system. Relevant aspects include:
   - Task and workflow design
   - User experience design
   - Answer validation
   - Task assignment and contributor performance
   - Crowd learning, feedback, and tutorials
   - Gamification and rewards
2. An online community perspective that considers social and other communication and interaction activities that support task-centered efforts. This is related to quantitative and qualitative approaches for content and community analysis, including:
   - Analysis of discussion forum and chat activity
   - Social network analysis
   - Interplay with other community spaces such as social media
   - Analysis of community trajectories
   - Lurker behavior and more general contribution patterns
   - Conflict and collaboration
   - Surveys of motivation and incentives
3. An open science perspective that focuses on citizen science as an emerging model of collaborative research. In particular:
   - Open, participatory approaches to all stages of the scientific lifecycle
   - Crowdfunding for science
   - Open access publishing of research ideas and outcomes
   - Openness in data acquisition and sharing
   - Participation of young volunteers in citizen science activities
   - Scientific publishing for crowdsourced science

In order to discuss the design and engagement of citizen science, the workshop consisted of a number of talks and working groups that incorporated these differing perspectives throughout.

4 Overview of Keynote Talks

4.1 Crowdsourcing for Smart Cultural Heritage: Harnessing Human Semantics at Scale

Lora Aroyo (Vrije Universiteit Amsterdam, NL)

The state of the art in machine learning and information extraction has advanced the detection and recognition of concepts and objects, like people, locations, and various other types of named entities. However, still there is various types of human knowledge that cannot yet be captured by machines, especially when dealing with wide ranges of real-world tasks and contexts. The key scientific challenge is to provide an approach to capturing human knowledge in a way that is scalable and adequate to real-world needs. Human Computation has begun to scientifically study how human intelligence at scale can be used to methodologically improve machine-based knowledge and data management. My research focuses on understanding human computation for improving how machine-based systems can acquire, capture and harness human knowledge and thus become even more intelligent. In this talk I will present use cases related to smart culture, e.g. enrichment of cultural heritage collections of artworks, videos, newspapers, etc. I will show how the CrowdTruth crowdsourcing framework http://crowdtruth.org facilitates data collection, processing and analytics of human computation knowledge. Processing real-world data with the crowd leaves one thing absolutely clear - there is no single notion of truth, but rather a spectrum that has to account for context, opinions, perspectives and shades of grey. CrowdTruth is a new framework for processing of human semantics drawn more from the notion of consensus then from set theory.

4.2 Open Citizen Science: What, how and with whom?

Claudia Göbel (Museum für Naturkunde Berlin/ European Citizen Science Association, DE)

The presentation focused on unpacking relations of open and collaborative aspects in Citizen Science. On a conceptual level, I identified synergies and tensions between Citizen Science and Open Science by mapping agendas from the European research policy discourse against each other. What is done in Citizen Science practice was explored by looking at findings of an international stakeholder analysis on Citizen Science data interoperability and examples from other areas of Open Science. Finally, I sketched some cornerstones of an analytical perspective focusing on stakeholder networks and organizations for further analysis.

4.3 The state of the art of OpenStreetMap: technologies, community and research challenges

Maurizio Napolitano (FBK - Centre for Information Technology/ Open Street Map, IT)

OpenStreetMap - OSM is known as the free world map created on a voluntary basis. But OSM is not just a map, it is much more: one of the biggest geo-referenced open data resources, a community of people who are able to give voice to a territory, an important resource for entrepreneurial initiatives, and much more. This talk introduced the project from its history and technological aspects in order to highlight what challenges science could help solving and the benefits deriving from it.

5 Overview of Tutorials

5.1 Citizen Science as a New Way To Do Science

Marisa Ponti (University of Göteborg, SE)

Main reference Marisa Ponti, "Citizen Science as a New Way To Do Science", SocArXiv, 2017.
URL http://dx.doi.org/10.17605/OSF.IO/KGXSQ

Citizen science has received increasing attention because of its potential as a cost-effective method of gathering massive data sets and as a way of bridging the intellectual divide between layperson and scientists. Citizen science is not a new phenomenon, but is implemented in new ways in the digital age, offering opportunities to shape new interactions between volunteers, scientists and other stakeholders, including policymakers. Arguably, citizen science rests on two main pillars: openness and participation. However, openness remains unexploited if we do not create the technical and social conditions for broader participation in more collaborative citizen science projects, beyond collecting and sharing data with scientists. "Public participation" has too often accounted for the assumed ease with which hierarchies in science can be horizontalized, and economic and geographic barriers can be removed. However, public participation is a contested term, which should be problematized. The Scandinavian tradition of participatory design can help explore conceptually the challenges related to participation and to design for participation.

5.2 Citizen Science as a Social Machine

Elena Simperl (University of Southampton, GB)

Tim Berners-Lee envisaged the web as a platform for large-scale participation. Social machines are socio-technical constructs that enable all this. They define and support processes where people contribute creative work and algorithms carry out the more routine, predictable or "engineer-able" tasks to bring the results of that work together. Citizen science is itself an

example of a social machine: bringing groups and communities of volunteers together to help advance science by collecting or analysing data relevant to scientific experiments. To date, lots of the research has focused on showcasing how and where citizen science can help or studying specific properties of the systems using a variety of methods from social computing, HCI, or scientific communication. We need to appreciate that citizen science can be studied from several angles, and that more research is needed in understanding how citizen science can move away from being just an approach useful for the professional scientist to a fully-fledged social machine. In the same time, we need to design better tools and teach the scientists to understand how to improve and maintain community health and determine whether the volunteers are happy and producing enough data. This talk introduced a number of research questions and solutions, such as how to design platforms for meaningful contributions, how to manage the data produced by the volunteers, studying emerging communities, and designing incentives for participation. The ultimate question is 'what makes citizen science successful?' and we need a interdisciplinary, inclusive approach to solve it.

6 Overview of Working groups

Working groups were held on days 1 and 3 of the workshop in order to facilitate discussions on topics related to the areas outlined above. In the following sections we provide brief summaries of the main outputs of these working groups.

6.1 Working Group One - Sustainability

Reported by Alessandro Bozzon. This group discussed various dimensions of sustainability related to citizen science. These are presented in the table below, along with a series of challenges for each that must be overcome in order to ensure sustainable projects and practices. In addition the group discussed a number of developments that may help to increase sustainability, such as a 'graveyard' of projects that have run their course, addressed their hypothesis and wrapped up engagements with their community, along with a repository of workflows that have been annotated with what problems they've worked well on previously. This would allow other citizen science teams to see what has worked, and build from there. Other suggestions were a high-school level curriculum module on citizen science, to increase awareness of this from a younger age, and a re-usable set of personas to help in the design process.

1. Openness and Reuse of Data
   - Data and Metadata
     - Incentivising data sharing
     - Dealing with diverse standards
     - Licensing
     - Provenance and tracking
   - Technology
     - Creating a Github for open citizen science hardware and software
     - Designing projects for re-use
     - Describing workflows
     - Standardisation
     - Document deployment configurations and conditions

- Discovering existing initiatives
  - Creating a citizen science directory
  - http://firstmonday.org/ojs/index.php/fm/article/view/5520/4194
  - https://scistarter.com/
  - http://vgibox.eu/
  - https://www.citizensciencealliance.org/
  - https://en.wikipedia.org/wiki/List_of_citizen_science_projects
- Methodology
  - Co-creation and knowledge sharing for project design and implementation
  - People
  - Incentives
  - Tasks
  - Trade-offs
  - Lessons Learned
  - For example, repository of case studies, personas, design guidelines, etc.

2. Socio-economic aspects
   - Engagement/Participation
     - Limited number of potential participants
     - Reaching motivated participants
   - Fitness for purpose/availability
     - Evidence of required expertise
     - Trustworthiness
     - Spatial and temporal activity patterns
     - Responsiveness
   - Agency/Empowerment
     - Community tools for building citizen science
     - Co-creation of projects and research
   - Education
     - Accepted as curriculum
     - Hands-on and built on practice
     - Revisit pedagogical approaches
   - Social and Ethical goals
     - Encourage empowerment of under-represented groups (e.g., migrants)
     - Increase enthusiasm for science in lay public
3. Ecosystemics
   - Project Variety
     - Ensuring ecosystem is fit for different timescales and project cultures
   - Project Lifecycle
     - Funding - connecting funding for infrastructure and for doing science
     - End of Life - data archival, communities and connections
     - Creating bridges to other projects

6.2 Working Group Two - Measuring Success

Reported by Paul Groth. The success working group discussed a range of topics such as investigating what constitute 'success' in a citizen science context, the different types of outcome, and different frameworks for assessing various facets of success. The group then moved on to outline a range of grand challenges for this topic area. Each challenge was summarised with an impact statement of where, ideally, the state of play would be in ten years time.

Predict the success of a project before starting, based on a manual of best practices. For this challenge, the different criteria for different stakeholder groups were discussed, along with what features can be looked at to understand success. The scientists themselves, funders, and citizens may all have differing expectations about what success means in a citizen science project, and going even further there are planetary scale stakeholders that are concerned with the positive ecological impacts that many of these initiatives may have. Success may be measured using a huge rage of metrics and features, going beyond measures of the number of classifications or volunteers to look at how many visitors the project has attracted, the number of scientists and developers engaged in the process, the quality of the input data, the number of countries reached, etc. As a grand vision, in ten years time, we will be able to predict successful citizen science projects and tell you why.

Expand the diversity of contributors to citizen science. In order for citizen science as an approach to be truly successful, there needs to be a diverse range of volunteers who participate in projects to ensure that there is no cultural, socio-economic, age, educational or language bias recorded in the data. We identified a number of steps required such as identifying what the current bias is, and eliciting methods of motivating more people to participate, from different backgrounds. There were also discussions around what automatic adjustments to design could be made so that an existing project or tasks could easily be replicated or made suitable for a different demographic group. The proposed impact statement for this group was: in ten years time, we will have citizen science projects that reflect the demographic breakdown of the world.

Radically scaling up co-created citizen science. Citizen science already reaches large numbers of volunteers in some cases, with this an essential aspect of many projects. However the role that the volunteers play in the design process is less established, and one of the areas of further potential is a true collaborative, co-creation process at all stages of the scientific process (as opposed to the majority of current projects that incorporate volunteer involvement in the data collection and/or analysis phase). Foundational work needs to be carried out to understand what is currently done in co-creation process so that experiences and lessons learned may be documented. A testing and development process for components of the scientific method workflow will then allow the creation of processes that work across multiple scientific domains, before over time things are scaled up so that performance may be tested against a baseline to determine the successfulness of running truly co-created citizen science projects. If it is possible for a completely citizen-led project to match or even beat the performance of professional scientists' own research then citizen science will have reached a high level of success as an approach to executing the scientific method. In ten years time, a citizen science platform with over 100,000 people will beat the performance of professional scientists for the full scientific process.

6.3 Working Group Three - Community Engagement

Reported by Oscar Corcho and Christopher Kyba. Community engagement is crucial in order to ensure that volunteers continue to contribute to a project, as well as return to both the current project and future projects that may be of interest to them. This working group investigated two challenges as outlined below.

A framework and contextual conditions for citizen science projects

The current status-quo. Funding agencies have not adapted their funding strategies/calls to account for applying Citizen Science in any type of project. Either a project is submitted as a specific Citizen Science project for a Citizen Science call, or if it is a general project call then there is no clear space for the funding characteristics of Citizen Science projects.

- In some cases (e.g., some Ministries in Germany), there are no possibilities of having follow-up projects, and for a successful citizen science project this may not be adequate (you make a community orphan). The usual term for a research project (3 years) is not enough for building a successful community (which requires between 3 and 6 years).
- Private and public foundations (e.g., Wellcome Trust, Knight Foundation, World Development Bank, UN) in some countries are providing some smaller funding for this type of initiatives.
- There is not infrastructure-type of funding for Citizen Science projects, as it is done for other types of large-scale research infrastructure-related projects (e.g., ESFRI projects).

There are costs associated to Citizen Science projects that are not easy to claim for (e.g., creating a good brochure to reach a large set of citizens - 1 month of work instead of a couple of days needed for a scientific community -, to do scientific outreach/transfer) or that are not easy to get into the accounts of scientific organisations (e.g., problems with audits). How do you pay freelancers or crowdworkers?

Citizen Science projects are generally more risky (e.g., engagement with people may not work and they may not show up). We should be also more open to the fact that experiments may generate less (or more) data than originally expected, or with a different quality to the original expectations. So there must be more room for flexibility (even changing the hypotheses during the project execution), and some more knowledge inside research organisations of the characteristics of Citizen Science projects.

Research organisations have already some budget for scientific outreach (e.g., open days, education activities for secondary schools). However, Citizen Science projects are not normally considered as fundable under this budget, even if they cover some of these needs (creating awareness about Science for Society), but only fundable under the budgets for research projects.

Challenges. The following challenges were identified as necessary to overcome in order to advance this aspect of citizen science research:

- Change the funding (and accounting) mechanisms from funding agencies to adapt better to the funding needs for Citizen Science projects (e.g., to be more flexible)
- Expand the collection of organisations that can provide funding for these projects by providing a better explanation of what is done in Citizen Science projects and their potential benefits
- Connect Citizen Science to evidence-based decision making (e.g., fighting against fake news in newspapers)

- Create Citizen Science infrastructure helpdesks inside research organisations or at the national/international level.
- Create specific budgets inside research organisations for funding Citizen Science projects.
- Adapt administrative procedures - e.g. possibilities of paying volunteers for contributions or participation in the project; modalities of payment (paying upfront)
- Appoint Citizen Science coordinators in research organisations, so that they can help scientists to contact and engage citizens.

Actions. In the short term, we need to create a clear set of administrative procedures that can be applied by research institutions when paying volunteers for their contributions. To support this, helpdesks/support units could be created for Citizen Science projects at different levels, including online training and appointing Citizen Science coordinators.

In the medium term, there needs to be influence on public funding bodies in order to incorporate Citizen Science into their usual funding streams. In order to drive this, foundations should be engaged to disseminate information about the opportunities that citizen science can offer, and alongside this we should exploit the social corporate responsibility of many large companies as they may become potential funders. Future research should explore how citizen science could fit new business models for publishers, and there should be work to create budgets within organisations for funding citizen science projects. Finally, citizen science project involvement should be aligned to university curricula so that (for example) students can participate in projects and earn credits towards their degree.

Longer term, the group discussed the creation of citizen science 'wallets' where volunteers could collect remuneration for their contributions and develop a record of how they've helped in various science projects. Additionally, there is the need to create an ESFRI-like structure for large scale projects.

Upscaling and diversity

The current status-quo. As discussed in the diversity section of the 'Success' working group, citizen science project participants are not currently representative of societies in general. They are often drawn from higher educated groups. To some extent this reflects self-selection, but this may be because of lack of awareness in other demographic groups. Not all projects collect participant information, so we don't necessarily have the full picture of who is participating and aggregated statistical information about participant characteristics is not currently available or accessible. Many citizens don't realise that they could be involved in science projects such as these.

Furthermore, in some projects the data often has very strong clustering - with some areas having a lot of data and others having little or none. Participatory approaches have a very strong bias towards Europe and North America and reflect cultural differences: within Europe there is considerable diversity in the tools that are available, and approaches such as bioblitz [1], (popular in e.g. Germany and UK, but not taking place in e.g. Spain). There are cultural differences particularly between Eastern and Western European countries with regards to participatory research. Even within countries, there are differences in participation rates among the population. In Spain for example, participation varies strongly between northern and southern Spain. Language creates a barrier for expanding participation across European countries and also to countries outside of Europe.

Research questions in projects that are labeled Citizen Science are far more commonly initiated by academic researchers, rather that co-created although this is less of the case in "digital social innovation". Citizen science projects are however rarely being initiated by

citizens directly. Not all projects are being funded with a clear "closing strategy" for when the project will be handed off to e.g. a foundation, or how it will be properly ended (e.g. final communications, publications, etc.). As such, there is currently duplication of effort around these areas.

- Individual citizen scientists are very seldomly represented at discussions of citizen science
- Access to paywalled articles is strongly differentiated by class
- Participants are not necessarily getting enough feedback and education from the projects they participate in: missing a chance to expand understanding of critical thinking

Challenges. Subsequently, there are a number of challenges in terms of upscaling and increasing diversity in projects:

- Attraction of participants: how do you get them to know about your project?
  - Overcoming language barriers
  - There is no one "European media", so getting messages out is difficult
  - Finding translators and funding for citizen science projects (apps, lesson plans, official communication and feedback)
  - How do you find "gatekeepers" who can promote your project to members of their group?
- Community management is very difficult with an international participant group
  - There can also be cultural issues in how to communicate (e.g. with students versus elderly people)
- If you want to expand a project into a neighboring country, how do you find who you need to contact?
- There is a lack of funding for community management, funders don't realize that the true cost of a project is much larger than the cost of an app/platform.
- Citizen science funding is often put out in specific calls, and not more generally accepted within disciplinary funding programs
- How could you allow for citizen project initiation, while still dealing with questions of funding oversight?

Actions. To address these challenges, we first need to provide tools and resources to citizen scientists so that they can participate more easily in projects. Funding needs to be allocated to community management and towards a strong communication strategy in order to draw in a more diverse set of participants. To help with this further, online science literacy classes could be provided to give volunteers the skills they need in order to help out, while seminars could be given to scientists on how to interact with citizen volunteers. One larger-scale suggestion was to provide European funding for a citizen scientists conference where participants of projects can attend if they have demonstrated a level of participation in a project. There also needs to be efforts to share data among projects, and disseminate different engagement strategies.

In the medium term, this should go further and demographic information needs to be shared across projects so that we have a better understanding of what communities are being engaged. Citizen science practitioners also need to be up-to-date with what digital transformations are impacting citizen science.

Finally, longer term, there needs to be open access to research papers so that all interested parties can find out about the research opportunities and outcomes in citizen science. We discussed the idea of making citizen science part of the junior high school curriculum so that everyone becomes aware of its existence and so that students have the option to participate throughout their lives.

6.4 Working Group Four - Linking

Reported by Andrea Wiggins. The linking group discussed the complexities involved in linking communities across citizen science projects. Currently, many projects want to link and connect with others in order to enhance the wider citizen science network. However there are currently no incentives for such networking and it is not seen as innovation in citizen science design. There is little funding available for these activities and therefore most attempts generally start from the beginning rather than building on the records of previous attempts.

In the short term, the idea of networking communities within citizen science should be made clear to all projects so that existing and established communities may be used rather than always creating new ones. This should be detailed in the early stages of project formation, including the grant-writing process and project planning. There should also be an emphasis on sharing technology including the tools, user interfaces, data and profiles used within projects, and technologies such as OpenID that are easy to use and already have well documented APIs should be adopted.

In the longer term, there needs to be greater effort to network co-ordinators together. There needs to be an increase in funding allocated towards linking communities, consisting of tasks such as administration, arrangement of workshops and planning strategies. There should also be specific incentivisation mechanisms for linking so that this is not just tacked on to a project but instead planned appropriately. In order to encourage this, a change is required at the grant application stage so that policies and funding opportunities encourage linking to be included in project proposal and if this can be promoted internationally for cross-border sharing then the potential of citizen science will be realised faster.

6.5 Working Group Five - Quality

Reported by Elena Simperl. Data quality is a matter of primary concern for researchers who use citizen science. There are generally three stages at which data quality typically becomes a concern: the input data (e.g. the collection of objects or the results of a machine learning classifier), the data collected or produced by citizens, and the final results of the task. This working group considered mechanisms that projects employ to ensure quality in the data produced by these projects, and discussed frameworks and mechanisms for managing these - in particular a framework of 18 mechanisms presented by Wiggins et al. (2011) that are based on direct experience and existing surveys [3]. The working group took into consideration sources of error in the data, as well as different stages of the citizen science project process which may introduce such errors. Using this as a starting point for our discussions, we considered what the current debate in this area should be in citizen science.

There are different aspects of data quality that need to be considered. Relevance and trustworthiness are required in order to ensure that valid and reliable data has been collected. There then needs to be a representative presence covered by the data, which may be through the density of coverage (e.g. spatially). There also needs to be information to ensure that we know the degree to which a volunteer followed the research protocol - and therefore there needs to be a process of recording this so that it can be assessed. Finally, the data should be reproducible so that the same distributions of data can be obtained if the same process is followed again.

Common standards may help with ensuring a base level of quality across projects, but it is likely that there are too many different project types in order to develop a reliable standard in this area. As such it may be more appropriate to develop standards around the

tools, question types or methods used, and to also consider the domain specific concerns involved in specific projects as many of these already have existing standards. While some of these may be shared between domains and projects, there may also exist a conflict between what constitutes 'good' data.

Because of the complexity of devising a common benchmark for citizen science data quality, a number of issues were listed as things that need to considered in order to proceed. In terms of data quality, it must be established exactly how citizen science differs from conventional science - and whether or not this requires a new way of thinking compared to traditional approaches. Furthermore, the different technologies and tools used by different projects will introduce unique data quality problems, and therefore an understanding of these needs to be shared so that project designers may anticipate and prevent potential problems. Noisy data may be caused by inconsistent or inaccurate sensors (either physical, or human), or through ambiguity and bias in the process. This may originate from the task design, where specific labelling requirements or semantics mean that bias is introduced into the process. It is therefore crucial to understand these so that future projects can be designed to alleviate these problems. Because of the impact that the citizens themselves have on the data, we must also consider the impact from using different participatory or engagement frameworks, and how these in turn determine the resulting data quality.

6.6 Working Group Six - Manifesto for Citizen Science

Reported by Paul Groth, Edited by Neal Reeves A final working group developed a manifesto for the future of Citizen Science, intended to support different stakeholders in producing projects and in publishing scientific results. This document can be viewed at https://goo.gl/ojuZiR.

References

1 Cathy Lundmark. *BioBlitz: Getting into Backyard Biodiversity.* BioScience, JSTOR, 53;4, 329–329, 2003.

2 Alexandra Swanson, Margaret Kosmala, Chriss Lintott, Robert Simpson, Arfon Smith and Craig Packer. *Snapshot Serengeti, High-Frequency Annotated Camera Trap Images of 40 Mammalian Species in an African Savanna.* Scientific Data, Nature Publishing Group, 2, 150026, 2015.

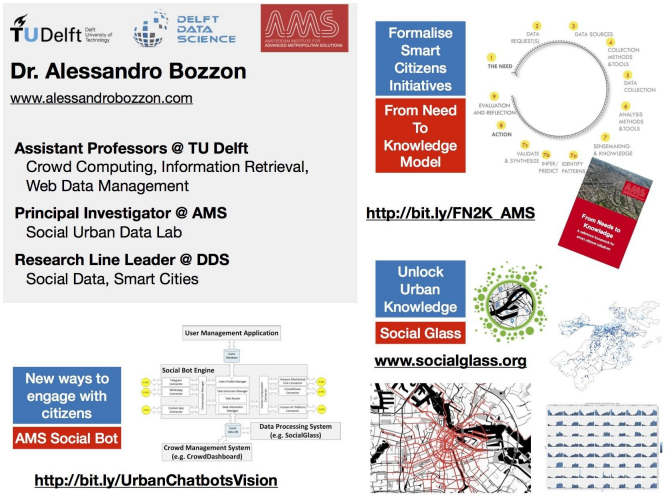
3 Andrea Wiggins, Greg Newman, Robert D. Stevenson and Kevin Crowston. *Mechanisms for Data Quality and Validation in Citizen Science.* Proceedings of the 2011 IEEE Seventh International Conference on e-Science Workshops (ESCIENCEW '11), IEEE Computer Society, Washington, DC, USA, 14-19, 2011.

Participants

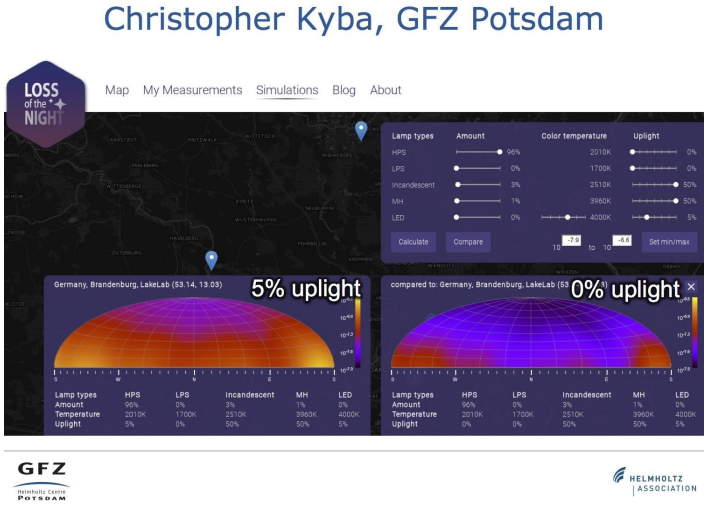
- Lora Aroyo
Free University Amsterdam, NL
- Alessandro Bozzon
TU – Delft, NL
- Maria Antonia Brovelli
Polytechnic University of Milan, IT
- Irene Celino
CEFRIEL – Milan, IT
- Oscar Corcho
Polytechnic University of Madrid, ES
- Dominic di Franzo
Cornell University – Ithaca, US
- Claudia Göbel
Museum für Naturkunde – Berlin, DE
- Esteban González Guardia
Technical University of Madrid, ES
- Paul Groth
Elsevier Labs – Amsterdam, NL
- Lynda Hardman
CWI – Amsterdam, NL
- Franz Hölker
IGB – Berlin, DE
- Tomi Kauppinen
Aalto University, FI
- Christopher Kyba
GFZ – Potsdam, DE
- Dave Murray-Rust
University of Edinburgh, GB
- Maurizio Napolitano
Bruno Kessler Foundation – Trento, IT
- Jasminko Novak
Hochschule Stralsund, DE
- Christopher Phethean
University of Southampton, GB
- Marisa Ponti
University of Göteborg, SE
- Lisa Posch
GESIS - Cologne, DE
- Gloria Re Calegari
CEFRIEL – Milan, IT
- Neal Reeves
University of Southampton, GB
- Sven Schade
EC Joint Research Centre – Ispra, IT
- Sibylle Schroer
IGB – Berlin, DE
- Elena Simperl
University of Southampton, GB
- Alice Verioli
BSDesign – Milan, IT
- Christopher A. Welty
Google – New York, US
- Andrea Wiggins
University of Nebraska – Omaha, US
- Amrapali Zaveri
Maastricht University, NL

8 Appendix A - Lightning Talk Slides

Alessandro Bozzon, TU Delft

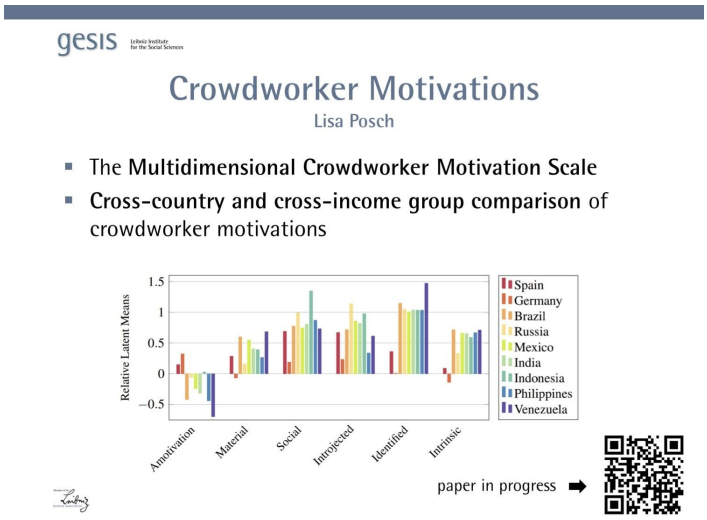


Christopher Kyba, GFZ - Potsdam



Maria Antonia Brovelli, Polytechnic University of Milan

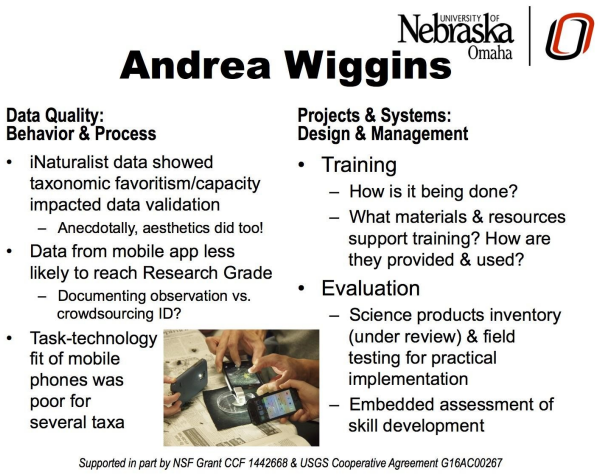
Lisa Posch, GESIS - Köln



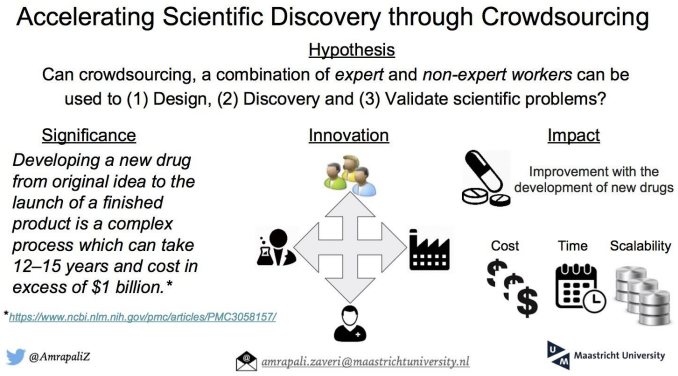
Sven Schade, EC Joint Research Centre

Sibylle Schroer, IGB - Berlin

Andrea Wiggins, University of Nebraska



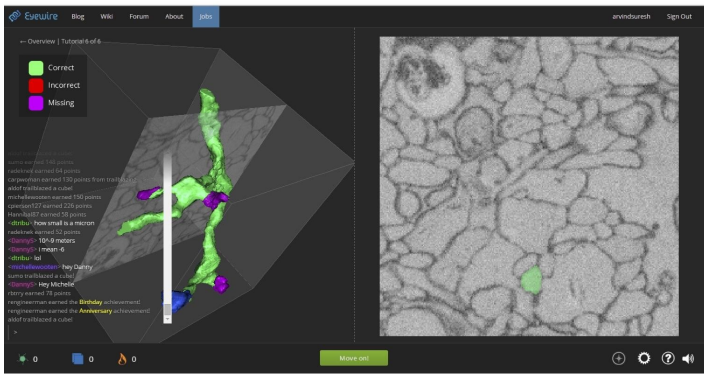
Amrapali Zaveri, Maastricht University



Tomi Kauppinen, Aalto University

Neal Reeves, University of Southampton

Neal Reeves – Sociality and Communication in Virtual Citizen science



Digital Economy
EPSRC
University of Southampton | Web Science DTC

Irene Celino, CEFRIEL

Give and Take in Citizen Science

Take from citizen scientists

- **Gathering** information about their surrounding environment (e.g. urban POIs in http://bit.ly/urbanopoly)
- **Ranking/filtering** information based on "popularity" (e.g. cultural heritage in http://bit.ly/indomilando)
- **Validation** of scientific data (e.g. Land Cover Game at http://bit.ly/foss4game)
- **Classification** of images (e.g. ISS photos in https://www.nightknights.eu/)

Give back to citizen scientists

- **Fun** (e.g. playing "pastime" games)
- **Engagement** (e.g. competition with leaderboard and prizes)
- **Recognition** (e.g. being part of a community, of something "bigger")
- **KNOWLEDGE, INSIGHTS, LEARNING**
  - About their surrounding environment
  - About science and "grand challenges"
  - About themselves (e.g. quantified self, awareness, self-consciousness)

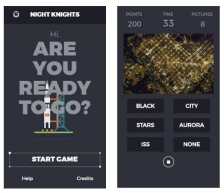
Human Computation, Gamification, Games with a Purpose, Serious Games, Data Analytics, …

Irene Celino – irene.celino@cefriel.com – iricelino.org/publications – ninjariders.eu

DAGSTUHL SEMINAR ON CITIZEN SCIENCE - JULY 2017

Cefriel

Gloria Re Calegari, CEFRIEL



STARS4ALL

Game With A Purpose for Image Classification

http://stars4all.eu/

Gameful Citizens Engagement app for city Co-Creation

https://ninjariders.eu/talkinpiazza/

ninja riders

Decision Making analysis in Mobility field using Gamified tools

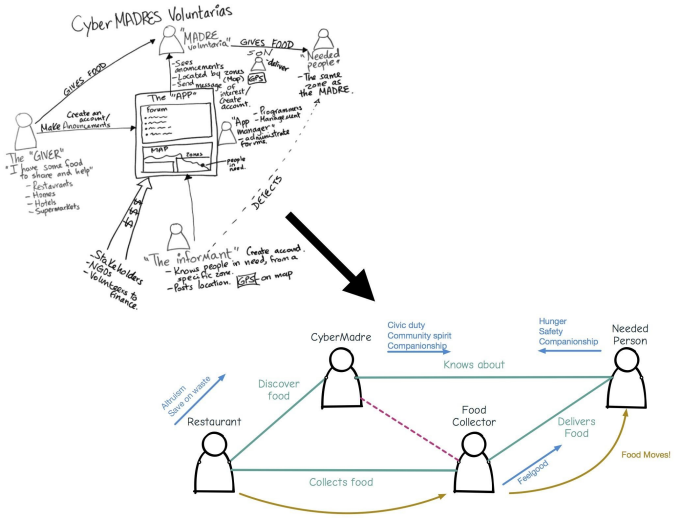
https://www.ninjariders.eu/

Data Analysis and Machine Learning

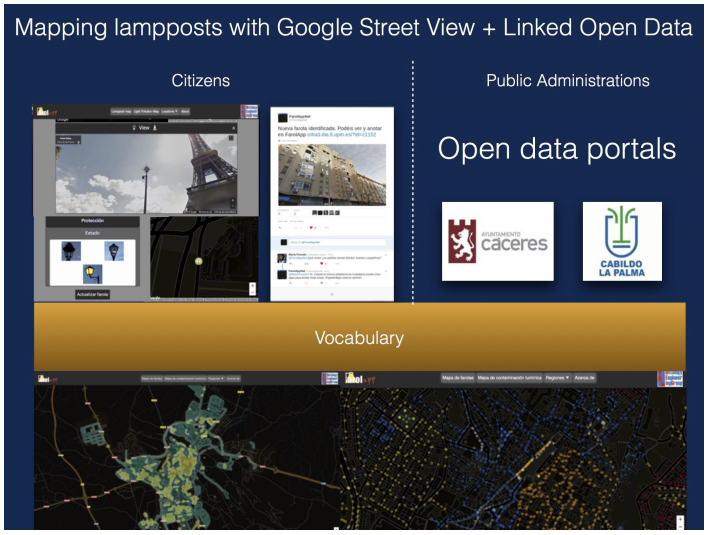
Gloria Re Calegari

Cefriel POLITECNICO DI MILANO

Dave Murray-Rust, University of Edinburgh



Esteban González Guardia, Technical University of Madrid



Chris Phethean, University of Southampton

Jasminko Novak, Hochschule Stralsund