CHAPTER 3

Open Online Knowledge Sharing

3.1 Background

This chapter addresses open online knowledge sharing, which some label as the "memory component" in CI. Several different examples will be presented to illustrate how this new culture of sharing is emerging. Before the time of the Internet, only a very small part of the population made their opinions and knowledge publicly available to others. The communication model was built around enabling experts to disseminate their knowledge to the rest of the population. Today, the situation has changed entirely, with a majority of the population publishing and sharing all kinds of information with each other through social media. The costs of producing and publishing both unimodal and multimodal content have almost disappeared, permitting anyone to publish almost anything. Individuals do not need to be passive recipients of the "wisdom" of certified experts, but they can now publish their own opinion, information or product. Consequently, there has been an enormous increase in people participating in the cultural production and public conversation through the online setting.

A decade ago, this development was regarded as an amazing new step towards a better society through a democratization of knowledge production processes (O'Reilly, 2005). Benkler (2006) claimed these new online networks strengthened individual autonomy and human freedom and represented a fundamental improvement in human life. Everyone with Internet access can now take a more active role than what was previously possible in the industrial information economy. In the online setting, individuals can produce their own cultural environment. They can do more by themselves and create their own expressions. If a person wants to publish something, one does not need help from others or a permit from a licensing body. Individuals are also free to continue to develop and build upon much of others' creative work. The invention of new license systems such as Creative Commons has also made it much easier for anyone to

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

share their work in a flexible way. In a range of different sectors like science, education and business, both amateurs and experts are now sharing more knowledge than ever before.

In this chapter, examples of open online knowledge sharing will cover both the domain of expert-produced scientific knowledge and the massive amounts of citizen-produced practical knowledge. Not surprisingly, the sharing of scientific knowledge has become much more effective with the Internet. When the costs of publishing are reduced, open access has become the new dominant trend that makes research accessible to everyone. Increased production of open textbooks gives a more readable access to scientific knowledge and reaches a much wider audience. In addition, scientific knowledge construction processes are becoming transparent. This includes the establishment of many more open digital databases that allow anyone both to make their own contributions and get free access to all the data (e.g. citizen science project like eBird). More of the knowledge construction processes are becoming open, including both advanced scientific discussions (e.g. Polymath Project) and the development of encyclopedic knowledge (e.g. Wikipedia). Furthermore, the recent decade has resulted in an enormous increase in amateur-produced practical knowledge, involving both the sharing of texts and videos. Enthusiasts share their skills and passions concerning any activity that might be of interest to other like-minded persons. It includes a wide range of content, including more sharing of political opinion through video publishing and argument mapping. Inspired by open innovation, even business has begun to share more of their knowledge openly instead of concealing it.

3.2 Open Sharing of Scientific Knowledge

3.2.1 OpenAccess Publishing

In the history of science, the sharing of scientific knowledge has been an essential part of how humans have advanced their collective knowledge about the world. However, in the world of pen and paper, it was expensive to produce and publish research papers. A published paper required extensive typesetting, layout design, printing, and hardcopies of journals had to be sent all over the world if scientists were to have access to each other's research. With the Internet, there is no need for printed versions, and it is easy and cheap to distribute scientific papers. As a result, there has been a gradual shift in the last 15 years from a pay-for-access model in scientific publishing towards more open access (OA) publishing. There is no consensus on the definition of OA, but the most influential definition, the

2002 Budapest Open Access Initiative (BOAI), highlights that content must be free to read and free to reuse. The long-term goal is to make all research results openly available because this is how science can work optimally. Access is important because new research should build on all previously established results that are relevant. This knowledge will also be freely available to others who can potentially benefit, such as companies, journalists and student (Piwowar et al., 2018; Schiltz, 2018).

However, this transition is not happening without resistance. Publication paywalls are still withholding a substantial amount of research results from a significant part of the scientific community and from the rest of society. Because the cost of subscriptions from the large publishing houses has increased, more universities and libraries cancel their subscriptions (Piwowar et al., 2018; Schiltz, 2018). Consequently, policy guidelines have been and still are pivotal in supporting this transformation towards more open sharing of knowledge products within science. A recent political milestone happened in 2016 when the EU Ministers of science and innovation decided that all European scientific publications should be immediately accessible by 2020.

Moreover, Plan S is a new policy that aims for full and immediate access of all scholarly publications from 2021, which are to be published with a Creative Commons Attribution license (CC BY). Major stakeholders (researchers, universities, libraries) and public funders of research in Europe are supporting the plan. Several American research-funding institutions have now also made OA publishing mandatory, including US National Institutes of Health, US National Science Foundation, and the Bill and Melinda Gates Foundation (Schiltz, 2018).

Interestingly, some studies also find an "open access citation advantage" (OACA), indicating that OA scientific papers maximize visibility and receive more citations than other papers (Piwowar et al., 2018). In addition, the Plan S guidelines strongly encourage the early sharing of research results and data through preprints. A "preprint" is the final draft of a scientific paper, which is ready to be reviewed by a scientific journal for publication. The publication of these preprints have increased the speed of knowledge sharing, and it is now common that scientists publish a preprint at a local institutional website, or through academic social networks like ResearchGate and Academia (Nielsen, 2011: 161).

3.2.2 Open Database Projects

Furthermore, digital databases are becoming increasingly important. One example is the National Cancer Registration and Analysis Service, which

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

links hundreds of thousands of cases of each year. It collects diagnoses, scans, images and past treatments. These data are then combined into tools that can help patients choose different treatment options and doctors in their daily work (Mulgan, 2018: 28). It has also become easier to let volunteers provide data to such online databases. There are examples from many different areas and contexts, like in environmental research (e.g. Luftdaten.info) and disaster management (Bhuvana & Aram, 2019).

One prominent example is the eBird project, a citizen science project initiated in 2002 by Cornell University's Laboratory of Ornithology. On this website, amateur birdwatchers share their observations: what species of bird they saw, when they saw it and where they saw it. Most contributors submit checklists that give a complete account of both the birds that were present and absent in the area. Still, doing this work primarily requires available time as a resource, rather than a very high level of expertise about birds. In addition, some organizations and federal agencies upload and share their data on eBird. In 2016, over 270,000 volunteers had provided over 280 million bird observations. At an aggregated level, all the submitted observations provide a unique overview of the world's bird populations.

The website offers intuitive graphics and maps that show the density of particular birds in different locations. These maps are useful in tracking how climate change influences bird populations. They can also be used to inform the public. In total, 120 scientific publications have used data from the site, showing that the database has produced a significant amount of scientific knowledge (Cooper, 2016: 44–49).

The volunteer birders will typically be motivated by a desire to help bird conservation. In one incident, the Nature Conservancy in the United States used eBird data to decide which "pop-up" wetlands to fund during bird migration through Central Valley in California. The Pacific Flyway is a migration route for shorebirds traveling the Artic to South America, and the Central Valley is the natural stopover site for migrating water birds. It supports 30 percent of shorebirds and 60 percent of waterfowl, thereby hosting the highest density of migrating waterfowl in the world. The problem is that more than 95 percent of the original wetlands have been lost, and because of extreme drought in the region, the migrating birds have even fewer stopover sites. In this situation, the Nature Conservancy decided to help these birds by renting land from farmers and creating artificial "pop-up" wetlands. The key to the project's success was about identifying the right acres to be flooded at exactly the right time. Here, the citizen science data in California are invaluable, with over 30,000 checklists of the area. With the help of these data and high-performance computing, the eBird team was able to forecast where birds were likely to be present. Farmers in specific locations temporarily filled their fields with a few inches of water during spring and fall migration, in periods of six and eight weeks. All 57 species of shorebirds and a total of 220,000 birds were recorded in these pop-up wetlands during migrations (Cooper, 2016: 44–49).

Members in eBird also become part of a global network of birdwatchers, with both amateurs and researchers sharing checklists. Many use the site to locate where birds are in a specific area when they are planning birding trips. While millions visit the site, only a very small percentage of these users submit the vast majority of bird sightings (Cooper, 2016: 44–49).

3.2.3 Open Textbooks

Open textbooks is a third emerging area, which shares scientific knowledge in a format more accessible to a wider audience. The digital version is made freely available with a license that usually also allows modification of the content. The print version will typically resemble a traditional textbook, but at a significantly lower price. In tertiary education, one challenge today is that the cost of textbooks prevents many students from buying them. In one recent study from a large private university in the US, more than half of the students said that they had not purchased a textbook because of cost (Martin et al., 2017). Likewise, in another study, Feldstein et al. (2012) found that only 47 percent of the students purchased the paper textbooks, but when they switched to an open textbook, 93 percent of students reported reading the free online textbook. The cost of textbooks is a barrier especially for students from lower socioeconomic backgrounds (Feldstein et al., 2012). Other studies also show that the use of open textbooks is as good as other alternatives concerning content quality and student performance (Delgado, Delgado, & Hilton III, 2019; Hilton III et al., 2019; Jhangiani et al., 2018; Pitt et al., 2019).

Some of the most successful projects have received both financial and political support. For example, in 2012, the Ministry of Advanced Education announced its economic support for the creation of open textbooks for the 40 highest enrolled subject areas in the post-secondary system. The University of British Columbia (BC) in Canada was responsible for running the project, and it resulted in 180 open textbooks during the five first years. In June 2019, the site estimates that over 100,000 students have saved a total of approximately ten million dollars, involving more than 500 Faculty at over 40 institutions (open.bccampus.ca). Another example is OpenStax, an open textbook publisher based at Rice University in Houston, which since 2012 have published 29 free, peer-reviewed, openly licensed textbooks for the highest enrolled high school and college courses. More than six million students have used these books. In 2018, 2.2 million students in 5160 institutions saved a total of $177 million by using free textbooks from OpenStax. This includes approximately half of all US colleges. In addition, many schools outside the US, as in the UK or Poland, use the textbooks (Ruth, 2018). In contrast to the BC textbooks, OpenStax is reliant on philanthropic funding. Authors are usually paid to produce curriculum-aligned textbooks, which are both peer reviewed and regularly updated (Pitt et al., 2019).

Until now, the usage of open textbooks has largely been confined to North America (Allen, 2018). Although the cost of textbooks is a more significant barrier among US students, there is, for example, a rising concern around student costs in UK higher education (Pitt et al., 2019). Therefore, an increasing number of institutions have now begun to fund the production of open textbooks. These books are used much more often than other forms of Open Educational Resources (OER). One likely reason is that it is easier to use these books in the same way as traditional textbooks, not having to change any part of the pedagogical practice. If the quality of the book is sufficiently good, the cost savings will motivate a change (Pitt et al., 2019). Another advantage with open textbooks is their availability in different formats, making the book readable on digital devices.

However, there are still significant barriers. First, it is a challenge to find the relevant high-quality open textbooks that meet users' needs. Although a large amount of content has been produced, it is archived in local repositories that are not necessarily connected with each other. Neither are all repositories well organized, making it difficult and time consuming to find the best open textbook (Al Abri & Dabbagh, 2018).

Second, quality assurance of open textbooks is important because people are still skeptical about the quality of free and open resources. Consumers often use price as a measure of quality if they do not have access to other measures of quality. A free textbook is assumed to be of inferior quality compared with a costly textbook (Abramovich & McBride, 2018). Therefore, textbooks and other OER materials will have to be peer reviewed because this is the most legitimate quality control processes in academia (Al Abri & Dabbagh, 2018). For instance, when The Open Textbook Network runs workshops at member universities, it encourages participants to review open textbooks. Open reviews also make the quality of the textbook transparent, adding an extra advantage to traditional textbooks (Pitt et al., 2019).

Third, the open license makes it possible to adapt or change the educational content, but people still lack an understanding of how this can be done. Nevertheless, this is important to ensure that the quality is maintained over time (Al Abri & Dabbagh, 2018).

3.2.4 Wikipedia

The online encyclopedia Wikipedia represents one of the largest knowledge-producing communities in the world. It has greatly extended our ability to provide "vast and complete" encyclopedic knowledge. It was established in 2001, and by 2020, the English edition of Wikipedia had more than six million articles ("Wikipedia:Size comparisons," 2020). Every article will usually also have a large number of internal links to other articles and external links to more relevant information on the web, and the complexity of the encyclopedia is also displayed through the enormous number of articles that are linked together. The sheer size, the open invitation to participate and the quality of the content have made many researchers claim that Wikipedia is the ultimate example of what CI can achieve in its attempt to support a more informed global society (Benkler, Shaw, & Hill, 2015; Bonabeau, 2009; Castells, 2010; Malone et al., 2009)

Common sense suggests that if amateurs without payment or ownership make millions of contributions, the quality of the work will be poor. However, studies have shown that the quality is comparable to traditional encyclopedias (Giles, 2005), and that vandalism and inaccuracies are often quickly reverted (Kittur & Kraut, 2008). Today, Wikipedia is one of the most important sources when looking for reliable and valid information on the Internet. It is the world's most frequently used source of medical information, not only used by patients, but also health professionals. For example, in 2017, the English language medical pages registered more than 2.4 billion visits, far more than websites like those of the World Health Organization (WHO). An article on pneumonia has 8,000 views a day. The popularity makes it even more important keep the articles updated with reliable information sources, so all stakeholders can access the same background information (Murray, 2019).

The production of articles introduces new types of collective writing. Articles are constantly modified and updated, and are in this sense never completely finished. With this as a premise, contributors only need to publish a draft version on an article and expect unknown others to continue the work on the article at a later point in time. Work on the articles also includes a range of different microtasks, such as keeping an article updated with new information, removing "nonworking" links, and adjusting the article to an encyclopedic format. Often, it will not be too difficult to find relevant secondary sources to use in a Wikipedia article, and a lot of the writing translates content between encyclopedias in different languages.

The writing process is special in that most articles can be changed by anyone at any time. Revisions continue until there is an informal consensus that the article has reached a sufficient level of quality. There is no hierarchical editorial process. If two people disagree on the content in an article, they are strongly encouraged to find a solution on the specific article's talk page. Here, anyone can discuss issues regarding a specific article, like shortcomings, improvements and even a proposed deletion of the article. Because everything written on Wikipedia needs to have a source, this is an essential component to all articles, and often a popular topic of discussion. Most of the editors have never met each other in real life (Carleton et al., 2017; Malone, 2018: 117).

For example, Wikipedia's medical pages require that all content refers to a high-quality secondary source which is regarded as being more reliable, with less content bias. One avoids primary sources because this information can be refuted. The articles aim to represent the current state of knowledge, presented in an impartial manner. Organizations with a mission of disseminating information, like Cochrane and Cancer Research UK, are therefore now collaborating with Wikipedia. Since the encyclopedia is widely used, increased engagement from health professions can provide better information to everyone about health issues (Murray, 2019).

Although the Wikipedia user community is without a centralized structure, it still depends on a range of different norms and policies that guide actions. Guidelines help contributors to write appropriate articles within the genre of an encyclopedia and resolve conflicts between contributors. Although anyone can participate and contribute to Wikipedia, many norms regulate online behavior. Instead of letting a central body monitor all behavior, the Wikipedians monitor each other (Carleton et al., 2017). The norms build on a general hacker ethos, and include sentiments such as "Be bold" and "Leave things better than you found them." The Wikipedia community resembles a participatory culture in its emphasis on behavioral guidelines like "civility," which refers to a social policy that encourages respectful and civil participation. Contributors should both try to understand others' positions and "strive to become the editor who can't be baited" ("Wikipedia:Civility," 2020). The guideline "Assume Good Faith" refers to the treatment of others as if they have good intentions and one should avoid accusing others of harmful motives without clear evidence ("Wikipedia:Assume good faith," 2020). If a disagreement is not solved, the debate can involve a third party (Algan et al., 2013; Carleton et al., 2017). These social norms are an important reason why the community manages to produce articles of high quality.

A major concern in open editing is that, when anyone can change an article, how can we trust that the information is correct? Wikipedia tackles this through the participation of a dedicated community of Wikipedians, volunteers who continuously monitor articles and receive automatic alerts when articles are changed. This makes it possible to quickly remove vandalism and restore the original article. Other controversial edits are discussed on the articles talk page until consensus is reached. The norms emphasize a civil, open debate in an attempt to produce unbiased objective content (Murray, 2019). An important technical feature in the wiki software is that it stores all edits permanently, making it possible to trace and restore previous versions of both articles and discussions. This makes the production environment very transparent because the complete decision-making process can be scrutinized by anyone at any later point of time. The success of the online community is reliant both on this transparent quality control mechanism and on specific social norms.

3.2.5 The Polymath Project

The Polymath Project, initiated by Fields Medalist Timothy Gowers in 2009, is another interesting example of open scientific knowledge construction processes. Inspired by web 2.0, Gowers wanted to explore if massively collaborative mathematics could be possible. In his personal blog, he invited anyone to join him in solving a mathematical problem through a virtual math team effort. The goal was to find a new proof for a theorem, which had previously only been proven in a very indirect and obscure way. The invitation was accepted by Terence Tao, another fields medalist working at UCLA, in addition to a number of other less famous colleagues, including both schoolteachers and graduate students. Although the project required a high level of mathematical skill, the participants were a mix of both researchers and hobby mathematicians (Michelucci & Dickinson, 2016; Tao, 2014).

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

The first Polymath project was solved successfully after approximately one month (37 days), involving contributions from 27 persons. The number of contributors in the projects are usually relatively small, typically not more than a few dozen persons. Although the outreach is large, and anyone can join, participation still requires a high level of background knowledge.

Newcomers also have to build on previous work in a sequential fashion by leaving comments on blog posts. In the early phase of the project, it was quite easy to keep an overview of the ongoing discussion. However, because of the popularity of the project, the number of comments grew quickly, eventually reaching 800 comments and 170,000 words. Although a wiki site was set up to extract the most important insights from the discussions, it was difficult for newcomers to join the project in a late phase because they had to read an increasingly large portion of previous contributions that had been made (Franzoni & Sauermann, 2014; Gowers & Nielsen, 2009; Nielsen, 2011: 51). Until 2016, there have been nine Polymath projects taking place over the course of several months to a year; three of them also resulting in published papers (Kloumann et al., 2016).

In the Polymath projects, the problems are usually at first presented as a unified whole, and any decomposition needs to arise from the collaboration itself (Kloumann et al., 2016). The disadvantage with this lack of initial modularization is that it becomes more difficult to let a very large group of mathematicians contribute (Nielsen, 2011: 51). For instance, the successful Polymath8 project had a much stronger modular structure with a problem that could be decomposed into separate pieces. This made it easier for people to contribute on one subtopic without necessarily being expert in all other areas. It was easier to measure progress in the project and there was a guaranteed end to the project (Tao, 2014). Another issue is if the modules or subtasks are relatively large, and require a significant amount of time and effort, the number of potential contributors will usually decrease (Franzoni & Sauermann, 2014).

Although most Polymath projects require some level of mathematical background knowledge, they do not require a lot of very specialized and technical mathematical expertise. This is important if one wants to recruit a large group of people to join the project. However, a consequence is that these projects have only made progress on problems where there has already been a number of promising ways to make progress. For the truly difficult mathematical problems, where some genuinely new insight is needed, it has not been proved that these projects have achieved more than what an individual mathematician could (Michelucci & Dickinson, 2016). The Polymath projects have been very good at solving minor technical or mathematical issues, like tracking down a little-known piece of mathematical folklore, or performing a tricky computation (Tao, 2014). In addition, the online setting has recruited people with relevant expertise who would never have heard about the project if it had been done in a traditional way.

Furthermore, Gowers not only describes the problem and the background materials, but he has also made a list of collaborative rules. These rules are important in creating a polite and respectful atmosphere during the informal discussions. One of these guidelines encourages participants to publish ideas even if they are not fully developed. It underlines the importance of sharing unfinished ideas, rather than thinking offline and waiting to contribute with a larger idea in a single comment.

At all stages of the research process, the comments are fully open to anybody who are interested. All the participants can follow the rapidly evolving conversation and jump in whenever they had a special insight. In the online setting, this is much easier to do. The project illustrates how a relatively large group can effectively harness each participant's special competence, "just-in-time," as the need for that expertise arises. In conventional offline organizations, such flexible responses are usually only possible in small groups. In larger groups, this will normally not be possible and participants will instead be focus on a preassigned area of responsibility (Nielsen, 2011: 34–35).

The blog is also interesting because it gives an insight into the minds of some of the world's leading scientists. When all posts are archived, they are left open for others to read afterwards, and leave traces of the knowledge construction process. The discussions follow a timeline, and provide a glimpse into the minute-by-minute communication between scientific partners. It is possible to observe how the best in the world struggle to extend our understanding of some of the deepest ideas of mathematics. It also shows how individual ideas are refined and further developed through open collaboration. A wide range of ideas is displayed, but not all are followed up. It is possible to read a record of the entire collective process that leads to the proof, giving a complete account of how a serious mathematical result is discovered. In this way, the Polymath Project makes both the scientific culture and the exploration of scientific problems more transparent (Kloumann et al., 2016; Nielsen, 2011: 167–168). The archived comments show how proposed ideas grow, change, improve and are discarded. It reveals that even the best mathematicians make mistakes and pursue failed ideas. False starts are an integral part of the process, but through the mistakes and wrong choices, the insight gradually emerges. The transparency surrounding the ongoing problem-solving process stands in contrast to how research results are usually proved in private and presented in a finished form. The Polymath Project illustrates how knowledge construction processes that have traditionally remained tacit in scientific research can be openly shared with others (Tao, 2014).

The discussions of mathematics in the blog are different from a face-to-face conversation in other ways, too. In the online setting, most comments in the Polymath Project focuses on only one point in a relatively sharp way. This is usually not possible in offline academic conversations because someone will become confused, it will be necessary to backtrack, while others will leave the discussion. However, asynchronous communication let everyone read the comments at a suitable time, and they can even do so several times before they write their own comment. In complicated mathematical problem solving, this can be a significant advantage. It is not necessary to take an immediate stance to a problem, which will usually be the case in a conversation in an offline setting (Nielsen, 2011).

Furthermore, in the online environment, it is easy to have a quick look and ignore irrelevant comments. In the project, there were a small number of contributions of low quality, but it was relatively easy for well-informed participants to ignore them. This is often a major concern in other open online environments because of trolls, spammers or even people who are just plain unpleasant. In the Polymath Project, the strategy was simply not to give these participants the same amount of attention. In comparison, when this situation occurs in an offline setting, you may have to stay and listen to a person speaking about something irrelevant for a longer period before you can move on. In the blog, you can more freely choose between what ideas you want to continue to work with. In addition, one can easily return to previous comments at a later point in time because they are archived and can be retrieved through search engines (Nielsen, 2011). The Polymath Project illustrates the potential of scaling up the number of participants in academic discussions, but it is more uncertain if such projects are sustainable without coordinators who have the main responsibility.

3.2.6 Galaxy Zoo Quench

The Galaxy Zoo Quench project is interesting because it aimed to be more ambitious than most other citizen science projects. Citizen scientists were invited to be involved in the complete research process, not only classify images, but also analyze data, discuss the findings and write a research paper (Franzoni & Sauermann, 2014). In the first phase, the participants classified galaxies independently from each other, following a common coding system in the Zooniverse platform. This task was quite simple and was completed successfully. However, the difficulties began already in the next phase, when the volunteers were assigned to create a dataset suitable for analysis. This was the first collaborative task. A suitable sample of galaxies needed to be included in an unbiased way, but because the volunteers refined the data differently, they did not manage to reach a decision together. The lack of academic background knowledge made it difficult to know what selection criteria were appropriate in making the dataset ready for analysis (Crowston et al., 2018).

In the data analysis phase, the volunteers struggled even more in coordinating the collective work. They were uncertain of the most relevant set of results to include in a research paper. The lead scientist encouraged the volunteers to "play" with the data and try to find some interesting trends, but they did not receive any specific advice. They found it difficult to do these explorations on their own because they had not written scientific papers before. As a result, the volunteers did different analyses independently of each other. Because they had limited scientific domain knowledge, they did not know what data would be interesting for publication. Therefore, the project never reached the writing phase. In the evaluation, the volunteers suggested that the lead scientist should have coordinated more of the work and provided more feedback. At the same time, collaborative writing of a paper requires much more complex interdependent work and it is not certain whether volunteers can be trained in developing these skills over a short period (Crowston et al., 2018). This project shows the importance of also examining limitations in volunteer contributions to scientific knowledge.

3.3 Open Sharing of Practical Knowledge

3.3.1 Open Sharing of Videos

If we look at the scale of online knowledge sharing in recent decades, videos arguably represent the most important contribution to human collective memory in its production of amateur content. YouTube is the dominant media platform in the world, and in 2017 it had over 800 million unique visitors each month (Lee et al., 2017). The company website claims that their billion users are watching a billion hours of content each

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

day (Burgess & Green, 2018). Unlike social media platforms like Facebook, the user engagement on YouTube is centered around the sharing of content, and the video in itself is regarded as the primary vehicle of social communication (Klobas et al., 2019). Established back in 2005, more than a decade ago, YouTube became an instant success, making it easy for anyone to share and stream videos with standard web browsers and modest Internet speeds. Videos could be rated or commented, and the website also became popular because of new social features like the automatic receiving of other video recommendations, the possibility of embedding video and the sharing of comments through email links (Burgess & Green, 2018). Already from the beginning, the content contributors were a diverse group with multiple interests, including large media producers, major advertisers, small-to-medium enterprises, cultural institutions, artists, activists and amateur media producers. All had their own separate aims, looking for a cheap distribution alternative. With the exception of violent and sexually explicit content, users could upload whatever content they wanted. This turned YouTube into a dynamic cultural system (Burgess & Green, 2018: vi–vii, 3).

YouTube's popular culture is still characterized by its own two "native" genres, the clip or quote, and the vlog. Early YouTube contained a wealth of short video quotes, snippets of material that captured the most significant part of a program, shared by ordinary users. The quotes are edited selections of TV shows, news, sketch comedy, music videos or movies uploaded informally by ordinary users, highlighting a particular moment from a favorite television show or sporting match. This quoting is very different from sharing a complete TV program. It is similar to how GIFs on Facebook and Twitter are used as visual annotations or reactions. The quotes give information about what engages the audience, but some also express particular identities, like footage from soccer matches, edited to include pictures of fans and a certain theme highlighted throughout the season. Although these clips may attract many viewers, they do not necessarily trigger a lot of discussions (Burgess & Green, 2018: 50, 75, 81, 129).

Furthermore, the "vlog" (short for videoblog) genre is one of YouTube's most central cultural forms, dominating the "amateur" videos and vernacular creativity from the early years of the platform. The vlog only requires a webcam and is technically easy to make. The emphasis is on good storytelling and a direct, personal address, typically presented as a monologue delivered directly to a webcam, including home movies and personal photography. The topic can be anything from comedy, celebrity gossip, political debate to the mundane details of everyday life. It is a mode of individual self-expression and everyday aesthetic experimentation that not only wants a large audience, but invites feedback in a direct face-to-face address to the viewer. It is a genre of communication that invites critique, debate and discussion, with direct response, through comments or video response, being at the core of this type of engagement. Early vlogs were frequently responses to other vlogs, directly addressing comments left on previous vlog entries (Burgess & Green, 2018: 39–40, 81, 127). The vlog builds on live performance traditions and resembles the vaudeville tradition of the late nineteenth and early twentieth centuries, with a wide range of short memorable acts, usually under 20 minutes. Without directors, actors in this tradition chose their own emotional material and adjusted their performance based on direct audience feedback. Like in the vlog, the emphasis is both on immediacy and conversation (Burgess & Green, 2018: 80–82, 87).

From the perspective of knowledge sharing of societal value, the vlog is relevant in how it transforms everyday life into more "public" debates around social identities, ethics and cultural politics. Existing assumptions are questioned through the presentation of intimate and vulnerable moments, making it possible to promote a public discourse about uncomfortable, or difficult topics that other media avoid. For instance, the sharing of "coming out" videos have become important "social media rituals" for LGBTQ YouTubers, displaying stories about difficulties and how one overcomes them (including homophobic bullying). It illustrates how popular culture becomes a part of political participation and citizenship, especially for woman, LGTBQ people, and religious or ethnic minorities (Burgess & Green, 2018: 124, 127–128).

A major difference today is that the scale and complexity of its commercial practices has increased, providing content watching for a large number of users. However, the informational content still includes user-created newscasts, interviews, documentaries that resemble the vlog genre, in that they frequently critique popular media through commentary or visual juxtaposition and commentary. Many music artists also preface their work through a discussion of their motivation, attempting to establish a more intimate relationship with the audience by responding directly to suggestions and feedback (Burgess & Green, 2018: vi–vii, 3, 22, 81, 87, 94, 126; Klobas et al., 2019). The highly invested content creator is not only a media company, but also professional "amateurs." On the one hand, online video businesses are working to professionalize previously amateur YouTubers. But on the other hand, the vlog and the vernacular aesthetics is often held up as the gold standard of the YouTube brand. There still remains a cultural logic of community, openness and authenticity that highlights ordinary people's active participation (Burgess & Green, 2018: vi–vii, 22, 87, 94, 126; Klobas et al., 2019).

Furthermore, educational videos are the third most commonly viewed type of content, after music and entertainment videos, including videos made by both professionals and amateurs (Klobas et al., 2019). Auto-captioning and translation of YouTube videos have also increased the potential audience that can watch a video (Lee et al., 2017). All this video content can support students' learning. For example, in one study in medical education, the vast majority of students report using internet sources, with 78 percent using YouTube as their primary source of anatomy-related video clips (Barry et al., 2016). Many universities publish video lectures, also in combination with Massive Open Online Courses (MOOCs) that offer more affordable education to a global community that would otherwise not have access to this kind of content (Lee et al., 2017).

Furthermore, a rich mix of knowledge providers outside of the traditional higher education institutions also produce and publish short clips that attempt to explain complex in a simple way (e.g. health issues). For example, science channels are made by media companies (e.g. National Geographic), science journalists (e.g. Periodic Videos) and science educators (e.g. SciShow), while other videos are made by hobby amateurs who have a passion for science. Many videos aim to be both educational and entertaining at the same time, targeting both children and adults. A typical video will explain a particular issue in just a few minutes, with music and sound animations; some will also include funny scenes from everyday life (Rosenthal, 2018; Schneider et al., 2016). One example is a video demonstrating the Magnus effect with a back-spinning basketball dropped from a very high point, which has been viewed more than 40 million times (Rosenthal, 2018; Veritasium, 2015).

In this genre, there are millions of amateur-produced clips that intend to help users with everyday tasks just about any subject, craft or skill – guitar-playing, cooking, dancing, maths, repair work or computer games. These instructional videos are especially effective in supporting procedural learning, and in principle, anyone can teach others a skill by creating a video. For examples, gamers will often show in-game achievements by showing and talking about what they are doing in the game. This is both a way of sharing knowledge as well as "showing off" one's own competencies. These clips are often made by private persons in their leisure time and illustrate

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

how people want to share their passion and knowledge for hobbies with others who have the same interest. This peer learning is both about making your own knowledge explicit, and letting others learn from what you know (Burgess & Green, 2018: 125–126; Lee et al., 2017). Studies show that videos on YouTube are used to support both formal learning and self-directed learning, offering individuals a large degree of autonomy and control regarding what and how to learn (Lee et al., 2017).

Note that YouTube is not only a massive repository of video content but also a constantly growing record of the popular culture of the Internet. Users from all over the world have created a diverse and disordered public archive of contemporary cultures. Major music labels have contributed videos from their catalogues and TV channels such as HBO and BBC. Today, a majority of viewers go to YouTube to listen to music they are already familiar with. Adults can listen to old music videos or watch old clips from TV series, as a way of recapturing memories from their childhood or young adulthood (Burgess & Green, 2018: 135–136).

3.3.2 Open Sharing of Geographical Resources

Another interesting open database project is OpenStreetMap (OSM), founded in 2004 by Steve Coast. He wanted to make a local map but became frustrated with all the restrictions on traditional maps because of copyright and excessive royalty payments. Therefore, he bought a GPS and started collecting tracks around his local area of central London. The data were then displayed openly, and when he presented his work at a conference, many people wanted to join the project. Within 16 months, there were 1,000 registered users, and after five years, the number had grown to 100,000. Although the coverage varies, OSM has continued to grow. The data sources are free of charge and allow anyone to reuse the data as they like (Chilton, 2009). Local maps have been created to serve different purposes, such as skiing, hiking or public transportation. The Wheelmap project is one example of how maps can be tailored to wheelchair users or visually impaired pedestrians, utilizing haptic feedback. Another example is how the maps have been successfully used to produce and distribute free mapping resources in disaster management ("Humanitarian OSM Team," 2020; Neis & Zielstra, 2014).

In 2020, the OSM project had more than six million registered members. Most of the information about the project information is shared in the official OSM wiki. This includes information about usable software and tutorials for beginners on how to map an area. In the past, volunteers could only report an error in the map data in the form of a note, but now they can make direct modifications or corrections in the map. This "wiki-solution" has strengthened the collective effort of the project. Like in the eBird project, only a small percentage (1.6 percent in 2013) of volunteers contribute on a regular basis. A few individuals will usually collect most of the data from one specific area. Although contributors can communicate with each other on internet relay chats (IRCs) or mailing lists, most of the collaboration is purely incidental, as most work is done by individuals separately (Neis & Zielstra, 2014).

3.3.3 Open Sharing of Corporate Knowledge

Moreover, open sharing of knowledge has increased in sectors that traditionally have kept their knowledge secret to others. In the business sector, some companies are changing their strategy and emphasizing open sharing of knowledge to a larger degree. According to Bogers et al. (2018), there are two important kinds of open innovation: outside-in and inside-out. As mentioned in Chapter 2, crowdsourcing, or the outside-in part of open innovation, is about integrating external inputs. In addition, the inside-out innovation requires organizations to allow underutilized ideas to go outside the organization for others to use. The basic assumption is that openness can be useful for process innovation (Bogers et al., 2018). According to Chesbrough (2017), this type of innovation is inspired by open source methods from software communities. Usually, innovation activities are concealed because they are a source of competitive advantage that should not be shared with anyone.

As counterintuitive as it may seem, Von Krogh et al. (2018) find that most companies can build greater advantage by following a policy of open process innovation. One strategy is to open up the organization internally as much as possible. By sharing innovative practices and success stories, this increases the likelihood that the best ideas become part of the overall corporate program, thus improving the operational performance. It is often easier to implement new ideas within the same organization because the different factories will usually be comparable. In one example, a Volvo Group remanufacturing factory were forced to think harder about their current practices when they learned about the best practices from other units. Companies can also improve if they use ICT to share practices more systematically (Von Krogh et al., 2018). In another example from Xerox, the technicians were usually alone while they repaired a copier, but the time they spent together at breaks was a critical resource for open sharing of their work. There, they discussed how to fix important work-related problems not written in the official manuals. Partly because of this work, Xerox later created an online tool called Eureka that technicians could use to share tips with one another across the company (Malone, 2018: 118–119). Likewise, the Volvo group collects best practices from factories and shares them in a global online database. In addition, global online knowledge-sharing conferences are held ten times a year, with a couple of hundred persons attending. The conference slogan illustrates the core idea behind this intracompany open process innovation strategy: “Everyone has something to teach; everyone has something to learn.” The best-in-class factories also develop their own expertise by teaching others about what they do. The better you are, the more you can gain by opening up. For instance, in a Volvo Group truck assembly, the customer fairs moved to the factory site. In this way, customers could question blue-collar operators working directly on the line, and received passionate answers. In addition, the operators learned firsthand what customers really wanted from Volvo trucks (Von Krogh et al., 2018).

The key issue here is to put more emphasis on the pace of the process innovation. Protecting innovation processes will give a competitive advantage for a limited time only. In the end, it will be a losing strategy because competitors usually catch up. Instead, it is important to compare your own practices with someone else’s practices. This exposure motivates both managers and employees to speed up problem solving and idea generation. The key is not to be better, but faster than competitors at process innovation (Von Krogh et al., 2018).

3.3.4 Open Sharing of Political Arguments

Regarding CI in the political domain, there is today an increasing disappointment with lack of informed political debate in the online setting. Currently, popular social media produce little deliberation, large volumes of highly disorganized and low-quality content, toxic interactions, and in some cases, clique formation amplifies extreme political points of view (Fujita, Ito, & Klein, 2017). From a technological perspective, part of the problem can be due to limitations in the communication technology. For example, in time-centric tools like blogs or discussion forums, the contributions are organized according to when a post is submitted. When the number of contributions increase, posts about the same topic will often be widely scattered, and it will be increasingly time consuming to identify all relevant issues, ideas, and arguments in a debate. As this becomes more

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

difficult, so the likelihood of redundancy increases. There will be a lot of repetitions, digressions and people talking past each other (Klein, 2012, 2017).

Collective argument mapping represents an interesting technological alternative that attempts to avoid these problems by letting a large group co-construct the bigger picture of an issue from multiple perspectives. This is done through the collective production of a coherent argument map (e.g. Deliberatorium, Kialo). User contributions are organized through the construction of a tree structure consisting of specific issues, potential solutions, and pro and con arguments. This structure provides a better overview through easy navigation, rating and collaborative editing of the map. The goal is to produce a well-organized map with nonredundant, high-quality content for complex controversial problems. The map intends to support deliberation, long and careful discussions where groups of people identify possible solutions for a problem, evaluate these alternatives, and select the solution or solutions that best meet their needs (Fujita et al., 2017; Klein, 2017).

In the map, the arguments are captured as topically organized tree structures where arguments comprise questions, possible answers, arguments or statements in favor of an answer or argument. All relevant arguments and subarguments within the same topic are organized hierarchically in the same branch of the tree. The map can grow collaboratively from a simple seed question into a large range of ideas that represent a single, coherent, meaningful structure. With the visual support of a multidimensional map structure, all participants in a community can bring forward any question or issue on a topic, and the community can evaluate the content together (Bullen & Price, 2015; Klein, 2017).

In political discussions in large groups in an offline setting, many perspectives will easily be ignored. Typically, small groups of people will outline a policy, and then attempt to engage wider support for their preferred options. The large majority will not be involved in formulating alternative solutions. If the problem is complex, many important ideas may be ignored. Therefore, the map aims to offer a group a comprehensive overview of a problem that supports more informed deliberations that can lead to better collective decisions (Bullen & Price, 2015; Klein, 2017).

Today, several different collective argument-mapping tools support large-scale discussions. One example is the Deliberatorium, a software developed by Mark Klein and associates, which mediates complex collective discussions with a large number of persons involved. The objective is to facilitate deliberation that is more effective (Fujita et al., 2017). In one experiment, 220 masters students discussed biofuels in Italy over a period of three weeks. During that period, the students posted over three thousand ideas and arguments and 1,900 comments into one single argument map (Klein, 2012). About 1,800 posts were eventually certified, 70 percent without any changes. It demonstrated that most authors were able to create properly structured posts. This community of nonexperts were able to create a comprehensive map of the current debate on biofuels, with references to technology and policy issues to environmental, economic and sociopolitical impacts. Klein (2012) compares the collective work with gathering 200 persons to write a book together on a complex subject over a period of a couple weeks where no one is in charge.

Another argument map is DebateGraph. This tool also supports complex policy topics in different fields like education, health, conflict resolution and policy dialogue (Bullen & Price, 2015). Participants explore problems together by first breaking down the subject under discussion into discrete ideas. These ideas are displayed as thought boxes, and can be enriched with videos, images, charts, tables, documents, as well as being cross-connected to other relevant maps. Arrows and colors signal different types of relationships between the ideas in the map. In addition, both the ideas and the relationships between them are visualized in the map structure. This makes it easier to explore and get an overview of clusters of interrelated ideas. When the understanding of a topic evolves, the participants revise both the map and the interrelationship between the ideas. All members can add new ideas and information, or edit and rate existing ideas (Bullen & Price, 2015).

In a deliberative process, there are at least five advantages with using argument maps. First, the map can provide a very good overview of all the arguments in a discussion. If it is well organized, the argument will appear at only one place in a coherent map system (Klein, 2017). If we assume that ideas have a Gaussian distribution, widely known points will be submitted frequently from multiple sources, and the valuable "out-of-the-box" arguments will be far less common. Consequently, the number of ideas will grow much more slowly when the number of participants increase. The goal is to avoid some of the redundancy problems that large groups face in online discussion fora (Klein, 2012).

Second, when all the content is co-located in a hierarchical tree structure, it will be easier to identify what has and has not already been said. It becomes easier to work towards a more complete coverage when everyone has a better overview of the discussion. Argument mapping increases users' chances of "finding their tribe" or other person who have the same

https://doi.org/10.1017/9781108981361.003 Published online by Cambridge University Press

interests. In comparison with an online discussion forum, the benefits of contributing to an argument map will increase as the community scales up in size. It is much easier to place your own contribution and identify other relevant contributions in a tree structure. You only have to pick the correct top-level branch, and the right subbranch, until you reach the place where your argument belongs. This does not require a lot of extra work, and the overall costs of participation are therefore relatively low even when the size of the community scales up. In comparison, in unstructured online discussions, the high volume and redundancy decreases the likelihood of actually finding other relevant posts (Klein, 2012, 2017).

Third, every argument becomes more valuable when being part of a wider argumentative context. Participants can freely choose to engage with one particular aspect of the map or the totality of it. Before making a new contribution, it is also necessary to read existing views and opinions in the map. The process of placing an argument in the map will automatically enhance the participant's understanding of the topic. Instead of just adding to free-flowing online discussions, individuals will ideally be exposed to all parts of the logical structure of the argument: What decisions must be made? What are the arguments for and against each option? Critical thinking is stimulated in the process of making the map (Bullen & Price, 2015; Klein, 2017).

Fourth, idea sharing and equal participation is important in order to avoid extreme opinions. The map offers a greater diversity of ideas by letting every voice be heard. Compared with discussion in an offline setting, a much larger number of participants can be involved. The tree structure might also reduce balkanization by visualizing all competing arguments right next to each other. It offers a more intuitive access to the complexity of an issue, and aims to challenge both readers and contributors to overcome the constraints of groupthink and homophily (the tendency for people to associate with others who share the same beliefs) (Bullen & Price, 2015; Klein, 2017). In many other online discussions, it is also a problem that some people intentionally ignore others and try to "win" a discussion by repeating the same arguments many times. Consequently, potentially promising ideas from smaller groups or less vocal individuals will easily get lost. These individuals may feel overlooked and reject the final decision. In contrast, the argument map can more easily integrating all positions in a debate (Klein, 2017).

Fifth, the quality of the arguments may improve. If many persons can provide multiple independent verifications, this will, according to the many wrongs principle, reduce the number of errors or cancel out the bias (Klein, 2012). The large group size will also increase the diversity of perspectives. Some participants may be better at proposing ideas; others will be good at finding practical solutions. Some may be more critical and better at finding counterarguments. The sharing of all these ideas in the same map environment can also potentially stimulate synergistic solutions that build on combinations of existing ideas (Klein, 2012, 2017).

Traditional online discussions seldom elicit such win–win solutions that maximize the collective outcome for all participants. They often only elicit solo ideas or "dream choices" of individual participants, and seldom provide support or incentive for members to work together to collaboratively develop new ideas. Participants tend to push their own ideas rather than collaboratively try to find new ideas that might give both parties most of what they want. Collective decision-making typically follows a zero-sum frame where competing cliques will stick to their original solutions. A collective solution will be decided either by voting or through a bargaining process where both parties make concessions. While negotiations where parties meet in the middle can produce optimal agreements for simple decisions (i.e. with a few independent issues), this is not the case for complex decisions which often involve many interdependent issues (Fujita et al., 2017; Klein, 2017). Although argument maps are not mainstream, they represent a promising new way of enhancing political deliberation in large groups.

3.4 Summary

The examples in this chapter illustrate the growth in open online knowledge sharing. A major trend is the enormous increase in *complete knowledge products* of various size and formats. Both open access research and open textbooks show how scientific knowledge products are more available today. In addition, practical knowledge products are shared at an unprecedented scale, particularly "know-how" videos on open platforms (e.g. YouTube). These amateur-produced instructional videos obviously vary a lot in quality, but represent a new type of knowledge product that centers on passionate contributions from enthusiasts. Videos represent an important knowledge format that can inform and educate viewers in new ways because of the level of detail in the content. On the one hand, some of these products like online videos and open access research papers will typically be reused but remain unchanged. On the other hand, content modification has become much easier with Creative Commons licenses. One example is open textbooks that make it possible to produce new versions adapted to local context.

Another major trend is that *knowledge construction processes* have become more available and transparent in the online setting. Within the scientific knowledge domain, this includes open scientific discussion (e.g. Polymath – scientific knowledge production) and encyclopedic knowledge production processes (Wikipedia). Both in Wikipedia and the Polymath Project, people do not need to be formal experts, demonstrating that scientific knowledge production today is not only restricted to professional researchers. In addition, a range of new, open digital databases allow anyone to both make their own contributions and get free access to all the data. Volunteers or informal experts are invited to make important contributions in different citizen science projects. Argument maps also make it easier for a large group to participate in political discussions.

Although the knowledge construction processes are different, they show how individual contributions are part of a larger collective work, whether it is a database, a Wikipedia article or a comment in an argument map. For example, in the eBird project, volunteers collect and upload data from many different areas, which provides a much larger value on an aggregated level. In a collective argument map, new contributions will add to existing contributions, and the complete argument map will provide an overview of the collective knowledge. However, with the exception of Wikipedia, most advanced collective writing projects have failed. One example is the Galaxy Zoo Quench project, which challenged a large group of amateurs to write a scientific paper. These failures are important in understanding the limitations of amateur contributions.

Both knowledge products and knowledge construction process can be regarded as important parts of the memory dimension in collective intelligence. Most knowledge products provide long-term sharing in an online setting (e.g. research databases or YouTube). Therefore, the target group of the knowledge sharing can both be universal and directed towards a specific local context at the same time. For example, a published video can target one specific local community or area, but the information may also be relevant for others in another context at a later point in time. When knowledge is shared more rapidly, whether as corporate or scientific knowledge, this amplifies collective knowledge advancement in the society as a whole.

Furthermore, this new openness illustrates the value of transparency. In large-scale deliberation, this transparency gives the group the opportunity to make choices that are more informed. Knowledge is not only reused but can easily be improved by new contributors. For example, in Wikipedia, it is common to translate and adjust articles to many different language versions on the same topic.

Many of these new knowledge products, including both unimodal and multimodal formats, build on what some label as a peer production model (Benkler, 2006; Benkler et al., 2015). This production model, building on CI, involves open creation and knowledge sharing in an online setting. Groups will work in a decentralized manner, set goals together and typically have nonmonetary motivations. Knowledge products are typically common property and build around participatory, meritocratic and charismatic organizational models of governance. It is arguably the most significant organizational innovation that has emerged from the Internet, being an alternative to competition models in more traditional, market- and firm-based approaches. The peer production model is also different from crowdsourcing, which to a larger degree is built around centralized control and external predefined formulation of problems (Benkler et al., 2015). These issues will be further analyzed in the forthcoming chapters (see particularly Chapter 4).