# BEYOND OPEN SOURCE SOFTWARE: FRAMEWORK AND IMPLICATIONS FOR OPEN CONTENT RESEARCH

Okoli, Chitu, Concordia University, 1455 de Maisonneuve Boulevard West, Montreal, Quebec H3G 1M8, Canada, chitu.okoli@concordia.ca

Carillo, Kevin Daniel André, Toulouse Business School, 20 Boulevard Lascrosses 31068 Toulouse, France, k.carillo@esc-toulouse.fr

## Abstract

*The same open source philosophy that has been traditionally applied to software development can be applied to the collaborative creation of non-software information products, such as books, music and video. Such products are generically referred to as open content. Due largely to the success of large projects such as Wikipedia and the Creative Commons, open content has gained increasing attention not only in the popular media, but also in scholarly research. It is important to investigate the workings of the open source process in these new media of expression. This paper introduces the scope of emerging research on the open content phenomenon beyond open source software. We develop a framework for categorizing copyrightable works as utilitarian, factual, aesthetic or opinioned works. Based on these categories, we review some key theory-driven findings from open source software research and assess the applicability of extending their implications to open content. We present a research agenda that integrates the findings and proposes a list of research topics that can help lay a solid foundation for open content research.*

*Keywords: Open content, Open Source Software, Wikipedia, Creative Commons, Free software, Open knowledge, Free cultural works*

# 1   Introduction

The open source philosophy traditionally used for software development has been applied to the collaborative creation of numerous non-software information products. Most notably, Wikipedia has emerged as a comprehensive general encyclopaedia comprising over 18 million articles in over 250 languages, built using the "open source" model of mass collaboration under the framework of a legal license that permits liberal reuse and extension (Okoli et al. 2012). It is now the world's leading Web reference source (Perez 2007). Numerous other examples exist that are having increasingly visible impact on society at large: open educational resources are revolutionizing the academic landscape (Downes 2007); open access journal publication is disrupting the scholarly publishing world (Collins 2011); open music is providing new channels and business models for music delivery. This extension of the open source approach to apply to non-software information products is generally called "open content" (Pfaffenberger 2001). As open content begins to develop in its breadth and depth of applications, the possibilities of new community-created media products are endless: open books, maps, video, poetry, recipes—indeed information products of just about any form (see http://freedomdefined.org/Portal:Index for a directory of projects).

The information age has introduced the digitization of virtually every medium of human recorded expression. In this article, we refer collectively to such works as copyrightable works, referring to books, recorded music, videos, articles, and all other works that fall under the scope of copyright protection, as distinct from other kinds of human expression (such as ideas and processes) that are not thus covered. With the advent of the Internet, such works can now be developed using Web-based information systems that enable geographically- and temporally-distributed content developers to collaborate on their creation. Whereas numerous studies have provided guiding frameworks for research on open source software (OSS) (Aksulu and Wade 2010; Feller and Fitzgerald 2000; Jin et al. 2007; Krogh and Hippel 2006; Lerner and Tirole 2001, 2005; Nelson et al. 2006; Niederman et al. 2006; Rossi 2004; Scacchi 2007), there is little research that has attempted to encompass all forms of open content in a comprehensive framework such as to facilitate the learning of fundamental principles that apply in this broader phenomenon, including OSS among other forms. This article presents a framework that can serve as a theoretical base for the scholarly study of the nascent and rapidly growing phenomenon of open content. With the experience from OSS, much has been learnt and yet needs to be learnt; however, the applications to diverse media are much broader and far-reaching than just to software.

Because many categories of works are sometimes labelled as "open content", it is necessary to specify a precise definition of our usage of the term. We define open content in this article as any copyrightable work that is distributed in a format that permits and facilitates legal use and redistribution of the work; it may or may not permit derivative works and it may or may not include a copyleft or share-alike clause. "Copyrightable work" indicates that international copyright law is the reference to the genres of content (however, even if the work in question has since entered the public domain, we still consider it a "copyrightable work"). Thus, for example, patented or trademarked works cannot be considered "open content", nor can ideas such as scientific theories or ideas of inventions or new business opportunities, as long as they are not recorded in a fixed medium of expression. The "format" of distribution refers to the fact that the work must be delivered in a way that facilitates use and distribution; thus, for instance, books encrypted with digital restrictions management cannot be considered "open content". "Open content" might or might not include a "copyleft" clause that requires derivative works (if authorized) to be licensed under an open content license; the Creative Commons Attribution-ShareAlike (CC-BY-SA) license is a popular such license.

In this article we distinguish open content from the following related phenomena: user-generated content, where users contribute content to an Internet service and the license might or might not authorize redistribution outside the service; social media, referring to Internet services that facilitate

social interaction; and open standards, referring to royalty-free implementation of a technical specification. These categories are outside our scope of consideration.

We also distinguish between our definition of open content and the narrower subset of works defined by the Open Knowledge Definition (OKD) and the Definition of Free Cultural Works (DFCW). These definitions try to parallel the liberties accorded to users of open source software and free software, respectively, to the broader range of all copyrightable works. While we recognize the intentions of these definitions to maximize the reuse of open content, our definition is much broader and more closely corresponds to the range of works licensed under Creative Commons licenses. We recognize that the range of issues and concerns that open content creators might have is much broader than that of licensors of free and open source software.

This article proceeds as follows: having motivated the need for open content research and defined its scope, we proceed in the next section with a general review of the scholarly research to date that has studied open content. In the following two sections, we develop a framework for categorizing copyrighted works that is particularly relevant for understanding the implications of research on open content. Based on these categories, we then consider the applicability of some implications of significant theory-relevant findings from open source software research for open content. Finally, we derive a research agenda that introduces important research topics for future open content research.

## 2 Overview of scholarly research on open content

Considering the recent nature of the field of research, the bulk of scholarly work thus far conducted on open content has been the extensive body of research on Wikipedia (Okoli et al. 2012). For this article, we focus only on reviewing open content other than Wikipedia; according to our knowledge such a review has never as of yet been conducted. We searched the citations and abstracts of some relevant scholarly databases (EBSCO, Sage Publications, ProQuest, and ISI Web of Knowledge) for the keywords "open content", "free content" and "free cultural works". After removing Wikipedia-related studies and those that did not meet our open content definition (for example, the "Open Content Alliance" is a digital archive initiative, but is not actually "open content"), we only found 9 peer-reviewed journal articles; we summarize their contents here.

Cheliotis (2009) rigorously investigated the implications of open content licensing based on the characteristics of the works. We will later discuss his open content categorizations in detail, as well as some of his suggestions for licensing various kinds of content. In an empirical study of photographers' Creative Commons license choices in the photograph-posting sharing site Flickr, he found that license choices largely reflect content creators' anticipation of future commercial revenues from their works.

A few scholars have discoursed legal aspects of open content licensing. Liao (2006) traces some of the intellectual property reforms that Western societies have been experiencing (including questioning the "property" aspect of "intellectual property"), and encourages Chinese society to aim towards a "creative da-tong" (state of societal prosperity through creativity) by incorporating the emergent open cultures, including open content creation, into its culture and copyright reforms. This advocates a focus on open content for the sake of promoting the resultant products, beyond just the economic returns to creators. Armstrong (2010) discusses the legal problems in the United States with open source and open content licenses in light of the American legal restrictions on copyright transfers. Chiao (2010) experimented on optimal team size modularity in open content teams, concluding that non-modular projects should restrict team size to be small, and that projects with large team sizes should be organized modularly.

Certain studies describe specific websites or projects that implement open content models, rather than treating the open content methodology or philosophy. Such projects include the webgis system for an urban planning geographical information system (Budoni et al. 2007); and the Digital Universe, an open content encyclopedia that uses named experts to create its content (Korman 2006).

The other open content studies we found are not directly related to its development: Schweik et al. (2005) suggested that the open content development model could be used to reform various academic practices such as peer review and the collaborative development of scholarly work, which they demonstrate in modeling change in land-use. Ven et al. (2008) presented libOR, a collection of open content data sets for operations research. In addition to discussing open access, Ballantyne (2009) briefly discusses the beginnings of open content licensing in the International Rice Research Institute.

Our review shows that, other than Cheliotis' (2009) study on licensing implications for copyrightable works, there is little structure or general direction for the development of open content. In the following section, we develop a framework for categorizing copyrightable works in a way amenable to theorizing about their development when they are licensed using open content models.

# 3 Classification of open content works

Open source software is a very particular kind of work, and has some fundamental characteristics that might not be applicable to all kinds of open content. To help understand the nature of these characteristics of open content, we draw from some past thinking on the subject. In a discussion of the role of traditional copyright laws in contemporary society, Richard Stallman, founder of the free software movement, described three categories of works with fundamentally different characteristics that would affect the most appropriate copyright licensing terms (Stallman 2002). First, he described "functional works", where the goal is to produce a useful product; "this includes recipes, computer programs, manuals and textbooks, reference works like dictionaries and encyclopedias" (2002 p. 141); this is the category that includes open source software. Second, he described "works whose purpose is to say what certain people think" (2002 p. 142), that is, statements of people's subjective opinion, such as essays. Third, he described "aesthetic or entertaining works" (2002 p. 142) such as novels, music, and non-documentary films.

In a study of the implications of various Creative Commons licenses, Cheliotis (2009) broadly categorized works as either functional or cultural goods. He described a functional good as one whose goal is to fulfil a consumer's practical needs, whereas a cultural good is one that mainly serves to entertain the consumer. Partially based on Stallman's three categories and Cheliotis' two, we identify two orthogonal dimensions along which works can be classified, yielding four distinct categories, with pertinent implications for open content development. For our classification, we consider works primarily from the perspective of how the consumers of the works might assess or judge them.

## 3.1 Two dimensions for consideration: relativist-universalist versus objective-subjective

The first aspect of our classification involves two distinct dimensions based on the fundamental perspective of how reality is observed or evaluated. First, reality can be assessed ontologically as being more or less relativistic or more or less universalistic—that is, whether there are multiple realities, each depending on the relative perspective of the individual, or if there is a universal reality that applies to all people. The second dimension concerns how the value or merit of a work is assessed, whether objectively based on more or less the same criteria for all people, or subjectively varying for each individual. The relativist-universalist dimension concerns the nature of the reality that is being observed. In the objective-subjective dimension, there is no consideration per se of whether the work is "real" or "true", but rather a consideration of if it is valuable, useful, or worthy of appreciation.

### 3.1.1 Relativist versus universalist works

The first dimension borrows from Järvinen's (2008) taxonomy of information systems research. He distinguishes between the "value-laden" design science paradigm, where certain outcomes are considered more valuable than others, and the "value-free" natural/social science paradigm, where the

goal is to ascertain the real and actual state of affairs in the world, with no preference given to any particular outcome. We frame these distinctions based on the philosophical duality of relativism versus universalism. Relativism holds that truths or values are not absolute; they depend on factors both intrinsic and extrinsic to individuals. Universalism holds that truths or values are universal and absolute, irrespective of the subject or context. In this dimension, we present two contrasting approaches to classifying copyrightable works based on how their quality is judged (that is, a value judgement): works can either be judged based on a relativist assessment of how valuable or worthy of appreciation they are (Järvinen's value-laden category), or based on a universalist assessment of how they conform to some universally-held standard (analogous to Järvinen's value-free category). For example, software and fine art are judged relativistically: they are not judged as "right or wrong", but rather as "useful or beautiful", which are relativistic standards. In contrast, maps and news reports are judged universalistically: what matters is whether they are accurate and faithful to the real facts of the situation. Although for purposes of classification we present this dimension as a duality, it should be properly considered a spectrum, with works containing characteristics that might be more or less relativist or universalist.

### 3.1.2    Objectively-evaluated versus subjectively-evaluated works

The second dimension involves the consideration of whether the value or quality assessment is generally considered an objective judgement (based on highly objective characteristics that do not vary significantly regardless of who is making the evaluation) or if the assessment is generally considered to be subjective (where it is understood and accepted that different people would evaluate the work differently based on their own ideologies or personal preferences). Some works are judged based on their artistic or aesthetic merits. Paintings, musical compositions, poetry and fiction are common examples of works that are judged not so much in terms of whether they are right or wrong as much as by whether they are beautiful, ugly, or plain. We label these "subjectively-evaluated works". The other category consists of works that have an objective criterion for determining their quality: the degree to which they are accurate according to some extrinsic standard, or to which they achieve some independently defined goal or purpose. Textbooks, software, and encyclopedias are examples of what we call "objectively-evaluated works". Works from this category are judged on their accuracy, usefulness, practicability, and other quantifiable, objective criteria of merit or value. As with the relativist-universalist dimension, the objective-subjective dimension should be considered more of as a spectrum than as a discrete categorization.

## 3.2    Four categories of open content

Based on these two dimensions, we have four categories of works. Although these categories apply to all copyrightable works, we will discuss them particularly in the context of open content.

### 3.2.1    Utilitarian works: Objectively-evaluated and relativistic

First, we have utilitarian works, which are objectively evaluated according to how well they attain a relativistic value goal. Open source software falls in this category, since a software program has a definite utility goal, and is judged based on how well it does the job. Even art-oriented software such as drawing software (e.g. GIMP or Inkscape) or video-editing software (e.g. Kino or Blender) is utilitarian; it is not judged based on the aesthetic value of the resulting works; the software is judged on how well it enables artists to carry out their creative visions. Other examples in this category are cooking recipes (Foodista), how-to manuals (WikiHow), engineering designs (Appropedia), taxonomies (WikiSpecies). These all share the characteristics of their value goal not being to achieve some sort of universalist "truth", but rather an attempt to be valuable according to some relativistic criteria. Nonetheless, their evaluation of whether or not they attain these relativistic criteria is based on concrete objective criteria.

### 3.2.2    Factual works: Objectively-evaluated and universalist

The second category is factual works, which are objectively evaluated primarily according to how universally true the work is. For such works, there are some absolute truth claims, and the works are evaluated according to these criteria. Wikipedia is in this category, as exemplified by its Neutral Point of View doctrine: the project explicitly forbids statements of opinion, and only authorizes content that can be externally, objectively documented. "Universal truth" in the case of Wikipedia refers to the existence of respectable citations external to Wikipedia of the statement in question.

**Value assessment**

*Objective* ←—————————————————→ *Subjective*

| | **Utilitarian** | **Aesthetic** |
|---|---|---|
| *Examples* | Software, recipes, how-to manuals, engineering designs, architectural blueprints, taxonomies, typologies | Fine art, music, literary fiction, poetry, song lyrics, non-documentary films, photographs, drama, games |
| *Open content development characteristics* | Large number of contributors; high incentives for businesses to contribute; projects with high modularity and usually large granularity | Contributions in small teams; low incentives for businesses to contribute; projects with very low modularity and large granularity |
| *Preferred licenses* | FDL, BY, BY-SA, BY-NC, BY-NC-SA | FDL, BY, BY-SA, BY-ND, BY-NC, BY-NC-SA, BY-NC-ND |
| | **Factual** | **Opinioned** |
| *Examples* | Textbooks, dictionaries, encyclopedias (e.g. Wikipedia), maps, educational materials, datasets | Essays, editorials, blogs, commentaries, scientific articles, news reports and historical records, product reviews, religious and philosophical texts |
| *Open content development characteristics* | Large number of contributors; low incentives for businesses to contribute; numerous applications of OC development; projects with high modularity and small granularity | Contributions by individuals or in very small teams; few examples of OC development; projects with high modularity; granularity ranges widely from very small to very large |
| *Preferred licenses* | FDL, BY, BY-SA, BY-NC, BY-NC-SA | BY-ND, BY-NC-ND |

The vertical axis is labeled **Truth perspective**, ranging from *Relativist* (top) to *Universalist* (bottom).

*Table 1.        Categories of works with open content licensing implications*

Of Stallman's three categories, his first category of "functional works" encompasses both the utilitarian and factual categories we describe here. Among his list of examples cited earlier, recipes and computer programs are utilitarian works, since their value is based on how well they attain some objective standard of being "valuable"; documentation, textbooks (Free High School Science Texts), and reference works (Citizendium) are factual works, since their quality is based on reference to some external, universal standard of what is "true" or "factual". We will see that these two categories have similar licensing implications, which is probably why Stallman considered them as a unitary category.

Referring to what we call utilitarian and factual works, Stallman (2002 p. 143) says: "For all these functional works, I believe that the issues are basically the same as they are for software and the same conclusions apply. People should have the freedom even to publish a modified version because it's very useful to modify functional works." In other words, he believes that it is appropriate to apply a standard open content license, such as the GFDL, or CC-BY or CC-BY-SA, to such works. Cheliotis (2009) seems to agree that functional goods are generally best served by a CC-BY or a CC-BY-SA license. However, unlike Stallman, he also recognizes that some authors prefer to only permit non-

commercial use, modification, or redistribution; thus, he includes CC-BY-NC and CC-BY-NC-SA as suggested licensing terms.

### 3.2.3  Aesthetic works: Subjectively-evaluated and relativistic

The third category we identify features aesthetic works (Stallman's "aesthetic or entertaining works") where beauty is in the eye of the beholder; these works are subjectively evaluated based on an evaluator's relativistic preference of what is valuable or beautiful. This includes music (ccMixter, Kompoz), works of fiction, and fiction movies (Blender project). Concerning open content licensing of aesthetic works, even Stallman recognizes that there are significant complexities involved:

> Now for these works, the issue of modification is a very difficult one because on the one hand, there is the idea that these works reflect the vision of an artist and to change them is to mess up that vision. On the other hand, you have the fact that there is the folk process, where a sequence of people modifying a work can sometimes produce a result that is extremely rich. ... It's a hard question what we should do about publishing modified versions of an aesthetic or an artistic work, and we might have to look for further subdivisions of the category in order to solve this problem. (Stallman 2002 p. 144)

Cheliotis (2009) argues that for aesthetic works (which he calls "cultural goods"), the copyleft provision is not as meaningful: even though an artist may want to permit modification and redistribution of his or her work, it is often not meaningful to reincorporate these modifications into the original work. The artist would usually be content with receiving attribution as the original source of the aesthetic idea, but would generally want to retain the original work intact. Moreover, Cheliotis further argues that the non-commercial restriction is more meaningful for aesthetic works because they are more often exploited without modification than are utilitarian or factual works. A commercial exploiter's options are limited if a utilitarian or factual work is protected by a copyleft provision; in this case, although they might exploit the work commercially, they would be required to provide any modifications they might make for free, thus limiting any possibilities of monopolizing someone else's work. On the other hand, it is easier for a marketing organization to exploit an aesthetic work better than the artist can themselves; the non-commercial provision assures that the artist shares in the profits. As a result of these complexities, we would expect to see the widest range of licensing options in place with aesthetic works: any of the eight Creative Commons licenses might make sense, as well as the FDL.

### 3.2.4  Opinioned works: Subjectively-evaluated and universalist

Finally, we have opinioned works, which make universalistic claims, but such claims are understood to be subjective without an inordinate attempt to objectively evaluate such claims. These include essays, editorials, commentaries, blogs, comments on blogs, scientific publications, product reviews, religious and philosophical texts. The common theme with these kinds of works is that although they put forth theses or statements that can only be evaluated subjectively, it is of great importance that the work be presented accurately as a faithful representation of the author's beliefs or opinions.

It is somewhat non-intuitive that we consider essays, political theses, and religious texts as universalist works. This is certainly not because we hold their contents to be universally true, nor necessarily even based on whether the author considers the contents as universally true. Rather, we are focused on a very particular universalist truth claim: whether or not the work is an accurate expression of the author's personal opinion, view or perspective. On one hand, the works are subjectively-evaluated: it is up to the author, or those who agree with the author, to determine whether they agree with or like the work. On the other hand, the truth or validity or reliability of the work is not based on the relativist criteria of if the work is useful or beautiful; it is based on the criteria of fidelity to the author's actual beliefs or opinions. From another perspective, an author is only satisfied that the work is completed when it accurately and fairly (universalist criteria) expresses what he or she truly feels, thinks or believes (subjective evaluation).

Whereas scientific publications can easily be seen as universalist works, it might not be so obvious why we consider them subjectively-evaluated, and thus opinioned, works. To understand this classification, it is important to distinguish a scientific discovery as an abstract concept from a concrete article that reports the discovery. A scientific discovery is not a copyrightable work; ideas and general knowledge are not copyrightable, not even by the person who originates or discovers them. Only the article that records and reports the discovery is copyrightable as an expression of scientific knowledge. Thus, while the discovery itself is an objectively-evaluated universalist (thus factual) concept, such as could be described in Wikipedia, the article that originally reports it is an expression of the author's chosen interpretation and mode of presentation—both of which are subjective expressions. Another illustration of this distinction is seen in the move towards "open data", which attempts to release the factual scientific data via open content licenses or public domain dedication for unrestricted reanalysis and repurposing, versus the "open access" movement, which aims to permit redistribution of scientific articles, without any attempt to permit modification of these opinioned works.

| Key findings from OSS research | L A | Applies | | | |
|---|---|---|---|---|---|
| | | U | F | A | O |
| **Implications for all four open content categories** | | | | | |
| OSS peer review is a means to ensure better bug discovery and better bug solving flexibility. The possibility to integrate additional OSS security functionalities enables high software security. But binary-only programs are less vulnerable and add peer review provides visibility to potential attackers. (Payne, 2002) | A | H | H | H | H |
| OSS communities are combinations of gift giving and scientific knowledge sharing cultures. Highly skilled programmers collectively develop software. Loosely coupled communities kept together by strong common values in line with the hacker culture. (Ljunberg, 2000) | G | H | H | H | H |
| OSS communities rely on a gift giving culture where ideas and products freely circulate. The giver is given a power from giving away that guarantee the quality of the released ideas or products. Quality is assessed through peer review. (Bergquist and Ljungberg, 2001) | G | H | H | H | H |
| Framework to develop hybrid-OSS communities for software development relying on of community building, governance, and infrastructure. (Sharma, 2002) | G | H | H | H | H |
| OSS is increasingly used and developed by SMEs (Lundell et al. 2010) | O | H | H | M | M |
| Software companies must gradually re-orient their business models towards more profitable activities (Vitari & Ravarini 2009) | O | H | H | M | M |
| The quality of the proprietary software decreases as the quality of OSS increases (Jaisingh et al. 2008) | O | H | H | M | M |
| In software markets with high network effects, a proprietary software offering can survive only if it is more usable than the competing commercial version of the OSS (Sen 2007) | O | H | H | M | M |
| **Implications for utilitarian, factual and aesthetic works, but not for opinioned** | | | | | |
| Non-restrictive OSS licenses and sponsored projects attract greater user interest. (Stewart et al., 2006) | I | H | H | H | N |
| Sustained participation is associated with the co-evolution of situated learning and identity construction. Initial conditions to participate do not predict long-term participation. (Fang and Neufled, 2009). | I | H | H | M | N |
| Using an incremental life cycle is highly motivating and supports learning. The incremental approach may raise difficulties during development of complex new features. (Jorgensen, 2001) | G | H | M | M | N |
| Affective trust positively related to team size and effort; cognitive trust was not; freedom ideology negatively related to cognitive trust; team effort and communication quality positively related to task completion; team size was not. (Stewart and Gosain, 2006) | G | H | H | H | N |
| OSS project effectiveness (in terms of team size, team effort and team's level of completion) is affected by expertise integration. Expertise integration was found to have an impact on team size and team effort, which in turn were found to jointly influence task completion. (Chou and He, 2010) | G | H | H | M | N |
| A software firm can benefit from giving away software due to the accelerated diffusion process and increased net present value of future sale (Jiang and Sarkar, 2009). | O | H | M | H | N |
| Key: Level of Analysis (LA): A = Artefact; I = Individual; G = Group/Project/Community; O = Organization <br> Degree of applicability (Applies): U = Utilitarian; F = Factual; A = Aesthetic; O = Opinioned <br> H = Highly applicable; M = Moderately applicable; N = Not applicable; ? = Uncertain applicability | | | | | |

*Table 2a.       Proposed implications of some OSS research findings for open content*

This identification and categorization of opinioned works has important implications for their open content development and corresponding licenses—in this case, they are most likely to be licensed in ways that inhibit open content development. Stallman (2002 p. 142) argues: "The whole point of those works is that they tell you what somebody thinks or what somebody saw or what somebody believes.

To modify them is to misrepresent the authors; so modifying these works is not a socially useful activity. And so verbatim copying is the only thing that people really need to be allowed to do." The equivalent Creative Commons licenses are CC-BY-ND or CC-BY-NC-ND.

# 4 Implications of research on open source software for open content and research agenda

An important application of our framework is for developing theoretically-grounded propositions. Though there is an abundance of OSS research (especially in the software engineering literature); we do not attempt to conduct an exhaustive review of the literature here. Rather, we highlight key findings from some of the leading information systems journals, which tend to be more theory-oriented in that they select topics of inquiry and research approaches that facilitate drawing wider implications to similar and related phenomena—in our case, open content. We searched the abstracts of six leading IS journals (European Journal of Information Systems, Information Systems Journal, Information Systems Research, Journal of AIS, Journal of MIS, and MIS Quarterly) as of October 2010 for the binary search string < "open source" OR "free software" OR FOSS OR FLOSS >. (FOSS is an abbreviation for "free and open source software", and FLOSS for "free/libre/open source software"; both are widely used synonyms for the phenomenon.) Our search resulted in 27 articles (article references are listed in detail in a separate article (Okoli and Carillo 2013)).

| Key findings from OSS research | L A | Applies | | | |
|---|---|---|---|---|---|
| | | U | F | A | O |
| **Implications for utilitarian and factual works only** | | | | | |
| OSS vendors released patches faster (Arora et al., 2008) | A | H | H | N | N |
| The quality of code produced by OSS is lower than that which is expected by industrial standards. To a certain extent, the average component size of an application facilitates OSS development and is also negatively related to user satisfaction. (Stamelos et al., 2002) | A | H | H | N | N |
| In presence of OSS alternative, willingness to pay for proprietary software dropped (Raghu et al., 2009) | I | H | H | N | N |
| Non-market sponsors attract greater development activity; restrictive licenses did not attract greater development activity. (Stewart et al., 2006) | G | H | H | N | N |
| Only a small core of developers is responsible for most project outputs. Only a small number of programmers work together on a same file. (Koch and Schneider, 2002) | G | H | H | N | N |
| Trust is not that important, OSS communities rely on forms of social control and self-control. (Gallivan, 2001) | G | H | H | N | N |
| Opensourcer must not seek to dominate the process; must provide business expertise; must help establish trusted ecosystem. OSS community must have democratic authority structure; must have responsible and innovative attitude; must help establish sustainable ecosystem. (Ågerfalk and Fitzgerald, 2008) | O | H | H | N | N |
| There is no overall difference in terms of evaluation criteria for proprietary or open source enterprise application software; implementation factors such as ease of implementation and support are much more crucial in the evaluation of OSS enterprise application software (Benlian and Hess, 2010) | O | H | H | N | N |
| **Implications for utilitarian and aesthetic works only** | | | | | |
| Intrinsic motivation of challenge (problem solving) is associated with the developers' preference for licenses with moderate restrictions, while the extrinsic motivation of status (through peer recognition) is associated with developers' preference for licenses with least restrictions. (Sen et al., 2008) | I | H | N | H | N |
| Developers prefer joining a project when they have past relationships with initiator, and other members are more experienced. (Hahn et al., 2008) | I | H | N | H | N |
| OSS has been transformed to a commercially viable business strategy (Fitzgerald, 2006) | O | H | N | H | N |
| **Uncertain as to applicability of implications beyond software** | | | | | |
| Shared macroculture and collective sanctions facilitates the coordination of exchanges in open source service networks; collective sanctions facilitate the safeguarding of exchanges (Feller et al. 2008) | O | ? | ? | ? | ? |

*Table 2b.        Proposed implications of some OSS research findings for open content*

We classified the reviewed OSS articles according to the following five distinct levels of analysis: software artefact, individual, team/project/community, organization, and society (Niederman et al. 2006), though no article in our sample of 27 adopted a societal perspective. Since we restricted our examination to just a few leading IS journals, we do not attempt to make any inferences about the total number of articles with implications for open content. For example, had we searched software

engineering journals and conference proceedings, we would surely have identified far more articles where the software artefact and where the developer as an individual would be the primary level of analysis, probably with a different distribution of implications from what we have found here. Nonetheless, we expect our IS sample to be more amenable for deriving theoretical implications.

| | **All four open content categories** | **Utilitarian, factual and aesthetic works only** | **Utilitarian and factual works only** | **Utilitarian and aesthetic works only** |
|---|---|---|---|---|
| **Artefact** | ▪ Influence of peer review practices on OC product quality<br>▪ Impact of OC product visibility on open content product quality | ▪ Relationship between OC product licence and product quality<br>▪ Influence of an incremental life cycle for OC production on OC product quality | ▪ Assessment of OC work quality using industrial standards<br>▪ Determining specific standards to evaluate OC product quality<br>▪ Differences of production patterns between OC and proprietary works<br>▪ Influence of the average component size of an OC product on its development process | ▪ Relationship between the degree of license restrictiveness of an OC work and its quality<br>▪ Relationship between peer recognition mechanisms and OC product quality |
| **Individual** | ▪ Relationship between OC product visibility and individual contribution quality<br>▪ Influence of peer review on contributor performance in OC projects | ▪ Relationship between OC product licence and contributor interest<br>▪ Do sponsored OC projects attract greater contributor interest?<br>▪ Applicability of legitimate peripheral participation theory to explain sustained participation in the OC context<br>▪ Role of initial conditions in predicting long-term participation in OC projects | ▪ Impact of the existence of OC product alternatives on willingness of a consumer to pay for proprietary products<br>▪ Relationship between average component size of an OC product and motivation to join an OC project | ▪ Relationship between contributor intrinsic/extrinsic motivations and the degree of license restrictiveness of an OC product<br>▪ Influence of past relationships on joining an OC project<br>▪ Influence of past experiences on joining an OC project |
| **Group /Project /Community** | ▪ Study of the shared cultural components that characterize OC communities<br>▪ Study of the peer review mechanisms governing OC communities | ▪ Impact of the incremental OC production approach on the motivation and learning of product contributors<br>▪ Using a trust perspective, determining the factors affecting team size, team effort, and task completion in the OC context<br>▪ Relationship between expertise integration and OC project effectiveness/ team size/team effort | ▪ Organization and structure of OC communities<br>▪ Evolution of the number of contributors working on the same OC product component<br>▪ Role of trust-building mechanisms in OC production<br>▪ Instances of social control and self-control mechanisms in OC production | ▪ Study of the peer recognition mechanisms that govern OC communities<br>▪ Impact of OC group characteristics on group effectiveness<br>▪ Study of the problem-solving mechanisms that characterize OC production |
| **Organization** | ▪ Potential for SMEs to use/develop OC products<br>▪ New forms of business models engendered by the spread of OC development<br>▪ Quality difference between OC and proprietary products<br>▪ Role played by network effects in markets where OC and proprietary products compete | ▪ Benefits for a company to release a product under an OC license<br>▪ Diffusion of a product released under an OC license by a company<br>▪ Impact on net present value of future sales when releasing a product under an OC license | ▪ Description of the different forms of ecosystem that govern OC production<br>▪ Role of democratic authority structures in helping OC communities function and perform better<br>▪ Determining evaluation criteria that can be used to evaluate an OC product solution | ▪ Determining whether OC production can become commercially viable<br>▪ Study of the different types of OC production-based business strategy |

*Table 3.        Research agenda and research topics sample for open content research*

For each of these articles (see Tables 2a and 2b), we carefully considered the key findings and analysed the possible implications that could apply to open content other than OSS. We discuss the implications of each study in detail in a separate article (Okoli and Carillo 2013), but here we only present general implications from the ensemble of studies sampled.

We note, however, that we categorized the 27 articles only into the first four levels of analysis, as no article in our sample studied OSS from a societal perspective. This level of analysis would also pertain to issues related to the diffusion of the free software philosophy to non-software areas that involve the licensing of intellectual property, which addresses the open content phenomena itself.

Based on the findings generated by the literature review and using the framework we introduced earlier, we have derived a research agenda (see Table 3). According to open content work category and level of analysis, the agenda suggests a sample of relevant research topics in an attempt to provide some meaningful and sound directions for open content research.

## 5 Conclusion

In this article, we have argued that it is important to frame the theoretical knowledge base that contributes understanding to the open content model. The existing research on open source software has only considered its traditional role in software development. However, the same open source philosophy can be applied to the collaborative creation of non-software information products, such as encyclopedias, books, and dictionaries. In these areas, the open content development model might have much greater potential of significant societal impact because it is not restricted to the domain of highly-trained specialists, as in the case the software developers who contribute to open source software. On the contrary, for example, any literate person with an Internet connection can contribute to Wikipedia. Whereas certain kinds of open content will always require special skills for contribution (such as open music and open video), such skills are much more widely dispersed among the general population than are software development skills, which promises much broader participation.

Open content is important as a new direction in the availability of information products, and has only recently started receiving serious scholarly attention (Cheliotis 2009). Just as open source software has become such a dominant force in the software landscape, open content is increasingly becoming a significant market factor in all the media it touches. Both subjectively-evaluated content, such as open music, poetry, fiction, and video, and objectively-evaluated content, such as maps, GPS navigation systems, and courseware, are gaining increasing importance. We encourage researchers to focus attention on this increasingly important domain of our information society.

## References

Aksulu, A., and Wade, M. 2010. "A Comprehensive Review and Synthesis of Open Source Research," *Journal of the Association for Information Systems* (11:11).

Armstrong, T. K. 2010. "Shrinking the Commons: Termination of Copyright Licenses and Transfers for the Benefit of the Public," *Harvard Journal on Legislation* (47:2), pp. 359–423.

Ballantyne, P. 2009. "Accessing, Sharing and Communicating Agricultural Information for Development: emerging trends and issues," *Information Development* (25:4), pp. 260 –271.

Budoni, A., Maurelli, P., De Bonis, L., Federici, P. A., and Temperini, M. 2007. "Integration of webgis and open content environments for self-empowering e-governance," *Urban and Regional Data Management UDMS 2007 Annual*, pp. 105.

Cheliotis, G. 2009. "From open source to open content: Organization, licensing and decision processes in open cultural production," *Decision Support Systems* (47:3), pp. 229–244.

Chiao, H. F. 2010. *Essays in information economics* (Doctoral Dissertation), United States -- Michigan: University of Michigan.

Collins, M. 2011. "Open Access Literature Review 2008–9," *Library Resources & Technical Services* (55:3), pp. 138–147.

Downes, S. 2007. "Models for Sustainable Open Educational Resources," *Interdisciplinary Journal of Knowledge and Learning Objects* (3), pp. 29–44.

Feller, J., and Fitzgerald, B. 2000. "A framework analysis of the open source software development paradigm," In *21st International Conference on Information Systems*ICIS '00Brisbane, Australia: Association for Information Systems, pp. 58–69.

Järvinen, P. 2008. "Mapping Research Questions to Research Methods," In *Advances in Information Systems Research, Education and Practice*IFIP International Federation for Information Processing, D. Avison, G. M. Kasper, B. Pernici, I. Ramos, and D. Roode (eds.), (Vol. 274), pp. 29–41.

Jin, L., Robey, D., and Boudreau, M.-C. 2007. "Beyond Development," *Information Resources Management Journal* (20:1), pp. 68–80.

Korman, K. 2006. "Exploring the digital universe," *NetWorker* (10:1), pp. 26.

Krogh, G. von, and Hippel, E. von. 2006. "The Promise of Research on Open Source Software," *Management Science* (52:7), pp. 975–983.

Lerner, J., and Tirole, J. 2001. "The open source movement: Key research questions," *European Economic Review* (45:4–6), pp. 819–826.

Lerner, J., and Tirole, J. 2005. "The Economics of Technology Sharing: Open Source and Beyond," *The Journal of Economic Perspectives* (19:2), pp. 99–120.

Liao, H.-T. 2006. "Towards creative da-tong," *International Journal of Cultural Studies* (9:3), pp. 395–406.

Nelson, M., Sen, R., and Subramaniam, C. 2006. "Understanding Open Source Software: A Research Classification Framework," *Communications of the Association for Information Systems* (17:1).

Niederman, F., Davis, A., Greiner, M. E., Wynn, D., and York, P. T. 2006. "A Research Agenda for Studying Open Source I: A Multi-Level Framework," *Communications of the Association for Information Systems* (18:1).

Okoli, C., and Carillo, K. 2013. "Beyond Open Source Software: An Introduction to Researching Open Content," *SSRN.*

Okoli, C., Mehdi, M., Mesgari, M., Nielsen, F. Å., and Lanamäki, A. 2012. "The people's encyclopedia under the gaze of the sages: A systematic review of scholarly research on Wikipedia," *SSRN eLibrary.*

Perez, J. C. 2007, February 17. "Wikipedia Breaks Into U.S. Top 10 Sites," *PCWorld.*

Pfaffenberger, B. 2001. "Why Open Content Matters," *Knowledge, Technology & Policy* (14:1), pp. 93.

Rossi, M. A. 2004. *Decoding the "Free/Open Source(F/OSS) Software Puzzle" a survey of theoretical and empirical contributions* (Department of Economics University of Siena No. 424), Department of Economics, University of Siena.

Scacchi, W. 2007. "Free/Open Source Software Development: Recent Research Results and Methods," In *Architectural Issues* (Vol. Volume 69)Elsevier, pp. 243–295.

Schweik, C., Evans, T., and Grove, J. M. 2005. "Open Source and Open Content: a Framework for Global Collaboration in Social-Ecological Research.," *Ecology & Society* (10:1), pp. 1–25.

Stallman, R. M. 2002. "Copyright and Globalization in the Age of Computer Networks," In *Free Software, Free Society: selected essays of Richard M. Stallman* (1st ed, )Boston: Free Software Foundation, pp. 133–154.

Stallman, R. M. 2009. *Free Software, Free Society: Selected Essays of Richard M. Stallman*, (J. Gay, ed.)CreateSpace.

Ven, K., Sörensen, K., Verelst, J., and Sevaux, M. 2008. "Stimulating information sharing, collaboration and learning in operations research with libOR.," *International Journal on Digital Libraries* (8:2), pp. 79–90.