Crowdsourcing

2.1 What Is Crowdsourcing?

In 2006, Jeff Howe coined the term crowdsourcing as a way to capture how large groups of people in the online setting were coming together to solve different types of problem. By combining "crowd" and "outsourcing," the new term emphasized how organizations made open calls in an online setting to outsiders who could help them solve tasks that they had previously completed within the organization. Instead of "outsourcing" the task to one specific external expert or company, the new call invited anyone to contribute. Today, crowdsourcing tasks vary significantly, and can be anything from the design of a new product to a scientific problem, but the problem is usually formulated in advance. The most important advantage with inviting a large group of people to contribute is that the outreach and the number of contributions offer more diversity and can therefore potentially offer a better solution. The contributors are typically unknown to each other and will have many different types of backgrounds (Innocent, Gabriel, & Divard, 2017). Another aspect of crowdsourcing is the emphasis on volunteering and self-selection of tasks. Although many people receive an invitation, only the individuals who think they have the skills and the time to contribute will participate (Brabham, 2013).

Today, crowdsourcing is receiving increased attention from a wide range of stakeholders, like businesses, scientists, policymakers and funding agencies. Crowdsourcing is part of open innovation, a new paradigm that expects organizations to use external ideas to advance their innovations. In open innovation, outsiders are valued because they can contribute to new and unexpected ways of solving a problem. Before the invention of the Internet, this type of innovation would typically happen at fairs, conferences, exhibitions or through joint projects (Von Krogh, Netland, & Wörter, 2018). The basic assumption is that knowledge will always be widely distributed in the economy, or in more popular terms, "most smart

people work for someone else" (Bogers, Chesbrough, & Moedas, 2018). More specifically, crowdsourcing resembles "outside-in" open innovation; a strategy that involves direct use of ideas and knowledge from external stakeholders outside the organization. By reaching out to new potential problem solvers, the aim is to utilize a larger degree of cognitive diversity (Chesbrough, 2017). Today, numerous businesses and other organizations recruit outsiders to help them solve different type of organizational challenges (e.g. Innocentive, IdeaConnection), design challenges (e.g. Threadless), scientific problems (e.g. Foldit), IT challenges (e.g. Topcoder), financial challenges (e.g. Kickstarter) or broader societal challenges (e.g. Climate CoLab). The goal will often be to find outsiders who can think "outside the box" and utilize unconventional sources of knowledge. In addition, crowdsourcing covers a range of simpler tasks or routine activities, like classifying images in science (e.g. Galaxzy Zoo). Although the methods vary, all crowdsourcing strategies assume that they can harness unique human knowledge in a way that machine intelligence is not capable of (Franzoni & Sauermann, 2014; Innocent et al., 2017).

In an attempt to better understand the basic collective problem-solving mechanisms in crowdsourcing, this chapter will cover a broad range of examples from both open innovation and citizen science, two of the perhaps most interesting new areas in relation to the potential societal benefits of crowdsourcing. In open innovation, the two examples are from innovation intermediaries who host online innovation contests (IdeaConnection and Climate CoLab). In citizen science, Zooniverse (e.g. Galaxy Zoo) and Foldit are selected as two prominent examples that will be introduced and analyzed in detail. Note that the examples chosen are relatively successful case stories, and not failed examples. This is with the intention of identifying in this chapter the basic problem-solving mechanisms in crowdsourcing. In addition, the selection of topics and case stories reflects those areas where it was possible to find relevant in-depth information and relevant research studies.

2.2   Online Innovation Contests

2.2.1   Background

Organizations have always tried to use external expertise when they have been unable to solve their own internal problems. However, because of the easy access to a large number of competent individuals in a global online setting, online innovation contests have increased in popularity in

recent years. In these contests, a solution-seeking organization will host an open challenge to solve a specific problem. The host can be a company, a public organization or a nonprofit organization. Solvers will usually win prize money, ranging from a few hundred dollars to several million dollars depending on the complexity of the challenge. Some large organizations host their own innovation contests (e.g. Cisco and Starbucks). One of the most well-known examples is the Google Lunar prize contest that received wide media attention in 2007. Contestants could win US$20 million in prizes if they managed to land a robot on the Moon, travel more than 500 meters on the surface and send back high-definition images and video (Innocent et al., 2017). However, in recent years it has become more common to use intermediaries that help the solution-seeker in organizing and hosting the innovation contest (e.g. marketing, answering questions, selecting winners). Some intermediaries have been around for more than a decade, with InnoCentive (founded in 2001), IdeaConnection (2007) and Topcoder (2001) being among the first. While most platforms are orientated toward research and developmental work, others, like eÿeka, focus on marketing. The innovation intermediaries usually offer a "package" of support, like guidance in formulating an appropriate challenge, connecting seeking companies with problem solvers, finding relevant technology, or help strengthening innovation networks. Several intermediaries host hundreds of innovation contests every year for their clients (Agogué et al., 2017; Terwiesch & Xu, 2008).

Both Innocentive and IdeaConnection host contests in similar areas such as chemistry, life sciences, medical science, engineering, IT and business. The Topcoder Community specializes in IT and covers areas within visual design, code development and data science projects. They offer both innovation contests and paid crowd work to its over one million members. Today, some of the intermediaries also address issues on social innovation. For example, in November 2019, one public challenge on polio eradication sought proposals on how to tackle anti-vaccination propaganda on social media in Pakistan. There were three prizes of $10,000 USD each and 316 active solvers working on a proposal (Innocentive, 2019b).

Only solvers who provide successful solutions will receive the money, transferring the risk of failure from the organization to the solver. Many contests also have a winner-takes-all competition, where the likelihood of being paid is relatively small. However, the size of the reward varies a lot. In IdeaConnection, the public challenges will usually have prizes that range from a few thousand to several hundred thousand dollars. More

prize money will usually attract the most competent experts, while in the low-prize contests; there will be fewer contestants, but a greater chance of winning. In Topcoder, some of the prizes are very small (as low as $20), because the tasks are relatively simple and have been split into many small tasks through modularization. Here, many contests also have more than one winner (e.g. first and second prizes) (Topcoder, 2019a). The financial awards are typically larger because this challenge requires more work (Innocentive, 2019c). Most intermediaries also use a fixed-price reward structure, which is known in advance. The solution-seekers will therefore know the innovation cost, and will only pay the prize money if the solution is acceptable. Therefore, more companies today consider this innovation strategy to be interesting because it can reduce innovation costs.

Another reason why innovation intermediaries are popular is that the seeker can choose to remain anonymous throughout the solving process. However, the degree of anonymity depends on the specific challenges and the intermediary. For example, in Topcoder, the winning submission is shared with the other finalists (Shafiei Gol, Stein, & Avital, 2018). After the seeker has paid for the solution, the intellectual property is transferred from the problem solvers to the solution-seeker. The solvers agree to this before they begin working on the challenge (IdeaConnection, 2019b, 2019e). The intermediaries are important because they have expertise in dealing with legal issues concerning the transfer of the intellectual property of winner solutions (Hossain, 2018; Innocent et al., 2017).

All the intermediaries are reliant on some basic requirements. They need a large and diverse pool of talent which can connect with the solution-seekers. Topcoder, which both arranges contests and offers paid crowd work, has more than a million members. Another example is Innocentive, which has 400,000 solvers with nearly 60 percent educated to Masters Level or above (Innocentive, 2019a). Most of the solvers are highly skilled, with both a relevant educational background and working experience in the field (Hossain, 2018; Innocent et al., 2017). However, in the public challenges, anyone can submit a solution and, in principle, participation is independent of age, gender, location, skill level, education or experience. Solvers are not only professionals in work, but "amateur scientists" or "garage scientists," motivated by financial reward. For instance, in the case of IdeaConnection, the solvers will also include students, retired scientists and scientists not in full-time work (Hossain, 2018).

The innovation intermediaries depend on their members bringing into play the untapped expertise from around the world. The large number of potential

solvers is necessary because solutions must be produced within a short period, both detailed proposals and working prototypes. If there are more experts in the member database, this increases the probability of reaching a potential solver with the optimal solution at that exact point in time. Because many of the challenges require advanced creative skills, it will be an advantage to recruit experts from many different fields, which increases the probability of arriving at an unusual but relevant solution (Innocent et al., 2017).

The solving rate appears to have improved significantly over the last decade. For example, Innocentive claims to have run over 2,000 Premium Challenges, with a total payout of over $20 million. And in 2016, 80 percent of the prizes that year were awarded (Innocentive, 2019a). This is a radical increase from the 30 percent solving rate that Jeppesen and Lakhani identified ten years earlier (Lakhani, Jeppesen, Lohse, & Panetta, 2007).

There may be many reasons. As time has passed, the pool of expert members has increased and the intermediaries have also improved their ability to formulate challenges in a more precise way, thus increasing the likelihood of finding the correct problem-solver. In the first phase of the problem-solving process, it is important to give precise information about the challenge. Members need to assess whether they are capable of solving the problem quickly. This increases the likelihood of solving the problem. Therefore, the innovation intermediary will often guide the solution-seeking organization in describing the problem in a format that is motivating and easy to understand. If the solutions to a problem already exists, it is essential to describe the problem in such a way that it is possible to identify the already-available solutions and customize it to fit with a seeker's problem (Hossain, 2018; IdeaConnection, 2019b; Innocent et al., 2017).

Another probable reason why the solver rate has increased is that some of the challenges have become easier to solve. For example, a technology scouting challenge invites professional searchers to locate critical technology that the seeker lacks. This challenge only requires that solvers identify existing technology that can be reused in a new context. One solver also states that some challenges primarily require laborious work, "I think this particular challenge was rather straightforward but laborious. And this is the trend I am seeing on IdeaConnection – rather than seeking 'innovation' per se, companies find this an easy place to crowdsource a lot of very cumbersome literature plowing."[1] Some of the work is more time consuming than creative, although some element of expertise is still required.

Still, although the innovation contests are organized in different ways, they will typically require that solvers come up with proposals within a relatively short period. Challenges range from a week (IdeaRally in IdeaConnection) to a few months (Confidential challenges in IdeaConnection). Because of the time constraints, the seeker will want a very specific solution. Either solvers can work individually or in a group, but in recent years, teamwork has become more common. Even Innocentive, which originally organized only individual challenges, now also offer team challenges where individuals can form their own teams.

2.2.2   The IdeaRally: Rapid Problem Solving in Large Groups

Recently, IdeaConnection have also introduced the IdeaRally, an interesting new crowdsourcing method that allows dozens and even hundreds of experts to participate in a one-week-long intensive idea-building process. By increasing the group size, it is assumed that a quality solution can be developed even within a very short problem-solving period. The large group produces and refines a much larger number of ideas compared with what a small team manages (IdeaConnection, 2019c). One solver describes it as a brainstorming process, "I think people are much more creative together absolutely because you can't just think of everything. With other people, their comments and ideas can lead you off into other areas. So brainstorming with multiple people is definitely advantageous." The brain-power of the large group is underlined, as well as how the group manages to coordinate their action so they can pursue particular ideas. In one specific IdeaRally, more than five hundred researchers participated during a period of only a week. A solver describes it as a "great learning experience":

> My first reaction to the IdeaRally® was the big surprise of having to encounter so many people with so many ideas which were mostly interesting. Now, my task became more complicated since I needed to put up some ideas which were different from others. However, I soon discovered that I do not really need new ideas all the time but could develop ideas from others or build on others' ideas . . . Building on the ideas of others is useful to both individuals and also the sponsor. Philosophically, it is by such collaboration, we all can move forward in life rather than an unhealthy competition.

When so many ideas are produced, the solver discovers that he does not have to invent new ideas but can instead build on others' ideas. It illustrates that it is possible to create online innovation contests that

synthesize and refine ideas and not only rely on individual competition between the contributors. According to the intermediary, often hundreds of ideas are being discussed, and participants are challenged to criticize, defend and expand upon ideas. A peer review process let participants vote and rank ideas, and also move them to particular strings so they can be discussed separately (IdeaConnection, 2019c). One solver describes the voting as an important part of the discussions because it makes it easy to ignore bad ideas: "If it was a bad idea it would get down voted and ignored. What was nice was that there was a lot of active discussion on some really good ideas in terms of what's doable and we know about, what hasn't been explored yet, and how do we build on things that have been explored." The participants vote on ideas, and this helps them move the discussion towards the most realistic solutions. Ideas can both be virgin ideas or a novel take on some already known ideas.

The design of the IdeaRally is interesting in that it makes it possible for participants in a large-scale innovation contests to move beyond the production of superficial ideas, a typical critique of different crowdsourcing methods that build on aggregation of ideas. A solver illustrates this by expressing excitement about this idea development process, "I was most impressed with how an idea could evolve from something very simple to one with several add-on features, simply by including suggestions and ideas from the scientific community." The solver underlines how an idea moves forward rapidly from a simple to a more complex format through the large-scale collective work. One explanation is that most of the contributors are like-minded people who work in scientific communities. The same solver also underlines how the discussion included a broader multidisciplinary group that usually do not communicate with each other:

> With global online discussions such as the Crop Yield IdeaRally®, it is so important that we can collaborate with people in such diverse fields, people we don't typically have the opportunity to work with, or even talk to. It is rare that we can work together globally and reach consensus on a single issue, but an IdeaRally® creates a platform for scientists to interact in a timely manner; it allows us to have an exchange of ideas that crosses boundaries of normal modes of scientific interaction.

There are significant diversity benefits in multidisciplinarity, but the solver also experiences a global platform that offers a type of scientific communication that is unusual in its boundary-crossing mode. In this specific IdeaRally on crop yield, the solver reports that bioinformaticians, molecular biologists, and agricultural and social scientists were all working together. In addition, the strict deadline forces the group to rapidly reach

consensus on a single issue. A solver is stunned by the amount of valuable information that was produced, "The breadth of expertise was extraordinary and the seeker received a huge amount of valuable information from people who were knowledgeable in many areas." Large-scale collective work can produce a richer solution because of the "breadth of expertise." Another solver thinks this crowdsourcing method is ideal in providing a better overview of a complex area, "I think that when you're applying so many minds to something you have a better chance of teasing out important trends or important themes in the data that can be extended into the future or that can have possibility for innovation." By including more people, the probability of identifying the most important trends increase. This may be particularly important in large research areas where it is difficult to be updated on all the published research, "especially in biological sciences now, we have this massive database of published literature. But any one single person can really only mine so much of that data or literature on their own to get a background on their research topic or what they're trying to solve."

In most innovation contests, there are several different types of challenge. For example, at Innocentive, the Ideation challenges aim to produce a breakthrough idea, whether it is a technical problem or a new commercial application for a current product. Theoretical challenges involve the production of detailed description that can bring a good idea closer to becoming an actual product, technical solution or service. Practical challenges require physical evidence that proves the solution will work according to the predefined requirements (Innocentive, 2019c).

In a typical contest, the challenges will be announced to a large pool of members with potentially relevant expertise, and they will then be given a relatively short period to solve the problem (e.g. weeks or months). Note that the IdeaRally as a type of large-scale collective problem solving involves hundreds of motivated solvers who have the opportunity to join the project within a short time period. The solvers will also join and contribute with quite different approaches, adding up to the necessary cognitive diversity. For example, one solver contributed to the IdeaRally by focusing "on the things I knew about." He did not read any extra sources, but only engaged in "the things that I wanted to talk about." By using the knowledge he already possesses, this makes participation time efficient because he does not need to do extra research. Another solver is more of a "knowledge synthesizer," explaining that he contributes with a breadth of his understanding, and "being able to put together information from many different areas." In addition, the solution-seeker organization can invite individuals who they think should be part of the process.

In this way, the IdeaRally is different from other crowdsourcing methods in how it mixes both internal company members and external expertise. Members from the internal organization can engage in the discussion or just highlight ideas that solvers should explore in greater depth. At the end of the Rally, the seeker receives a document with all the ideas and discussion. Although this type of contest involves many persons, it can still be confidential. Typically, there are thousands of dollars in prizes and awards offered each day to sustain motivation. In the end, those who have provided the most valuable ideas also receive a significant award (IdeaConnection, 2019c).

Another interesting characteristic of the IdeaRally is that solvers enjoy being part of this type of online community. One solver states that being in one project with scientists from all over the world made "a deep impression on him." This setting enables all expertise in one field to meet and discuss ideas. Another solver enjoyed the comradery of the group and the feeling of being connected with people from all over the world that one might not otherwise have met.

Several solvers also highlight the learning experience. One solver emphasizes the value of being in a transparent online environment where one can access other ideas. He likes to "read everyone else's ideas." and describes it as a learning experience:

> I learned a lot. What I really liked was learning how other people would put things together. How they would come up with their solution and the different ways that people have of looking at the same problem. . . . There were all sorts of neat ideas that people had about parts of the plant, like improving parts of the plant to improve yield that I had not thought of. So I liked that a lot.

The solver enjoys the richness of perspectives raised when many look at the same problem together. Likewise, another solver values the access to others' ideas: "I think it was interesting from an intellectual perspective to see some information from other people's areas. It gave me some extra depth in an area, and I actually came up with a potential invention, which was in a related field. So that was an unexpected benefit." This solver claims that the access to "other people's areas" triggered his own creativity and was the reason why he came up with a "potential invention." Likewise, another solver emphasizes the excitement of building ideas in this way, providing insights into other possibilities, "Working and building ideas with one's peers is very exciting and pushes one's curiosity to a good level. Beside this, ideas from other contributors can give you a great insight into other possibilities in the science world. Reading and arguing about others'

ideas is very exciting and thought provoking. It also builds on your knowledge." This solver also describes how the collaboration gave "insight into other possibilities" and had a major impact because it was "thought provoking."

Furthermore, all the ongoing activities in the IdeaRally require that facilitators keep an overview of the collective work. One solver explains how several facilitators helped the large group to move forward with some ideas:

> Having facilitators was a benefit as they helped the participants to move forward in the right direction by asking right questions or directing them to what they need to do and what not to do. Personally, I benefited from a few instances where they brought to my attention that a similar idea was posted by another elsewhere. This could help me collaborate with that person.

In this specific case, the facilitator "matched" the solver with another solver who was interested in the same idea, but had been working in another area in the online environment. Another solver also mentions that the number of ideas that are produced, risk fragmenting the debate, "I appreciated having a facilitator onsite during the Rally. Having different perspectives on one side opens up the discussion to out-of-the-box ideas, but at the same time, diffuses the focus of the debate. The Facilitator helped in keeping the focus on the matter that is discussed in the Rally and avoided tangential discussions that would derail from the scope." Here, the facilitator helps keep the focus on the matter at hand. The facilitator encouraged solvers to seek more in-depth information and not emain at a superficial level.

2.2.3   The Climate CoLab: Transparent Innovation Contests

Furthermore, transparent online innovation contest is another crowdsourcing method that let contestants build on others' work in previous contests. For example, in Topcoder, software development contests will usually be modularized and split into smaller transparent pieces. Solvers often develop a specifications document with a detailed system requirement. Afterwards, the winning specifications might become the basis for a new contest during the problem solving, solvers also ask the seeker questions in an open web forum, which makes this information visible to all competitors. (Malone, 2018: 190–191). While most innovation contests have a limited degree of transparency, the MIT Climate CoLab platform is also different in how it allows all contributions and reviews

to be open and visible to others. The Climate CoLab, a nonprofit organization affiliated with Massachusetts Institute of Technology (MIT), was established in 2009. The contests invite people from all over the world to develop proposals on what to do about climate change, including both technical, economic and political issues (Malone et al., 2017). Anyone can join, and by early 2018, the Climate CoLab community had over 100,000 participants. In total, more than 2,000 proposals have been submitted and evaluated (Malone, 2018: 171–172). The goal of Climate CoLab is to harness the collective intelligence of people from all around the world to address global climate change as a complex societal problem. By engaging a broad range of scientists, policymakers, businesspeople, practitioners, investors and concerned citizens, the aim is to develop plans that can achieve global climate change goals that are better than any that would have otherwise been developed (CoLab, 2020; Malone, 2018: 179–180).

In the first three years (2009–11), the CoLab-activities organized a set of annual online contests that addressed general topics like "What international climate agreements should the world community make?" Some proposals were interesting, but most of them tended to focus on some narrow part of the overall global problem. They were limited in supporting the development of complex solutions. Therefore, the contest format was revised in 2013. The problem of climate change was divided into a family of a dozen contests that all were related to each other, but they focused on a different aspect of the same problem. For instance, there were separate contests on how to reduce emissions in transportation, buildings and electricity generation, how to change public attitudes about climate and how to put a price on carbon emissions. With this new way of organizing the contests, the proposals were more detailed and interesting. For instance, in 2013, the winning proposal came from a nonprofit organization in India, describing how small Indian farms could replace their expensive, emission-intensive diesel irrigation pumps with cheaper and more environmentally friendly foot-powered treadle pumps (Malone, 2018: 179–180).

However, the proposals were still limited because the contestants did not look to each other's work and try to combine ideas. To address this issue, integrated contests were introduced in 2015 with a new prize (currently $10,000) awarded to contributions that integrate and combine existing proposals. This contest type aims to motivate the creation of solutions that can address larger parts of the whole problem, because entries from previous contests have to be reused. Some of these integrated contests cover climate plans for the whole world, while others are

orientated plans for the largest emitting regions (like US, Europe, China, India) (Malone et al., 2017). Compared with the other innovation contests, the openness and transparency is much larger in the Climate CoLab. Contestants are given access to others' work, and have to assess and review this work in order to submit a proposal. The contest aims to utilize a better mix of competition and cooperation in the same online environment. For instance, the Popular Choice global winner in 2015 began originally with work done by the user "biocentric stayathome mom." After posting her original proposal, several authors contacted her, and agreed to make a global proposal that eventually included over 25 authors. Many of these members did not know each other, and are now actively working together to raise money for the ideas in their proposals (Malone et al., 2017). Here, the contestants had to contact each other and collaborate in the production of a new solution that combined pieces of previous work done by others. A complex problem like climate change requires a multitude of different types of knowledge about what to do in different places around the world. By enabling more people with diverse backgrounds to combine their knowledge, this increases the likelihood of producing better solutions.

2.3   Online Citizen Science

2.3.1   Zooniverse: Online Citizen Science Platforms

Citizen science is research conducted by amateurs or individuals who do not necessarily have a formal science background. They voluntarily contribute time, effort and resources toward scientific research in collaboration with professional scientists or alone. Many citizen science projects build on a long tradition of environmental research, but today they involve most other scientific fields (Hecker et al., 2018). The last decade, the interest in online citizen science has also increased significantly, and there are now thousands of projects worldwide. On one hand, the digitization of information (e.g. low-cost sensors) provides an opportunity to collect massive amounts of data that need to be analyzed. On the other hand, the Internet and smart phones has made it much easier for volunteers to engage in citizen science in new ways. Individuals cannot only collect data themselves, but they are also involved in analyzing data that researchers have collected.

As a result, citizen science is both becoming more institutionalized with the establishment of practitioner organizations in Europe (European Citizen Science Association – ECSA), and the US (Citizen Science

Association – CSA), and increasingly recognized as a distinct field of research. In 2016, the first scientific journal dedicated specifically to citizen science was established (Hecker et al., 2018). Compared with traditional research, citizen science differs in the openness both of project participation and intermediate inputs such as data or problem-solving approaches, which are widely shared (Franzoni & Sauermann, 2014). From this perspective, online citizen science is an example of a new CI practice that is of significant societal value.

If we look at online citizen science more specifically, it is worth mentioning that there are several different online platforms that have strengthened its visibility and accessibility. Many countries have their own citizen science portal, such as Scotland (Scottish Citizen Science Portal), the US (e.g. SciStarter, Zooniverse, CitSci.org) and Australia (Atlas of Living Australia). These platforms have made it easy to create new projects and also establish networks among participants and prospective stakeholders (Hecker et al., 2018). Today, Zooniverse is the largest citizen science platform in the world, with more than 550 million classifications done by 2.2 million registered volunteers, as of December 2020. The online platform hosts a range of different science projects that invite volunteers to analyze and interpret large datasets. Anyone can start a Zooniverse project by uploading data to the platform. The projects cover anything from counting penguins and drawing diseases in nuclear cells to the digitization of historical records. Initially, most of the projects were in astronomy. Before 2010, seven out of eight were astronomy projects, while this only includes three out of ten projects in the period afterwards (2010–14). Projects now involve a broader suite of ecology and humanities subjects, and the amount of new users and projects have increased steadily by around 30 percent a year (Graham et al., 2015). In December 2020, volunteers could choose from 75 ongoing projects on the site. In total, researchers have published more than two hundred articles using data produced by these projects.

Originally, Zooniverse grew out of the Galaxy Zoo project. In 2006, a spacecraft collected samples of interstellar dust particles from the comet Wild 2. The particles in the sample were extremely small and NASA had to take 1.6 million microscope images. However, because computers are not particularly good at image detection, volunteers were instead given the task of visually inspecting the material and reporting candidate dust particles. The project, known as Stardust@Home, received a lot of interest from astronomers all over the world (Michelucci & Dickinson, 2016). Because of this success, the researchers created the online platform Galaxy Zoo the

year after. Volunteers were invited to investigate millions of astronomical images collected by the Hubble Space Telescope, the Sloan Digital Sky Survey and others. Building on basic human pattern recognition, the image detection tasks were quite simple, and anyone could therefore join the project. Individuals were asked a number of questions about the shape of a galaxy captured in an image (e.g. the number of spiral arms or how round or elliptical they are). The project received 70,000 classifications per hour within 24 hours of its initial launch and more than 50 million classifications within its first year. Because of the positive media attention, this also strengthened the public engagement (Crowston, Mitchell, & Østerlund, 2018; Graham et al., 2015). Concerning accuracy and reliability, the quality of the work was ensured by letting multiple volunteers repeat the same classification task. Because there are a small number of possible results, a simple consensus rule is usually sufficient to merge the classifications. This reduces the need for coordination, nor is it necessary to have any information about the image or volunteer (Crowston et al., 2018). Because of this success, it was decided to establish a cooperation with other institutions in the UK and USA (the Citizen Science Alliance) to run a number of projects on an online platform "The Zooniverse" that involved other fields such as marine biology, climatology and medicine (Franzoni & Sauermann, 2014).

If we look at the overall mission of citizen science, the production of *scientific knowledge* and publications is still vital, with peer-reviewed scholarly publications being the most important indicator of scientific success. Likewise, the first main objective in the online Zooniverse platform is to make scientific contributions. Usually, volunteers are involved in scientific problem solving by transforming a huge amount of labor-intensive data into a manageable "data product." The data are usually not possible to analyze with computer algorithms, and the tasks are still simple enough for volunteers to do without any need for specialist knowledge or a formal background in science. In a few cases, citizen science contributors have also been included as coauthors in a scientific publication. In Zooniverse projects, such instances have only been observed in astronomy-related projects; specifically, variants of Galaxy Zoo, Planet Hunters and Solar Stormwatch. The most common reason is that a citizen scientist has made particularly significant and unusual discoveries when visually inspecting datasets (Graham et al., 2015). For example, a citizen scientists found Hanny's Voorwerp, a novel astronomical object (Crowston et al., 2018). However, while volunteers do classification tasks within the present knowledge domain, it is more uncertain how effective they are in noticing unknown objects outside the predefined classification schemes.

Although volunteers seldom participate in the complete research process, most researchers agree that they can make substantial contributions to data collection and data coding. While there have been concerns about the data quality, one of the most successful examples is eBird, which lets volunteers use an online checklist program to report bird observations. The eBird project was initiated by the Cornell Lab of Ornithology in 2002 and has resulted in more than one hundred peer-reviewed papers. The success builds on a substantial collection of data across both time and geographical areas. When most volunteers also use the same observation scheme, it is much easier to do rigorous data analysis afterwards and publish findings in scientific journals. Since the data are Open Access, more researchers have also become interested in the project and this has strengthened the scientific impact (Hecker et al., 2018) (see more information in Section 3.2. Open Sharing of Scientific Knowledge, Open Database Projects). However, not all projects end up with scientific publications. Graham et al. (2015) find that almost half the projects in the sample (7/17) from the Zooniverse platforms have not produced any publications to date. The projects with most scientific publications are primarily "early" projects within astronomy (e.g. Galaxy Zoo). Another interesting new trend is that some projects now offer video analysis of animal behavior (e.g. ape behavior in their natural habitat).

The second overall mission with citizen science is to strengthen the *public understanding and trust in science*. The scientific engagement emerges both through the informal learning of the volunteer work, and through activities arranged by the educational system and museums. Most citizen projects aim to recruit participants with various backgrounds in an attempt to empower citizens to make scientific contributions. Citizen science is also part of a policy that aims to create a more transparent government system. For example, most projects incorporate open source software, open hardware, open data and Open Access publications (Hecker et al., 2018). If we look at the online Zooniverse platform, many projects use blogs, Twitter and Talk pages as a way of communicating with the outside world. The projects also aim to educate and change public attitudes towards science by offering opportunities of learning. One example is that volunteers receive information about the scientific publications that are a result of their project participation (Graham et al., 2015).

While large public engagement has primarily happened in astronomy projects, one exception is Snapshot Serengeti. This project studies migration and behavior patterns for a range of species in the Serengeti. Snapshot Serengeti has a median number of 4.3 hours of sustained engagement per

volunteer versus an average of just over 30 minutes for all other projects. It has a median of 61 classifications provided by each volunteer, compared to 21 classifications in other projects (three times as many classifications). A potential reason for this variation may be due to the different lengths of time it takes to complete a single classification. Other better performing projects tend to be in the area of astronomy, like Galaxy Zoo projects and Planet Hunters. Overall, these measures show a significant contrast between projects that have strong project appeal and those that do not. A typical challenge in most projects is a high incidence of users leaving the project after supplying a low number of classifications. (Graham et al., 2015).

2.3.2   FoldIt: Citizen Science Games

Online games are also becoming more popular in citizen science projects (e.g. EteRNA, Eyewire, Cancer Research). One important reason is that gamification designs motivate participants to contribute over longer periods and attract individuals with more time available (Hecker et al., 2018). Today, the protein-folding game Foldit, a collaboration between the Center for Game Science and the Department of Biochemistry at the University of Washington, arguably stands out as the most successful project. The online puzzle game is designed to enhance our knowledge of protein structure and shapes, an area that scientists have struggled to understand. This is important because a lot of biological research is reliant on figuring out the three-dimensional shapes into which the molecules in a protein chain will fold. These specific shapes explain how proteins function and interact with other proteins and cells.

However, since the configuration possibilities are endless, the most common strategy has been to make computers identify the three-dimensional movements of the protein chains. The disadvantage is that the computation is extremely intensive. Therefore, back in 2005, volunteers were allowed to help by sharing computational power from their personal computers. By chance, the screensaver was designed with a visual interface that showed proteins as they folded. To the surprise of the researchers, some volunteers began posting comments that suggested better ways to fold the proteins. This spurred the idea that human visual ability could supplement computers in doing protein modeling in a more efficient way (Franzoni & Sauermann, 2014).

In 2008, Foldit launched an online multiplayer game that aimed to combine human visual skills with computer power. Any person could join

the game and attempt to solve protein-structure prediction problems by constructing three-dimensional structures. Players compete against each other in the lowest free energy of a protein model (Koepnick et al., 2019). Because players did not need any background knowledge in biochemistry, the game became an instant success, with several thousand users signing up.

The basic gaming principle in these protein-folding puzzles is that proteins fold to their lowest free-energy state. Computer power can automatically calculate this energy level (Koepnick et al., 2019). The players use the mouse to move and rotate the chain branches of proteins in an attempt to find the most stable, low-energy configuration. A high score indicates that the protein shape has a low energy state according to the computerized energy function. The gamers use their spatial reasoning ability to manipulate three-dimensional shapes in space (Cooper, 2016: 120). This special cognitive skill does not require any background knowledge from biochemistry. Nor can computers do it effectively (Franzoni & Sauermann, 2014). The game let the players create their own scripts or short programs that automate game tasks. These scripts can improve a fold or identify the part that needs to be improved. Hundreds of such scripts have been publicly shared. All the collective work is also informed by the computerized game score, which provide precise feedback on the most useful strategies. If one high-scoring player shares a strategy, other players pay attention (Nielsen, 2011: 147).

From the very the beginning, the players showed that they were good at solving several difficult problems, and some players even outperformed the best structures designed by the computer (Cooper, 2016: 120). Some Foldit players even competed in the 2008 and 2010 worldwide competition of biochemists, using computers to predict protein structures, and they performed as well as protein-folding experts (Nielsen, 2011: 147). Because of this initial success, Foldit players were in 2011 given a challenge that had puzzled scientists for over a decade. They were to figure out the folded shape of a special type of protein associated with AIDS in monkeys (Mason-Pfizer monkey virus). Astonishingly, two teams managed to develop the most likely fold of the protein in only three weeks. The refined structure provided new insights for the design of antiretroviral drugs. These teams were also credited as coauthors in a paper published in the journal *Natural Structural and Molecular Biology* (Cooper, 2016: 120; Malone, 2018: 183). It is regarded as the first instance in which online gamers solved a longstanding scientific problem (Khatib et al., 2011). Another success came in 2012 with the remodeling of a computationally

designed enzyme (the Diels-Alderase enzyme) so it could increase its ability to catalyze chemical reactions. A typical problem with such designed enzymes is that they have significantly lower catalytic efficiencies than naturally occurring enzymes. The enzyme became 18 times more efficient after the players had improved the shape (Cooper, 2016: 124–125; Eiben et al., 2012).

The most recent trend in Foldit is de novo design of an entire protein. In the first years, this challenge was considered too difficult for amateur gamers. This is because the creation of a plausible protein backbone that could be the lowest energy state of some amino acid sequences is an extremely open-ended problem. In principle, there will be a practically unlimited number of solutions, so computers cannot do this work. In a recent experiment, Foldit players were repeatedly only given a week to design stably folded proteins from scratch. Based on the results, the game design was improved several times. Initially, most top-scoring designs were not good enough, but after many iterations of model improvement, both the top-scoring solutions and the game design improved (Koepnick et al., 2019).

Most of the protein designs were exceptionally stable, including 56 of the 146 Foldit player designs. The protein designs are comparable in quality with those of expert protein designers, and the diversity of these structures is unprecedented in de novo protein design, representing 20 different folds – including a new fold not previously observed in natural proteins. These results are impressive especially because de novo protein design is a completely new research area. The 56 successful designs were also created by as many as 36 different Foldit players (the most prolific player created ten successful designs); and 19 designs were created collaboratively by at least two cooperating players. It shows this is an achievement made by the entire Foldit gaming community and not just one or two exceptional Foldit players (Koepnick et al., 2019). Because of the diversity of contributions in the community, the players used more varied and complex exploration strategies than computer-automated design protocols. Although the players lack formal expertise in protein modeling, they have acquired a high level of knowledge and intuition just from playing the game. It illustrates that human game players can be exceptionally capable at finding and exploiting unanticipated solutions that are otherwise unexplored by experienced scientists. One possible reason is that gamers approach the problem in a different way than the researchers, because they aim to get the best high score, not only solve a scientific problem (Koepnick et al., 2019).

During the years, players have also regularly made suggestions on new automatic tools that could improve the game. The game has been modified several times based on player feedback and observations of player activity. Initially, most of the tools in the game did not exist, and the game design has adapted to players' best practices (Cooper et al., 2010). For example, one player strategy, called "Bluefuse" involved wiggling a small part of a protein, rather than the entire structure. It outperformed "Fast Relax," a piece of code that the researchers had worked on for quite a long time (Khatib et al., 2011; Nielsen, 2011: 147).

Most of the active players are part of a team. While some players work independently, most successful solutions come from larger teams which have developed solutions collaboratively by building on each other's ideas (Franzoni & Sauermann, 2014). The successful teams consist of a mix of players with different expertise who specialize in different parts of the puzzle. For instance, some players will concentrate their efforts on the start phase, while others are best at the end stages. The finishers or the "evolvers" are usually highly skilled and at the top of the rankings. They will complete puzzles that others haven't been able to finish. The players in a team also switch between being in a competitive and collaborative mode. In one team, three or four evolvers would first compete against each other in finishing a puzzle. Afterwards, they share their results with each other and collaborate in the design of the final structure. The players become better by studying each other's solutions (Cooper, 2016: 124).

The game includes several scoreboards that lists players' performance, both individually and in teams. Many players form teams to improve the rankings. In addition, there is an online community between the gamers. Gamers communicate with each other in a forum, and they share information about strategies in a wiki (Nielsen, 2011: 147). To attract a large audience and prolonged engagement, the game designers have attempted to develop a diverse reward structure, including short-term rewards like game score and long-term rewards like player status and rank. Gamers also motivate each other in chats and forums. Although players are motivated by the competition, a survey of player motivation shows that the ability to contribute to science is the most motivating factor. Social interaction is also important, as well as the feeling of being immersed in the game (Cooper et al., 2010).

Like in many other global online communities, a small group of enthusiasts is vital in the Foldit community. There are many more registered participants than active participants. About two to three hundred players actively attempt to solve most of the puzzles. Many drop out early

because of the mandatory training period; new players need to complete a series of 32 tasks (already solved) as part of a tutorial (Cooper, 2016: 121). Furthermore, only 20–30 persons comprise the core who discuss the game on forums, dominate in-game player statistics, write content for the game wiki and mentor new players. In this group, participation is a very important part of their leisure time activities. One survey shows that most of these gamers have been playing for more than two years, spending about 15 hours per week. They enjoy being part of scientific activities. One player illustrates this point, "the real point is that Foldit simply allows us folks without the proper CVs, and [who] would crawl over broken glass to participate given half the chance, an opportunity to do this stuff. It's that simple" (Cooper, 2016: 121). Most players emphasize that the game requires skills such as patience, dedication and scientific inquisitiveness (Cooper, 2016: 121).

The active players also have a similar background profile. Nearly 80 percent are male, and 70 percent in this group are over 40 years old. Interestingly, the large majority of these players have no interest in other computer games (Cooper, 2016: 121). Although training matters, one should be aware that some young people might have better visualization skills than adults. For example, one of the best players is a 13-year-old American boy. When thousands of people tried the games, the people who were good at playing returned to the game. The broad outreach is important in an attempt to recruit the few persons who possess great intuitive visualization skills. They are often difficult to find, because the persons may not even be aware that they have these special skills (Malone, 2018: 182–183).

2.4   Summary

In relation to CI, both innovation contests and citizen science projects represent promising new ways in which large groups can help solve problems of societal value. All the examples in this chapter illustrate how outsiders or unknown others can make significant and valuable contributions within the framework of a predefined challenge. The formulation of a specific problem makes it possible to bring a group of problem solvers together, whether this is an innovation intermediary, a game challenge in Foldit or a micro task in Galaxy Zoo. As mentioned in the introductory chapter, the power of the group size is about crowd production of cognitive and informational diversity, which leads to better or more accurate decisions. However, if we compare the online innovation contests

and citizen projects with each other, there are also significant variations in the collective problem-solving process, concerning both the type of skills that are used and the more specific crowdsourcing methods (crowd contests vs crowd community).

2.4.1   Crowdsourcing Skills

In most crowdsourcing projects, the outreach is broad and anyone can join. Most of the online communities have many more registered members than active participants in a specific project, making self-selection of tasks an important part of the process. The different examples show various use of different human skills.

First, in some of the citizen science projects, the tasks are simple and the contributions require only a very small amount of work. These projects typically utilize visual perception skills that most people have by analyzing images. Although the pattern recognition tasks are simple for humans to do, computers have until now not been able to do such work effectively. Project like Galaxy Zoo and Snapshot Serengeti shows that amateurs can participate successfully in providing metadata to images that researchers have already collected (Michelucci & Dickinson, 2016).

Second, some crowdsourcing projects aim to utilize special skills or special interests that only a few persons have. For example, in the citizen science game Foldit, the best gamers have exceptional spatial reasoning skills that they may not even be aware. Such three-dimensional pattern-matching skills are required to solve challenging scientific problems in the game. Computers have not been good at performing such tasks because the task also requires some degree of human intuition. Good gamers are more likely to have these skills than good researchers are. In their struggle to achieve the highest score, the gamers follow a logic that motivates them to find "unanticipated solutions that are otherwise unexplored by experienced scientists" (Koepnick et al., 2019). Not so differently, the Climate CoLab aims to identify local solutions that would perhaps not otherwise have been made public. In the open database eBird, volunteers can also contribute with local information about birds. Here, passion and interest in the topic will be more important than expert skills.

Third, online innovation contests will typically recruit highly skilled expertise. Participation in such contests may take weeks or months of work and will often require advanced expert skills. Innovation contests within science and IT will require a significant amount of specialized background knowledge or skills. Participants also know that the competition is fierce,

with no guarantee of winning any prize money. This makes intrinsic motivational factors more important, like passion for the work or learning something new (Baltzersen, 2020).

2.4.2 Design of Crowdsourcing

In designing crowdsourcing, the examples show that crowds can either be organized to aggregate a collection of contributions, compete against each other, or collaborate and share ideas. First, several of the crowdsourcing projects build on *crowd competition*, including both individual and team competitions. In both Foldit and in several types of innovation contests (e.g. InnoCentive) members create their own teams. While Foldit is built around a game design with leaderboards that include a ranking of everyone, the online innovation contests are centered on winning the first prize by coming up with the best solution. In Foldit, there are no economic rewards because gamers to a larger degree are intrinsically motivated. Depending on the tasks foldit also displays many types of different leaderboards. In innovation contests, economic rewards will be more important. However, since the basic principle in innovation contests is that "the winner takes it all," solvers must also be intrinsically motivated to sustain participation (Baltzersen, 2020). The size of the economic reward depends on the size of the tasks. If the online contest and the tasks are highly modularized like some challenges in the Topcoder community, the prizes will be small. If the contest requires advanced skills, the prizes are typically higher.

Second, several of the crowdsourcing projects aim to build a *creative crowd community.* These crowds share knowledge openly, even when the main activity is organized as a competition. In the IdeaRally, a large group shared ideas as part of the competition. This environment produces many ideas because of the large number of participants. The participants play a more important role in evaluating the ideas, when they comment and vote on them, as a part of the ongoing work. With the support of facilitators, the community selects a few of the most promising ideas that they continue to work with.

The integrated contests in Climate CoLab represent another example of how a challenge invites contestants to combine and build on previous winner solutions. The basic assumption is that Climate Change is a wicked systemic problem that does not have any quick fix, but requires complex solutions. Proposed solutions are part of a contest web that provide an overview of a large number of contests and proposals that are interlinked

with each other. This transparent contest environment aims to amplify the sharing and development of new ideas.

Many Foldit players also share problem-solving strategies with each other, and this might be easier when there is no prize money to top performers. The recent experiment in de novo protein folding illustrates that the achievement should be regarded more as a community effort than a specific individual or team performance. The community of players use more varied and complex exploration strategies than both computer automated design protocols and the small group of top-performing enthusiasts. Some players even give advice in the further development of the game design (Koepnick et al., 2019).

When problems are complex, ill-defined and unknown, it is likely that such community approaches will be more effective. All these examples illustrate that transparent crowdsourcing methods can be successful by letting everyone produce, reuse and combine solutions that others have already made. In these ideagoras, proposed solutions are commented on, evaluated and enriched in a continuously iterative process. The process of sharing appears to utilize the "many eyes" principle in a different way that permits a much larger degree of synthesizing efforts than the competitive mode.

Crowdsourcing has only been around for two decades and is still a new and immature way of solving problems. Because of the online setting, it is evident that this type of collective problem solving can be both a time-efficient and cost-effective way of including a large number of contributions. The examples in this chapter illustrate that crowdsourcing can both encompass simple and complex creative tasks. New crowdsourcing methods are likely to be invented in the near future. This topic will be further examined in the forthcoming chapters (see especially Chapter 5).

Note

[1] This and the following quotations in this chapter are selected excerpts from 80 interviews of top solvers published on the IdeaConnection website. See more information in Section 1.3 Top Solvers in Online Innovation Teams.