THIS.Institute

The Healthcare
Improvement
Studies Institute

Citizen science: crowdsourcing for research

Catherine Lichten
Rebecca Ioppolo
Camilla D'Angelo
Rebecca K Simmons
Molly Morgan Jones

UNIVERSITY OF
CAMBRIDGE

THIS.Institute

Contents

Authors
Catherine Lichten, RAND Europe
Rebecca Ioppolo, RAND Europe
Camilla D'Angelo, RAND Europe
Rebecca K Simmons, The Healthcare Improvement Studies Institute
Molly Morgan Jones, RAND Europe

THIS.Institute The Healthcare Improvement Studies Institute

RAND EUROPE

Summary

Crowdsourcing draws on a large pool of people to gather inputs such as ideas, funding or labour. It can be used in citizen science research projects, where 'citizens' – usually members of the public – provide inputs and valuable contributions despite not being formally trained experts in the topic of study. This learning report provides a practical overview of the use of crowdsourcing in scientific research projects.

Researchers have used crowdsourcing successfully in a range of projects. The most obvious benefit of crowdsourcing is the ability to collect or analyse data on a much greater scale. A large crowd creates efficiency gains in terms of speed, throughput and cost. As an added benefit, citizen science projects based on crowdsourcing approaches promote dialogue between researchers and citizens. This can help members of the public become more involved in research and make research more democratic.

Researchers interested in incorporating crowdsourcing into their work can make use of a range of existing online platforms and tools. With these tools, researchers can crowdsource data gathering, image classification, systematic reviewing, innovative ideas and funding. New tools continue to be developed, as researchers in diverse disciplines – including healthcare improvement – expand the use of these approaches to tackle new research questions.

There are a number of challenges to consider when involving a large group of participants with diverse backgrounds in research projects. In this report, we offer practical tips to maintain data quality and scientific rigour, and to motivate, retain and reward participants. Researchers should carefully consider Intellectual Property and data ownership issues, and if resources permit, evaluating citizen science projects is strongly recommended.

1. Introduction

Crowdsourcing draws on a large pool of people to gather inputs such as ideas, funding or labour. The term, coined in 2006, has been defined as 'the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call'.[1] Crowdsourcing can be used in citizen science research projects, where 'citizens' – usually members of the public – provide inputs and valuable contributions despite not being formally trained experts in the topic of study.[2]

Crowdsourcing and citizen science offer ways to generate ideas, solve problems, and carry out tasks – all of which can contribute to research processes. Citizen science projects based on crowdsourcing approaches also promote dialogue between researchers and citizens, which can help members of the public become more involved in research and make research more democratic.[3] Crowdsourcing is particularly useful for labour-intensive tasks, making it possible to do work that would be prohibitively time-consuming or resource-intensive for a small group of researchers. Digital technology and the internet have significantly expanded the reach and potential of crowdsourcing and citizen science initiatives.

The aim of this learning report is to provide a practical overview of the use of crowdsourcing for research. It covers good practice, useful tools and illustrative examples and should serve as a useful and up-to-date practical reference for the design of a research programme that makes use of crowdsourcing and citizen science.

2. Methods

We gathered evidence for this learning report by conducting a rapid review of the literature and interviewing crowdsourcing experts. We used Google and Google Scholar to search academic literature, policy reports and non-peer reviewed studies, using search strings with keywords 'citizen science', 'crowdsourcing', 'ethics', 'science' and 'good practice'. After reviewing 460 titles, we eventually included 53 articles in the report, selecting articles written in English, published from 2012 onwards, reporting primary research, and covering the use of citizen science and crowdsourcing in research contexts. We interviewed four people, who were chosen because they are involved in managing different types of projects or online tools that use crowdsourcing for research:

- Daniel Lombraña González, Founder and CEO, Scifabric
- Anna Noel-Storr, Information Specialist and Cochrane Crowd co-lead, Cochrane
- Andy Paterson, Solution Architect, Cancer Research UK
- Helen Spiers, Post-doctoral Researcher in Citizen Science and Medical Research, Zooniverse.

This report is not intended to be a comprehensive overview; the platforms, tools and examples we present are not necessarily representative of the full landscape of crowdsourcing projects. Rather, the report covers key relevant literature, tools and examples and should help to inform the design of research projects that make use of crowdsourcing.

3. Findings

In the sections that follow, we first introduce how crowdsourcing can be used in research, then outline specific applications of crowdsourcing in research. We present a brief overview of tools and platforms that have been used to enable large groups of participants to engage in research projects. Finally, we summarise important considerations related to good practice, and ethical and legal issues that arise when using crowdsourcing in research.

Using crowdsourcing in research

Crowdsourcing and citizen science offer ways to make use of collective intelligence and networked communities to contribute to research processes. In citizen science projects, crowdsourcing has the potential to bring benefits to both those who run the projects and those who participate in them.

The most obvious benefit of crowdsourcing is the ability to collect or analyse data on a much greater scale. A large crowd creates efficiency gains in terms of speed, throughput and cost. In addition, crowdsourcing offers new ideas for research questions and ways to solve problems – drawing on a wider range of perspectives. Crowdsourcing activities have been used in a range of disciplines, including astronomy, ecology, history, medicine and political science. One example that illustrates how a crowdsourcing approach can bring scientific benefit is Foldit, an online game where players contribute to research by solving 3D puzzles. The game attracted tens of thousands of players, who were able to outperform computer algorithms at predicting protein structures and helped solve the structure of a protein that is important for AIDS research (the Mason-Pfizer monkey virus (M-PMV) retroviral protease).[4] Before a team turned to crowdsourcing for the solution, researchers had struggled to work out the structure of the protein – information useful for developing drugs that target it.[5]

Citizen science projects also present an opportunity to involve non-researchers in the scientific process and for researchers to interact with the wider community. Involving the public in research can help to improve scientific understanding and literacy, and enhance public trust in science.[4] It can also help researchers better understand the perspectives of patients or members of the public and ensure these perspectives are used to shape research and policy decisions.[6]

Involving a large group of participants with diverse backgrounds creates a number of challenges when running citizen science projects. Attracting and retaining participants is crucial to the success of crowdsourced research, but can be difficult to achieve and is resource intensive. Participants lack formal training in research methods and will have variable knowledge and competence. Researchers must carefully consider ethical issues, intellectual property rights and copyright agreements. Finally, poorly executed projects can have detrimental impacts on public trust in science and on the willingness of non-researchers to engage with research.[7]

Several of these challenges could have serious implications for the quality and quantity of participation and contributions, and thus for the research itself. As discussed below, it is therefore important to design feasible tasks, check the accuracy of contributions, and provide training and feedback.

Applications of crowdsourcing in research

Crowdsourcing can be used in various phases of the research process, including data collection, processing and interpretation.[8] In line with trends towards widening participation and increasing the involvement of the public in research,[9, 10] crowdsourcing can also be used to shape research priorities and project design. Table 1 outlines a number of applications of crowdsourcing in research and provides relevant examples.

Platforms and tools

Researchers interested in incorporating crowdsourcing into their work can make use of a range of existing online platforms and tools. In this section, we highlight examples of citizen science project platforms, platforms for crowdsourcing systematic literature reviews, labour market platforms and others.

Citizen science platforms

Citizen science project platforms are websites that host or link to a range of projects. They can serve as a useful resource for members of the public to discover projects and may also offer resources for researchers to create projects. Two citizen science platforms that offer both project building and hosting capabilities are Zooniverse and Scifabric.

Zooniverse is free to use, while Scifabric is a commercial platform. Zooniverse offers an online project builder that anyone can use to develop their own citizen science project. Project leads can then apply to have their project shared publicly on the Zooniverse platform, a process requiring review by both the Zooniverse team and volunteers, or they can share their project privately with their own crowd. Zooniverse projects have included Galaxy Zoo and Cancer Research UK's Cell Slider (case study on page 9). Scifabric has developed projects for London Zoo and the British Museum, among others.

Table 1 Applications of crowdsourcing in research	
Application	**Description and examples**
Data processing	
Image classification	This is a common application of citizen science in crowdsourcing. It is useful for images that humans can interpret better than computers can, such as images of landscapes containing wildlife. It is used in many projects on the Zooniverse platform.
	Examples: CRUK's Trailblazer (case study on page 9); Zooniverse's Galaxy Zoo; and testing of a computer-aided detection system for detecting polyps in CT colonography images.[11]
Classification of data other than images	Data types include audio, handwritten text and scientific abstracts.
	Examples: Cochrane Crowd (case study on page 10) and the crowdsourced testing of a clinical decision support system (case study on page 11).
Data collection	
Data gathering	Data gathering is a common application of crowdsourcing in citizen science. Crowds can collect large volumes of data covering many geographical locations or moments in time.
	Examples: Flu Near You,[12] uBiome[8] and other microbiology studies (MetaHIT, Pathomap[4]); MapMyHeart challenge to map cardiac defibrillators.[13] the Flint Water Study, where citizens concerned about their water quality worked with university researchers who tested samples provided by Flint residents;[14] the Great Backyard Bird Count, one of the largest citizen science projects in the US.[15]
Problem solving	
Contests and prizes	Participants, who are often not traditional experts in the subject area, compete to solve problems or develop novel ideas. This approach enables benefits by drawing on diverse perspectives.
	Examples: Gates Foundation 'Reinvent the Toilet Challenge'; Computational biology algorithm contest.[16]
Puzzle games	'Games with a purpose' may encourage participation because they are fun and do not require knowledge of the underlying research questions.
	Examples: Foldit, an online game where players contribute to research by solving 3D puzzles to predict protein structures; Fraxinus, a Facebook game where players help to analyse genetic data on the ash dieback disease that affects ash trees; Phylo, a puzzle game where players help to align genomic sequences – a task that is difficult for computers.[4, 17]
Shaping research priorities	
Agenda setting by citizens	Communities of stakeholders (e.g. members of the public or patients) may identify areas of importance that they feel should be addressed in research.
	Examples: Flint Water Study (see 'Data gathering' above); PatientsLikeMe's Amyotrophic Lateral Sclerosis (ALS) study.[18]
Idea generation	
Collaborative community	Idea generation and management platforms are a type of crowdsourcing platform that offer a digital, social space to generate, discuss, refine and evaluate ideas,[19] e.g. Crowdicity,[20] IdeaScale.[21] Organisations or individuals can use them to create online spaces where communities of stakeholders can gather to share and rate ideas in real time. Ideas can be collaborated on, voted on, and researched by participants. Top-ranked ideas can then be adopted by organisations, or taken forward in other ways.
	Examples: LSE Constitution Project (case study on page 12); Wikipedia.
Consensus building	
	A variety of methods exist for building consensus. The most well-known is Delphi, a technique for collecting and synthesising expert opinions developed at RAND Corporation in the 1950s and 1960s.[22] Studies based on Delphi methods are now commonly carried out using online tools, e.g. ExpertLens.[23] Such approaches have been extensively used in healthcare research; for example, the Robert Wood Johnson Foundation used ExpertLens with 119 researchers and practitioners to identify definitional features of continuous quality improvement in health care.[24]

The Zooniverse team and researchers using the platform have conducted meta-research about online citizen science projects, covering topics such as methods for measuring success[25], assessing data quality[26] and increasing engagement.[27] Two other citizen science platforms, CitSciBio and SciStarter, do not host projects; they mainly function as project databases to help prospective volunteers find projects to work on and as networking platforms.

Systematic review platforms
Dedicated platforms that combine crowdsourcing and systematic reviewing include Cochrane Crowd (see case study on page 10) and Mark2Cure.[28] These platforms are the subject of a separate THIS Institute learning report.[29] While this field is in its infancy, crowdsourcing for systematic reviews has the potential to help investigators to save time and money, as well as reducing the workload for expert reviewers.[30–34]

Labour market platforms
Though not designed with scientific applications in mind, researchers often use distributed labour markets (e.g. Amazon MTurk[35], Upwork[36], ShortTask[37] and CrowdFlower[38]) to annotate text, complete surveys, or other tasks.[39, 40] These tools allow managers of any type of project to recruit a large group to complete a specific task without needing to maintain them as part of a permanent workforce.[41]

Other platforms and tools
Other online tools include: research and development prize platforms (e.g. InnoCentive[42] and Topcoder[43]), which promote the development of innovative solutions; platforms for structured discussion and crowdsourcing ideas (e.g. Crowdicity[20]); patient data platforms (e.g. PatientsLikeMe[44] and Open Research Exchange[45]); and a role-playing game that incorporates a citizen science platform (EVE Online[46]).

Good practice in citizen science
Several organisations in Europe, North America and elsewhere have issued guidance, principles and toolkits for developing and using citizen science and crowdsourcing. These organisations include the European Citizen Science Association (ECSA),[47] the U.S. Office of Science and Technology Policy (OSTP),[48,49] the German Citizen Science Association,[50, 51] the League of European Research Universities (LERU),[7] and the EU-funded Socientize project.[52,*] The ECSA's 'Ten Principles of Citizen Science' provides a useful high-level summary of good practice in citizen science.[47] These ten principles are generally consistent with suggestions and recommendations from the wider literature (**Box 1**).

Box 1 Adapted from The European Citizen Science Association's Ten Principles of Citizen Science[47]

1. Citizen science projects actively involve citizens in scientific endeavour that generates new knowledge or understanding.

2. Citizen science projects have a genuine science outcome.

3. Both the professional scientists and the citizen scientists benefit from taking part.

4. Citizen scientists may, if they wish, participate in multiple stages of the scientific process.

5. Citizen scientists receive feedback from the project.

6. Citizen science is considered a research approach like any other, with limitations and biases that should be considered and controlled for.

7. Citizen science project data and meta-data are made publicly available and where possible, results are published in an open access format.

8. Citizen scientists are acknowledged in project results and publications.

9. Citizen science programmes are evaluated for their scientific output, data quality, participant experience and wider societal or policy impact.

10. The leaders of citizen science projects take into consideration legal and ethical issues surrounding copyright, intellectual property, data sharing agreements, confidentiality, attribution, and the environmental impact of any activities.

Data quality
A key challenge for crowdsourcing projects is maintaining data quality and scientific rigour.[53] There are some concerns that persist in the scientific community about the validity of data and analyses from projects involving non-professional scientists.[54–56] Resnik et al. (2015) emphasise the importance of data quality, suggesting that the risk is greater in citizen science projects as participants lack research and data management training, might be liable to make systematic errors, or may have motives to falsify data (e.g. a conflict of interest).[55] Researchers should therefore take a number of steps to ensure the highest standard of data quality.[53,55]

---

* A more complete list has been compiled by the ECSA at https://ecsa.citizen-science.net/blog/collection-citizen-science-guidelines-and-publications

- Provide training and supervision to participants to reduce heterogeneity in how tasks are carried out.

- Provide training to participants on responsible research conduct.

- Monitor and review how data are gathered.

- Cross-check the consistency of observations across participants, and with existing literature or observations made by professional researchers.

- Simplify or adapt tasks to ensure they are achievable by participants.

- Anticipate sources of errors and bias.

- Use statistical checks to identify and address errors, poor quality contributions, or other issues related to the quality of data or analyses provided by participants

Crowd motivation and participation

Motivated and engaged participants are critical to the success of any crowdsourcing project. People join crowdsourcing and citizen science projects for a variety of reasons, including personal interest, curiosity and a desire to learn and discover; commitment to the issue being addressed; altruism and collectivism; and a desire to teach others.[56, 57] Rewards in the form of material or professional gain are not usually provided to participants. Rather, benefits to participants tend to come in the form of personal gratification, a gain in skills or knowledge, or fulfilment of interest. Researchers need to understand why participants choose to take part in their research project and consider their expectations.[58]

Motivations for participation can affect both data quality and quantity.[59] One study found that four factors affected how much a participant engaged in a project: belief in the aims of the project, reputation, expectations of others, and personal enjoyment or interest.[59] However, just the first two factors – belief in the aims and reputation – enhanced the quality of each person's participation. The authors of that study concluded that it is important to: (i) ensure that participants are committed to a given project and its goals; (ii) consider which factors are likely to be most important for a given project, and how that could affect participant motivation; and (iii) incorporate ways for participants to become more involved and challenged over time.[59]

As well as having different motivations for participating in crowdsourcing projects, individuals also engage to varying degrees.[56, 60] Generally, a small number of participants contribute a high proportion of the overall effort.[5] Participation has been characterised by the '90-9-1 rule', which states that 90 per cent of users observe without actively participating, 9 per cent contribute at a low level,

and 1 per cent contribute the majority of the content. This rule has been observed consistently in digital health social networks.[61]

Researchers have found that more engaged participants tend to be motivated by social factors or competition while the 'dabblers', who represent the majority of participants, have more solitary experiences.[60] Dabblers are often driven by curiosity, working on multiple projects at once and fitting their contributions in around other life activities.[60] To cater to these participants, researchers can offer opportunities for independent work and small tasks that can be completed quickly, and send out regular project updates.

Other factors that motivate participants include interactions with other participants (e.g. through online discussions)[57,58] and competition. Turning tasks into games (so-called 'gamification') can make tasks more fun and competitive,[59] but individuals with a strong interest in a topic may be less motivated to engage in a game.[57] Competition may help motivate some participants but should be introduced with care. Participants must feel that the competition is fair, the competition element should not reduce data quality or accuracy, and less competitive individuals should not be discouraged from participating.[57]

Feedback is another powerful tool for motivating participants and it is essential for training participants to complete more complex tasks. Positive feedback, where participants are informed that their contributions were valuable and correct, can be very encouraging, while a lack of feedback has the opposite effect. Feedback can be provided immediately to a participant after they make a contribution or in the form of participant rankings. Participants can also be acknowledged through credits in publications.[57] It takes time to provide detailed feedback; while it can be effective, it can also be costly.[57]

Researchers using crowdsourcing have made some surprising discoveries about factors that encourage participation. In the Snapshot Serengeti project, where participants were asked to view photographs and identify the species, number and behaviour of any animals in each picture, researchers found that participants stayed engaged for longer when they were shown more pictures that had no animals in them.[62] In Cancer Research UK's Trailblazer project, the project team discovered that participants preferred not to be shown information about basic cell biology as part of their training for a task involving analysis of cell images.

Evaluation

Researchers involved in citizen science generally agree that it is important to evaluate projects involving the public.[48, 50, 63] Evaluations can show how well the project has worked, in terms of delivering benefits for science and for the participants, and provides evidence to ensure that the value of these activities is recognised appropriately.[7, 50]

Researchers working with the Zooniverse platform have published a useful, high-level framework for evaluating the success of citizen science projects (Table 2). The framework is tailored to the types of projects that Zooniverse hosts, which mainly involve volunteers contributing to data classification tasks to make the analysis and interpretation of large datasets more efficient.

The framework places equal emphasis on contributions to science and public engagement, reflecting the dual aims of the Zooniverse platform and of the team's projects. In applying their framework to a set of 17 Zooniverse projects, the Zooniverse team showed that projects that scored well on one element tended to score well on the other. They thus recommend that project organisers ensure both elements are well supported, for instance by putting together project teams with the necessary mix of skills.[25]

While useful, not all of the framework indicators below will be suitable for all types of projects. Definitions of success will vary in line with project aims. Beyond benefits to science and project participants (e.g. through gaining knowledge or skills), researchers should also consider capturing wider outcomes and impact, such as improved public awareness of research, impacts on policy and the empowerment of citizens to influence decisions about issues of concern.[48, 64]

While the measures suggested in the framework from Cox et al. (2015) could be assessed at the end of a project, it may be more useful to monitor them throughout the life of the project, where feasible. For instance, continuous evaluations could be used to assess the effects of any changes introduced, such as how training is delivered, or to check whether certain communication activities boost participation.

Table 2 Citizen science success framework, adapted from Cox et al. (2015)		
Outcome	**Performance indicator**	**Measurement**
Contribution to science	Data value	• Number of publications produced. • Academic impact (citations). • Whether the analysis is scientifically and statistically valid.
	Project design and resource allocation	• Resource savings realised (considering the time that would be required for a professional researcher to carry out the project tasks). • Equity of the distribution of effort across participants. • The proportion of volunteers who were successfully trained (as indicated by the volunteers going on to complete at least one task after doing the tutorial).
Public engagement	Dissemination and feedback	• Collaboration – the number of papers that include at least one citizen scientist as an author. • Communication – the amount of communication activity that took place (e.g. blog posts, tweets). • Interaction – the number of interaction events (e.g. blog post comments and replies) that took place between the researchers and citizen scientists.
	Participation and opportunities for learning	• Project appeal – the number of contributors. • Sustained engagement – the median time period over which each participant engaged with the project. • Public contribution – the median number of classifications completed by each participant.

Ethical and legal issues

The use of distributed workforces or crowds raises some ethical and legal issues. Conventional principles and policies regarding scientific integrity and research ethics generally apply to citizen science and crowdsourcing projects. Specific issues for crowdsourcing projects relate to data quality and the responsible conduct of research (discussed above); data sharing and intellectual property; and exploitation of and risks to participants.

Data sharing and intellectual property

It is generally agreed that citizen science projects should be conducted in an open way, using a fully transparent methodology, and that resulting data, publications and software should be made openly available.[7, 47, 48] Researchers should inform participants about rules and procedures for sharing data and usual care should be taken to protect confidential information.[55]

When many anonymous members of a crowd contribute to a project, it can be difficult to maintain the integrity of the project owner's intellectual property (IP) rights. Scassa & Chung (2015, 10–11) provide a useful typology of IP in citizen science projects, with an assessment of potential issues that might arise. There are two issues particularly pertinent to crowdsourcing. The first is where copyrighted material, such as photographs, videos and written text, may be contributed by study participants. The second is where participants conduct problem-solving or data management; in this situation, some participants may qualify as having invented ideas or products, or having co-authored the research.[65]

IP and data ownership issues can generally be mitigated through clear communication and discussion before and potentially during the project to ensure an agreement has been reached that is acceptable to the parties involved.[55] It is also important for researchers to carefully draft terms of participation for projects, and consider whether participants may make contributions for which they have IP rights, or may be able to claim IP rights on the output.[65]

Risks to participants

The issue of participants' exploitation relates to the sharing of benefits that can result from the work, particularly when participants are volunteers. Like any scientific study that involves human participants, researchers need to provide opportunities for participants to ask questions if they do not understand a specific instruction and allow them to leave the project if they do not consent to any of the project's tasks or terms.[66] In addition to ensuring reasonable agreements are reached and respected regarding data and IP ownership, it is important that other benefits are provided as appropriate, such as formal recognition for contributions, education related to the study conducted, financial compensation, acknowledgement in an article, or authorship credit.[55]

While most citizen science projects do not involve financial compensation, crowdsourcing projects that involve paid tasks generate further ethical considerations. For example, it is unclear how much participants should be compensated for their time. Participants often come from different areas of the world where there are different standards for appropriate compensation.[66]

Catherine Lichten, Rebecca Ioppolo, Camilla D'Angelo, Rebecca K Simmons, Molly Morgan Jones

Case study 1. Cancer Research UK's Trailblazer

Trailblazer is one of five online citizen science projects developed by Cancer Research UK, a major cancer research and awareness charity in the UK. In these projects, participants contribute to aspects of cancer research by analysing images. The analytical tasks are sometimes presented in the form of a puzzle. The projects' popularity[67] has shown that many members of the public have a strong interest in contributing to cancer research through this type of activity.

In the first project, Cell Slider, released in 2012, participants analysed images of breast tissue samples by categorising cells in the images as being cancerous or non-cancerous, and estimating levels of oestrogen receptor present in the cells. Nearly 100,000 citizen scientists participated during October 2012 to June 2014, assessing 180,000 images and providing accurate oestrogen receptor data.[68] However, they struggled to differentiate and accurately count the cancerous cells. In Trailblazer, the team tackled this issue by iteratively refining the user interface and training to improve participants' performance and test hypotheses about user interaction and tutorial design.

The Trailblazer project showed that members of the public can, with a limited amount of training, act as a reliable and accurate workforce to support aspects of cancer research. The project helped its developers better understand how to design a platform for participants to analyse images of tissue samples.

The crowd's task
Users completed tutorials to learn how to analyse images of tissue samples by marking off areas of cancerous cells. They were then presented with images and asked to determine if cancer cells were present, and if so how many of them were 'stained' with a biomarker.

Platform and tools
Trailblazer was developed by the citizen science team at Cancer Research UK, in collaboration with academics conducting cancer research. Citizen science platform developer Scifabric provided the core citizen science platform. A UK-based digital agency was involved in an iterative process of prototyping, testing and development to refine the Trailblazer tutorial and user interface.

The crowd
Cancer Research UK's citizen science projects have collectively engaged more than 500,000 people worldwide. They have made more than 11 million contributions. About 1,000 volunteers were involved in Trailblazer (as of Spring 2016).[69]

Timeline
The team developed the initial Trailblazer release in just eight weeks. They then carried out iterative development for a six-month period, during which there were 15 releases.[69]

Lessons learned
Trailblazer showed that citizens without advanced science training could learn to accurately identify cancer cells. After 25 minutes of training that included feedback on their performance, volunteers achieved a 95 per cent level of accuracy compared to cancer research experts. They were also able to detect clinically relevant features of the cells with 90 per cent accuracy.[69]

Trailblazer's developers found a 'lean start-up' approach – an approach based on developing and testing hypotheses iteratively – to be particularly valuable. Instead of making assumptions about what training approaches would be most effective and then proceeding to build the platform, the team tested and evolved their approach with guidance from participants.

Figure 1. A screenshot from Cancer Research UK's Trailblazer project
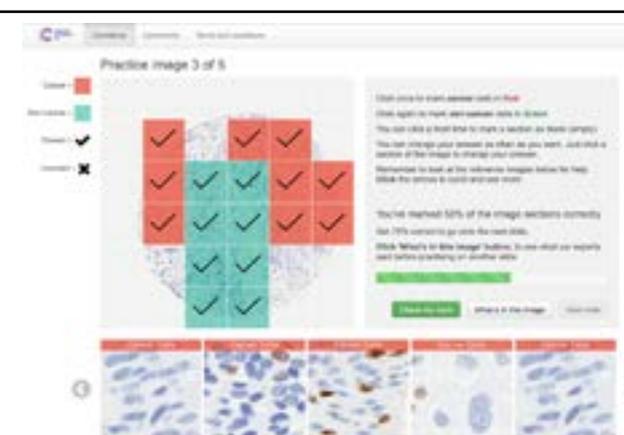
cancer.pybossa.com

Case study 2. Cochrane Crowd

Conducting and updating systematic reviews requires searching through thousands of search results. Given the increasingly rapid pace at which new research is published, Cochrane are using an online crowd to help review research articles. Cochrane is a network of researchers, healthcare workers and patients who collaborate to produce independent, comprehensive reviews of health information.[70] The organisation developed Cochrane Crowd, an online platform where members of the public can contribute to its systematic reviews. Crowdsourcing, combined with algorithms to safeguard accuracy and improve efficiency, has brought a significant, much-needed boost to their systematic review activity.

The crowd's tasks

Contributors do not carry out all steps of a systematic review, which include article screening, information extraction and quality assessment. Rather, they focus on binary classification – determining whether or not each article describes, for instance, a randomised controlled trial (RCT). The Cochrane team found that involving a crowd in the full systematic review process was very challenging. They broke the process down into manageable tasks where accuracy could be assessed to ensure that results were reliable.

Platform and tools

The Cochrane team developed their own platform, which integrates with the other parts of the organisation's IT infrastructure. An important part of the system is an algorithm that assesses whether agreement has been reached about a particular record. A training feature helps participants learn how to screen and a text highlighting function helps them to quickly find relevant information.[71]

The crowd's activity also generated a large dataset, which has been used to train machine-learning algorithms to automatically identify RCTs. These algorithms work better than search filters for identifying RCTs[72] and significantly reduce the number of articles that require human screening.[34]

The crowd

While the platform is open to anyone, most contributors have medical experience or expertise. Volunteers are required to train on 20 practice records before they begin.[34] Screeners are ranked at different levels (novice, expert, resolver) and unlock rewards and tasks as they complete more classifications. As of June 2017, the Cochrane Crowd platform had close to 6,000 contributors who had screened more than a million articles.

Lessons learned

Evaluations have shown that the crowd performs very accurately, meaning 'non-traditional' reviewers can provide valuable contributions to the systematic review process. The crowd's sensitivity (true positive rate) and specificity (true negative rate) were both more than 99 per cent compared to the performance of an information specialist and a systematic reviewer.[71]

The Cochrane Crowd team discovered that crowdsourcing helps to improve efficiency but requires significant investment to ensure it runs effectively. They have encountered challenges with participant drop-out. Time is required to monitor the crowd and its activity, to engage with participants (e.g. through social media), to encourage their continued involvement, and to run evaluations.

In addition to article classification, the Cochrane Crowd team are developing additional tasks for the crowd related to the identification of basic research study information such as Patient, Population or Problem, Intervention, Comparison and Outcome (PICO).

Figure 2. The Cochrane Crowd homepage

http://crowd.cochrane.org/index.html

Case study 3. Testing the accuracy of a clinical decision support system

A research team developed a prototype clinical decision support system (CDSS) to help clinicians make point of care decisions about whether patients should be screened for cervical cancer. They used a crowd of clinical staff to create a reference dataset of realistic clinical examples, which were then used to test, validate and improve an automated tool. If the tool was reliable, it could help clinicians adhere to complex clinical guidelines on cervical cancer management, resulting in fewer suboptimal or inappropriate decisions about how to manage patients with abnormal screening results.[73, 74]

The researchers compared the prototype's decisions to those of actual clinicians, finding flaws in the prototype. Once those flaws had been corrected, the CDSS generated accurate recommendations in all but one of the 169 test cases. By comparison, clinicians provided suboptimal recommendations, relative to guidelines, in about one third of the test cases.

The crowd's task

Each user was shown a one-minute training video, then asked to review patient records and make recommendations about patient surveillance and screening. Each participant completed seven cases.

Platform and tools

The CDSS was developed using a business rule management system (Drools).

The crowd

The crowd included 25 people, who were consultants, residents and nurse practitioners working in family medicine, internal medicine, and obstetrics and gynaecology. All crowd members worked at the researchers' institution and were invited by email to participate.

Timeline

The web interface was open during a three-week period in Spring 2012.

Lessons learned

Crowdsourcing was used to successfully create a reference dataset of realistic clinical examples, to test and improve a CDSS, and to identify the types of decisions that clinicians find most challenging.[74]

Case study 4. Crowdsourcing the UK constitution

The web is increasingly used by governments to engage citizens in democratic and policy-making practices.[75] In particular, crowdsourcing can be used to gather a large number of inputs from citizens, which can feed directly into decision-making processes at a local, national and international level.[76] Therefore, crowdsourcing tools could offer revitalisation to democratic processes experiencing poor levels of participation.

The Constitution UK project drew together members of the general public to debate the content of a hypothetical constitution for the UK. Run by researchers at the London School of Economics (LSE), the project sought to encourage an open process whereby citizens, regardless of their expertise, could have their say on the UK's system of governance. It also aimed to trial an informal learning experience.

The project used both live events and an online platform[77] to crowdsource ideas and promote discussion. It demonstrated that crowdsourcing is a useful means for citizens to engage in public policy debates and showcased a novel approach to learning. Experts at LSE's Institute of Public Affairs put together a draft constitution using all the crowdsourced data that received a positive vote on the platform. They presented it to MPs in 2015 and posted it on the Crowdicity platform.[78]

The crowd's task

Using an online platform, participants discussed and voted on a set of 11 topics. On the platform, they could also view videos that provided a basic introduction to constitutional law and watch videos of the live events.

Platform and tools

The project used Crowdicity, a commercial web- and mobile-based platform that enables companies and organisations to build collaborative communities and where participants can gather, share, and rate ideas in real time.[79] A group of 20 facilitators moderated the online discussions and debates.

Social media feeds embedded in the platform kept users updated with real-time news and information related to each topic. The project team also used social media channels (Facebook, Storify and Twitter) to encourage discussion and communicate with users on a daily basis.

The crowd

A community of around 1,500 people from the UK, other countries in Europe, the US, Canada and Australia shared their ideas, comments and votes. The project generated tens of thousands of online interactions and participation increased as the project progressed. A large increase in participation occurred during the last two weeks; nearly half of participants remained engaged at the end. The leading contributors to the online community were invited to a constitutional convention at LSE.

Lessons learned

The project managed to build a large online community while avoiding the high attrition rate seen in other online learning environments, such as the massive open online courses (MOOCs) that have become more prevalent in recent years.[80] Members of the project team suggested that the highly flexible and informal aspects of the learning environment, including the lack of lectures, formal teachers and assignments, were important as participants chose when and how to interact with the project. In surveys, the majority of participants stated that they had gained new knowledge (80 per cent) or skills (70 per cent) through the project. Half stated that working with others directly contributed to their learning experience.[81] Following on from the success of Constitution UK, the team is working on a roadmap of projects to draw on the potential of crowdsourcing to solve problems, address key issues and engage in digital citizenship.

Figure 3. The online Constitution UK crowdsourcing platform, run by Crowdicity

https://constitutionuk.crowdicity.com

Conclusion

Researchers have used crowdsourcing successfully in a range of research projects. Crowdsourcing offers efficiency gains and helps promote dialogue between researchers and citizens. Researchers interested in incorporating crowdsourcing into their work can make use of a range of existing online platforms and tools. There are also a number of practical approaches to maintain data quality and scientific rigour, and to engage, retain and reward participants. Researchers should carefully consider IP and data ownership issues, and if resources permit, evaluating citizen science projects is strongly recommended.

Acknowledgements

References

1. Howe, J. 2010. 'Crowdsourcing: a definition'. Crowdsourcing. As of 16 February 2018: http://www.crowdsourcing.com/cs/

2. Swan, M. 2012. 'Crowdsourced Health Research Studies: An Important Emerging Complement to Clinical Trials in the Public Health Research Ecosystem.' *Journal of Medical Internet Research* 14 (2): e46. doi:10.2196/jmir.1988

3. Del Savio, L., B. Prainsack, & A. Buyx. 2016. 'Crowdsourcing the Human Gut. Is Crowdsourcing also "Citizen science"?' *Journal of Science Communication* 15 (03): A03.

4. Garbarino J., & C.E. Mason. 2016. 'The Power of Engaging Citizen Scientists for Scientific Progress.' *Journal of Microbiology and Biology Education* 17 (1): 7–12.

5. Khatib F., F. DiMaio, Foldit Contenders Group, Foldit Void Crushers Group, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska, I. Pichova, J. Thompson, Z. Popović, M. Jaskolski, & D. Baker. 2011. 'Crystal Structure of a Monomeric Retroviral Protease Solved by Protein Folding Game Players.' *Nature Structural and Molecular Biology* 18: 1175–1177.

6. Weiner, M. 2014. 'The Potential of Crowdsourcing to Improve Patient-Centered Care.' Patient 7 (4): 123–127.

7. Grey, F., D. Wyler, J. Fröhlich, & K. Maes. 2016. 'Citizen Science at Universities: Trends, Guidelines and Recommendations.' League of European Research Universities Publications. As of 16 February 2018: https://www.leru.org/publications/citizen-science-at-universities-trends-guidelines-and-recommendations

8. Afshinnekoo, E., S. Ahsanuddin, & C.E. Mason. 2016. 'Globalizing and Crowdsourcing Biomedical Research.' *British Medical Bulletin* 120 (1): 27–33.

9. Anderson N., C. Bragg, A. Hartzler, & K. Edwards. 2012. 'Participant-Centric Initiatives: Tools to Facilitate Engagement in Research.' *Applied and Translational Genomics* 1: 25–29.

10. Vayena, E., R. Brownsword, S.J. Edwards, B. Greshake, J.P. Kahn, N. Ladher, J. Montgomery, D. O'Connor, O. O'Neill, M.P. Richards, A. Rid, M. Sheehan, P. Wicks, & J. Tasioulas. 2016. 'Research Led by Participants: A New Social Contract for a New Kind of Research.' *Journal of Medical Ethics* 42: 216–219.

11. McKenna, M.T., S. Wang, T.B. Nguyen, J.E. Burns, N. Petrick, & R.M. Summers. 2012. 'Strategies for Improved Interpretation of Computer-Aided Detections for CT Colonography Utilizing Distributed Human Intelligence.' *Medical Image Analysis* 16 (6): 1280–1292.

12. Smolinski, M.S., A.W. Crawley, K. Baltrusaitis, R. Chunara, J.M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, & J.S. Brownstein. 2015. 'Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons.' *American Journal of Public Health* 105 (10): 2124–2130.

13. Merchant, R.M., D.A. Asch, J.C. Hershey, H.M. Griffis, S. Hill, O. Saynisch, A.C. Leung, J.M. Asch, K. Lozada, L.D. Nadkarni, A. Kilaru, C.C. Branas, E.M. Stone, L. Starr, F. Shofer, G. Nichol, & L.B. Becker. 2013. 'A Crowdsourcing Innovation Challenge to Locate and Map Automated External Defibrillators.' *Circulation: Cardiovascular Quality and Outcomes* 6 (2): 229–236.

14. Verhulst, S. 2016. 'Citizen Science and the Flint Water Crisis.' GovLab Digest, The Governance Lab @ NYU. As of 16 February 2018: http://thegovlab.org/citizen-science-and-the-flint-water-crisis/

15. Feldkamp L. 2016. '7 Citizen Science Projects for Bird Lovers.' Cool Green Science. As of 16 February 2018: https://blog.nature.org/science/2016/03/30/7-citizen-science-projects-bird-watching-lovers-birding-birds/

16. Lakhani K.R., K.J. Boudreau, P-R. Loh, L. Backstrom, C. Baldwin, E. Lonstein, M. Lydon, A. MacCormack, R.A. Arnaout, & E.C. Guinan. 2013. 'Prize-Based Contests Can Provide Solutions to Computational Biology Problems.' *Nature Biotechnology* 31: 108–111.

17. Williams, S.C.P. 2014. 'Biology: A Game for a Crowd.' *Biomedical Computation Review* Spring: 4–6.

18. Wicks P., T.E. Vaughan, M.P. Massagli & J. Heywood. 2011. 'Accelerated Clinical Discovery Using Self-reported Patient Data Collected Online and a Patient-matching Algorithm.' *Nature Biotechnology* 29: 411–414.

19. Saldivar J., M. Báez, C. Rodriguez, G. Convertino, & G. Kowalik. 2016. 'Idea Management Communities in the Wild: An Exploratory Study of 166 Online Communities.' In *2016 International Conference on Collaboration Technologies and Systems* (CTS), 81–89.

20. Crowdicity Idea Management Software. Crowdicity. As of 16 February 2018: https://crowdicity.com/

21. IdeaScale. 2014. Innovation Management Software. IdeaScale. As of 16 February 2018: https://ideascale.com/

22. Dalkey N.C., & O. Helmer. 1963. An Experimental Application of the Delphi Method to the Use of Experts. *Management Science* 9: 458–467.

23. RAND Corporation ExpertLens. 'An Online Approach to Expert Elicitation and Stakeholder Engagement.' Santa Monica, Calif.: RAND Corporation. As of 16 February 2018: https://www.rand.org/pubs/tools/expertlens

24. Khodyakov D., S. Hempel, L. Rubenstein, P. Shekelle, R. Foy, S. Salem-Schatz, S. O'Neill, M. Danz, & S. Dalal. 2011. 'Conducting Online Expert Panels: A Feasibility and Experimental Replicability Study. *BMC Medical Research Methodology* 11: 174.

25. Cox, J., E.Y. Oh, B. Simmons, C. Lintott, K. Masters, A. Greenhill, G. Graham, & K. Holmes. 2015. 'Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects.' *Computing in Science Engineering* 17: 28–41.

26. Kosmala M., A. Wiggins, A. Swanson, & B. Simmons. 2016. 'Assessing Data Quality in Citizen Science.' *Frontiers in Ecology and the Environment* 14 (10): 551–560.

27. Jackson, C.B., C. Østerlund, G. Mugar, K.D. Hassman, & K. Crowston. 2015. 'Motivations for Sustained Participation in Crowdsourcing: Case Studies of Citizen Science on the Role of Talk.' In 2015 *48th Hawaii International Conference on System Sciences*, 1624–1634.

Catherine Lichten, Rebecca Ioppolo, Camilla D'Angelo, Rebecca K Simmons, Molly Morgan Jones

28. Mark2Cure (homepage). 2018. As of 16 February 2018: https://mark2cure.org/

29. Strang, L & R.K. Simmons. 2018. 'Crowdsourcing for Systematic Reviews' THIS Institute.

30. Mortensen, M.L., G.P. Adam, T.A. Trikalinos, T. Kraska, & B.C. Wallace. 2017. 'An Exploration of Crowdsourcing Citation Screening for Systematic Reviews.' *Research Synthesis Methods* 8: 366–386.

31. Nama N., K. Iliriani, M.Y. Xia, B.P. Chen, L.L. Zhou, S. Pojsupap, C. Kappel, K. O'Hearn, M. Sampson, K. Menon, & J.D. McNally. 2017. 'A Pilot Validation Study of Crowdsourcing Systematic Reviews: Update of a Searchable Database of Pediatric Clinical Trials of High-Dose Vitamin D. *Translational Pediatrics* 6 (1): 18–26.

32. Noel-Storr, A., C. Struthers, S. Cullum, R. McShane, S. Creavin, D. Davis, & K. Huckvale, K. 2013. 'Many Hands Make Light Work – Or Do They? Results of Two Pilot Studies Looking at the Effects of Crowdsourcing.' In *Better Knowledge for Better Health | Un meilleur savoir pour une meilleure santé.* Abstracts of the 21st Cochrane Colloquium. Québec City, Canada: John Wiley & Sons.

33. Sun, Y., P. Cheng, S. Wang, H. Lyu, M. Lease, I. Marshall, & B.C. Wallace. 2016. 'Crowdsourcing Information Extraction for Biomedical Systematic Reviews.' arXiv:1609.01017 [cs].

34. Wallace B.C., A. Noel-Storr, I.J. Marshall, A.M. Cohen, N.R. Smalheiser, & J. Thomas. 2017. 'Identifying Reports of Randomized Controlled Trials (RCTs) via a Hybrid Machine Learning and Crowdsourcing Approach.' *Journal of the American Medical Informatics Association* 24 (6): 1165–1168. doi:10.1093/jamia/ocx053

35. Amazon (2017). 'Amazon Mechanical Turk – Welcome.' As of 16 February 2018: https://www.mturk.com/mturk/welcome

36. Upwork (homepage). 2018. As of 16 February 2018: https://www.upwork.com/

37. ShortTask (homepage). 2018. As of 16 February 2018: http://www.shorttask.com/

38. CrowdFlower (homepage). 2018. As of 16 February 2018: https://www.crowdflower.com/

39. Khare R., B.M. Good, R. Leaman, A.I. Su, & Z. Lu. 2016. 'Crowdsourcing in Biomedicine: Challenges and Opportunities.' *Briefings in Bioinformatics* 17: 23–32.

40. Gleibs, I.H. 2017. 'Are all "Research Fields" Equal? Rethinking Practice for the use of Data from Crowdsourcing Market Places.' *Behavior Research Methods* 49 (4): 1333–1342.

41. Baba J. & K. Rodgers. 2016. 'The Value of Crowdsourcing: A Public Sector Guide to Harnessing the Crowd.' Deloitte.

42. InnoCentive (homepage). 2018. As of 16 February 2018: https://www.innocentive.com/

43. Topcoder (homepage). 2018. As of 16 February 2018: https://www.topcoder.com/

44. PatientsLikeMe (homepage). 2018. As of 16 February 2018: https://www.patientslikeme.com/

45. Open Research Exchange (homepage). 2018. As of 16 February 2018: https://www.openresearchexchange.com/

46. EVE Online. EVE Online – One Community. Countless Journeys. As of 16 February 2018: https://www.eveonline.com/discovery-proteins/

47. European Citizen Science Association. 2015. 'Ten Principles of Citizen Science.' London.

48. Holdren, J.P. 2015. Memorandum to the Heads of Executive Departments and Agencies on the Subject of 'Addressing Societal and Scientific Challenges through Citizen Science and Crowdsourcing'. Washington DC: Office of Science and Technology.

49. US Government. 2017.'Federal Crowdsourcing and Citizen Science Toolkit.' In CitizenScience.gov. As of 16 February 2018: https://crowdsourcing-toolkit.sites.usa.gov/

50. Bonn, A., A. Richter, K. Vohland, L. Pettibone, M. Brandt, R. Feldmann, C. Goebel, C. Grefe, S. Hecker, L. Hennen, H. Hofer, S. Kiefer, S. Klotz, T. Kluttig, J. Krause, K. Küsel, C. Liedtke, A. Mahla, V. Neumeier, M. Premke-Kraus, M.C. Rillig, O. Röller, L. Schäffler, B. Schmalzbauer, U. Schneidewind, A. Schumann, J. Settele, K. Tochtermann, K. Tockner, J. Vogel, W. Volkmann, H. von Unger, D. Walter, M. Weisskopf, C. Wirth, T. Witt, D. Wolst, & D. Ziegler. 2016. G*reen Paper Citizen Science Strategy 2020 for Germany.* Helmholtz Centre for Environmental Research (UFZ); German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig; Museum für Naturkunde Berlin; Leibniz Institute for Evolution and Biodiversity Science (MfN); Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Berlin.

51. Pettibone, L. K. Vohland, A. Bonn, A. Richter, W. Bauhus, B. Behrisch, R. Borcherding, M. Brandt, F. Bry, D. Dörler, I. Elbertse, F. Glöckler, C. Göbel, S. Hecker, F. Heigl, M. Herdick, S. Kiefer, T. Kluttig, E. Kühn, K. Kühn, K. Oswald, O. Röller, C. Schefels, A. Schierenberg, W. Scholz, A. Schumann, A. Sieber, R. Smolarski, K. Tochtermann, W. Wende, & D. Ziegler. 2016. C*itizen Science for all – A Guide for Citizen Science Practitioners.* German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Helmholtz Centre for Environmental Research (UFZ), Leipzig; Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB); Museum für Naturkunde (MfN); Leibniz Institute for Evolution and Biodiversity Science, Berlin.

52. Socientize EC. 2014. 'White Paper on Citizen Science for Europe'. As of 18 February 2018: http://www.socientize.eu/sites/default/files/white-paper_0.pdf

53. Garcia-Soto, C., G.I. van der Meeren, J.A. Busch, J. Delany, C. Domegan, K. Dubsky, G. Fauville, G. Gorsky, K. von Juterzenka, F. Malfatti, G. Mannaerts, P. McHugh, P. Monestiez, J. Seys, J.M. Węsławski, & O. Zielinski. 2017. 'Advancing Citizen Science for Coastal and Ocean Research.' In *Position Paper 23 of the European Marine Board*, edited by V. French, P. Kellett, J. Delany, & N. McDonough. Ostend, Belgium.

54. Aguilera, H., M.M. Ladeira, & M. Slanova. 'Crowdsourcing in Medicine and Health: Challenges, Risks and Future.'

55. Resnik, D.B., K.C. Elliott, & A.K. Miller. 2015. 'A Framework for Addressing Ethical Issues in Citizen Science.' *Environmental Science and Policy* 54: 475–481.

56. Everett, G. & H. Geoghegan. 2016. 'Initiating and Continuing Participation in Citizen Science for Natural History.' *BMC Ecology* 16 (Suppl 1): S13.

57. Dunn, S. & M. Hedges. 2013. 'Crowd-sourcing as a Component of Humanities Research Infrastructures.' *International Journal of Humanities and Arts Computing* 7: 147–169.

58. Morrow, A. 2013. 'The Impact of Citizen Science Activities on Participant Behaviour and Attitude: Review of Existing Studies.' The Conservation Volunteers.

59. Nov O., O. Arazy, & D. Anderson. 2014. 'Scientists@Home: What Drives the Quantity and Quality of Online Citizen Science Participation?' PLOS ONE 9: e90375.

60. Eveleigh A., C. Jennett, A. Blandford, P. Brohan, & A.L. Cox. 2014. 'Designing for Dabblers and Deterring Drop-Outs in Citizen Science.' In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 2985–2994. New York: ACM.

61. van Mierlo, T. 2014. 'The 1% Rule in Four Digital Health Social Networks: An Observational Study.' *Journal of Medical Internet Research* 16 (2): e33.

62. Bowyer, A., V. Maidel, C. Lintott, A. Swanson, & G. Miller. 2015. 'This Image Intentionally Left Blank: Mundane Images Increase Citizen Science Participation.' *Human Computation and Crowdsourcing: Works in Progress and Demonstrations. An Adjunct to the Proceedings of the Third AAAI Conference on Human Computation and Crowdsourcing.*

63. Tweddle, J.C., L.D. Robinson, M.J.O. Pocock, & H.E. 2012. *Guide to Citizen Science: Developing, Implementing and Evaluating Citizen Science to Study Biodiversity and the Environment in the UK.* Wallingford: NERC/Centre for Ecology & Hydrology.

64. Shirk J.L., H.L. Ballard, C.C. Wilderman, T. Phillips, A. Wiggins, R. Jordan, E. McCallie, M. Minarchek, B.V. Lewenstein, M.E. Krasny, & R. Bonney. 2012. 'Public Participation in Scientific Research: A Framework for Deliberate Design.' *Ecology and Society* 17 (2): 29. doi:10.5751/ES-04705-170229

65. Scassa, T. & H. Chung. 2015. 'Typology of Citizen Science Projects from an Intellectual Property Perspective: Invention and Authorship between Researchers and Participants.' Washington DC: Wilson Center.

66. Chandler J. & D. Shapiro. 2016. 'Conducting Clinical Research Using Crowdsourced Convenience Samples.' *Annual Review of Clinical Psychology* 12: 53–81.

67. CRUK. 2016. 'The Projects' in Cancer Research UK Citizen Science. As of 16 February 2016: http://www.cancerresearchuk.org/support-us/citizen-science/the-projects

68. Candido dos Reis, F.J., S. Lynn, H.R. Ali, D. Eccles, A. Hanby, E. Provenzano, C. Caldas, W.J. Howat, L-A. McDuffus, B. Liu, F. Daley, P. Coulson, R.J. Vyas, L.M. Harris, J.M. Owens, A.F.M. Carton, J.P. McQuillan, A.M. Paterson, Z. Hirji, S.K. Christie, A.R. Holmes, M.K. Schmidt, M. Garcia-Closas, D.F. Easton, M.K. Bolla, Q. Wang, J. Benitez, R.L. Milne, A. Mannermaa, F. Couch, P. Devilee, R.A.E.M. Tollenaar, C. Seynaeve, A. Cox, S.S. Cross, F.M. Blows, J. Sanders, R. de Groot, J. Figueroa, M. Sherman, M. Hooning, H. Brenner, B. Holleczek, C. Stegmaier, C. Lintott, & P.D.P. Pharoah. 2015. 'Crowdsourcing the General Public for Large Scale Molecular Pathology Studies in Cancer.' EBioMedicine 2: 681–689.

69. Doherty, J. 2016. 'Blazing a Trail in Cancer Diagnosis - Scifabric.' As of 16 February 2018: https://scifabric.com/blog/2016/04/20/Cancer-Research-UK-Trailblazer.html

70. Cochrane. 2017. 'Our Vision, Mission, and Principles.' As of 16 February 2018: http:// cochrane.org/uk/about-us/our-vision-mission-and-principles

71. Thomas J., A. Noel-Storr, & J. Elliott. 2015. 'Human and Machine Effort in Project Transform: How Intersecting Technologies Will Help Us to Identify Studies Reliably, Efficiently and at Scale.' In *Cochrane Methods* edited by J. Chandler, J. McKenzie, I. Boutron, & V. Welch. *Cochrane Database of Systematic Reviews* Suppl. 1: 37–41.

72. Marshall I., A.H. Noel-Storr, J. Kuiper, J. Thomas, & B.C. Wallace. 2018. 'Machine Learning for Identifying Randomized Controlled Trials: An Evaluation and Practitioner's Guide.' *Research Synthesis Methods* 2018: 1–12.

73. Wagholikar K.B., K.L. MacLaughlin, M.R. Henry, R.A. Greenes, R.A. Hankey, H. Liu, & R. Chaudhry. 2012. 'Clinical Decision Support with Automated Text Processing for Cervical Cancer Screening.' *Journal of the American Medical Informatics Association* 19: 833–839.

74. Wagholikar K.B., K.L. MacLaughlin, T.M. Kastner, P.M. Casey, M. Henry, R.A. Greenes, H. Liu, & R. Chaudhry. 2013. 'Formative Evaluation of the Accuracy of a Clinical Decision Support System for Cervical Cancer Screening.' *Journal of the American Medical Informatics Association* 20: 749–757.

75. UN Department of Economic and Social Affairs. 2016. 'UN E-Government Survey 2016.' New York.

76. Bruno, E. 2015. 'Co-deciding with Citizens: Towards Digital Democracy at EU Level.' European Citizen Action Service Brussels.

77. Constitution UK. 2015. ConstitutionUK - Crowdsourcing the UK Constitution. As of 16 February 2018: https://constitutionuk.crowdicity.com/

78. The Constitution. 2015. The Constitution – LSE Institute of Public Affairs. As of 16 February 2018: https://lsedesignunit.com/theconstitution/html5/index.html?page=1

79. Anon. 2017. 'Crowdicity.' In *PCMag Business Software Index.* As of 16 February 2018: https://www.pcmag.com/business/directory/idea-management/452-crowdicity

80. Clow, D. 2013. 'MOOCs and the Funnel of Participation.' In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 185–189. New York: ACM.

81. Bryant P., C. Fryer, & D. Moon. 2015. 'Harnessing the Power of the "Massive": An Innovative Approach to Participation, Digital Citizenship and Open Learning On-line.' In *Learning with MOOCS II,* 2–3 2015, New York, USA.

At THIS Institute we aim to strengthen the evidence-base for improving the quality and safety of healthcare.

Co-created by two exceptional organisations – the University of Cambridge and the Health Foundation – THIS Institute is founded on the guiding principle that efforts to improve care should be based on the highest quality evidence.

Our work is defined by a highly inclusive approach that combines academic rigour with the real concerns of patients and staff. We're open, transparent, and we do not shy away from difficult subjects.

Together we're creating an evidence base to improve healthcare.

THIS Institute is made possible by the Health Foundation.

THIS.Institute

The Healthcare
Improvement
Studies Institute

THIS Institute
School of Clinical Medicine
University of Cambridge
Cambridge Biomedical Campus
Hills Road
Cambridge CB2 0AH

thisinstitute.cam.ac.uk
@THIS_Institute

UNIVERSITY OF
CAMBRIDGE

Made possible by

The
Health
Foundation