From open source to open content: Organization, licensing and decision processes in open cultural production

Giorgos Cheliotis

Keywords: User-generated content, Free culture, Decision process, Licensing, Creative commons, Social media

1. Introduction

The development of software in an open and collaborative manner has received a great deal of attention lately, especially after it became widely known through a great number of publications in the press and online campaigning that some open source software (OSS) tools are increasingly used for mission-critical business tasks and for scientific inquiry, sometimes taking the place of well-known proprietary alternatives [17]. This was unexpected, as OSS projects depend largely on the loosely coordinated efforts of volunteers, a model of production that, until recently, was not seen as a possible competitor to the highly organized teams of engineers and developers that would usually be in charge of commercial software projects [31]. Thanks to the success of OSS there is a growing interest in both academic and industry circles in open innovation in general, i.e. the process whereby innovations of any kind are shared and jointly developed by more than one person. At the center of such activities is the user who is not any more constrained to the role of the consumer of a product, but is able to actively participate in the innovation process. This is why we talk about the democratization of innovation [29]. This process of democratization is not unique to OSS, but finds its most ardent proponents and most enthusiastic practitioners within the OSS community of developers. Note that in this article we will often use the term OSS to refer to both free/libre and open-source software (I will also briefly discuss the origins of this distinction).

Following up on the democratization of software development, in recent years we are also witnessing the democratization of another form of production, that of goods of cultural or entertainment value, as exemplified by the hugely successful Wikipedia, Flickr, YouTube and many more open communities, which aim at the production and communication of media content rather than being solely focused on social networking (like Myspace or Facebook). Similarly to OSS, such communities are driven by user contributions and by the increasing ability of artists, whether professionals or amateurs, to communicate directly with their audience online, bypassing traditional marketing and distribution channels [8].

Building on the growing body of literature on OSS, the literature on the management and economics of intellectual property, and some of the first empirical evidence on licensing, sharing and reuse behavior in communities producing media content, this article aims at distilling the main features that differentiate the production of cultural goods from that of functional goods. To this end we compare and contrast the objectives of open software development vis-à-vis the development of new media content, the policies and organizational forms that have emerged in respective online communities, and the role that licensing plays in the production of ‘functional’ vis-à-vis ‘cultural’ goods.

© 2009 Elsevier B.V. All rights reserved.
2. Drivers of supply and demand

The success of specific OSS projects has led to wide recognition that the open availability of code and the ability of users with various skills and backgrounds to contribute to that code can lead to better software. This is because both skills and knowledge of needs are distributed in the population of users [28]. Similarly, the democratization of cultural production is driven by two complementing factors: the distribution of skills in the population and the distribution of tastes.

2.1. Distribution of skills

It is impossible to imagine how OSS would have succeeded without the long-distance large-scale collaboration enabled by the Internet and the formation of online communities. There is no single user, or indeed, professional developer, with full knowledge of all the programming techniques, API’s, software interfaces and programming languages that would be required to build modern, complete, complex and robust software solutions. To address this issue companies involved in the development of such complex software products hire and organize very large numbers of developers in distributed global teams, so as to break down the development effort into manageable tasks. Yet, even with the acquisition of formidable engineering and software design talent, such companies will sometimes see flagship products competing head to head with high quality OSS alternatives. The situation is exacerbated for commercial vendors by the fact that OSS is available free of charge and is being continuously improved by teams of volunteers, some of whom are very capable software developers. Also, through the open availability of source code these developers are able to learn from each other and thus continuously improve their skills. It is also likely that this openness contributes to a spirit of constructive competition among OSS developers who take pride in making contributions that are adopted by the community and made available for all to see.

Similarly, creative talent is not uniformly distributed in the population, nor is it concentrated in only a small number of individuals. The same is generally held to be true for original ideas. Certain people are more ‘gifted’ than most, in the sense that they can achieve more with less effort, as well as reach a higher creative ground than most with the same amount of effort. Nevertheless, the surprising popularity of user-generated videos featuring practical jokes, amateur karaoke performances and quirky satire is testament to the power of amateur creativity and the sheer enthusiasm that is driving the production of user-generated content. Apparently, there is significant demand for content that some of us would have otherwise deemed worthless. As the Internet is proving to be the most efficient means yet for matching supply and demand, an ever increasing number of users utilize social networking sites and media-specific communities to find an audience, as well as to seek peers with skills that may complement one’s own talents and know-how.

Websites like YouTube, Revver or Friendster most likely do not need to be introduced here, but a good example of this trend of utilizing the Web to match creators, is the Kompoz.com community. In Kompoz, users upload individual tracks, e.g., a guitar track that they have recorded, and seek other musicians who may be interested in adding more tracks, such as a bass, drum or vocal track to the mix. Similarly to the spirit of constructive competition in OSS communities, some users will attempt to produce better versions of existing tracks, thus helping bring an idea to fruition, and a rough original sketch to a complete musical mix. In this manner Internet users cannot only construct a virtual surrogate for the perennial rock band, but can also potentially enlist hundreds of virtual band members who will contribute different parts of the music to the best of their abilities. Several more music-related communities, like EM411, last.fm, MP3.com or Soundclick, and online independent labels, like CD Baby, Magnature or Jamendo, have emerged over the years, with some communities putting greater emphasis on facilitating the creative reuse of content, in a manner similar to Kompoz: Splice, ccMixer, Yourspins, Jamglue and Acid Planet are but a few examples of online platforms aimed primarily at the facilitation of remixing and online musical collaboration. Remixing, or more generally, the practice of the creative reuse of elements of previous works, is a widespread practice in the arts (sometimes referred to as collage), in science and in technology [11], and is certainly not the domain of only these few online communities, but these online initiatives do present us with a unique opportunity to study the process of creative reuse in near real-time and at a much finer scale than ever before, due to the unprecedented availability of content and metadata on the content and the authors.

However, it is not only the production process and volume that have changed. The move from traditional broadcast (TV, radio) and physically distributed media (CD, DVD) to online distribution has also changed the nature of the medium itself. Actors who may never achieve fame on the big screen do achieve the online equivalent of this fame through short video episodes which are virally distributed through word-of-mouth on online forums, blogs and social bookmarking or user-submitted news sites like Digg.com. The shift from full-length feature films and professional camera equipment to short low-resolution video clips, webcams and mobile phone cameras has allowed some talented individuals to discover that they are capable of producing very attractive, entertaining and original short movies which would not have reached their intended audience before the age of Web 2.0. In other words, the opening up of cultural production through UGC has generated demand for new ‘skills’; or, more accurately, has given Internet users the opportunity to showcase skills which would not have been suitable or would not have been deemed attractive enough for the box office and other more established forms of cultural consumption. This is not unlike the satisfaction that some OSS developers must have felt when they first realized that they too could contribute to a large and important software project from the comfort of their homes.

We are only now beginning to understand what perhaps talent agencies, publishers and record labels have long known by virtue of experience: success in the creative industries is as much a matter of talent as it is a matter of luck and of achieving a critical threshold of viewership [23]. It follows again that those among us who are not successful writers, musicians or actors may nonetheless have the capacity to produce something that is of value to a particular audience. The problem often has been to find that audience that will value the work highly enough to help in its promotion, as well as potentially support the work financially, either directly, e.g., through donations and micro-funding, or indirectly, through the generation of advertising revenue.

2.2. Distribution of needs and tastes

As mentioned earlier, it is not just the accumulation of various skills within OSS communities that is essential for their success. Knowledge of user needs is also highly distributed. Though it may be true that many users will have similar expectations with respect to a particular software tool, many more will not be fully satisfied by the features that a commercial vendor may choose to offer in a specific version of the software. Extensive market research and user feedback can assist a vendor in producing an appealing product mix, with a basic version that can satisfy a broad base of consumers and with additional features for power users, but the software will generally not satisfy, at least not fully, ‘niche’ users. In fact, vendors may not even have full knowledge of the diverse needs of these users, and even if they did, it would be difficult or simply not cost-effective to satisfy these needs. Mass customization, programmable environments and user toolkits may alleviate some of these issues, but will shift the
burden of implementation, to various extents, from the vendor to the user. In cases where open-source alternatives provide similar functionality and often greater extensibility at zero cost, it may be hard to justify the cost of purchasing the commercial product, if it does not readily satisfy a sufficiently large share of a user’s needs off the shelf.

Does open cultural production differ in this respect? We have discussed how user expectations of the functionality in a particular software tool may differ, yet it is reasonable to assume that at least at the most basic functional level all users will have a certain minimum set of expectations in common, with respect to what a particular type of software should provide. This is the most likely reason why there are only a handful of internet browsers or a handful of word processors in the market, even if the acts of browsing and writing are extremely common. Let us try to imagine for a second the surely overwhelming number of documents produced every day around the world, for the management and reporting of daily tasks, or perhaps documents expressing diverse personal opinions or political views, or attempting to capture the subtle nuances of the different cultural and ethnic backgrounds of the authors, arguing for or against known positions on a variety of topics or providing different accounts on historical events, or on aspects of everyday life in various countries around the world: all that, and more, can be produced using one software product, and most often these days this will be a version of Microsoft’s Word software. It is certainly true that producing another browser or word processor is not a trivial task, yet given the very wide audience for such tools, it would be feasible to produce more than a handful of versions, if there were sufficient demand for tools performing essentially the same function.

This, however, is not the case. In fact some of the best known OSS projects in this space, e.g., Open Office and Firefox, aim (the former more than the latter) at providing a free alternative to established proprietary tools rather than attempting to re-imagine and (to borrow a term more common in the arts) ‘reinterpret’ the tool. Interestingly enough, this does not always hold true for more specialized applications with a smaller user base. For example, Wikipedia lists 90 different open-source and commercial Content Management Systems (CMS), while INFORMS lists 52 tools for statistical analysis and 91 optimization tools at the time of this writing. Such tools are not less complex when compared to web browsers or word processors, so we will assume that the difference lies primarily not in the inability of the marketplace to produce more browsers, but rather in the fact that users of more specialized tools share a much smaller set of common expectations and needs. Such users will generate demand for a larger number of solutions, which may share some features, but will often target different ‘niche’ market segments.

The situation as we reviewed it for functional information goods, i.e. software, is in sharp contrast to what we observe in the creative industries. Recorded music for example may serve many subtly different purposes, but it is primarily consumed in the same ways by millions of people around the world, in the form of CD’s, broadcasts, or downloadable MP3’s, and is therefore (and perhaps also because of the lack of better models) usually treated as just one of a handful of information goods in the economics and Information Systems (IS) literature. Yet, while the iTunes Music Store lists only 29 music genres, Wikipedia lists an impressive 1568 music styles (including many traditional and ethnic music styles) and the “How Much Information?” project at UC Berkeley estimates total world production of original audio CDs for 2002 to be around 90,000 units [16]. The same report states that an estimated 10,432 original film and video productions were released worldwide in 2001, citing IMDB as the source. Although these figures are slightly outdated and only rough estimates, they are indicative of the annual number of professional productions in the entertainment industry. Each of these units is an individual cultural product vying for consumer attention among many close substitutes (in the sense of offering functionally equivalent, interchangeable options for the satisfaction of the same needs, e.g., the need to be entertained, the desire to watch a movie, etc.). This does not even include user-generated content, which would be harder to classify exhaustively in the manner of the aforementioned studies, but whose sheer volume and diversity serves to amplify the distinction between functional and cultural goods, at least in terms of volume and diversity of output.

Why this abundance of close substitutes in the creative industries? New and cheaper technology has certainly made it easier to produce media content, but the same can be said of producing new software functionality. The explanation then must lie, at least in part, in the differences of taste in the population, which generate demand (in the form of user-creation or in the form of simple consumption of the products of others) for many different musical genres and sub-genres, as well as for different artists. Similarly to software vendors, publishers and (especially brick-and-mortar) distributors of cultural goods would typically focus on products that would target mainstream consumers, however with the advent of e-commerce the full extent of the long tails of both supply and demand distributions for information goods has been revealed [1] and it appears to be much larger for cultural goods than for functional goods. The last statement needs to be qualified with the caveat that there is still a lack of conclusive empirical studies comparing long tails for these types of goods and there are more differences in demand-side externalities and consumer utility between cultural and functional goods which could account for some of the observed difference in the quantity and diversity of supply.

2.3. Network effects

Demand for information goods is known to be subject to network effects, a form of externality whereby the utility of a good for a user depends on the utility that other users derive from the same good. This has been well documented for software and other functional goods [27], but there is less agreement with respect to the strength of such demand-side network effects in the consumption of cultural goods. The disagreement stems in part from the use of claims for the existence of such network effects in debates surrounding digital piracy. If there are strong demand-side network effects in music for example, then piracy of digital music may serve to generate more demand for legally purchased music through word-of-mouth and through the use of music in social settings [20]. But critics argue that there is no substantial empirical evidence in support of the hypothesis that such network effects, if they exist, can be significant enough to offset the negative effects of piracy, and sometimes they see in such arguments a weak attempt to present piracy in a more positive light [15]. Yet it is common experience for most of us and the experiments in [23] also confirm that word of mouth is important in driving demand for cultural products, so we can assume that such effects do exist also for cultural goods, but their intensity may be lower, when compared to similar effects for functional goods.

This is a point that requires more attention and is somewhat counterintuitive from a historical perspective: historically, functional goods were seen as possessing much less of a social aspect in their consumption, compared to cultural goods, as the latter would often be ‘consumed’ within social networks such as a circle of friends, or a group of fans of an artist [11]. Magnetic and then digital recording technology changed this for cultural goods, which can now be consumed in complete isolation, even if social networks still matter. As for functional goods, it may well be true that network effects are weak for goods which do not depend on the existence of a network for their continued consumption and enjoyment, however with more consumption being directed towards technologically advanced goods with complex designs, network access and higher cognitive demands for their consumption and maintenance, network effects for functional goods have become stronger. Network effects for cultural goods, when they
exist, are more indirect and have more to do with enhancing the experience, rather than being essential for consumption. For example, music or a movie can be a social experience or consumed in isolation (and with digital distribution, can also be purchased without the need for a distribution network), whereas a user of an instant messaging client can only communicate with users of the same or a compatible client. Weaker network effects mean demand (and supply) can spread over a wider range of products (winner does not take all), so this may account for the greater diversity of output in cultural production.

2.4. Transient utility

Another contributing factor for the production of many close substitutes in the cultural industries is the transient nature of consumer utility for cultural goods. This is shown schematically in Fig. 1. After an initial learning period, user utility is maximized for new software and will be maintained at a high level for a relatively long time, until new technological developments and competition start to negatively impact this utility (which nevertheless may take a very long time to reach zero). In contrast, cultural goods can be consumed immediately after purchase, and thus the full utility will be realized quickly, but their value for the average consumer will also drop relatively quickly after consumption. The durability of such goods depends on the particular medium, and will be generally lower for movies, as opposed to books or music (the former because consumption of books is stretched over a longer period of time, and the latter because of music’s replay value), but in any case will be much lower than the durability of software, because it does not entail a strong functional component which would have lasting value for the average consumer long after the thrill of the first encounter.

The above means that consumers are much more likely to purchase more than one close substitute of a cultural good over the course of a year, rather than switch to new software (irrespective of whether this software is offered free of charge or not). This generates strong demand for both a large number of mainstream productions with similar characteristics, as well as for a substantially large (and, thanks to the efficiencies of online distribution, profitable) long tail of niche cultural products. Consumer choice and diversity are therefore even more important for cultural goods than they are for functional goods, such as software. This has implications for the organization of production in these areas, especially in the context of online communities, where there is more room for experimentation.

3. The organization of production

As discussed, we generally do not consume or appreciate functional and cultural goods in the same manner. Society at large will derive little if any value from the existence of several variations of the same functional component of an innovation, whereas the existence of several interpretations of a popular song, or of multiple perspectives on a topic of interest is generally desired (within some limits again). In other words, although both open innovation and open culture are the products of collaboration and greater choice can benefit consumers of both types of products, pluralism is generally more desirable in the latter. Moreover, the nature of the production process is qualitatively different. To produce high quality tools we need developer communities sharing a common vision of the functionality to be embedded in these tools and an agreement on how individual components will work together. In cultural production, it is exactly the differences in the creative visions of the producers that matter.

We may take this a step further by saying that open innovation is about facilitating know-how sharing and component reuse whereas open culture is to a large extent about facilitating serendipity [26]. Serendipity here refers to the fortunate accidental discovery of a new theme, a new melody or sound effect that can result from willful experimentation or from the unintended and incorrect setting of a parameter in a digital effects unit, a new painting technique that may be accidentally discovered when the painter will splash paint over the canvas in a moment of anger or simply by mistake, a new subject for a video clip which may seem unlikely at the time of production but may become a runaway hit on the Internet, an old and forgotten musical theme that may trigger something in the mind of a DJ who may then remix it to great success, and so on. Serendipity also refers to the nature of improvisation that is so common today in the performing arts and is also taken up by the ‘YouTube generation’ with thousands of mostly improvised videos posted online, often capturing unlikely moments of everyday life, such as a baby that is dancing to the music of the artist Prince (an actual case and the subject of a recent lawsuit on the part of Universal that owns the copyright to the song). As one new media artist put it in a personal communication: “engineers are interested in controlling for accidents, whereas in art we are interested in the accidents”. It follows that online communities focusing on cultural rather than software production will need to operate in a different manner, even if all such online communities have some characteristics in common. We will now examine some of the differences and similarities, with a focus on the organization of the means of production.

3.1. Branching, merging and forking

In [5] a framework is developed based on Coase’s theory of the firm to explain how open-source-like models of organizing production can occupy a space between markets and firms. In this article we dive deeper into the characteristics of what that work calls “commons-based peer production”, to uncover the finer differences between the production of functional and that of cultural information goods. In any sizeable software project, whether open-source or not, the coordination and optimal utilization of multiple resources are keys to success. It is thus common practice to branch some of the source code from a particular version of the software, so that it can be developed in parallel, and to some extent independently of the core of the software, with the intention to merge this code again at an appropriate point in time, in order to produce a new complete version of the software. This is shown graphically in Fig. 2.

Since OSS projects make the entire source code publicly available, it is possible for someone to take a version of the software and fork it
into an independent project. Forking is in fact an essential characteristic of OSS. Forking may happen because of strong differences of opinion within an OSS project, because of competition among developers, or because some developers may simply wish to produce a new derivative product which will offer new functionality tailored to particular needs, albeit building on some of the same components as the original project. It is still possible for the two projects to coexist and interact, as long as the contributions of the new project are licensed in a manner that is compatible with the license of the original project. Copyleft licenses like the GNU GPL are designed with the prevention of forking under a different license in mind, so as to ensure that valuable contributions can always be exchanged between projects. Forking is often feared or frowned upon by the developer community, as it can result in duplication of effort (which, as discussed, is generally wasteful for a project aiming at producing functional components) and to (sometimes bitter) rivalry among developers. However, arguments have also been made in support of forking, as it can promote competition and release developers from the drudge of slow consensus-driven processes and occasional forking, as it can promote competition and release developers from the drudge of slow consensus-driven processes and occasional stalemates in some OSS projects. Fig. 3 provides a simple illustration of the forking process leading to independent parallel projects.

In order to compare and contrast these basic models of parallel production to the open production of cultural products, which is often not organized in well-structured collaborative projects, we produce a more detailed view of the branching process in Fig. 4, showing the interactions between the main project and specific user contributions/modules. Fig. 4 is based on one of several possible branching patterns [2], where the project maintains a stable mainline, or trunk of source code which is continuously improved upon by a core team of developers, while individual components (e.g., A, B and later C) are contributed by developers and some are branched out again, to be developed in parallel (e.g., B), while others (e.g., A) are maintained in the trunk. This, and similar models used in OSS development, are the result of the need for coordination among developers, so as to ensure that all software components can be adequately tested and will work well in tandem when a new version of the software is released. Some have argued that in spite of this need, the organization of OSS development resembles more that of a bazaar rather than a cathedral [22], although the validity of the bazaar analogy has been questioned by other commentators [6].

The need for coordination is also present, as we will see, in cultural production, although it is not necessary and it may even be desirable to promote the uncoordinated production of cultural goods in some cases. This may at first seem counterintuitive, but will become clearer in the following section.

3.2. Challenges in organizing open cultural production

Open software production is facilitated by the following factors:

1. the public release of the executables and their wide availability free of charge
2. the release of the source code in an open format
3. open access to the code and documentation
4. the wide availability of open and free tools to manipulate the code
5. the use of software licenses which will ensure that this ‘openness’ is preserved

We will now examine to what extent the production of cultural goods emulates that of OSS. We will focus on audio material, but the discussion that ensues has parallels in all media types. With respect to the first two factors, the situation with digital media content is somewhat more complicated, as there exist multiple competing formats for each media type (text, audio, photographs, CGI, video), as well as for individual elements of media types (e.g., for document typesetting information, for the description of audio or video filter parameters, etc.). For audio specifically, MP3 and AAC are emerging as the de facto standards in online distribution. They are both adopted as open standards, by ISO/IEC and MPEG respectively, although (unlike with the Ogg Vorbis format) some aspects of the technology are still patented, so developers of codecs (necessary to create or play files in these formats) need to obtain a license from the respective owners. In spite of this relative shortcoming, their standardization and wide support mean that users can share audio content in these formats and be certain that other users will be able to play back the files. Similar formats exist for most other media types.

The above, however, refers only to the distribution of the ‘executable’, i.e. the audio file itself. What if we want to have the ‘source code’ to this executable distributed in an open format? This is clearly not as simple for cultural goods, as it is for software. First let us note that even with software there are aspects that the source code and documentation cannot reveal to the user. These relate to the domain expertise and creative thinking that goes into every line of code. We can read the instruction, but we can only assume why a particular sequence of instructions was chosen over a functionally equivalent alternative, or why a particular object-oriented design was favored over another. The same is true for cultural goods and for music in particular. Even if we had the source code, in terms of the key, tempo, notes, orchestration, reverb and effects data, understanding the music piece in terms of understanding its structure, intent, harmony and other factors, would still be very challenging, and would require a significant amount of domain expertise on our part. Hence, we are not concerned with such aspects when talking about “open-sourced media”. We are rather interested in releasing the instructions encoded in the final mix, from the musical notation to the actual performance information (e.g., as encoded in MIDI velocity data), to the individual instrument tracks comprising the mix, effects parameters, etc.

Unfortunately, for most media other than text (where Open Document, HTML, XML and Postscript can be of great assistance in
ensuring readability, free editing and portability), there are no widely
used open standards for the transfer or manipulation of all this
information. At best, what creators do can is include metadata
describing the media file, the date of recording, or the equipment used
to produce it (e.g., using ID3 tags in an MP3 audio file, or EXIF-
formatted metadata in the header of a JPEG photograph). As most
online content today is produced digitally and with the help of
software, it is possible to distribute the software project files
together with the rendered final product, and some users are doing just this,
although, as most such software is commercial and using proprietary
file formats, only users of the same software will be able to read and
manipulate the ‘source code’. In communities where specific com-
mercial software solutions are widely used, the inclusion of the
project files is a poor but nevertheless somewhat effective surrogate
for the ideal of totally free and open access.

In spite of the complexities involved in open-sourcing digital
media, as long as a community is open to submissions from any
individual and digital content is published in a standard format
(whether de facto or official standard) that at will at least allow for
widespread consumption and sharing of the content, if users are in
possession of at least some basic tools for the production and
manipulation of content, open cultural production can become a
reality. This is what we are witnessing on popular websites like
YouTube. YouTube is not geared towards online collaboration and
open standards-based production, but it is successful because content
is easy to upload, easy to view using a Macromedia Flash player (a
proprietary but common format), and easy to share among users in
the form of clips embedded in HTML pages. This is the most
rudimentary form of democratized cultural production. It is demo-
cratized in the sense of wide participation by the public (both in
producing and in reviewing the content), but not yet in the sense of
the public having the ability to self-organize and control the means
of production as well as exercise full control over the evolution and
ultimate fate of the products.

Some online communities, like the now famous Wikipedia and the
lesser known Kompoz music community, require some form of
coordination among members for the production of content. In
Wikipedia, the entire production process takes place online and hence
open tools and open formats are required. It is virtually impossible to
imagine Wikipedia without the collaboration and collective editing
software that supports it. In Kompoz on the other hand most of the
production takes place offline and the community provides simple
tools for users to communicate and exchange individual contributions.
Acid Planet is similar in that respect, with the only difference that it is
run by Sony Corporation and is built around a popular software
platform for remixing, Sony’s ACID software, which offers some
degree of integration between the offline software client and the
online community. Jamglue, Splice and Yourspins, which are also
collaborative music production and remixing communities, go a step
further by providing a rudimentary online audio mixer and sequencer
for arranging individual recorded tracks, setting their relative levels,
etc., within the browser. The main difference between these
communities is that Jamglue and Splice are based on the remixing
of user-submitted content, whereas Yourspins is aimed more at the
remixing of music from popular artists, which has implications for the
licensing of the resulting tracks, as we shall see later on when we visit
the issue of licensing in open cultural production. While each
community goes to different lengths in their efforts to open-source
the creative process, from providing just a collaboration platform, all
the way to providing simple editing tools, Kompoz, like Wikipedia, s
utilizes a variant of the branching model we discussed earlier for
software (see Fig. 4). That is because it is project-oriented, with each
project having an owner and requiring a structure that will ease
collaboration among members. Typically the output of any collabora-
tion will also be evaluated or moderated in some form by a
community-appointed editor or by the project owner, i.e. the initiator
and leader of the project (in addition to the moderating effect of user-
submitted reviews and community discussion boards). Splice, Jamglue,
Yourspins and ccMixter (yet another online music community,
focused on remixing) are not project-oriented for the greatest
part and thus exhibit different patterns of organization.

Forking also becomes an issue for open-sourced media. Any
member can take a version of a Wikipedia page or an audio file from
one of the music-oriented communities (assuming that file is easy to
copy or download in its entirety, which is not always the case), and,
barring possible licensing restrictions and authorship issues, start
a new project with the same raw material. Like with OSS development,
forking of media content can both become the source of rivalry and
and discontent and have a positive effect in terms of providing more
opportunities for good ideas and talent to surface. But if forking may
sometimes be acceptable as a practice in OSS, it is potentially of
much higher value in the production of open content, because it is
may be desirable to produce many close substitutes if these can satisfy
different consumer tastes, as well as to let users steer the material of a
project in unexpected creative directions.

This is again because of the arguments presented earlier in the
article about the distribution of tastes in the consumption of cultural
products, the likely weaker network effects in such markets, the
transient consumer utility, and the concept of facilitating serendipity,
which may be sometimes difficult to grasp for the engineer or the
computer scientist but is almost second nature to the artist. I do not
claim here that there is unlimited demand for multiple takes of the
same song (though this is not too far from reality for songs with wide
and lasting appeal). Rather we should think again about the example
given already on word-processing software, where a single product
can be used to author a very large range of articles, books, essays,
short pieces, novels, poems, commentaries, editorials, speeches, and so on. The
same argument could be made with a typewriter in lieu of a word-
processor or even by thinking about all the different musical genres
that can be played on the piano, which can be viewed as a cultural
good itself, but actually also serves a very practical purpose for the
player or composer, that of aiding in the composition and performance
of a great many different musical works. Like the word-processor and
the typewriter, and like its digital siblings which are often imple-
mented in software, it is a tool that can be put to many different uses,
and in fact digital tools tend to be even more ‘general purpose’ than
the equivalent products of earlier periods which were often built for a
more specific purpose and with only limited room for repurposing.
The various uses of these tools are the products of our (artistic,
technological, scientific, political, religious) culture. These uses are
much greater in number and more varied than the tools used to
facilitate them.

3.3. Ad-hoc and other models of cultural production

Given the great variety of artistic (and other creative) visions and
the role of serendipity in cultural production, completely ad-hoc and
unmoderated, decentralized networks of cultural production may be
just as good, if not better-suited for the production of content with
artistic value or aspirations (as opposed to content of primarily
functional or educational value, like Wikipedia’s encyclopedic pages;
though educational content in particular poses some challenges in
terms of our comparison as it is not a tool per se but can be said to have
both functional and cultural utility). One online music community
that is organized around mostly ad-hoc content generation is the
ccMixter community, where users submit complete ‘mixed down’
tracks or individual audio samples at will and other users may use
these as sources to produce new content that may or may not
resemble the originals. This is shown schematically in Fig. 5. In this
model there exists no central authority and no “trunk” from which to
branch out or fork new projects. The Jamglue, Acid Planet and Splice
communities are similar in this respect. Each individual contribution
to the community is an independent micro-project, usually the work of a single creator or, more rarely, of a small team, e.g., a music band. Production takes place entirely offline, or partly online, if simple web-based tools are available, but either way there is no requirement for coordination between members (although the community website is designed such as to ease the search for raw material, and the tracking of reused/remixed material from one contribution to the next).

The natural outcome of such a system is the creation of many close substitutes and also of many works of doubtful value to anyone besides the creator. Such seeming waste of resources might be unacceptable for OSS projects (though it is akin to the bazaar analogy of OSS development), but it could well be a suitable model for the generation of new musical ideas. In an ad-hoc cultural production model each micro-project may be independent, but we can regard the entire community as one large project, whose ultimate goal is not to produce an item of specific purpose, but rather the generation and realization of diverse creative visions, with some common ground in terms of the raw material used and the reliance on open collaboration. This is where open-sourced cultural production distinguishes itself more strongly from the open-sourced production of functional items, such as software. Every remix in the ccMixter community is a ‘fork’ of a micro-project, and is not only desirable; it is in fact the very reason for the existence of the community.

However, even in ad-hoc environments of loose or no coordination among participating members, certain forms of organization or emergent patterns occur. An example of this is what we will here call the relay model of cultural production, whereby each new contributor builds on top of an earlier contribution by someone else, with no prior agreement between the two on the direction of the new contribution. If we take a moment to think about this, we will realize that in fact most of our cultural heritage is based on the partly organized and partly ad-hoc creative reuse of prior art, as creativity always builds upon the past, through the liberal borrowing of elements of earlier works and ideas, without the need for a grand plan dictating an ultimate desirable outcome. This is also true of sequential innovation by multiple parties [7]. When such sequential production of new goods is carried out independently and with no common ‘trunk’ serving as the common repository of sequential versions of a product, we have a process of mere chaining. When the process is coordinated and aiming at a pre-determined final objective, it is more akin to relaying. In online communities this process of relaying can be taken to new extremes, as was done for example by some members of the Soulseek file sharing community, who produced a highly original music album whereby each track is produced by a different musician, but with each track having to build on the previous track in the sequence. Without a central plan apart from the relay sequence, or a prescribed artistic direction, one might expect the album to be ultimately lacking in substance and cohesiveness, yet, surprisingly, this is not the case (though the loose structure of the album guarantees that the listening experience will not be an ordinary one). In a similar vein several online communities have been set up around collaborative online storytelling, often following the same relay model.

Other common structures that can be observed in otherwise uncoordinated cultural production networks are the seeding and synthesis patterns, shown in Fig. 6 alongside the relay and chain patterns. We call ‘seeding’ here the generation of a number of derivatives (i.e. alternate takes, remixes, or reinterpretations) from one original work, while we use the label ‘synthesis’ for the reverse process, whereby a number of independently produced works are compiled into a new cohesive whole. It is useful in the context of this article to avoid naming these ‘branching’ and ‘merging’ respectively, as these terms are used in software production to refer to two stages of the same process, where specific modules are branched out with the intention of being merged back to the trunk of a project. Seeding and synthesis on the other hand do not require coordination and are driven by a different set of needs and conditions.

Remix contests for example, which have become a staple of online music communities, are seeding processes, as they can lead to the very
Compilations and mash-ups produce additional value for the consumer and/or the creator. The rapid generation of hundreds of new works, all using elements of the same source material, but each being a unique work which can be produced by anyone, for any purpose, without the need for prior agreement or consensus between authors (bar licensing constraints, which we will review later in the article). Moreover, ad-hoc seeding and synthesis are again as desirable as forking, relay and chaining, as they can lead to great increases in the volume and diversity of the output of the community. In fact, preliminary findings from our ongoing investigation of the ccMixter online community suggest that remixing (ad-hoc and through contests) can at least double the creative output of a community. As a final note on this matter, it is interesting that the production of functional information goods is lately veering towards a direction that borrows elements from cultural production, in the form of loosely coupled web services and API mashups. An examination of this more recent trend in software development would further complicate our analysis, so it is left for future research.

**4. Incentives for participation**

We have examined the general social and economic drivers of supply and demand for open collaborative production, but one of the most common questions researchers, the media and practitioners have been asking with respect to the incentives of individual actors who choose to contribute rather than free-ride is: “Why do they do it?” The question can be divided into two parts:

a) Why would users be contributors instead of spectators/freeloaders?

b) Why would users reveal their work and allow others to build upon it?

To answer the first question one needs to understand why users would be willing to invest time and effort into a venture with no direct financial reward, sometimes just to replicate functionality which may be already available in commercial software or to provide an alternate take to the theme of a commercially available movie. Answering the second question requires that we weigh the benefits that a user will reap from revealing his/her work as opposed to keeping it secret and utilizing it for personal advantage.

A popular assumption in the OSS literature is that users who contribute to the community obtain additional private rewards, while freeloaders do not [30]. This builds partly on the principles of collective action, first formulated by Mancur Olson [19]. The most obvious reward for participation comes in the form of learning and consequently being able to produce works satisfying one’s own and unique needs. It is a common observation in the OSS universe that contributors of new modules are often motivated by their own need for functionality which an OSS project and other available software contributors are not providing. After the innovation is revealed, other similarly talented users, so there would be little benefit in not sharing. If, on the other hand, users choose to reveal their works and allow others to build upon them, it is likely that some will improve on the work, to the benefit of both the original creator and the subsequent adopters (see [7,22] for software). Also, the original creator’s reputation in the community will grow, thus producing positive network externalities for the author or inventor (see [29] for functional goods and [10] for how cultural goods may benefit from the exposure on file-sharing platforms). The only difference here between functional and cultural goods is the fact that functional goods can become revenue-generating without ever being revealed, as they may serve as inputs to production. Cultural products on the other hand, unless specifically designed to be inputs for larger projects (e.g., sound effects that can be later utilized in a movie project), are not of any financial value to the producer unless released to the market. Whether the producer will opt for a closed or a more open licensing model when releasing to the market is another matter altogether and will be looked at in more detail in the next section, when we examine licensing options.

Copyright licensing is a matter that all authors of cultural products need to deal with sooner or later (whereas producers of functional products may choose to ignore patenting if they so wish). Copyright law will automatically grant exclusive exploitation rights to each author, and for a very long period (author’s life plus 70 years in most jurisdictions), so open-sourcing and third-party improvements/contributions to one’s own work will not even be legally possible, unless explicitly allowed by the original author through a special license, and even then such improvements by third parties may not be exploitable by the original author if the parties producing the improvements do not license their contributions in a similar manner as the original author. So, even if licensing is also an important consideration for OSS projects (not the least because software is also copyrightable), it is the sine qua non of open cultural production.

Nevertheless, licensing considerations are not the only or sometimes not even the main factor determining sharing modalities in online cultural production. Many authors are active contributors of communities utilizing many different forms of licensing, whose details or importance they may not even understand fully, because in any case they derive an additional benefit from being actively engaged in the creative process, a benefit which freeloaders will not realize through mere consumption of the products of the authors’ labor. Active contributors to online communities like YouTube, Flickr, ccMixter and others derive great personal pleasure from their contributions and the praise that these may receive. Such social ‘hedonic’ rewards are a form of private reward that is discussed in [5] and they are a strong factor of intrinsic motivation that is also a driving factor in OSS development, although it has not received much attention in the IS literature (with some exceptions, e.g., in [3]). Other likely factors that receive relatively little attention in the IS literature, perhaps because of the inherent difficulties in assessing them, are a user’s ideological motivations (a strong motivating factor behind some OSS developers [18]) and altruism (potentially also a significant factor). Users may elect to freely share their innovations or creative work, even when in full knowledge that they will not receive
any form of compensation in return (financial or otherwise). A suitable licensing agreement of course can help create the conditions under which non-financial private, altruistic or ideological motivations may be translated to monetary rewards, depending also on the market value of the product. This is discussed in [9] for cultural goods and in the context of Creative Commons licensing, which we will review and contrast to OSS licenses in the next section, with a focus on analyzing the decision process of individual producers.

5. The role of licensing models

The nature and extent of the rights of producers of innovations, or of artistic and literary works, are governed primarily by Intellectual Property (IP) law. Producers of innovative products of functional value (as opposed to artistic, educational or entertainment value) may seek to patent their inventions, so as to exclude imitators, for the duration of the patent. Patent offices may grant inventors this privilege of exclusive exploitation of an invention because it is believed that forming such a temporary monopoly for new technology will increase incentives for innovation. The scope of what is ‘patentable’ has been gradually extended so as to include software and business models. The existence of patents in a particular domain is thus an issue for OSS development, but as patents are not directly relevant to open culture (‘cultural goods are not patentable’) we will focus our attention here on copyright law, which affects both.

5.1. Open software licensing

As the objective of OSS proponents is to enable the open collaborative production of software, based on open access to the software as well as to the source code, the development of licenses tailored to this purpose is of paramount importance. Copyright is enacted automatically at the birth of a work and will exclude anyone from making copies or modifications to the work for the duration of the copyright term. Creators wishing to define a different range of exclusive rights can license their work under a public license that redefines the scope of these rights, thus also allowing them to opt out of some of the restrictions of copyright. Software, like all expressions of ideas, is copyrighted at birth.

Several different licenses have been proposed and are currently in use for OSS projects, though perhaps the most influential is the GNU General Public License (GPL), developed and promoted by the Free Software Foundation. The GPL is aimed at ensuring the propagation of the basic 4 freedoms of free software, as articulated by Richard Stallman and promoted by the FSF: the freedom to use the software for any purpose, the freedom to share the software, the freedom to change the software, and the freedom to share changes to the software.

The FSF has long held the view that these freedoms should be propagated over all future generations and derivatives of a free software module. The GPL reflects this view by including a so-called copyleft restriction, i.e. a requirement that all software building on GPL-licensed code must also be licensed under GPL or an equivalent license. This will ensure for example that any project ‘forks’ will also have to license their contributions under GPL, so that, even if developed as independent projects, their contributions will still be free (as defined above), and moreover, contributions from a fork can be incorporated back into the original project at will, when and if that is desirable. So, there is a very real and pragmatic argument in support of such a constraint in the license.

The ideals of freedom embodied by the GPL, their promotion by a respected personality such as Richard Stallman, and the use of GPL in some iconic projects, such as the GNU/Linux distributions, have led to the very wide use of the license by many software projects, as evidenced by the large number of projects listed on the Free Software Directory maintained by the FSF. However, besides these contributing factors we can examine some more practical reasons for the popularity of copyleft licenses, based on our earlier discussions on the organization of production in online communities. Software, including commercial software, unlike most commercial cultural goods, is not static. It is continuously improved, so as to fix ‘bugs’ in the source code, improve computational efficiency, and add new functionality, to address shortcomings of older versions or to adopt to the changing needs of users. The evolvability of software makes it all the more important that future contributions, no matter where they come from, can be incorporated back to the trunk of the project. Duplication of effort is thus not entirely avoidable, and we discussed how it may even be desirable in some cases. In short, licensing does matter and there is some evidence in the IS literature in support of the role that licensing can play in contributing to the success of an OSS project [25].

Some licenses utilized for OSS do not include the copyleft restriction of the GPL, but can still qualify as ‘open source licenses’ according to the definitions of the Open Source Initiative (OSI), which are somewhat more relaxed than those of the FSF [21]. Still, copyleft and the distinction between free and open-source that is often made by supporters of the GPL, are some of the most passionately debated license design and nomenclature issues in the OSS movement [24]. The same issues are very relevant for open cultural production, but with some additional considerations that are specific to cultural goods.

5.2. Open content licensing

The success of OSS has been a great inspiration for the more recent developments in open-sourcing cultural production. It is then not surprising that one of the popular licenses for text documents is the GNU Free Documentation License (GFDL), also designed by the FSF. Many more licenses exist for documents and other media forms, e.g., the Free Art License, the EFF Open Audio License, the OpenMusic licenses, etc. These licenses vary in terms of the extent to which they fully support the freedom ideals embodied in the GPL, the media types they are intended for, and other factors. We will collectively refer to them as Open Content Licenses [14].

In the last couple of years one licensing model is starting to dominate the production of open content: the Creative Commons (CC) licenses. These were first proposed by Lawrence Lessig, as a means of promoting the growth of a free culture [13]. In a departure from FSF’s philosophy, CC provides not one license, but a menu of licenses for authors to choose from. Firstly, authors may choose whether they wish to allow for the creation of derivative works, and if they do, whether they wish to enforce a copyleft constraint on such derivatives. Secondly, authors are able to limit uses of the work to non-commercial only. In short, the main CC licenses are the following:

- **BY** Requires that users of the work give attribution to the author
- **BY-SA** Requires that derivatives be licensed under the same license
- **BY-ND** Forbids the creation of derivatives
- **BY-NC** Same as BY, but permitting only non-commercial use
- **BY-NC-SA** Same as BY-SA, but permitting only non-commercial use
- **BY-NC-ND** Same as BY-ND, but permitting only non-commercial use

In the above list BY stands for “By attribution”, SA for “Share-Alike”, ND for “No Derivatives” and NC for “Non-Commercial”. CC licenses have become extremely popular, and although it is extremely difficult to count the number of items licensed under CC, estimates generated by a combination of web search methods and data extracted from specific online communities known to host a large number of CC-licensed work indicate that the size is in the order of millions of items [9]. In addition, it has been shown that the popularity of CC spans multiple media types, online communities, and geographical regions, making CC licensing a truly global phenomenon.
The rapid and global growth of CC has attracted media attention as well as a fair share of criticism with respect to the design of the licenses. The inclusion of the copyleft ‘SA’ restriction as an option, instead of a requirement, has drawn significant amounts of criticism from supporters of free software. Yet, if we think about our earlier discussion on the differences between the production of functional versus cultural goods, we will see that the copyleft constraint is not as valuable for open culture. We noted earlier in this section that copyleft ensures that future changes to the source code will remain free and thus can be incorporated back into any GPL-licensed software. This is clearly valuable as functionality is backwards-transferable. There is however little value, if any, in transferring parts of a remix of a music piece back to the original work. The same can be argued for any work that is primarily a work of artistic expression, as opposed to works that serve primarily an educational or documentary purpose, although there are other, less pragmatic, ideological and perhaps also social psychological reasons why the reciprocal behavior engendered by Copyleft is also popular in Creative Commons licensing, as we shall see. On the other hand, we have already noted that forking is desirable for cultural content as it promotes diversity, pluralism and serendipitous creation. In fact, the encouragement of forking in open cultural production may help alleviate some of the negative effects of the commercialization and homogenization of information production, wrought by the intensification of IP restrictions in business and society [4].

Copyleft limits forking to some extent by nature of excluding authors who would be willing to produce a derivative work, but do not wish to use the same license, or simply cannot use the same license due to contractual or other obligations which may be imposed by an employer or publisher. Copyleft restrictions are also known to introduce incompatibilities between licenses and make the licensing of some works infeasible, when these works utilize sources with mutually exclusive copyleft restrictions [12]. In cases where copyleft may make sense, e.g., where content is being produced for educational or documentation purposes (typically in a more coordinated manner and based on textual or visual media), authors can select a CC license which incorporates the SA constraint. In these cases the content has a strong functional component and it is therefore more important to ensure that future contributions can be added back to the resulting product (although the issue of mutually exclusive copyleft constraints may still inhibit the inclusion of otherwise valuable source material).

Another point where CC marks a significant departure from the definitions and philosophy of the FSF is the inclusion of the NC constraint as an option. When authors select a license with the NC restriction, only non-commercial copying and derivation are permitted. As the licenses are non-exclusive, commercial uses are still possible, but only after prior negotiation with the author. GPL and all similar free licenses do not make that distinction and indeed it can be challenging to define what exactly constitutes commercial use, yet we notice that NC licenses are actually very popular among users of CC, with about 2/3 of all CC-licensed items including the NC constraint (see [9] and the exposition later in this article). Lawrence Lessig has publicly defended the decision to include an NC option in CC by explaining that copies of cultural goods can have significant off-the-shelf value, whereas OSS will typically require modifications to be incorporated into a commercial solution, and these modifications will always trigger the copyleft restriction. This will force vendors of commercial solutions based on GPL-licensed software to make their products freely available under the same license and this in turn means that the entire free software community can benefit from the functionality introduced by the vendor. However, in the case of cultural goods, a CD distributor may make copies of a CC-licensed musical work as-is and sell these for a profit without giving anything back to the community whence that work came. The original artist would rightfully feel cheated. Yet, as we already mentioned earlier in the article, such distribution may prove very valuable for the author of the work after all, in terms of generating interest in the author’s work, so there is clearly a trade-off involved between maximizing a work’s exposure through more liberal licensing and preserving one’s rights to commercial exploitation.

Another potential issue when not using the NC restriction is the possibility that an advertising or public relations agency will utilize the work in advertising, where it is irrelevant whether the advertisement itself will be licensed under CC or not, as it is only a vehicle for the promotion of a separate commercial product. Arguably the advertisement may help boost the popularity of the work and consequently, the author, but the author may or may not approve of the use of the work for advertising purposes and will have no recourse against such use (at least not within the framework of copyright law), unless the license includes an NC constraint (or otherwise a specific ban on use in advertising). This issue is unique to cultural goods and there is no direct equivalent in OSS. Hence, an argument can indeed be made for the distinction that CC licenses make between commercial and non-commercial use. Finally, to complete our exploration of the similarities and differences between open-source software and open-sourced media, we need to understand the scale and scope of the aims, impact and intended uses of the licenses.

5.3. Scale and scope

It is not only the nature of the products of cultural versus functional/software production that distinguishes the two. It is also a matter of scale and scope. The Free/Libre/OSS movement has been very prolific and has inspired many initiatives beyond software, including Creative Commons, yet the scope of the core licenses is limited to software and the licenses are developed for and by software developers who believe in the value of collaborative, user-driven development. This is a specific segment (OSS developers) of a very small subset (software developers) of the population. Open culture on the other hand concerns all of us, as it has implications for the future of artistic expression, the creative industries, education, and the political process, as most public speech is captured and propagated in some digital media format. Also, the licensing of cultural production concerns the authors of multiple media types, including writers, journalists, academics, photographers, musicians, filmmakers, painters, new media artists, and many others.

Therefore, even if open culture largely builds on the precedent set by open source software and is equally driven and enabled by advancements in information technology, its aims, potential impact and intended audience are much broader. The implication is that CC licenses need to accommodate a wider spectrum of needs and their designers need to be sensitive to a wider range of viewpoints with respect to what constitutes freedom or openness. Returning to the discussion on the use of the NC restriction, some authors may feel that the reputation gains they can realize through the promotion of their work on commercial distribution channels are sufficient to justify using a license without the NC restriction. Others however may not be willing to make that concession, even if they are otherwise supportive of free sharing of cultural goods. In addition, for many artists, especially musicians and photographers, licensing of works for advertising or illustration purposes is a major source of income, so even if some of these artists would not object to the commercial distribution of some of their works in verbatim form, they might be unwilling to allow for the free creation of commercial derivative works.

It would be difficult to satisfy such a diverse set of users with just one license and one policy. The pragmatic Creative Commons approach of providing options for authors (as opposed to the FSF approach of dictating a specific stance based on not only pragmatic, but also deontological arguments [18]) has thus encouraged a very wide and diverse set of authors to embrace CC and engage in the formation of a free culture, even if the needs, aspirations, and understanding of this vision of a free culture may differ from author to author. Some of the data that I and my collaborators in the CC-Monitor
project have been able to collect in the last years lay testament to this diversity, as it is reflected in the licensing decisions made by CC users.

The last section of this article reviews some of the main findings of this work with an eye on the lessons that can be learned from this empirical study regarding the nature of cultural production and the decision processes of individual authors engaging in the opensourcing of digital content.

6. Lessons from CC usage patterns

The decisions that an author considering CC as a licensing model needs to make can be organized as shown in Table 1. All CC licenses allow for non-commercial copying and distribution, so an affirmative answer to D1 is the first requirement for an author to consider CC as a licensing model. Then, as CC licenses are naturally divided by the NC restriction into two otherwise identical groups of options, the first group being [BY, BY-SA, BY-ND] and the second being [BY-NC, BY-NC-SA, BY-NC-ND], the next most important decision for a rational author to make regards the inclusion of the NC restriction (D2). After making this decision, the author needs to consider whether or not to permit the creation of derivative works (D3), and if so, whether or not to enforce the copyleft restriction (D4). An individual author may not make decisions in this sequence, but the definition of a logical and ordered decision sequence will be helpful for analysis, without the loss of generality.

Collecting data on the use of Creative Commons licenses is not easy and any such data is bound to be incomplete, as the numbers of CC-licensed items worldwide are in the millions and there exists no central repository or registry of CC-licensed works. The only means and any such data is bound to be incomplete, as the numbers of CC-licensed items worldwide are in the millions and there exists no central repository or registry of CC-licensed works. The only means

<table>
<thead>
<tr>
<th>Decision</th>
<th>Description from author’s perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>Do you permit non-commercial copying and distribution of your work?</td>
</tr>
<tr>
<td>D2</td>
<td>Do you permit commercial uses of your work?</td>
</tr>
<tr>
<td>D3</td>
<td>Do you permit the creation of derivative works?</td>
</tr>
<tr>
<td>D4</td>
<td>Do you want derivative works to be licensed in the same way as your work?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Results of Web Search methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>YBL</td>
</tr>
<tr>
<td>GBL</td>
</tr>
<tr>
<td>YCC</td>
</tr>
<tr>
<td>D-YCC</td>
</tr>
<tr>
<td>(St. D.)</td>
</tr>
<tr>
<td>D-GCC</td>
</tr>
<tr>
<td>(St. D.)</td>
</tr>
</tbody>
</table>

Based on our definition of D1–D4, we can construct a decision tree to model author decisions. Then, based on the data we have collected on CC usage we can infer the probabilities of each outcome in the decision tree (by using the relative popularity of each license as a rough estimate of the probability of occurrence of this license and constructing the user decision process as a sequence of individual decisions relating to each of the defining properties of CC licenses: NC, ND, SA). Table 2 shows the results we obtained with several measurements in the period Feb–May 2007: a set of Yahoo backlink counts (YBL), a set of Google backlink counts (GBL), a metadata-based Yahoo CC-Search (YCC) with no search terms, as well as a set of Yahoo and Google CC-Searches with 1,000 randomly selected English words as search terms (D-YCC and G-YCC respectively, with the mean and standard deviation $\sigma$ displayed in consecutive rows). Arriving at a total per license using backlink counts amounts to the execution of a query on the index for each version of the license (multiple incremental and jurisdiction-specific versions exist per license, e.g., BY-NC-ND – Spain – version 2.5). Also note that CC-Search methods do not allow for the filtering of results according to the SA restriction, so these methods can produce only 4 groups of licenses instead of 6.

As expected, not all data collection methods returned the exact same results (due to the differences in the methods, as discussed in the previous section). Based on the range of values in the above table and the decomposition of the decision process as proposed, the combined result is shown in Fig. 7.

For completeness we have added decision point D0, which represents the decision to invest the time and effort it takes to discover, review and evaluate alternatives to full copyright restrictions. Some authors will simply publish their work under the existing copyright regime because this requires no additional effort on their part, or because they are not even aware of the alternatives. With some probability $X_0$, an author will consider and evaluate alternative licensing options. We cannot observe $D_0$ but the mere fact that in the population of authors there is imperfect information on CC and similar licenses (combined with the fact that copyright protection is automatic) suggests that $D_0$ must exist and that $X_0$ depends on the search costs associated with obtaining this information. Authors who
do choose to incur these costs will in turn choose to adopt a model like CC in decision point \( D_1 \) with probability \( X_1 \) or will prefer full copyright protection with probability \( 1 - X_1 \). Similarly to \( X_0 \), we cannot observe \( X_1 \), because of the automatic application of copyright law at the birth of a work. By combining the picture in Fig. 7 with our previous discussion of author incentives we can make the following empirical observations:

\( D_2 \) (choice of NC constraint): An author who permits non-commercial distribution of his/her work is, as we have already noted earlier in the article, almost twice as likely to forbid commercial distribution of the work rather than allow it. It follows that for the average CC user (an artificial construct, but one that is useful for analysis) the potential losses from permitting commercial distribution without compensation seem to outweigh in most cases the potential benefits of additional reputation gains and positive network effects or any ideological/altruistic/hedonic motivations that the author may have for licensing more liberally (unless the author’s ideological stance towards liberal licensing is limited to non-commercial uses and thus does not welcome commercial exploitation).

\( D_3 \) (choice of ND constraint): We notice that most authors choose to allow for the creation of derivative works, but interestingly authors are 3 to 6 times more likely to forbid the creation of derivatives when the scope of derivatives is already limited to non-commercial only. This is interesting because the authors who stand to lose more financially from the uncontrolled creation of derivatives are those who also allow for commercial uses. The apparent paradox can be resolved if we assume that authors who allow for commercial uses of their work are not financially motivated and act mostly on conviction or altruism, or at most with the expectation of reputation gains through the wider distribution of their work. Some of these authors might still oppose the creation of derivatives for various reasons, but at least the foregoing of a potentially lucrative licensing deal would be less of a concern for these authors. On the other extreme, authors who only allow for non-commercial uses of their work appear to be concerned about competition and potential missed financial opportunities and are thus more keen to forbid non-commercial derivatives, even if the potential impact of those derivatives on the market is doubtful and would be only indirect; such derivatives would not be commercially exploitable (unless licensed under a separate agreement), but would potentially compete for attention with the original work and any licensed commercial derivatives. So, in summary, the existence of two ‘poles’ within the population of CC users, which are distinguished in terms of their sensitivity towards financial rewards, can explain the observation that the preference for NC appears to be positively related to the choice of ND. Of course we should note at this point that true author preferences are not fully revealed in this dataset because CC licenses do not allow the user to make a distinction between commercial copying and commercial adaptation of a work. Also, lacking a survey of user intentions and given the early stages of empirical work in this area, we can at best only provide alternative explanations for our empirical observations.

\( D_4 \) (choice of SA constraint): Interestingly, even if our analysis of the relation between \( D_2 \) and \( D_3 \) signifies the existence of two different poles or mindsets in the community of authors, when it comes to the share-alike constraint, the two groups not only ‘share alike’, but also seem to think alike, in terms of favoring the application of the SA constraint. We cannot say with confidence.
exactly how much they favor share-alike, as we are observing the combined effect of original works being licensed with share-alike and derivative works being licensed as such due to their re-use of at least one original work incorporating that restriction. It is nevertheless interesting to note that the appeal of Copyleft (share-alike, or SA) appears to be strong inside the CC community, in spite of the relatively weaker applicability of the pragmatic arguments for the use of such constraints, as compared to the production of functional goods. This is hinting at the strength of ideological and social psychological motivations for enforcing reciprocity in licensing, rather than the pragmatic concern of being able to merge contributions back to an existing project, as the latter is not as critical for cultural goods.

6.2. Synthesis of findings and interpretation

With respect to the decisions on derivatives (D3), we should note that the assumption of the existence of two author groups, one being more strongly motivated by financial gain and the other by more altruistic/ideological/hedonic and in any case non-financial incentives, is a sufficient explanation of observed behavior only in the case where authors in the two groups have the same (rational or irrational) expectations of the commercial viability of their work. If authors have different perceptions of commercial viability, we would expect some to be more protective of their content than others. In both cases (of homogeneous or of differing perceptions) it is correct to say that the data points towards the existence of two poles with different sensitivity towards monetary rewards, but the source of that difference may not be so much a general disregard for the potential monetary rewards of licensing deals on the part of some authors, as much as the result of the (subjective) expected value of such rewards being very low (respectively high) for some authors.

It may therefore be that some of the authors who opt for more liberal licensing of their work are not so much motivated by ideology or altruism but rather have a lower expectation of the financial returns that their work can generate and are thus keen to license it more liberally, to benefit from increased distribution and, consequently, increased visibility and reputation. These would be the quintessential amateurs in the CC user population, not necessarily in the sense of lacking the skills to produce high quality content, but rather in the sense of the content they produce not being their primary source of income. For example, a video of a political campaign may be of great value as a means of communicating the campaign's message, but it may not have significant, if any, market value as a product in itself.

However, it is wrong for such authors to assume that allowing for commercial derivatives is only an additional means of increasing visibility without any financial loss. Commercial derivatives will only be created if there is financial incentive to do so, which would imply that the work has commercial value, if not as an original work, than as raw material for derivative works (for example, music and special effects material can be much more valuable when incorporated into the soundtrack of a successful movie or ad campaign rather than as stand-alone works). It follows that to explain the observed preferences we need to make the assumption that users who permit commercial uses (and generally also allow commercial derivation) have either:

- a low expectation of the commercial value of their work as-is and as a source for derivative works, or...
- are strongly ideologically motivated, or...
- act mainly on altruistic/hedonic motivations

In all of these cases they will choose the most liberal licensing possible under this framework (i.e. BY and BY-SA), with no particular reason to choose BY-ND.

On the other hand, those who choose to permit only non-commercial uses have either:

- a high expectation of the commercial value of their work (as-is or as a source for derivative works), or...
- are ideologically motivated but consider commercial exploitation to be incompatible with this motivation

The second group will have no reason to apply the ND constraint, but the first group will be more divided when it comes to their preference for BY-NC-ND, as the value they place on the ND constraint will depend on their (subjective) expectation of the value of the work as re-usable material for the creation of derivative works and their assessment of the potential for non-commercial derivatives to impact (positively or negatively, including network effects) the market of the original work and that of licensed commercial derivatives.

It is worth noting that a common argument brought forth for the use of open content licenses, that of increased reputation gains through increased distribution of the work (in its original form or in derivative form), does not seem to be such an important factor in the choice of a specific CC license (although it is likely one of the main reasons why an author would opt for an open licensing model in the first place). This stems from the realization that we are able to provide explanations for the observed preferences without needing to resort to this argument. First of all, the choice of NC is not relevant because allowing for commercial uses (an action that could increase the work's visibility), will only matter if the work is commercially valuable. A rational decision maker who is sensitive to monetary rewards would not permit such uses without a commercial licensing deal. But is perhaps the choice of the ND constraint influenced by such expectations of increased reputation through wider adoption? Works licensed under ND might receive less attention as their reuse in derivative works will require prior permission by the author and this might discourage such reuse. However, potential producers of commercial derivatives are more likely to invest time and resources in negotiating a deal with the author of the original work, so the imposition of the ND constraint should not discourage commercial derivation. It will likely only discourage non-commercial derivation and this may indeed limit somewhat the spread of a work. Authors wishing to promote themselves through wide distribution and use of their works would therefore not favor the ND constraint.

But there are also other reasons not to favor ND, as we have seen. When are reputation gains clearly a relevant factor in the choice of specific license terms beyond the basic premise of all CC licenses which allow for non-commercial copying (which is a potentially powerful reputation-enhancing mechanism in its own right)? Reputation gains (rather than the factors discussed earlier in this section) will drive authors' licensing decisions within the CC framework in the very specific case where authors have a high expectation of the commercial value of their work as-is or as a source for derivative works (i.e. they choose NC), but they assess the impact of non-commercial adaptation on that value to be acceptable, or the work is treated as a loss leader for the promotion of the author. Such authors would opt for a BY-NC or BY-NC-SA license. The number of examples for this case is increasing lately in the music industry where prominent artists will sometimes license popular tracks, or, more rarely, entire albums, in a manner which allows for (often non-commercial) derivation, so as to energize and expand their fan base through remix contests and other such activities.

There is only one more factor that is known to influence (negatively) the decision to allow for the uncontrolled creation of derivatives (whether commercial, or non-commercial): the fact that some authors wish to preserve the work, its composition, character and message, as intended in its original form. This 'integrity concern' may be an important consideration for some authors, is orthogonal to the factors discussed above, and is discussed more in the context of specific online communities in the next section.
6.3. Influence of the community

We will now examine the influence of the community (and the medium type, which is usually tied to the community) on individual author choice, by examining author preferences in more detail in a very popular online community and reviewing some of the licensing choices in more communities mentioned in the article. We would generally expect the attitudes and preferences of the community to exert some influence on the individual decisions of members given the diverse needs of different user groups and the social nature of open cultural production, as already discussed. For this purpose we collected data on the use of CC in popular online communities known to host a lot of CC-licensed content, with a focus on the Flickr photo sharing community, which, unlike most of the communities we looked at so far is not music-oriented, but has been an early and very prominent adopter of CC licenses. The use of CC licenses in Flickr exhibits a somewhat different pattern compared to general CC licensing preferences. The resulting decision tree, based on the relative frequency of use of each license, as reported by Flickr, is shown in Fig. 8.

We notice again a preference for prohibiting commercial use (which is even stronger in the Flickr community than for the average CC user) and the same pattern for derivative works, though again with a slightly stronger tendency to forbid free adaptation. This leads us to believe that Flickr users license content similarly to the average CC adopter, but more conservatively. One reason for this can be privacy concerns. Unlike with music, which is still for all intents and purposes a means of entertainment and creative expression, photography is also used for documentation and moreover, many photographs may include images of the author or of relatives and friends, so the author may wish to protect this content from commercial exploitation and (potentially abusive) adaptation.

The relatively strong preference for BY-NC-ND may also be related to the fact that photographs are often not as commercially valuable as music or movies as stand-alone products, but rather as illustration material for magazine articles, books, advertisements or news reports. Photographers may therefore wish to limit the creation of derivative works as much as possible, to demand a fee for every derivative use. This may be a strong motivation behind the choices we observe, but is factually inaccurate as the inclusion of a photograph without modification in an article of any type does not constitute a derivative work according to CC guidelines. This is in direct contrast to the use of music for movie soundtracks, as typically the recording needs to be cut into smaller parts and synched to the relevant movie scenes, hence leading to a soundtrack which is a derivative of the original track (again, according to CC guidelines).

Interestingly, the relative preference for the application of the share-alike constraint appears to be somewhat weaker in Flickr, compared to that of the average CC adopter and it is not the preferred choice among members who allow for commercial derivatives. The explanation for this may lie in the fact that Flickr contains a large number of amateur photographers with low artistic or commercial aspirations (and presumably no privacy concerns, which would lead them to be more protective of the content), who may use Flickr as a shared repository for photos taken during an event or during vacation. When such authors choose a more liberal license it would be rational to choose one of the most liberal licenses possible, as their content has no commercial value, and little, if any value as material for derivative works. For these authors the application of the Share-Alike constraint is probably less meaningful as they do not expect their work to be used in any adaptations. Users who choose to apply the NC constraint also exhibit a weaker preference for SA, but they still opt for SA about 2/3 of the time.

Communities which are built primarily for the purpose of online collaboration and remixing (of which Flickr is not one), are naturally attracted to open licensing models like Creative Commons but will tend to avoid the use of the ND constraint. Some will also avoid the use of the SA constraint, as it may complicate the licensing of derivative works reusing multiple sources under different licenses [12]. In ccMixter for example, all user-submitted samples and a cappella vocal tracks (both of which are mostly valuable as sources for remixing and do not fit as-is) are licensed almost exclusively under BY and BY-NC. It is instructive that vocal tracks are licensed under BY-NC much more frequently compared to instrument samples. This is likely because good recordings of capable vocalists are scarcer than good instrument samples, so the opportunity cost associated with commercial licensing opportunities that would be foregone under a BY license may be significant. The Kompoz and Jamglue communities also utilize CC licenses exclusively for user-submitted content, with the exception again of licenses with the ND constraint, while Jamglue offers the additional option of dedicating a work to the public domain (whereby the author opts to give up all rights to the content). Slices offers the same options as Jamglue, but with the odd absence of BY-SA (BY-NC-SA is available as an option). Exceptions are sometimes made in these communities for popular artists who may make some of their works available in the context of remixing contests under special conditions which may require a custom license. Acid Planet, perhaps because it predates CC, Web 2.0 and all the more recent debates on intellectual property, does not provide specific options for CC licensing, but includes a text field where the author may enter licensing information in free form text.

Yourspins presents a different twist on the theme of open-sourced content, and is not so ‘open’ by most definitions, as all source material is made available by popular artists under specific deals with the company that operates the online service, and remixes can only be produced in the browser and reside on the website. The terms and conditions of the service state that all copyright resides with the artists who submit the original content for remixing and, while this leaves some room for interpretation regarding the ownership status of the remixes, the terms also state that users cannot distribute their remixes in any manner outside of the website. This very restrictive licensing model may be part of the reason why Yourspins has been able to attract a number of tracks by well-known artists to its platform: artists and their publishers can attract more attention to their music and ride on the popularity of the ‘social web’ without having to give up any of their property rights, as all user activities take place within the walled garden of the web service. Yourspins is thus

\[
\text{Fig. 8. Flickr CC user's decision tree.}
\]
only marginally relevant for open culture, but it is an interesting model for the engagement of an artist’s audience through the encouragement of the participative manipulation and reinterpretation of the artist’s work.

Similarly, online music labels having a policy of using Creative Commons licenses, such as Magnatune and Jamendo, are not examples of open cultural production per se and the choice of licenses within these labels is also instructive in this respect: all artists on Magnatune must provide their works under the BY-NC-SA license, while Jamendo may offer all CC licenses as options to signed artists, but most artists will nevertheless opt for BY-NC-SA (49%) or BY-NC-ND (27%), the two most restrictive licenses in the CC framework. In accordance with the analysis in this article, we see that artists whose works are posted online with the unambiguous expectation of financial returns will be more sensitive to such potential rewards and will opt for an NC license. However, it is encouraging for free culture proponents (and reflective of the fact that creative reuse is more established in music than in photography) that many artists use the BY-NC-SA license instead of BY-NC-ND. This grows the pool of content that is available to all for at least non-commercial remixing. The fact that BY-NC is not preferred by such labels and artists is indicative of the value that many artists place on reciprocal behavior: if they choose to allow for non-commercial derivatives, then they wish to make sure that those who will benefit from the availability of content thus licensed, will use the same license for their derivative creations.

We thus see that a combination of predetermined policies and the patterns that emerge through members’ licensing decisions jointly determine the popularity of different licenses in an online community, which in turn place a frame of constraints on the organization of production. Understanding the motivation for and implications of such choices requires a prior understanding of the incentives of open cultural production and the ways in which it is different to the production of functional information goods. The examination of specific communities in light of the unique characteristics of open cultural production is worthy of a more in-depth exposition, but here we have limited ourselves to the review of only some interesting features of such communities. I have presented empirical evidence for the diversity of needs and expectations in the population of CC users and how both community-specific and medium-specific considerations may influence licensing decisions. While it may be true that OSS projects can also vary greatly in scope and aspirations (and this may partly explain the proliferation of open software licenses), the exposition in this article has hopefully provided enough evidence in support of the argument that open/free culture is a much broader undertaking than open/free software and thus must be treated separately and with at least as much attention as OSS has received in the literature. Fig. 9 summarizes the factors that we have examined in this article with respect to license choice in an open cultural production setting, such as that defined by the use of CC licenses.

### 7. Conclusion

In this article we have reviewed some of the literature on open source software with an eye on comparing the needs and modes of organization of software production versus the production of cultural goods. The article provided empirical evidence and argumentation in support of the assertion that the aims, scope and scale of open cultural production are much wider than those of the open production of functional goods, even if they both share some of the same characteristics, and without neglecting the fact that Free Software has been an inspiration for Free Culture.

A central theme of the article has been the distinction between goods of primarily functional value vis-à-vis aesthetic, entertainment,

<table>
<thead>
<tr>
<th>License choice</th>
<th>Influence of community</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual author choice</td>
<td>Ideological conviction</td>
</tr>
<tr>
<td>Financial incentives</td>
<td>Altruistic/Social Hedonic</td>
</tr>
<tr>
<td>Expectation of market value</td>
<td>Medium type</td>
</tr>
<tr>
<td>Expectation of reuse value</td>
<td>Aims/ownership of community</td>
</tr>
<tr>
<td>Reputation through distribution</td>
<td>Membership</td>
</tr>
<tr>
<td>Tolerance for commercial use</td>
<td>Community policies</td>
</tr>
<tr>
<td>Expectation of degree of reuse</td>
<td>Community practices and norms</td>
</tr>
<tr>
<td>Expectation of market value of derivatives</td>
<td>Expectation of reciprocal behavior</td>
</tr>
<tr>
<td>Impact of NC derivatives on market</td>
<td>Medium sharing and reuse practices</td>
</tr>
<tr>
<td>Quality/maturity of work</td>
<td>Expectation of reciprocal behavior</td>
</tr>
<tr>
<td>Commercial potential</td>
<td>Intended use</td>
</tr>
</tbody>
</table>

**Fig. 9.** Factors impacting the choice of license within an open cultural production setting.
educational, or more broadly cultural value, i.e. what we call here cultural goods. The former are continuously evolving as developers contribute new functionality and iron out earlier versions, while the latter may have value as standalone goods and this value can be realized immediately upon consumption, even if arguably, culture as a whole is never static. Also, functional goods are designed for a specific purpose and their development therefore requires some coordination among contributing parties, so as to avoid confusion and massive duplication of effort. Cultural goods on the other hand, may be produced in a coordinated manner or in an entirely ad-hoc fashion by creators acting independently of one another, since duplication of effort and the creation of multiple versions of the same good may be desirable, in order to satisfy the diverse tastes of the population and to facilitate the process of serendipitous discovery that is an essential component of creativity. This, as we have seen, has implications for the design of suitable licenses that will accommodate the diverse needs and incentives of producers of cultural goods.

Acknowledgments

I am grateful to Mike Linksvayer, Jon Phillips and Asheesh Laroia of Creative Commons for their support with collecting CC usage data and disseminating the results of this work to the community, as well as to Giri Tayi, Warren Chik, Ankit Guglani and Clint Gono for their various contributions to the CC-Monitor project and the many valuable discussions.

References