Title:

Impact of Free and Open-Source Software on the Cultural Fitness of Scientific Research

Table of Contents

Chapters

Figures

Tables

1. Abstract & Key Words

1.1 Abstract

This study investigates the impact of Free & Open Source Software (FOSS) on the cultural fitness of scientific articles hosted on the *Public Library of Science* (*PLOS*). The concept of cultural fitness is defined as the rate through which a variant is reproduced across cultural generations and, in the context of scientific articles, is measured through three indices: citation count, number of social media posts on *Facebook* and *Twitter* that link the article, and number of user visits that its webpage attracts. Generally-speaking, FOSS and non-FOSS development present alternative traits, such as transparency vs secrecy, voluntarism vs commodification, and decentralisation vs centralisation, respectively. These as discussed from an evolutionary perspective through the lens of the concepts of cooperation and competition. Usage patterns of FOSS licenses across years, journals, and subjects are examined with chi-squared tests. With multiple linear regression analyses, the presence of FOSS in articles is found to increase the citation count and to decrease the number of social media posts referring to it, which were both statistically-significant results. This result is discussed with reference to current challenges faced by scientists, such as the replication crisis, scientific malpractice, and rising distrust for science from the public. Furthermore, this study suggests that there is a feedback loop between the evolution of science and the evolution of software, leading to their co-evolution. The entirety of the research project is run in the coding language *R*, with the source-code being made available on an online repository under a FOSS license.

1.2 Key Words

Cultural evolution; Scientometrics; Software License; Transmission Bias; Cooperation-Competition.

2. Introduction & Literature Review

2.1. The Research Question

The purpose of this project is to investigate the impact of the usage of Free and Open-Source software on the cultural fitness of scientific research. The concept of cultural fitness, analogous to that of genetic fitness, refers to the rate at which cultural information is reproduced across generations (Dawkins 2016; Boyd & Richerson 2005). Examples of cultural transmission include songs broadcast over the radio and memes shared over the internet, with each single song and meme being transmitted at a different rate compared to other entities of the same kind. In the context of scientific research, cultural fitness may represent the propensity of articles to be reproduced across a variety of media, thus reaching a wider public of readers and having a greater impact on the scientific community and broader society. Here, I take in consideration three indicators of the cultural fitness in the context of scientific research articles: 1) the number of citations received from other articles published on peer-reviewed journals; 2) the number of posts referring to it on two major social media websites, *Facebook* and *Twitter*; and 3) the number of visits from internet users that its webpage attracts. Together, these three indices — citation count, social media posts, and webpage visits — provide a fairly comprehensive measure of the cultural fitness of scientific research in the digital age.

The main explanatory variable tested to predict the cultural fitness of scientific articles is the license associated with the software used by their authors to conduct the research presented in the article, specifically, the software license which they quote in the article itself. Typically, this would be software for data analysis whose license is named in the 'Methods' section of the article. The reason why it is interesting to study the relationship between software license and cultural fitness in scientific publications is simple. Nowadays, software is virtually-omnipresent in science, being employed in the collection, management, analysis, and sharing of data (Hill *et al.* 2006), and it can be said that it has become virtually inseparable from the

scientific profession. Software often comes under a license stipulating its terms of use, which can differ widely across different types. Software licenses can be broadly divided into two categories, 'Free and Open-Source' and 'Proprietary' (Morin *et al.* 2012) (**Figure 1**). The label 'Free and Open-Source' refers to software that is 'free' as in 'free-speech,' meaning that it can be used, modified, and shared at liberty. A prerequisite for this is that its source-code, *i.e.* the information interpreted by the computer to run the program, is 'open,' namely made available to users for study. By contrast, the label 'proprietary' indicates that the software whose owners — *e.g.* individuals or companies — limit the ability of other users to interact with the software for commercial purposes. However, things are often not so clear-cut: in fact, there is Free and Open-Source software that is also sold for profit, while some Proprietary software is Open-Source (Fortunato & Galassi 2021).

Figure 1. Types of Software Licenses (from Morin *et al.* 2012)

	Name	Latest Version	Copyleft	Patent Grant[a]	Permits[b] Code Linking	Used by[c]
FOSS	BSD	2-Clause	No	No	Yes	Gabedit, Chemkit, SciPy
	MIT	1.0	No	No	Yes	Weblogo, APBS
	ECL	2.0	No	Yes	Yes	RCrane, Sakai Project
	Apache	2.0	No	Yes	Yes	Imagemagick, Autodock Vina, GenMAPP
	MPL	2.0	Partial	Yes	Yes	Firefox, Thunderbird
	LGPL	3.0	Weak	Yes	Yes	ClustalW/X, IMP, BioJava, Taverna Workbench
	GPL	3.0	Strong	Yes	No	R Project, Perl, Coot, OpenBabel, GROMACS
Proprietary	Traditional "bespoke"[d]		No	Varies	Varies	*Majority of scientist-created software*
	"Inspection only"[e]		No	Varies	Varies	*Satisfies minimum publishing & peer-review requirement*
	Commercial		No	No	No	MS Windows, iTunes, Acrobat
Hybrid	Any combination		Varies	Varies	Varies	Pymol, MySQL, BDB, Phenix

Notwithstanding, user experience typically differs markedly under different types of software licenses (*ibid*.), which could reasonably impact the cultural fitness of scientific research. For examples, generally-speaking Proprietary software (PS) can often be used only after incurring in a financial costs and the source-code is generally invisible to end-users, preventing them from knowing how the software works in detail — such as which algorithms it employs — and also from adding new functions to the program. By contrast, Free and Open-Source Software (FOSS) licenses allows every user to study and improve on the functionality of the program and also to reuse its components for different projects. Furthermore, some FOSS licenses ensure that all subsequent software deriving from FOSS retains its free, open-source status for all future generations of users, a feature called 'copyleft' (Morin *et al.* 2012). FOSS is therefore of great importance for the scientific community, which is founded on the

principles of freedom, transparency, collaboration, peer-review, replication, and reproduction in the pursuit of knowledge.

FOSS could increase the cultural fitness of a scientific study in several ways. First, the transparency of the open-source ensures that those reviewing a scientific study can follow the entirety of the investigative process step-by-step, including for the collection, management, analysis, and sharing of data. Second, the absence of paywalls means that everybody with a computer and an internet connection can attempt to re-analyse (*i.e.* examine the same data), replicate (*i.e.* conduct the experiment using the same methods) or reproduce (*i.e.* conduct a similar experiment) the study (Juristo & Vegas 2010) with fewer impediments. Third, its free status allows other researchers to re-use part of all the source-code in other research project. FOSS would therefore make research studies more interesting to both scientists and the public, increasing the number of times they are read, cited, and shared.

Furthermore, FOSS could provide an answer to three growing concerns within the scientific community. First, there is a replication crisis, namely the inability to replicate many of the results claimed by other studies (Munafò *et al.* 2017). Second, there is a reported increase of distrust in science from the general public, whom perceives it as at times unnecessarily hard-to-follow and as potentially colluded with political and/or economic interests (Funk 2017). Third, just this year a survey questioning 7,000 researchers in the Netherlands found that half of the respondents engaged in questionable research practices and a staggering 8% falsified data (de Frieze 2021). Consequently, it is imperative to devise ways to increase the reliability of scientific research. As a solution, scientific research employing FOSS is easier to access and replicate and also harder to fake. The switch to FOSS in science is therefore also supported from ethical and epistemological standpoints. Furthermore, if it is found that the use of FOSS in scientific research is positively associated with its cultural fitness, it would mean that scientists interested in their research being widely read and shared should consider using free open-source software and join the FOSS community.

Of course, there are also other reasons for why some scientific articles have a higher cultural fitness than others beside the type of software they employ. Evolutionary research indicates

that a number of transmission biases affect the reproduction of a cultural variant. One is the prestige of the people displaying it, intended as their reputation, social standing, and/or actual or perceived expertise (Boyd & Richerson 2005). In the case of scientific publications, prestige can be conferred by different entities: by the journals in which they are published, which have different breadths of readership and standards for publications (Roberts 2009); by their authors, who have varying degrees of popularity, experience, and skill; and by their institutional affiliation, such as university and department, which rank differently in terms of in research quality. Another relevant factor is the time of publication. In fact, if newer research had less time for being read and shared, it is also true that today there are more scientists and people interested in science, also as a consequence of the Covid-19 pandemic (British Science Association 2020). Therefore, newer research may have more readers and sharers, whereas, by contrast, older research had a smaller readership, fewer media outlets for dissemination, and could have followed procedures that today are considered substandard. Yet another factor is the field of enquiry (*e.g.* biology, engineering, *etc.*), which may employ field-specific methodologies and attract readerships of different sizes. Therefore, these other factors needs also to be explored with regards to the cultural fitness of scientific articles.

Alternatively, articles may have a higher cultural fitness simply because they contain information that is more interesting and/or more valuable. This of course is a factor that is harder to quantify. Indeed, the whole purpose of science, through processes such as peer-review, replication, and reproduction, is to provide such a measure. Even though citation count from peer-reviewed publications may be the best index we have to assess the quality of paper, it is not without its problems. First, it may be affected by the above mentioned biases. Second, popular research may not necessarily be true and/or useful, and conversely, true and/or useful research may not always be popular (Aksnes *et al.* 2019). Third, not all citations imply agreement between those who cite and those who are cited, as in some cases researchers quote each other in a critical vein. Fourth, the citation count may be artificially inflated through unethical practices such as excessive and/or out-of-context self-citation (Hutchins *et al.* 2016; Van Noorden & Chawla 2019). Notwithstanding, the citation count remains one of the most widely employed measure by academic institutions to assess the quality of research output (Aksnes *et al.* 2019). Here, I take an agnostic perspective on

whether citation count and other indicators of cultural fitness efficiently measure research quality. What interests me more in this case is the pattern of the spread of cultural information across journals, which is adequately indexed by citation count. Furthermore, citation count is a simple quantitative measure signalling at least interest from other researchers and it can be obtained for virtually any peer-reviewed article, making it a prime candidate for cross-article comparisons.

With the advent of new digital technologies, the way people access and share information has changed. Today, an increasing number of people, including scientists, do not read and write articles on paper but on digital devices such as computers, tablets, and smartphones. These technologies enable users to access, store, and analyse data at almost any time and in any place, generally in a more reliable and efficient fashion (Sackstein 2015). Furthermore, digital services such social networking websites greatly increase the speed and reach of communication, enabling users to share information with virtually anybody on the planet at the touch of a finger. Concurrently, scientific discourse has partially moved beyond the exclusive locus of peer-review journals and also occurs on social media (Allgaier *et al*. 2013). In fact, nearly all research institutions and many researchers have one or several social media accounts that they use to interact with each other and the wider public.

In view of these developments, the number of social media posts that refer to a scientific article is therefore an important measure of its cultural fitness. Unlike citation count, which can only be increased by other scientists, anybody on social media can affect this score, providing a better measure of the popularity of articles among the broader population. Furthermore, social media have a readership base that is many orders of magnitude greater than that of peer-reviewed journals. For example, the two major social media websites analysed here, namely *Facebook* and *Twitter*, have 2.85 billion and 3.30 million active users, respectively (https://www.statista.com). By comparison, *Science* and *Nature*, the two widest read peer-reviewed science journals in the world, have only 130 thousand (https://www.sciencemag.com) and 63 thousand (https://www.nature.com) subscribers. Today, social media debate is an important aspect of professional and amateur science (Allgaier *et al*. 2013) and it is likely that it will continue to gain relevance in future, thus making

the number of social media posts linking to scientific articles an important measure of their cultural fitness.

Another index of article cultural fitness is the number of user visits paid to the webpage hosting each article. This represents the number of people who have accessed and/or read the article from that digital source. It should be noted that this measure could be increased by multiple visits by the same user. Notwithstanding, a greater visit count is representative of greater interest for the article. Visit count is therefore the third proxy of cultural fitness selected for analysis here.

Overall, testing the association between measures of the cultural fitness of scientific research and the type of software license it uses, each with their specific affordances, will increase our understanding of the way in which science and software evolve and coevolve. It will shed light on the transmission biases that take place in the context of the diffusion of scientific information and on the utilisation and impact of different types of software on scientific research. This will lead us to appreciate science and software better and, hopefully, make us more conscientious researchers and software developers.

2.2. Evolutionary Perspectives on Science and Software

The two evolutionary processes that are the focus of this research project — the evolution of software and the evolution of science — are two subprocesses of a more general phenomenon, namely cultural evolution. Cultural evolution is a relatively new area of scientific enquiry that emerged in earnest at the beginning of the 1980s (Laland *et al.* 2014). In truth, already in the 19th century pioneer of evolutionary thinking Charles Darwin hypothesised that evolution may not only shape species but also other entities such as behaviour and language (Darwin 1871) However, the more recent and formal antecedents of cultural evolution can be traced back to the Modern Evolutionary Synthesis of the 1940s, which brought together insights from various disciplines such mathematics, genetics, epidemiology, and biology, in an effort to shed light into evolutionary questions (Brown &

Richerson 2014). This led to an explosion in the number of research programs that sought to collect empirical data and test evolutionary hypotheses. Some researchers started to study the transmission and evolution of non-genetic entities, such as of cultural beliefs and practices (Cavalli-Sforza & Feldman 1981). Subsequently, other researchers noted that cultural evolution could feedback onto other processes, such as onto genetic evolution, yet they mostly limited their analysis to that of mathematical models (Boyd & Richerson 1985). It was only at the beginning the 21st century that cultural evolution acquired mainstream recognition as a major force in the evolution of several species, especially our own (Laland *et al.* 2013). Today, Cultural Evolution is a thriving research field tackling phenomena as diverse as the evolution of paper planes (Caldwell & Millen 2008), marriage systems (Fortunato 2018), and political systems (Currie *et al*. 2010), among others.

At its core, all Darwinian evolutionary processes rely on three fundamental principles (Boyd & Silk 2014). The first, *the principle of reproduction*, states that evolutionary entities should be able to reproduce their characteristics across generations. For example, it is observed that biological offspring will be more similar to their parents compared to other members of their group (*ibid*.). The second, *the principle of variation*, requires mutations to be stochastically introduced in the evolutionary process, so that offspring will not always be exactly identical to parents. For instance, even though cells that have undergone mitosis — *i.e.* split into two as to produce two daughter cells — should have exactly the same genetic make-up of the mother cell, sometimes they presents one or more different nucleotides in their DNA sequence as a consequence of radiation or chemical agents (Zimmerman 1971). The third, *the principle of competition*, expects that the resources necessary for the continuation of life are limited, thus engendering a struggle between individuals for preferential access to these. For instance, male mammals, among others, are known to routinely fight each other for territory and mates (Hunt *et al*. 2009). As a consequence, the individuals fitter to face evolutionary challenges (*principle of variation*) will have access to more and better resource (*principle of competition*) and will leave more offspring, which will be similar to them (*principle of reproduction*). Over time, this will lead to the evolution of traits among populations by natural selection.

Evolutionary theory has been applied with success to the study of genetic inheritance. However, it appears that it holds explanatory power also at other levels of analysis, such as in ontogenesis — the process by which genetic instructions are developed into biological structures. The development of neuronal networks in children, for example, has been observed undergoing a Darwinian process through which useful connections between areas are strengthened while those less utilised are removed, a phenomenon dubbed "synaptic pruning" (Paolicelli *et al*. 2011).

Furthermore, some animals — from cetaceans, to primates, to birds — also acquired the ability to develop a culture, namely, to transmit information between individuals in ways other than genetic inheritance, such as by observing others. Among *Pan* (which includes chimpanzees and bonobos) and *Tursiops* (bottlenose dolphins), two of the most cultural genuses, there are many inter-population behavioural patterns that cannot be explained by genetics alone (Sommer & Parish 2010). In *Homo*, our genus, the ability of culture reached a hitherto unseen level of sophistication and importance. Unlike other animals, we are able to faithfully imitate step-by-step the behaviours we observe in others, therefore generating stably-reproduced patterns of behaviour across time (*principle of reproduction*) (Whiten 2009). Variation occurs when individuals introduce a cultural novelty either by chance or by choice (*principle of variation*) (Mesoudi *et al.* 2004). Furthermore, cultural variants compete between each other for space inside brains, social institutions, and technological media (*principle of competition*) (Dawkins 2016). They also interact with the genetic fitness of the organisms that adopt them, making them more or less likely to reproduce and, therefore, to spread their specific cultural variants (*ibid*.). The presence of all the three evolutionary requirements explains why us and earlier *Homo* alone (Andersson *et al.* 2014) foster cultures that evolve, so that beneficial cultural variants are selected among populations over time.

Another key characteristic of evolution, which is made especially apparent by cultural evolution, is that multiple traits can affect each other's reproductive fitness, which leads to their co-evolution (Boyd & Richerson 2005). For example, the adoption of cattle-rearing and dairying, two cultural behaviours, by populations selected for genes granting the capacity to digest lactose even in adulthood (Holden & Mace 2009). These, in turn, made dairying more rewarding from an energetic perspective, thus consolidating the importance of cattle-rearing

practices. As a consequence of this and similar processes, genetic and cultural adaptations enter a positive-feedback loop. The mutual reinforcement developing between traits evolving across generations is generally known as co-evolution. The realisation that the effects of traits — whether they be genetic, ontogenetic, cultural, *etc.* — can affect the fitness of a great number of other traits across the genetic, ontogenetic, and cultural domains has led some contemporary evolutionary scientists to regard co-evolutionary processes as the norm, rather than exception (Laland *et al*. 2014).

Compared to other evolutionary processes, cultural evolution is much faster, potentially allowing for novel behaviours to be adopted by a population within a single generation. In fact, cultural information not only travels in a vertical fashion from parent to offspring (as genetic information does) but also between unrelated individual of different generations (oblique transmission) and of the same generation (horizontal transmission) (Ram *et al.* 2019). However, the transmission of cultural information is not without its dangers: being faster than genetic transmission and allowing for a greater degree of freedom in terms of variation, it can quickly bring about the proliferation of novel solutions but also lead to the spread of maladaptive cultural variants in the short-term (Mace 2014). However, theoretically, the fitness-enhancing effects of some cultural variants leads to their selection of beneficial traits in the long-term.

Past cultural adaptations can also act as a baseline from which further cultural variation can emerge, thus allowing individuals to produce complex inventions that could not be otherwise achieved in a single lifetime. This process, unique to human beings, is called *cumulative cultural evolution* (Mesoudi & Thronton 2018). For example, while primates have been using unmodified stone hammers for millions of years, eventually early *Homo* improved over them by fashioning first one-faced stone tools between 3.3 and 2.6 Ma (millions-of-years-ago) (Semaw *et al.* 2006; Harman *et al.* 2015), then a bi-faced stone axes around 1.76 Ma (Moore & Preston 2006). In the last 10,000 years humans discovered how to melt bronze and later iron to create more durable and effective tools (Armit *et al*. 2014). Only in the last few centuries, we invented machines capable to harness sources of power such as steam and electricity. No single human individual could have achieved these complex invention on their own within a single lifetime, without the accumulated wisdom of past innovators.

Moreover, the speeding up of the rate of cultural evolution is probably due, beside to rising population densities that facilitate the emergence and spread of variants (Henrich 2004), also to what has been called the "cultural evolution of cultural evolution" (Birch & Heyes 2021). Once cultural evolution was put into motion by our sophisticated social learning abilities, further adaptions such as psychological attentional biases, first, and explicitly linguistically-encoded learning biases, after, improved the quality and the speed of the cultural information transmitted. Eventually, culture "self-assembles" (*ibid*.) into continuously more complex and adaptive forms through a meta-cultural evolutionary process.

As a case study, I will briefly analyse evidence for the "cultural evolution of cultural evolution" through the history of software, one of the main focuses of this dissertation. The invention of the digital computer is less than 90 years old, building on previous 19th century computational machines (Campbell-Kelly *et al.* 2019), through a process of cumulative cultural evolution. The first digital computers could only solve the mathematical problems they were built to solve such as, for example, trigonometry. In 1941, another cultural cumulative milestone was reached with the invention of the first programmable computer by Konrad Zuse (2010). Now the functionalities of machines could not only evolve by improving their hardware, but simply and more flexibly by updating their software, *i.e.* a set of informational inputs such as numbers and instructions that programs the operations performed by the computer. Inventing software and through it improving the functionality of hardware equates to a process of cultural evolution by cultural evolution (Birch & Heyes 2021), for which a cultural improvements speeds up the rate of all future cultural improvements. Since then, computers and software have entered almost every aspect of our lives, producing extremely fast advances in the fields of, for example, biomedicine, finance, and space exploration.

Since the mid-1980s, a new cultural trend resulted in a faster rate and new modality with which software was developed. If, until then, software was produced individually or by closed-off groups such as universities, governmental agencies, and private companies, some people, often strangers, started sharing software on the internet under Free and Open-Source licenses and working collaboratively on them (Raymond 2001). One the first and most famous projects developed in this way was the operating system LINUX, whose current updates are

still today highly favoured compared to other operating systems by computer specialists (*ibid*.). Today, many people share their source-code, often under FOSS licenses, and help each other improve it on platforms such as *Stackoverflow* (https://stackoverflow.com) and *Github* (https://github.com). This way of developing software is interesting to the evolutionary scientists for three reasons: first, it speeds up the rate of cultural evolution through a process of cultural evolution (Birch & Heyes 2021); second, strangers on the internet, biologically-unrelated and often without previous history of interaction continuously share resources with each other, providing relevant case-studies on the processes through which cooperation between individuals evolves (Nowak 2005; West *et al*. 2007); third, this mostly amorphous and decentralised movement managed to produce top-quality software in multiple instances (Raymond 2001), providing an opportunity to study the emerge of complex, adaptive design from bottom-up processes.

Beyond technology, cultural evolution has created many other constitutive elements of human life — from language, to morality, to science (Hill *et al*. 2009), with the latter being the other main focus of this research project. Science is an activity involving a set of epistemological principles, social institutions, and technologies marshalled towards the goal of attaining greater and better knowledge about the world. In its most basic components, the pursuit of knowledge is a basic human and more broadly animal instinct produced by the brain, including by our dopaminergic system (Asma & Gabriel 2019). Dopamine-induced knowledge-seeking, regarding topics ranging from the location of resources to the causal interrelationships between events (*ibid*.), generates proto-scientific behaviours much earlier than the emergence of what we today would consider 'science.' Over time, especially since the 18th century, these traits culturally-evolved and coalesced with other social and technical adaptations to produce an organised activity that produce objective knowledge with unmatched success (Hepburn & Hanne 2021). Key features of the scientific enterprise are the formulation of hypotheses, *i.e.* explicit postulations regarding how the world works, the design of experiments, *i.e.* controlled settings for the observation of phenomena and the corroboration of hypotheses, and the use of mathematical tools, such statistics, to analyse data (*ibid*.). Furthermore, scientists publish their results in journals, where first they are subjected by peer-review by the editorial team and later by the rest of the scientific community. According to philosopher of science Karl Popper (2005), observation, analysis,

and peer-review disproves those theories whose predictions do not conform to reality and favours those that do. In the language of Cultural Evolution (Mesoudi *et al.* 2004), scientific theories are variants (*principle of variation*) competing for the attention of scientists, the space in publications, and funding from institutions. Theories with less explanatory power are selected out over time through observation, analysis, and peer-review (*principles of competition*), while better theories are reproduced among individuals, institutions, and journals (*principle of reproduction*). Furthermore, new theories are built on the wisdom of previous successful ones, cumulatively leading to the formulation of more sophisticated models (Mesoudi & Thronton 2018). An illustration of this is the systems of theories known as classical physics, produced by, among others, Galileo Gailei and Isaac Newton, with previously unseen explanatory power for predicting, for example, the fall of objects and the behaviour of planets. This was later improved upon and superseded by the theory of relativity, formulated by Albert Einstein, which had even more accuracy in tracking the movements of planets (Renn & Sauer 2007).

In view of the evolutionary patterns observed in the development of science and of software, their evolution and coevolution are legitimate and important fields of enquiry for evolutionary scientists. Their study promises to understand scientific and software-developing processes better and possibly to improve on them — in the spirit of the cultural evolution of cultural evolution (Birch & Heyes 2021). Concurrently, there has been a growing interest for discussing the role of software in science and what constitute optimal scientific practice in the light of newly-emerging digital technologies. For example, in the context of a symposium titled "Discussion with Prof. Richard McElreath: Science as Amateur Software Development", in which I had the opportunity to personally participate in on the 8th of March, 2021 (which can be watched in its entirety on *Youtube*: https://www.youtube.com/watch?v=zwRdO9_GGhY), much of the discussion verted on the importance of educating the scientific community on being more transparent with data, workflows, and algorithms utilised in research, as to ensure the safeguarding of highest scientific standards. Investigating whether the use of FOSS is associated with a higher cultural fitness for scientific publications is therefore a timely and necessary endeavour.

Previous research on science and software, which has in part inspired this study, include a select number of recent papers. The study by Smaldino & McElreath (2016) explicitly treats the scientific community as an evolving population in a statistical simulation which models labs as reproducing organisms. They were able to show that if less rigorous science produces purported significant results more easily by lowering methodological standards, it reproduces at the expenses of harder-to-make but more reliable science. As discussed earlier, the ability of cultural information to spread quickly can and does produce cases of maladaptation. Yet, I would challenge the findings of this simulation by arguing, as an alternative hypothesis, that in the longer-term less rigorous science as it fails to be replicated is selected out through peer-review. From empirically reviewing half-a-century of scientific publications, they also found that statistical power — *i.e.* the ability to correctly reject false hypotheses, which is a function of sample size — has not increased over time (*ibid*.), which should be considered with due seriousness, as this increases the probability of false-positives and false-negatives, to the detriment of the entire scientific evolutionary process. Using a similar simulation, O'Connor (2019) found that labs employing more traditional but reliable methods would have an advantage over those attempting newer procedures with a highly bimodal success-fail outcome. If this is regrettable for the loss of opportunities from discoveries that riskier, original science promises, for the evolutionary scientist the finding may represent nothing new under the sun, as it is at times said that evolution is conservative. It may be the case that the evolution of science proceeds gradually as the evolution of species (Barnosky 1987). In another study, Smaldino & O'Connor (2020) suggested that reviewers are biased to assess favourably methods similar to those they themselves use and that interdisciplinary projects can be a remedy to stagnating methodological outlooks. Undoubtedly, in view of all these studies, in the last half-decade evolutionary scientists have set their sights on the study of science, especially through theoretical modelling.

More recently, other researchers have directly tackled the question of the co-evolution of science and software with empirical research projects, which often have resulted in more optimistic findings. For instance, Minocher *et al.* (2020) surveyed 63 years of social learning research, finding that data could be recovered either on the paper, on the web, or by contacting authors in 30% of cases and that this proportion has increased exponentially over time. Data-availability is extremely important for peer-review, allowing other scientists to re-

analyse and replicate results. In another study, Colavizza *et al.* (2020) scanned *PLOS* (*Public Library of Science*) and *BMC* (*BioMed Central*) with an algorithm that could classify studies according to whether they provided a link to a repository with research data. They found that 20.8% of PLOS and 12.2% of BMC articles did so and that this was also associated with a 25.36% (± 1.07%) higher citation count. Christensen *et al.* (2019) also studied the effect of data-availability on citation count, with the additional feature of treating as a natural experiment the introduction of journal policies that require the publication of data. They found that policies mandating data-sharing had no effect on citation count, perhaps as a consequence of poor compliance, but also that researchers publishing their data received on average 97 (± 34) citations more. It is clear that data availability is positively associated with citation count, which deserves further investigation as to the underlying mechanisms of this process. These results should engender further visibility to discussions over scientific practices in the digital age.

The research project described in this dissertation was partly inspired by this recent interest for evolutionary research on science and software. It attempts to investigate a previously unexplored facet of the question, namely the effects the type of software *license* used by researchers, with its consequences on the openness of the source-code and the freedom of utilisation of the software. Furthermore, after measuring its impact on citation count, it also expands the notion of cultural fitness for scientific papers to include two other avenues that in the last years have proven to be important context for scientific dissemination, namely social media posts and webpage access.

In the language of evolutionary science and game theory (Nowak 2006; Dawkins 2016), FOSS developers act as "co-operators" who pool resources together and increase the cultural fitness of each other's research. This behaviour, especially when it does not emerge from kinship, reciprocation, or top-down regulation, but instead between strangers on the internet, might be explained through "fitness interdependence" models (Aktipis *et al*. 2018), where all co-operators (or FOSS developers) have a stake into each other's welfare (or cultural fitness). Alternatively, the prestige gained by being known as a good software developer and co-operator might also justify individual investment if bouts are iterative and memory of authorship is conserved (West *et al*. 2007). By contrast non-FOSS developers, in game-

theoretic terms, “selfishly” produce exclusive software and may charge for its use. In other words, FOSS are “public goods,” whereas PS are “private goods” (Jaeggi & Gurven 2013). One question that interest the evolutionary scientist is whether FOSS is at risk of suffering the “tragedy of the commons,” where collective resources are overexploited and undernourished (Foster 2004). Another question is whether in the long run the best outcome in terms of software capabilities and research endeavours belongs to the free open-source community with its ethos of voluntarism, transparency, and decentralisation or to the proprietary community and its ethos of commodification, secrecy, and centralisation. This opposition in values were crystallised by early hacker and open-source proponent Eric S. Raymond (2001) in a classic essay where FOSS was metaphorically described as a “bazaar” and PS as a “cathedral.” Experience from evolutionary science indicates that similar tensions between cooperation and competition are at the heart of every evolutionary process (Nowak 2006). Analysing the impact of Free and Open-Source software on the cultural fitness of scientific research will shed light on 1) which among FOSS and non-FOSS leads to higher cultural fitness and 2) the modalities by which software and science co-evolve.

3. Methods

3.1 Tools for Scientometrics

As a field of empirical enquiry, the evolution of science and software, and more broadly of culture, is still in its infancy. Even though there has been much theoretical-mathematical research, the current challenge is devising research programs that bridge the gap between theory and empirical data on cultural and that can reliably measure and assess instances of cultural evolution.

One way in which science can be studied from an evolutionary perspective is by looking at publication and bibliometric data. This approach is generally referred to as bibliometrics (Pritchard 1969) or, more specifically in the context of science, as scientometrics (Mingers & Leydesdorff 2015). Scientometrics has drawn attention as it provides a tool to measure and predict research patterns across authors, institutions, countries, journals, and years, to name a few. It has also been used as a human resource tool, to screen potential researchers to hire, and as financial management tool, to decide which research projects to fund — resulting in a collaboration between scientometricians and economists (Diamond 2000). Here, however, the focus is not on researching how to hire better employers or fund more profitable research, but rather on how to make better science and understanding how scientific information is transmitted .

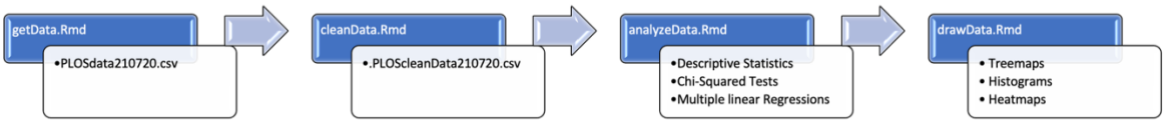
The infancy of both cultural evolution and scientometrics as disciplines means that it is not yet straightforward to find comprehensive analytical tools that suit the evolutionary researcher interested in the study of the evolution of science. Despite the existences of many libraries that allow to collect scientometrics, there is no single package in either *R* or *Python* that allows to gather all the information required by this and similar research projects. To investigate whether the presence of certain software licenses is associated with a higher citation count, social media presence, and webpage visit, it is necessary to interface between multiple scientometrics packages. After reviewing several options, including *bibliometrix* (Aria

& Cuccurullo 2017), *citationchaser* (Haddway & Grainger 2021), *fulltext* (Chamberlain 2021), and *easypubmed* (Fantini 2019), I eventually opted for utilising the libaries *rplos* (Chamberlain *et al.* 2021) and *rcrossref* (Chamberlain *et al.* 2020). The package *rplos* allows to gather publication data — such as article id, publication date, title, subject, journal, affiliation(s), author(s), *Facebook* posts count, and *Twitter* posts count, among others — as well as to search for specific text within the body of all the articles hosted on the *Public Library of Science* (PLOS), an online Open-Access publisher (https://plos.org). The package *Rcrossref* can fetch the citation count of papers from several different platforms, including *PLOS*.

3.2 Workflow of the Research Process

All the analyses in this study were performed with the coding language *R* within the programming environment *RStudio*, with the extension *R Markdown (.Rmd)*, and using some *R* libraries, which are all Free and Open-Source. In the spirit of FOSS, all the original code produced for this project is available under a Free and Open-Source license (GPL), which also ensures copyleft protection to all future software that derive from it. All source-code and the dataset are available on *GitHub*, a popular website for the sharing of software (link to repository: https://github.com/CEAdissertation/software_in_science). It follows a description of the workflow of the research process (**Figure 2**). By following these steps the researcher should theoretically be able to replicate the results of this study.

Figure 2. Workflow of the Research Process



The first source-code file to run is *getData.rmd*. This loads three libraries, the previously mentioned *rplos* and *rcrossref* which obtain publication data, and *dplyr* for general data management. The function *rplos::searchplos* is used to scan the *PLOS* archive for articles

referring to one of the seven major types of FOSS license, namely “BSD”, “MIT”, “ECL”, “Apache”, “LGPL”, “MPL”, and “GPL”— as found in Morin *et al.* (2012). The function is set to gather data regarding article id, publication date, journal, subject, number of webpage visits, number of *Facebook* and *Twitter* posts, affiliation, and author. Furthermore, two other variables were created, namely *license* (which one of the seven software license was mentioned in the articles) and *license_type* (in this case, all were FOSS). On the 22nd of July of 2021, the search yielded 29,716 articles referring to one of the seven FOSS licenses. The dataset was checked for duplicates — *i.e.* articles mentioning more than one license which would therefore be included twice in the dataset — and whose value for the variable *license* was intended to be set to ‘Hybrid FOSS.’ However, no duplicates were found.

The second search aimed to gather articles that did not mention any of the seven main FOSS license. Because *rplos::searchplos* regrettably lacks a random search function, it was necessary to search for articles still through a keyword. To produce reasonably random results, six of the most common words in the English language (Wikipedia 2021) selected because they can be habitually found in scientific publications were employed to simulate a random search, namely ‘when,’ ‘from,’ ‘one,’ ‘out,’ ‘about,’ and ‘than.’ The presence of these words should theoretically not impact citation count, software use, or other article characteristics. The function was set to obtain a sample of about the same size of the FOSS sample, namely 5,000 articles for each word for a total of 30,000 non-FOSS articles. Duplicates resulting from the various searches were deleted, which reduced their number to 24,294.

The citation count for each article was obtained using a different library because, puzzlingly, *rplos* has a function to obtain the number of social media posts linking to an article but not the number of scientific publications that cite it. The package *rcrossref* solves this problem through a function called *cr_citation_count*, which enables the user to obtain the citation count of articles by searching for them through their id. The potential replicator should be forewarned that fetching the citation count for all the articles from the internet through *rcrossref* was a relatively lengthy process, lasting about 4 hours. Only two citations counts could not be retrieved and therefore their articles were excluded from analysis. The total sample analysed, including FOSS and non-FOSS, included 54,008 articles. All the data was

ultimately aggregated within a single data-frame and stored into a *.csv* file called *PLOSdata210722.csv*.

The second step entailed the cleaning and tidying up of the data. To this purpose, the source file *cleanData.Rmd* performs several operations, including homogenising the names of the journals, because often the same journal was referred to with different spellings (*e.g.* "*PLOS Biology*" and "*PLoS Biology*"), and adding the number of *Facebook* posts to that of *Twitter* posts to produce a single count (*socialmedia_posts*). Furthermore, the number of days elapsed since the publication of each articles was obtained by subtracting the publication date from the download date. A new variable with the impact factor for each journal, obtained from consulting their respective *Wikipedia* pages, was also added. From the author column, which included all the co-authors of the article, only the first name was extracted for analytical purposes.

The subject and affiliation variables required a special treatment as they were constituted by two lengthy strings each with several subject and affiliation information for every article, which were also hierarchically-nested within themselves (*e.g.* for each row: "Subject1/Sub-subject1, Subject2/Sub-subject2 Subject3/Sub-subject3"; "Department1, University1, Country1, Department2, University2, Country2, Department3, University3, Country3,"). Examples of typical entries for subject and affiliate are shown in **Figure 3** and **Figure 4**, respectively.

Figure 3. An entry for article subject.

```
> SnS$subject[3]
[1] "/Medicine and health sciences/Cardiology/Heart rate,/Medicine and health sciences/Clinical medicine/Clinical
 trials/Randomized controlled trials,/Medicine and health sciences/Clinical medicine/Signs and symptoms/Edema,/Med
icine and health sciences/Clinical medicine/Signs and symptoms/Sepsis,/Medicine and health sciences/Epidemiology/M
edical risk factors,/Medicine and health sciences/Health care/Health care facilities/Hospitals/Intensive care unit
s,/Medicine and health sciences/Pharmacology/Drug research and development/Clinical trials/Randomized controlled t
rials,/Medicine and health sciences/Pharmacology/Drugs/Statins,/Physical sciences/Chemistry/Chemical elements/Oxyg
en,/Research and analysis methods/Clinical trials/Randomized controlled trials"
```

Figure 4. An entry for article affiliation.

```
> SnS$affiliate[6]
[1] "Pulmonology Institute, Department of Internal Medicine, Soroka University Medical Center and Faculty of Healt
h Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel,Clinical Research Center, Soroka University Med
ical Center and Faculty of Health Sciences, Ben Gurion University of the Negev, Beer-Sheva, Israel,Medical Intensi
ve Care Unit, Soroka University Medical Center and Faculty of Health Sciences, Ben-Gurion University of the Negev,
 Beer-Sheva, Israel,General Intensive Care Unit, Soroka University Medical Center and Faculty of Health Sciences,
 Ben-Gurion University of the Negev, Beer-Sheva, Israel,Duke Clinical Research Institute, Durham, North Carolina,
 United States of America,Clinical Biochemistry and Pharmacology, Soroka University Medical Center and Faculty of
 Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel,Institut de Pharmacologie et de Biologie
 Structurale, France"
```

I opted to extract only the first subject and the affiliation information before the first three commas (presumably the department, institution, and country of the first author) by parsing each string with the package *stringr.* This worked well for the *subject* variable, which was transformed into new variable, namely *main_subject*. However, it became clear that the output from parsing affiliation data would have not been utilisable in analysis, as very often the order 'department, institution, country' was not respected. Affiliation data was therefore dropped from further analysis. Furthermore, at a later stage of the analysis, it became apparent that the great majority of articles had a different first author, with a very small fraction of names appearing more than once. Analysing the effect of authors on the other variables for all 54,008 cases was therefore too computationally-demanding and eventually not very significant, which led to the dropping the *first_author* variable. Ultimately, the tidy data — including information for each article regarding their id, software license, type of software license, citation count, social media posts, webpage visits, journal, impact factor, days elapsed since publication, main subject, and publication date — was written into another file called *PLOScleanData210722.csv* (**Figure 5**).

The tidy dataset, containing 54,008 cases and 12 columns, was analysed with the source-code contained in the file *analyzeData.Rmd*. Descriptive statistics such as mean, median, standard deviation, counts, and relative proportions were obtained for each relevant variable. Furthermore, two-dimensional tables between categorical variables and other variables were computed. Furthermore, I performed four chi-squared tests between two couples of variables, namely either *license* or *license_type* with either *journal* or *main_subject*. In this way, I computed the difference between expected and actual values for the use of FOSS across journals and subjects.

Figure 5. Screenshot of the Tidy Dataset.

	id	license	license_type	citation_count	socialmedia_posts	webpage_visits	journal	impact_factor	days_since_pub	main_subject	year	publication_date
1	10.1371/journal.pone.0131329	Apache	FOSS	13	0	1945	PLOS ONE	2.740	2219	Biology and life sciences	2015	2015-06-25
2	10.1371/journal.pone.0152976	Apache	FOSS	13	198	5792	PLOS ONE	2.740	1919	Computer and information sciences	2016	2016-04-20
3	10.1371/journal.pone.0139374	Apache	FOSS	8	0	4448	PLOS ONE	2.740	2122	Medicine and health sciences	2015	2015-09-30
4	10.1371/journal.pone.0195628	Apache	FOSS	13	0	2068	PLOS ONE	2.740	1204	Biology and life sciences	2018	2018-04-05
5	10.1371/journal.pone.0109649	Apache	FOSS	21	0	3223	PLOS ONE	2.740	2484	Biology and life sciences	2014	2014-10-03
6	10.1371/journal.pone.0100514	Apache	FOSS	38	0	2698	PLOS ONE	2.740	2586	Biology and life sciences	2014	2014-06-23
7	10.1371/journal.pone.0234801	Apache	FOSS	1	1	1073	PLOS ONE	2.740	323	Computer and information sciences	2020	2020-09-02
8	10.1371/journal.pone.0196197	Apache	FOSS	15	0	2879	PLOS ONE	2.740	1176	Biology and life sciences	2018	2018-05-03
9	10.1371/journal.pone.0060561	Apache	FOSS	12	0	6728	PLOS ONE	2.740	3030	Biology and life sciences	2013	2013-04-05
10	10.1371/journal.pone.0066028	Apache	FOSS	15	0	3152	PLOS ONE	2.740	2974	Biology and life sciences	2013	2013-05-31
11	10.1371/journal.pone.0003226	Apache	FOSS	76	1	7188	PLOS ONE	2.740	4691	Medicine and health sciences	2008	2008-09-17
12	10.1371/journal.pone.0072792	Apache	FOSS	8	0	6594	PLOS ONE	2.740	2885	Biology and life sciences	2013	2013-08-28
13	10.1371/journal.pone.0162721	Apache	FOSS	16	4	8066	PLOS ONE	2.740	1757	Computer and information sciences	2016	2016-09-29
14	10.1371/journal.pone.0195445	Apache	FOSS	8	0	2600	PLOS ONE	2.740	1203	Biology and life sciences	2018	2018-04-06
15	10.1371/journal.pone.0226325	Apache	FOSS	6	0	2222	PLOS ONE	2.740	552	Biology and life sciences	2020	2020-01-17
16	10.1371/journal.pone.0051094	Apache	FOSS	16	0	3895	PLOS ONE	2.740	3149	Biology and life sciences	2012	2012-12-07
17	10.1371/journal.pone.0032286	Apache	FOSS	7	0	3113	PLOS ONE	2.740	3436	Biology and life sciences	2012	2012-02-24
18	10.1371/journal.pcbi.1008977	Apache	FOSS	0	0	0	PLOS Computational Biology	4.428	71	Biology and life sciences	2021	2021-05-12
19	10.1371/journal.pone.0221780	Apache	FOSS	3	0	573	PLOS ONE	2.740	675	Biology and life sciences	2019	2019-09-16
20	10.1371/journal.pone.0222974	Apache	FOSS	0	0	801	PLOS ONE	2.740	658	Biology and life sciences	2019	2019-10-03
21	10.1371/journal.pone.0237639	Apache	FOSS	0	0	965	PLOS ONE	2.740	337	Medicine and health sciences	2020	2020-08-19
22	10.1371/journal.pone.0209597	Apache	FOSS	11	0	2241	PLOS ONE	2.740	934	Biology and life sciences	2018	2018-12-31
23	10.1371/journal.pone.0141892	Apache	FOSS	12	77	11250	PLOS ONE	2.740	2030	Biology and life sciences	2015	2015-12-31
24	10.1371/journal.pone.0203642	Apache	FOSS	4	0	2818	PLOS ONE	2.740	1049	Biology and life sciences	2018	2018-09-07
25	10.1371/journal.pone.0097967	Apache	FOSS	9	0	3604	PLOS ONE	2.740	2603	Biology and life sciences	2014	2014-06-06
26	10.1371/journal.pone.0137721	Apache	FOSS	33	1	2715	PLOS ONE	2.740	2141	Biology and life sciences	2015	2015-09-11
27	10.1371/journal.pone.0246299	Apache	FOSS	1	0	490	PLOS ONE	2.740	164	Biology and life sciences	2021	2021-02-08
28	10.1371/journal.pone.0144259	Apache	FOSS	11	0	8961	PLOS ONE	2.740	2032	Biology and life sciences	2015	2015-12-29
29	10.1371/journal.pone.0135768	Apache	FOSS	8	2	1915	PLOS ONE	2.740	2164	Biology and life sciences	2015	2015-08-19
30	10.1371/journal.pone.0093687	Apache	FOSS	2	0	2952	PLOS ONE	2.740	2666	Biology and life sciences	2014	2014-04-04
31	10.1371/journal.pone.0098146	Apache	FOSS	16	9	6924	PLOS ONE	2.740	2605	Biology and life sciences	2014	2014-06-04
32	10.1371/journal.pone.0065283	Apache	FOSS	46	1	27127	PLOS ONE	2.740	2969	Biology and life sciences	2013	2013-06-05
33	10.1371/journal.pone.0191290	Apache	FOSS	16	0	4214	PLOS ONE	2.740	1261	Biology and life sciences	2018	2018-02-07
34	10.1371/journal.pone.0208563	Apache	FOSS	11	0	3746	PLOS ONE	2.740	961	Biology and life sciences	2018	2018-12-04
35	10.1371/journal.pone.0032328	Apache	FOSS	28	0	5810	PLOS ONE	2.740	3438	Biology and life sciences	2012	2012-02-22
36	10.1371/journal.pone.0146173	Apache	FOSS	2	0	2113	PLOS ONE	2.740	2023	Biology and life sciences	2016	2016-01-07

Showing 1 to 40 of 54,008 entries, 12 total columns

The main tests performed were multiple linear regressions. Multiple linear regression models were fitted on each of the three main variables, *citation_count*, *socialmedia_posts*, and *webpage_visits*. In the simplest models, each one of these variables were controlled with either *license* (eight levels: “BSD”, “MIT”, “ECL”, “Apache”, “LGPL”, “MPL”, “GPL”, and “Non-FOSS”) or *license_type* (two levels: “FOSS” and “Non-FOSS”). Each of the indices of cultural fitness were also controlled with each other. The more complex models also controlled each of the three cultural fitness indices for either *license* or *license_type* and for the time elapsed since publication and for journal impact factor. For the more complex models I employed the package *lme4*, which allows to fit a negative binomial model to offset the presence of many zeros or values close to it within the three variables indexing cultural fitness and to randomise the effect of *journal*. I also scaled each factor by subtracting its mean and dividing it by its standard deviation, to account for the widely differing scales of the values across variables, (*e.g.* relatively small and short-ranged *impact_factor* and widely ranging *webpage_visit*). All models controlling for the eight-level level variable *license* always used the level “Non-FOSS” against which the seven FOSS licenses were evaluated. Finally, I visually represented the descriptive and the inferential statistics with the source-file *drawData.rmd* using the libraries

treemap for expressing the relative prevalence of licenses, and *ggplot2* for histograms and heatmaps.

4. Analysis & Results

4.1 Descriptive Statistics

Out of the 54,008 scientific articles reviewed, 29,714 had FOSS licenses and 24,294 did not. FOSS-using articles therefore constituted 55.0% of the sample. The number of FOSS articles corresponded to all the hits that resulted from querying *PLOS* with each of the names of the seven main FOSS licenses (Morin *et al.* 2012). Among the FOSS licenses, ECL accounted for 49% of the sample, MIT for 39%, Apache for 4%, BSD for 3%, MPL for 3%, GPL 2%, and LGPL for less than 1% (**Figure 6**). Articles citing ECL were the most typical overall — even more than non-FOSS articles — from 2008 to 2012, before being overtaken by non-FOSS articles in 2013. Since 2017, MIT has been the most popular FOSS license. Overall, articles citing FOSS licenses were more prevalent than those who did not from 2003 — *i.e.* the very year of the creation of PLOS and of the publication of the earliest articles reviewed — to 2018, peaking in 2013. In the last few years this pattern was reversed, with non-FOSS articles being slightly more common (**Figure 7**).

Figure 6. Proportion of Licences in the Sample

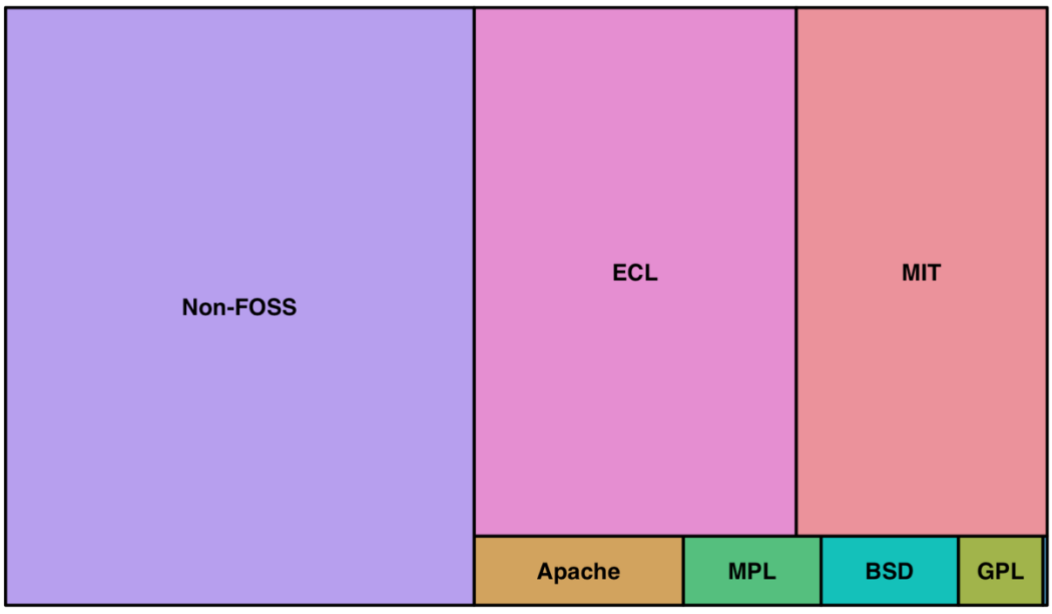


Figure 7

Licence Usage by Year

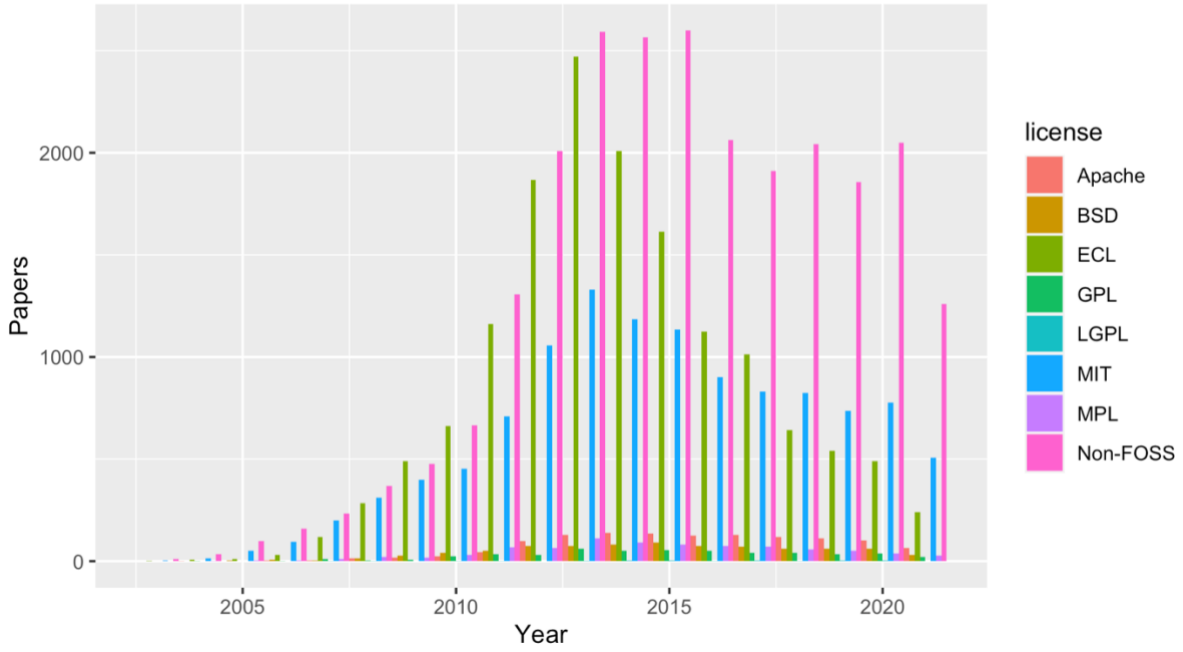
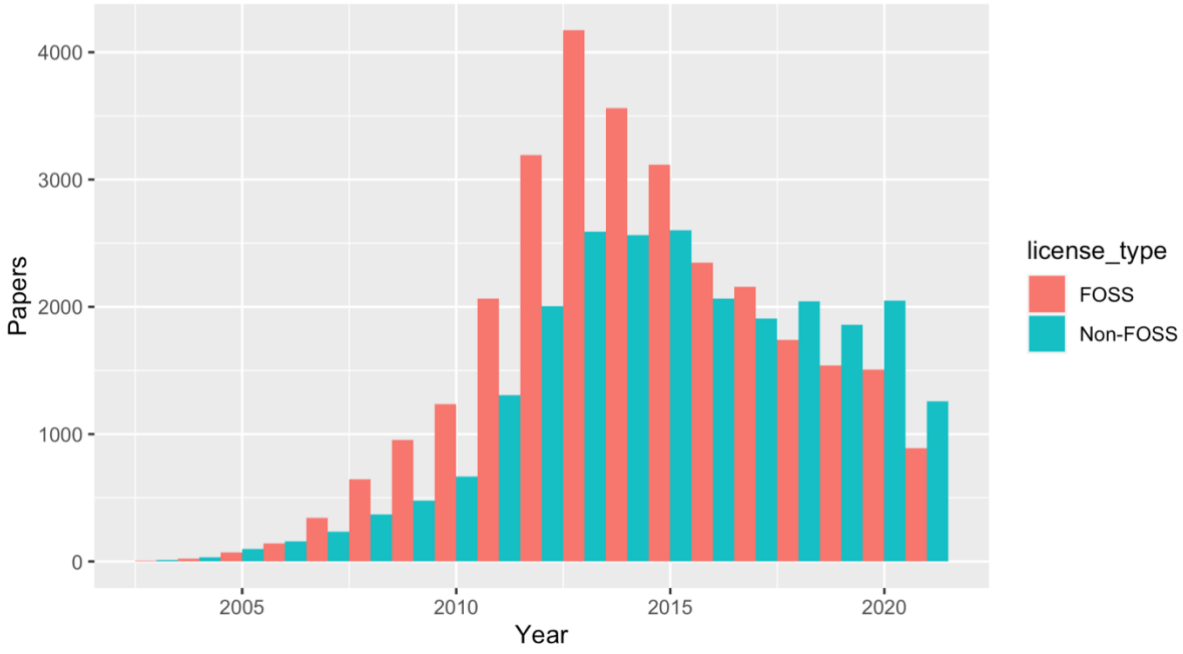


Figure 8

Licence Type Usage by Year

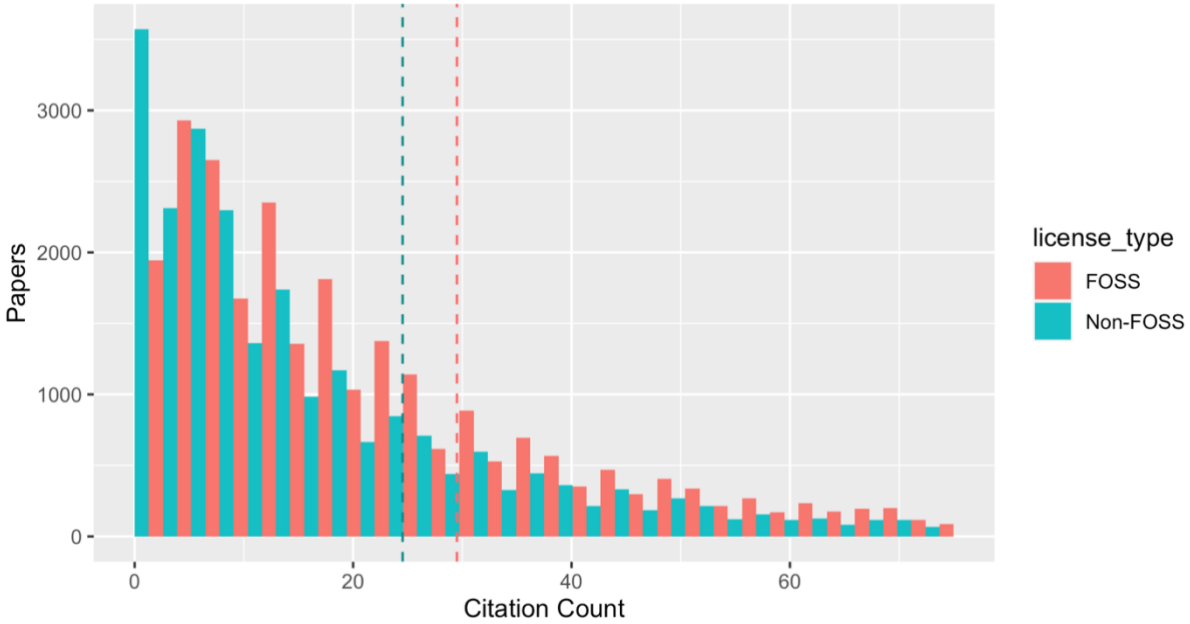


The citation count had a negative binomial distribution (**Figure 9**), taking only positive values and with the majority of them being equal or close to zero, with a maximum of 5,975 and a SD (Standard Deviation) equal to 68.8. The citation count of articles mentioning FOSS was less sparse (SD = 62.3) and had a higher mean and median — corresponding to 29.5 and 15, respectively — compared to non-FOSS articles — which sat at 24.5 and 11, respectively (SD =

75.9) . The license associated with the highest average citation count was LGPL (106.6), which was also the least common, followed by GPL at 36.6, BSD at 32.7, and MIT at 32.1 — which were instead the three most common licenses (**Table 1**).

Figure 9

Citation Count by Licence Type



The dashed lines represent the respective average citation counts for each group of papers.

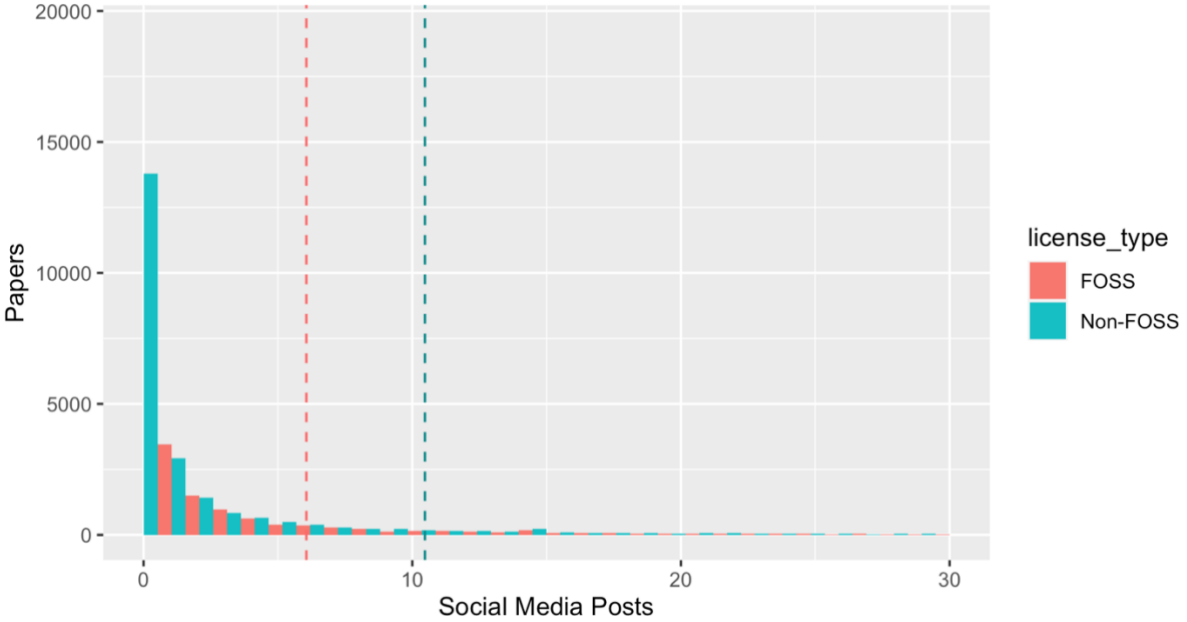
Table 1. Citation Count by License

License	Median	Mean	Max	SD
▪ **Non-FOSS**	**11**	**24.5**	**5,975**	**75.9**
▪ **FOSS**	**15**	**29.5**	**2,970**	**62.3**
○ Apache	12	25.9	922	56.1
○ BSD	15	32.7	1586	83.8
○ ECL	17	27.1	1019	35.2
○ GPL	13	36.6	2385	134.3
○ LGPL	17	106.6	1113	256.2
○ MIT	14	32.1	2970	80.0
○ MPL	16	31	638	55.5
▪ **All**	**14.0**	**27.2**	**5,975**	**68.8**

The number of social media posts had also a negative binomial distribution (**Figure 10**), with the most common value being 0 and reaching up to 7,184 (SD = 77.9). The number of social media was also less sparse for articles mentioning FOSS (SD = 58.5) but here the pattern for the mean was reversed: non-FOSS articles had a higher mean (10.47) than FOSS articles (75.9). The license associated with the highest mean number of social media posts was, in order, LGPL, MIT, Apache, and BSD. All medians by license sat at 0, expect for the relatively rare LGPL for which it was equal to 2 (**Table 2**).

Figure 10

Social Media Posts by Licence Type



The dashed lines represent the respective average citation counts for each group of papers.

Table 2. Social Media Posts by License

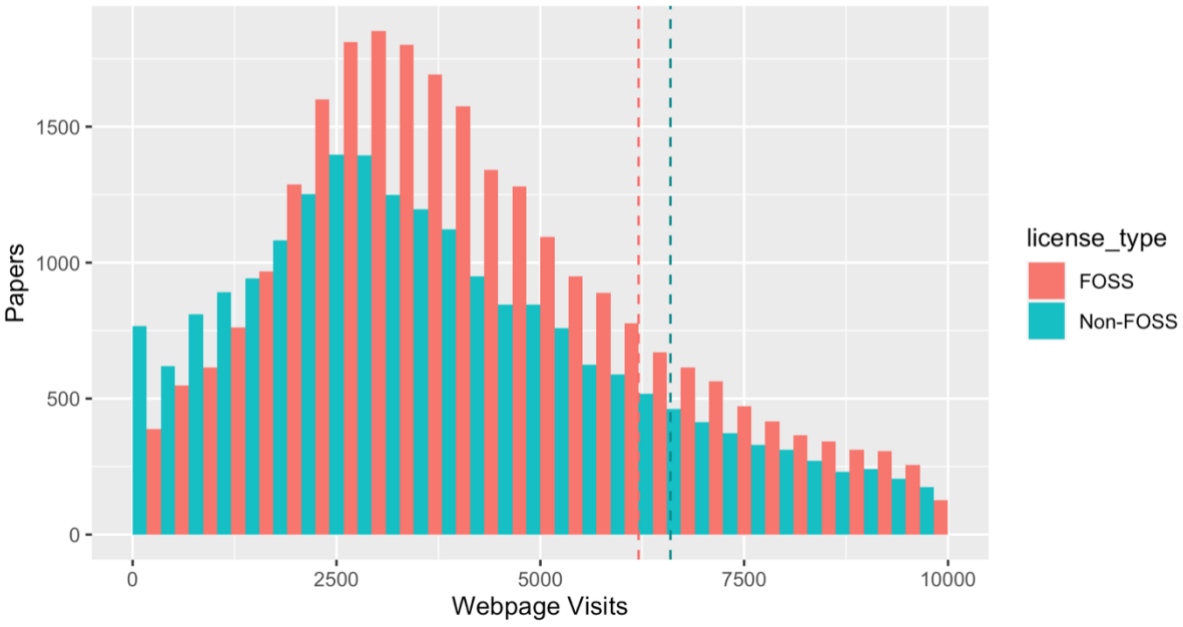
License	Median	Mean	Max	SD
▪ **Non-FOSS**	**0**	**10.47**	**7,184**	**93.6**
▪ **FOSS**	**0**	**6.1**	**5,207**	**58.5**
○ Apache	0	10.4	7184	96.3
○ BSD	0	9.3	713	34.6
○ ECL	0	2.1	1034	17.0
○ GPL	0	5.2	263	20.6
○ LGPL	2	24.8	379	77.6
○ MIT	0	10.8	5,207	88.1

○ MPL	0	5.3	906	42.6
▪ **All**	**0**	**8.0**	**7,184**	**77.9**

The number of webpage visits was right-skewed (**Figure 11**), with a median equal to 4,024 and an extremely long right tail reaching up to 2,629,673 (SD = 16,602). FOSS and non-FOSS articles had roughly the same average number of webpage visits (FOSS = 6,205; non-FOSS = 6,596), but the standard deviation for publications that did not mention FOSS was two times greater than that of those that did. The license associated with the highest mean webpage visits received was, again, LGPL (11,245), followed by MIT (7,475), with all the others sitting between 5,138 (ELC) and 6,886 (BSD) (**Table 3**).

Figure 11

Webpage Visits by Licence Type



The dashed lines represent the respective average citation counts for each group of papers.

Table 3. Webpage Visits by License

License	Median	Mean	Max	SD
▪ **Non-FOSS**	**3,775**	**6,596**	**2,629,673**	**21,395.33**
▪ **FOSS**	**4,186**	**6,205**	**691,519**	**11,255.97**
○ Apache	3,613	6,867	274,872	15,664.61
○ BSD	4,162	6,886	19,3562	12,256.96
○ ECL	4,131	5,138	269,804	5,074.857

○ GPL	4,118	5,915	57,987	6,692.526
○ LGPL	5,038	11,245	58,557	13,373.15
○ MIT	4,363	7,475	691,519	15,610.39
○ MPL	4,116	5,926	240,944	10,209.5
▪ **All**	**4,024**	**6,381**	**2,629,673**	**16,602.67**

4.2 Chi-Squared Tests

Among the journals published on PLOS, seven could be correctly identified and accounted for 97.5% of the articles, whereas for the remaining 2.5% the journal of origin was unknown (**Figure 12**). Across all journals, FOSS articles were more common — which could be explained by their slight prevalence in the total sample — except for *PLOS Neglected Tropical Diseases* where the pattern was reversed. By running a Chi-Squared test (X-squared = 1063.7, df = 7, p-value < 2.2e-16), the difference between actual and expected number of FOSS and non-FOSS articles across journals was computed. PLOS *ONE* (the largest journal)*, PLOS Medicine*, and *PLOS Neglected Tropical Diseases* were the only for which the number of FOSS articles was higher than expected, indicating a preference for using FOSS. by researchers publishing there. In *PLOS Genetics* and in *PLOS Pathogens,* articles mentioning FOSS licenses were instead considerably less common than predicted by random chance. Among individual licenses (X-squared = 3765.2, df = 49, p-value < 2.2e-16), in *PLOS computational biology* ECL (the overall most common license) was considerably more prevalent (expected: 698.1 less) whereas MIT was consistently avoided (expected: 773.1 more). In *PLOS ONE*, it was true the opposite: MIT was favoured (-782.7) while ECL was ignored (+172.2) (**Table 4**).

Figure 12

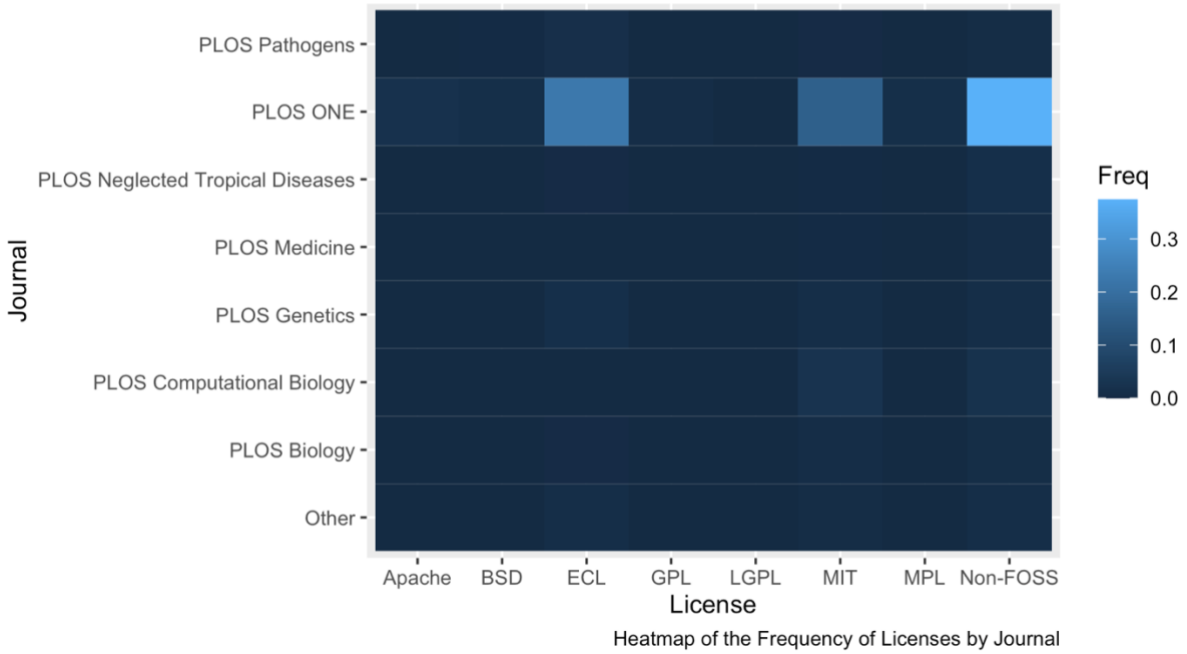


Table 4. Number of Articles by Journal and License Type			
Journal	**FOSS (Expected)**	**Non-FOSS**	**Total**
PLOS ONE	23,443 (-575.0)	20,212	43,655 (80.8%)
PLOS Computational Biology	1,585 (+92.8)	1,128	2,713 (5.0%)
PLOS Genetics	1,334 (+287.0)	569	1,903 (3.5%)
PLOS Pathogens	1,258 (+411.2)	281	1,539 (2,8%)
PLOS Neglected Tropical Diseases	436 (-241.2)	795	1,231 (2.2%)
PLOS Biology	675 (+64.8)	434	1109 (2.0%)
PLOS Medicine	124 (-140.0)	356	480 (0.8%)
Unknown	859 (+100.8)	519	1378 (2.5%)

The articles were categorised by PLOS as pertaining one or more of ten different subjects (**Figure 13**). Only the first subject mentioned was extracted for analytical purposes (see Methods section) Only for less than 0.1% of articles the subject could not be identified. The

great majority of articles (93.9%) were categorised as being about Biology & Life Sciences. It is not clear if this was an artefact of the way PLOS lists multiple subjects (with Biology & Life Sciences appearing first more than others — see Methods section) or this simply reflected an actual greater number of articles about biology on PLOS. Running a Chi-Squared test (X-squared = 1161.8, df = 10, p-value < 2.2e-16) on this dataset on RStudio displayed a message warning that results could be incorrect, possibly because the majority of subjects appeared very rarely (**Table 5**). This however does not invalidate the significance of any other analysis, as the variable subject was dropped from the multiple linear regressions.

The results of this test indicated that biologists preferred to use non-FOSS, while medics, in accordance with the results from the chi-squared test on journals employed FOSS more than expected, as well as computer scientists and engineers. The articles labelled as pertaining other subjects were too few too be deemed relevant.

Figure 13

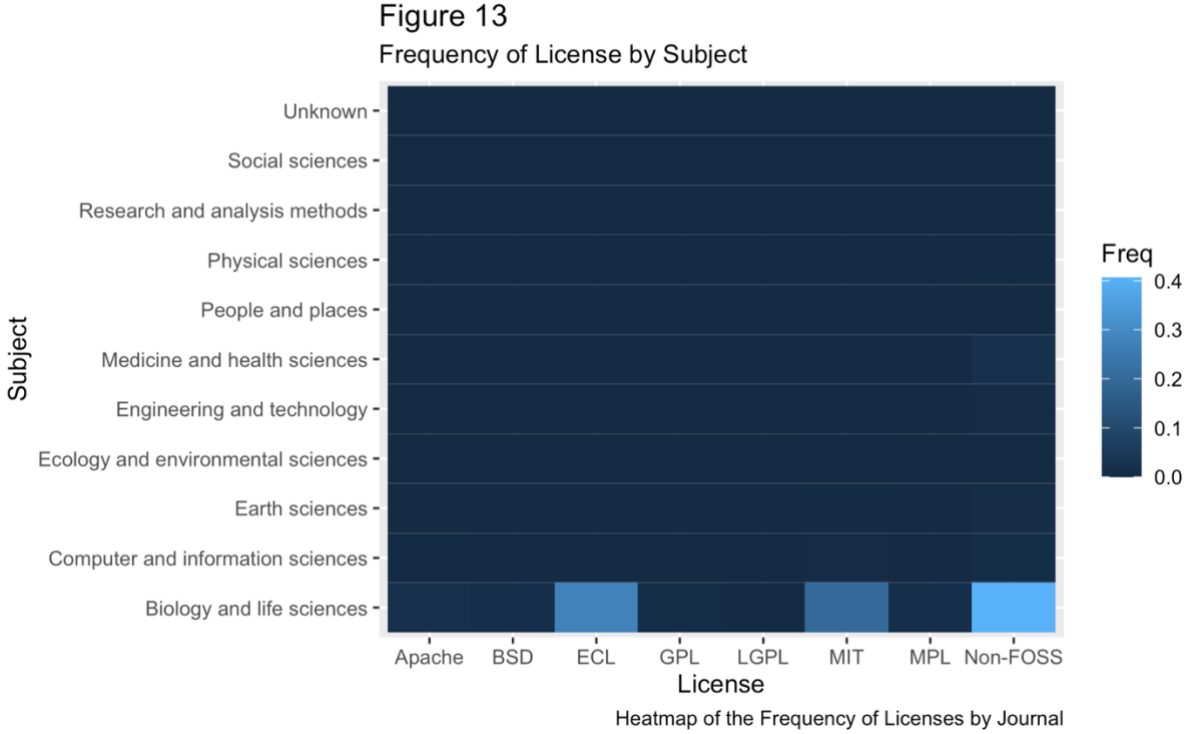


Table 5. Number of Articles by Subject and License Type

Subject	FOSS (Expected)	Non-FOSS	Total
Biology & Life Sciences	28,798 (+884.1)	21,938	50,736 (93.9%)

Medicine & Health Sciences	234 (-471.8)	1,049	1,283 (2.3%)
Computer & Information Sciences	389 (-132.5)	559	948 (1.7%)
Engineering & technology	84 (-78.3)	211	295 (<0.1%)
Earth Sciences	90 (+136.6)	322	412 (<0.1%)
Physical Sciences	41 (-27.2)	83	124 (<0.1%)
People & Places	37 (-16.4)	60	97 (<0.1%)
Ecology & Environmental Sciences	10 (-10.3)	27	37 (<0.1%)
Research & Analysis Methods	3 (+0.3)	2	5 (<0.1%)
Social Sciences	3 (+0.8)	1	4 (<0.1%)
Unknown	25 (-11.9)	42	41 (<0.1%)

4.3 Multiple Linear Regressions

In the simplest, two-variable linear models between indices of cultural fitness (citation count, social media posts, and webpage visits) and software license (as a binary FOSS/non-FOSS variable and as an eight-levels variable with licence names), the citation count was found to be positively correlated to FOSS use, while social media post was negatively associated with it. When testing individual licenses, non-FOSS articles were always used as reference level against which the effect of the FOSS licenses were measured.

In the linear model between citation count and the binary license variable, the use of FOSS increased the citation count of papers by 4.98 (±0.59) (p-value: < 2.2-16, read as negative sixteenth power of 2.2). Each individual FOSS license was found to be positively associated with citation count and — with the expectation of Apache — all the correlations were statistically significant. Among those with the greatest impact there was GPL (12.1579 ± 3.0848, p-value = 8.12e-05), BSD (8.2193 ± 2.4366, p-value = 0.000743), MIT (7.5887 ± 0.7780, p-value < 2e-16), and LGPL (82.1096 ± 13.7561, p-value = 2.40e-09). Plotting the residuals of either model showed that the model was homoscedastic, namely that they varied constantly with increasing citation count.

By contrast, in the models investigating the number of social media posts referring to the articles, this was found to be negatively associated with the use of FOSS, with non-FOSS having 4.4117(±0.6736) more posts on average (p-value = 5.84e-11). The only individual license that was found to have a significant adverse effect on the number of social media posts was ECL (-8.3585 ± 0.8119, p-value < 2e-16). Instead, MIT and LGPL were found to increase the number of social media posts, but not in a statistically-significant way. The residuals for these models were also homoscedastic.

The number of webpage visits paid to the article was also found to be negatively-associated with the presence of FOSS, with non-FOSS articles enjoying 391.18 (±143.60) visits more (p-value = 0.00645). About half of the individual licenses increased the number of webpage visits, while the other half decreased it, compared to non-FOSS. The only individual licenses with a statistically-significant impact on visits were the negatively-associated ECL (-1457.6 ± 173.0, p-value < 2e-16) and the positively-associated MIT (879.4 ± 187.7, p-value = 2.79e-06).

The three indices of cultural fitness were all positively correlated with each other. For example, each social media post referring to an article was found associated with 0.07097 (±0.00379) more citations from peer-reviewed source, whereas webpage exposure was increased by each citation by 79.5642 (±0.9798) and by each social media post by 99.7139 (±0.8104). All associations between indices of cultural fitness were highly statistically significant (p-values < 2e-16).

The full linear regression models, which controlled for the impact factor of journal and for the time elapsed since publications, also found a positive correlation between FOSS and citation count — although less marked — while both social media presence and webpage visits were negatively associated with it. It therefore mostly confirmed the results of the simple two-variable models, but with impact factors of FOSS and specific licenses on the indices of cultural fitness adjusted by the context and time of publication. The residuals of the full models fitted with the library *glmer* were heteroscedastic.

In the full model, citation count was found to be increased by using FOSS on average by 0.032514 ± 0.008711 (p-value = 0.00019). With each single day since publication, the citation

count grew by 0.837351 ± 0.005063 (p-value < 2e-16), and the impact factor of the journal increased it 0.101332 ± 0.057036, although with a significance level slightly above the commonly used .05 threshold. In the model testing individual software license, all FOSS licenses were found to be positively associated with citation count, with the exception of ECL which was negatively associated to it. Those that boosted the citation count the most were LGPL (1.62611 ± 0.19072, p-value < 2e-16), GPL (0.31262 ± 0.04399, p-value = 1.19e-12), and Apache (0.25379 ± 0.02865, p-value < 2e-16).

Controlling for time elapsed and journal impact factor confirmed that the social media presence of scientific articles is adversely associated with the use of FOSS, with non-FOSS research being linked by 0.56843 ± 0.02651 (p-value < 2e-16) more posts than its counterpart. Interestingly, while impact factor increased social media exposure, time elapsed since publications had instead a negative effect on it, albeit the latter result was above the .05 threshold level of significance. In the model with individual licenses, only LGPL was found to be associated with a rise in the number of social media posts, while the others were negatively associated with it.

In the full model testing the number of webpage visits, these were also found to be negatively correlated with the presence of FOSS (-0.165003 ± 0.008136, p-value < 2e-16). The number of visits resulted positively associated with journal impact factor was also associated (0.140620 ± 0.052844, p-value = 0.00779) and it grew by nearly half-a-visit per day (p-value < 2e-16). Individual FOSS licenses were mostly associated with a lower number of webpage visits, with the exception of LGPL (0.480206 ± 0.181888, p-value = 0.008288) and Apache (0.106924 ± 0.026576, p-value = 5.74e-05).

5. Discussion

5.1 Discussion of Results

In view of the results from the descriptive analyses run on the dataset, it is clear that the use of FOSS in science has gained momentum over the last two decades, likely as a reflection of the recent examination of the role of software in science (Minocher *et al.* 2020) and of the growing the concerns regarding the replication crisis (Munafò *et al.* 2017). In the very last three years, non-FOSS research seemingly gained instead the upper hand according to the results of this study. First, it should be noted that while the sample comprised every article on PLOS that cites one of the seven FOSS licenses, the number of non-FOSS articles analysed was arbitrary — chosen only to approximately match the size of the FOSS sample. Second, the causes dictating the relative proportion of FOSS and non-FOSS research are likely complex and multivariate, therefore I do not advance at this stage hypotheses as to why the trend seem, even though slightly, reversed for the last three years. Some individual FOSS licenses were clearly preferred by scientists compared to their counterparts, with ECL and MIT alone comprising 78% of the FOSS licenses. The preference may have been driven by whether they were copyleft licenses or not, as both ECL and MIT do not enforce the copyleft, whereas the licenses that were least used by scientists — namely MPL, GPL, and LGPL — were all copyleft licenses (Morin *et al.* 2012).

The incidence of FOSS was not equally apportioned across journals and articles. Regarding the use of FOSS in journals, the main publication outlet in the sample, namely *PLOS One*, and two of the journals that focus on health, namely *PLOS Neglected Tropical Disease* and *PLOS Medicine*, were more likely than expected to feature research employing FOSS. The other four journals plus those articles whose journal could not be attributed were less likely than expected to feature FOSS. The journals where the chi-squared test indicated that, all things being equal, FOSS was preferred indicated perhaps the presence of a self-selection process by which the scientists publishing in those avenues opted for FOSS more often.

Regarding the relationship of FOSS and main subject, especially articles regarding three subjects, namely Medicine & Health Sciences, Computer & Information Sciences, and Engineering & Technology, were much more likely to include FOSS than expected. This could potentially be explained by the fact that researchers in the fields focusing on technology may be more attentive to the software they use and the licenses stipulating their terms of use compared to their colleagues in other fields. In the case of those researchers dealing with computers and information, it may be the case that FOSS is more common because it allows to study their subject matter, *i.e.* digital technologies, in a greatly enhanced way. The reason why medical researchers were also more likely to use FOSS than others is not as easily answered, but could be related to the importance of clear and explicit analyses in studies that affect health related procedures, whose results generally have a deeper effect on human lives.

Each of the indices of cultural fitness, citation count, social media posts, and webpage visit, presented a negative binomial distribution with the majority of articles having zero or near-zero visibility and small minority forming extremely long tails. This is also called a Pareto distribution (Arnold 2014), where as a rule of thumb 80% of actors (which may represent articles, people, universities *etc*.) receive 20% of the resources (such as citations, posts, money, *etc*.) and the remaining 20% of actors amasses the other 80% of the resources. The emergence of this pattern of inequality could be due to previously discussed biases, such as those regarding prestige, or also by the accumulation of skills among more experienced actors. This finding is not surprising for the citation count, which is known to follow this distribution (Glanzel 2007). The Pareto distribution, incidentally, was also followed by the apportioning of articles across journals, where exactly 80.1% of articles are published in the main publication outlet hosted on *PLOS* out of 7 journals, namely *PLOS One*, thus in a ratio approximating the 20%.

The main hypothesis tested here, that the use of FOSS increases the cultural fitness of scientific articles, was supported by the linear regressions models testing some cultural fitness indices but not by others. In fact, interestingly, even though the three indices of cultural fitness analysed were all positively and strongly correlated with each other, they individually showed differential correlations in terms of directionality with the type of

software license that the articles presented. Citation count was found to be, as expected by the main hypothesis of this research study, positively associated with the use of FOSS licenses. By contrast, the number of social media posts and webpage visits was found to be negatively affected by FOSS. The high significance of all results indicates that this was likely not a product of random chance but that the use of FOSS affected the indices of cultural fitness of articles bidirectionally, with citation count going towards one side and social media posts and webpage visits going towards the other. This may be the case because citers, sharers, and readers belong to different, if partly overlapping, populations: the first is comprised by peer-reviewed scientists, the second by social media users, and the third by internet users. It is plausible therefore that scientists from one side and social media users and web surfers from the other were affected differently by the use of FOSS in research.

The reason why scientists were more likely to cite in their own papers research that employed FOSS could reflect, as hypothesised in the introduction, the fact that this makes studies easier to understand, re-analyse, and replicate. This increases the trust in the conclusions of the research articles employing FOSS, which would then result in more citations by colleagues. Social media and web users, which were not necessarily scientists, were instead less likely to share or read research including FOSS, which was a statistically significant trend. As a working hypothesis, articles containing FOSS may be more technical than their counterparts, hence being harder to read or having a too narrow focus to capture the interest of a population with a greater proportion of non-scientists.

Furthermore, the fact that FOSS increased citation count but decreased social media exposure and webpage visits suggests that the measures of cultural fitness studied index slightly differ phenomena and that scientific information is transmitted differently according to the medium where it is shared. Regarding this relationship several factors were also noteworthy. First, not all models predicted that FOSS negatively affected the number of times articles were shared on social media and visited. Second, in the mid-range, roughly between 1,300 and 10,000 visits, FOSS-using studies actually attracted more webpage visits than their counterpart, as shown in **Figure 10**. It might be the case that the cause for its negative association with FOSS resides in the extremely long righthand tail of the distribution of webpage visits. The same pattern, albeit slighter, can be seen in the distribution of the

number of social media posts, where between the range 1 to 3 posts it is actually FOSS articles that have the greater number of social media posts (**Figure 11**). It is not unlikely that a few large studies using non-FOSS might have gained popularity among non-specialists after being showcased in the news, greatly inflating the number of times they were shared and read. The curve that preferred sharing and reading FOSS articles at lower ranges might instead represent a subset of those populations, namely scientists on social media and on the web, the same that preferred to cite FOSS publications in journals.

It is interesting to note how the effects of time over the various indices of cultural fitness differed. While being an older publication predicted a greater number of citations — as well as more webpage visits — it was instead associated with a smaller number of social media posts referring to it, even though the timespan covered by this research project (2003-2021) is mostly overlapping with the advent of the two social media websites analysed — *Facebook*, founded in 2004, and *Twitter*, founded in 2006 (*Wikipedia*). This may be a reflection of the fact that the popularity of science and the phenomenon of sharing and debating science on social media has grown in more recent years and/or that the greater populations of social media users prefers to share newer research.

The impact factor of journals was instead always positively associated with all indices of cultural fitness. If this may be readily expected to increase the citation count, which is the product of a population of scientists whose reading and writing habits may be heavily shaped by the impact factor of journals, it may come as somewhat surprising that the impact factor was always found positively associated with the number of social media posts and webpage visits, which also indexed the preferences of non-scientists. This might mean that either 1) being published on journals with higher impact factor ensures that articles reach a wider readership which in turns cites, shares, and access the article more often or that 2) journals with a higher impact factor have higher standards for publication and/or publish more interesting material, which increases their spread rate. Without further evidence, the first hypothesis should be preferred because more parsimonious.

The linear regressions featuring individual licenses indicated that GPL and LGPL, among the licenses less used by researchers, gave instead the greatest advantages in terms of citations

received by colleagues. These were all copyleft licenses. It may be the case that a few high-impact studies, possibly developed by large teams, are more likely to feature copyleft software, as they are compelled to do so if they re-use parts of software previously published under copyleft terms. The license MIT was the only to consistently increase all indices of cultural fitness, including the number of social media posts and webpage visits. This could be due to the great popularity of the MIT license or point to a limitation of this study, for which the MIT sample was infiltrated by studies referring to the research institution called MIT rather than the license with the same name, as both instances would be returned by searching for "MIT" on *PLOS*.

The results on the linear regression featuring citation count and type of license were confirmatory of the initial hypothesis and also in agreement with other research that tackled similar topic. Like the studies previously reviewed indicated (Christensen *et al.* 2019; Colavizza *et al.* 2020), more open science leads to a higher citation count. This is encouraging and hints of a way forward in view of the challenges faced today by scientists, such as scientific malpractice (de Frieze 2021), the replication crisis (Munafò *et al.* 2017), and the growing distrusts for science from the public (Funk 2017). It furthermore empirically-counters the conclusions of more pessimistic simulations predicting that research with laxer standards maladaptively spread at higher rates (Smaldino & McElreath 2016). FOSS science is instead more reliable and the effect seems large enough to produce a statistically-significant increase in the number of citations that other scientists grants FOSS researchers.

5.2 Considerations on the Co-Evolution of Software and Science

The existence of a positive association between FOSS and the cultural fitness of scientific articles published in peer-reviewed journals means that Free and Open-Source Software and scientific studies establish — as cultural evolutionary entities — positive-feedback loops between each other, which leads to their co-evolution. The use of FOSS increases the exposure of studies in the scientific community, which in turn make FOSS more known and used by more researchers.

From a practical perspective, this raises the question as to whether scientists should employ FOSS in their studies to increase the prospective citation count of their publications. This may indeed be the case, but it is also important to understand exactly why. When using FOSS, research can count on the collective work and wisdom of the generations of software developers of the past — a past which, in contemporary digital contexts, is constantly updated second by second. How to correctly utilise specific FOSS application — such as the scientometric libraries *rplos* and *rcrossref* employed in this research project — is often clearly explained and discussed through guides, wikis, and user comments freely available on the internet, facilitating the work of developers. Furthermore, FOSS ensures the possibility to study and modify how the source-code works, leading to a better understanding of its functionality and an increased potential for customisation, compared to non-FOSS software. This allows for much greater and faster cumulative cultural evolution (Caldwell & Millen 2008) and for the emergence of cultural evolution of cultural evolution of software (Birch & Heyes 2021).

More importantly, when scientists license their own software under one of the FOSS licenses, they contribute to this shared repository, which in evolutionary terms corresponds to participating in a pool of shared resourced or "public goods," whereas by contrast proprietary software are "private goods" (Jaeggi & Gurven 2013). In this way, FOSS developers both receive a benefit from the shared source-code and guides and also participate to their growth. As previously noted, the contrast between cooperative and competitive tendencies is at the heart of every evolutionary process — from the origin of multicellular organisms from unicellular organisms, to the institution of complex societies by unrelated individuals (Nowak 2006). At least for some industries — such as, as indicated by this study, science — it seems that it is the cooperative attitude, as exemplified by the use and sharing of FOSS, to gain the upper hand and generate the best returns, which are also enjoyed at the individual level in the form of an increased citation count, visibility, and prestige.

In other industries, possibly those involving the sale of a product, it may pay off better to keep source-code secret as to maintain exclusive control over it and avoid being imitated by competitors. Bowles (2009), who researched tribal warfare but may offer valuable insights to

our case, predicts that in-group (such as between employers of the same company) sharing of public goods will be high, even at the cost of considerable individual sacrifice, when between-group competition is high, such as it is found between commercial companies competing for the same market. In-group participation in public goods, especially when the group extends beyond kin, typically evolves and is maintained by reciprocal sharing and/or by the enforcement of an explicit system of rewards and punishments (West *et al.* 2007; Boyd *et al.* 2003). The how and why of the evolution of software across industries, however, needs to be ascertained by further research with a case-to-case approach.

Another factor to consider in perspective of future research is that here the evolutionary process of interest was cultural, *i.e.* studying the way scientific and software information spread, as inspired by pioneer work by Richard Dawkins (1976 [2016]), Luigi Cavalli-Sforza & Marcus Feldman (1981), and Peter Richerson & Robert Boyd (1985; 2005). Therefore the proxies for fitness analysed were chiefly cultural: citation count, social media posts, and webpage visits. Further research on the evolution of software as it is utilised, for example, by commercial companies may need to consider alternative proxies for fitness, including money (Scott-Philips *et al.* 2011).

In science, however, we find no evidence for systems of punishments enforcing FOSS (as scientists surely suffer no official punishment for not contributing to the Free and Open-Source pool), yet we find a positive correlation between creating and sharing FOSS and being cited more. Interestingly, such positive correlation disappears — and actually transforms into a statistically-significant, albeit slight, negative correlation — when the response variable is the number of social media posts or of webpage visits. Therefore, there are several things, at least two, to be explained in view of these findings in terms of evolutionary function and mechanism, namely 1) the ultimate function of a trait that increases evolutionary fitness, in this case cultural, and 2) the proximate mechanism that brings about the trait — as taught by Nino Tinbergen (1963; Zeifman 2001; Bateson & Laland 2013). One corresponds to the reason why scientists spontaneously invest in public goods, namely share FOSS. The other is how come scientists are induced to favour, consciously or unconsciously, colleagues and/or their research when they use FOSS.

In terms of ultimate-functional explanation, it is reasonable to imagine that even though there are no explicit or systematic rewards bestowed to FOSS developers and researchers, they may still gain less overt benefits including visibility, prestige, and respect from other researchers. Displaying costly and hard-to-fake behaviours, also known as "costly-signalling" or as "the handicap principle" (Zahavi & Zahavi 1996), is a well-known way to advertise one's own fitness and gain the respect of others. The typical example is the gazelle that in the African heat sneeringly jumps up and down in front of the lion, to display health and discourage predatory attempts. Costly-signalling theory works equally well to explain why researchers may be interested in publishing each single line of the source-code, fruit of their painstaking labour, in order to advertise honesty, transparency, and trust in their own methods and conclusions. In this way, the presence of FOSS works as a guarantee of genuine, hard-to-fake science, counteracting concerns over the fact that "bad," corner-cutting science may be easier to achieve and spread (Smaldino & McElreath 2016). Scientists interested in raising scientific standards and ensuring that "maladaptive" theories and methods are quickly selected out of the scientific evolutionary process should therefore seriously considering the benefits afforded by FOSS in terms of replicability, reproducibility, and consequent trust by colleagues and public.

Furthermore, if cooperative bouts (*i.e.* instances of code-sharing between developers) are iterative, cooperation may be selected by the institution of reciprocal relationships between actors. Cooperation can evolve through reciprocity in at least three ways (Nowak 2006). In direct reciprocity, each actor of a dyad can choose at every bout whether to act altruistically or selfishly, also taking in consideration the behaviour shown by the other actor in past interactions. This does not seem applicable because science is not an exclusively dyadic phenomenon. Indirect reciprocation models actors who are often asymmetric in their capability to act, such as scientists whom periodically have research to publish while other do not and *vice versa* or as developers that intermittently produce software to share on the internet. Indirect reciprocation between asynchronous actors is kept in place by factors such the good name and prestige gained by being known as a co-operator. In some cases, in our context, these may be even digitally recorded if the websites where developers share code allow users to rate, follow, or like other users' work. Furthermore, peer-reviewed articles and whether they employed FOSS or not are digitally recorded, by-passing the need to remember

the history of behaviour of other researchers, as it may happen in non-digital contexts. It is true that FOSS was not found to increase cultural fitness on *Facebook* and *Twitter*, but this is very probably only due to the non-specialised nature of these websites. Among specialists, *e.g.* software developers and/or science researchers, publishing quality Free Open-Source Software will be surely found to be appreciated, in the same way this study demonstrated it to increase citation count. The third way by which reciprocal cooperation can evolve is through "network reciprocity," whereby co-operators try to spatially distribute themselves among other co-operators and avoid defectors (*ibid*.). This explanation must be adapted to the characteristics of the context where the evolutionary processes studied occurred, namely the digital and interconnected nature of the Internet, where the rate of encounters between actors is not limited by factors such as spatial and chronological distance. Network reciprocity predicts then that we should find websites that are known for sharing FOSS and where users routinely submit their own code for review, such as we find in https://stackoverflow.com and https://github.com.

Beside individual-level explanations, the relationship between FOSS and citation count may even be explained at a level approaching the group. Marcel Mauss (2002) already widely discussed the transversal presence across societies, archaic and contemporary, of gift-economies, where individuals periodically give presents to extended kin, group members, and business partners as a way to maintain relationships and gain prestige. He even talks of instances of "competitive sharing" among Pacific Islanders and Native Americans, for which those who can afford to gift a greater quantity and quality of resources are selected as leaders (*ibid*.). Is science a gift-economy? In the context of software usage and as shown by this and other studies (Christensen *et al.* 2019; Colavizza *et al.* 2020), this appears to be the case. However, this assertion should be qualified: science may very well not be a gift-economy in terms of for example funding, which is often preferably awarded to remunerative research (Durand *et al.* 2007). Furthermore, citations and positions, for example, are not necessarily gifted but more often hard-earned. In the case of software, however, which falls under the category of the methods employed by researcher, it seems that freely-sharing quality resources — *i.e.* effective methods, made public by sharing one's source-code under a FOSS license — is rewarded with more citations from colleagues.

Whether the use of FOSS in science can be considered to be selected at the group-level is however questionable. Evolutionary research suggests that cultural group-selective processes of this kind (see Boyd & Richerson 2005) usually require a neat separation between groups (with no switching between FOSS and non-FOSS research by the same scientist) and a system of rewards and punishment to enforce rules — which in our case may be tentatively identified in rewards consisting of citations and prestige, but with no discernible trace of punishment beyond the criticism that scientists may individually chose to advance regarding non-FOSS research. The latter behaviour is called "altruistic punishment" (Boyd & Richerson 1985), whereby some individuals spontaneously enforce the respect of rules (*e.g.* use of FOSS) at a cost for self because without the support of an explicitly system of rules and organised judicial structure. A more parsimonious explanation therefore suggests that the spread of science and FOSS mutually-reinforce each other through a bottom-up process of indirect and network reciprocity.

Regarding the mechanical explanations for the association found between FOSS and a higher citation count, it is necessary to find at the proximate level the factor that inspires researchers to favour, whether they are conscious of it or not, the research of colleagues who employ FOSS. In the context of the evolution of cooperation through a process of genetic selection, scientists have hypothesised the existence of so-called "green-beards" (West *et al.* 2007), genes that develop into recognisable phenotypes ("green beards") and also induce possessors to selectively cooperate with other organisms with said green-beard. If the existence of such genes is purely hypothetical, this may not be the case in the context of cultural evolution. Tagging one's research with a FOSS label ("GPL," "ECL," "MIT" *etc*.) may act as a green-beard which at the same time produces the cooperative behaviour (licensing code as FOSS) and advertises the methodological ethos of the researcher. The idea of licensing as cultural green-beard could be the most straightforward mechanical explanation for the findings of this study regarding the spread of FOSS in scientific journal.

After concluding that it is likely that the adoption of FOSS in science 1) increases the visibility and prestige of scientists and 2) evolved through a process of indirect and network reciprocation through a cultural "green-beard" effect, some final remarks on some points raised in the course of the review of this study regarding the co-evolution of software and

science are due. First, both scientific research and software-development were treated as evolutionary processes, which through cultural transmission, often digitally-mediated, and selection, guided by the preferential imitation of adaptive variants (*e.g.* theories; libraries), leads to the evolution of consistently better scientific knowledge and software capabilities over the years. Second, if science and software initially evolved independently from each other, as science became increasingly software-based during in the last decades, their evolution has become entwined, engendering a science-software co-evolutionary system. Third, the cultural fitness of scientific articles was measured by their ability to be cited, shared, and accessed by others, therefore treating journal publications and social media posts as cultural generations. Fourth, it was empirically found by analysing scientometric data obtained from the *Public Library of Science* that the license under which software is released, which determines its terms of use, affects the cultural fitness of scientific articles among scientists by increasing their citation count. Fourth, the inverse relationship was discovered between the use of FOSS and being shared on social media, likely a testament to the differential way in which the population that writes on peer-reviewed journals and that posts on social media (even though these populations partly overlap) are affected by FOSS. It is likely then, if these evolutionary processes keep proceeding as they did in the last two decades, that in future we will see more FOSS in scientific research and less on social media. Fifth, the reported greater cultural fitness of FOSS in science may represent the solution to three problems that the scientific community faces, namely the replication crisis, scientific malpractice, and distrust for science — a solution which perhaps emerged though a process of cultural evolution.

Future research could keep studying the co-evolution of software and science by reproducing this research method in the context of other journals not hosted on *PLOS* to seek whether the same associations are found. Furthermore, it would be interesting to investigate other dimensions of software beyond data availability, which was already investigated in recent years (Christensen *et al.* 2019; Colavizza *et al.* 2020), and of the license, which was analysed here, such as for example the effects of the specific software or the coding language employed by researchers.

6. Addendum

5.1 Timeline

After choosing and refining the research question and reviewing the literature for a few months, the empirical study was actually completed in less than 3 months, which included selecting and producing the relevant software, collecting data, and analysing the results. The process of data collection through *rplos* took about 1 hour and through *rcrossref* about 3-6 hours.

5.2 Cost Effectiveness

The resources needed to run this research are 1) a computer, 2) electricity 3) an internet connection, and 4) some software applications. Computers, electricity, and an internet are relatively cheap and can also be used for many other tasks. All the software used (see Methods section) was free as in Free & Open-Source Software and also as in gratis. The research was extremely cost-effective.

5.3 Ethical Considerations

All the data analysed was already in the public domain, being information published on scientific journals. I declare no conflicts of interest beside a keen interest for Free and Open-Source Software.

5.4 Potential Publication Outlets

Being a study about science, software, and evolution, it could be potentially published on general scientific journals such as *Science* and *Nature*, meta-science journals such as the *Journal of Scientometric Research,* journals focusing on software such as the *Journal of Software: Evolution and Process*, or journals researching evolution such as *Evolution*, *Evolution in Human Behaviour*, and *Evolution, Mind & Behaviour.*

7. References

- Aksnes, D.W., Langfeldt, L. and Wouters, P., 2019. Citations, citation indicators, and research quality: An overview of basic concepts and theories. *Sage Open*, *9*(1), p.2158244019829575.
- Aktipis, A., Cronk, L., Alcock, J., Ayers, J.D., Baciu, C., Balliet, D., Boddy, A.M., Curry, O.S., Krems, J.A., Muñoz, A. and Sullivan, D., 2018. Understanding cooperation through fitness interdependence. *Nature Human Behaviour*, *2*(7), pp.429-431.
- Allgaier, J., Dunwoody, S., Brossard, D., Lo, Y.Y. and Peters, H.P., 2013. Journalism and social media as means of observing the contexts of science. *BioScience*, *63*(4), pp.284-287.
- Andersson, C., Törnberg, A. and Törnberg, P., 2014. An evolutionary developmental approach to cultural evolution. *Current anthropology*, *55*(2), pp.154-174.
- Armit, I., Swindles, G.T., Becker, K., Plunkett, G. and Blaauw, M., 2014. Rapid climate change did not cause population collapse at the end of the European Bronze Age. *Proceedings of the National Academy of Sciences*, *111*(48), pp.17045-17049.
- Arnold, B.C., 2014. Pareto distribution. *Wiley StatsRef: Statistics Reference Online*, pp.1-10.
- Asma and Rami Gabriel, 2019. *The Emotional Mind. The affective roots of culture and cognition.* Cambridge (Massachusetts): Harvard University Press.
- British Science Association 2020. *Young people are more interested in a scientific career as a result of COVID-19.* https://www.britishscienceassociation.org/blog/young-people-are-more-interested-in-a-scientific-career-as-a-result-of-covid-19
- Bateson, P. and Laland, K.N., 2013. Tinbergen's four questions: an appreciation and an update. *Trends in ecology & evolution*, *28*(12), pp.712-718.
- Boyd, R., Gintis, H., Bowles, S. and Richerson, P.J., 2003. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, *100*(6), pp.3531-3535.

- Bowles, S., 2009. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors?. *Science*, *324*(5932), pp.1293-1298.
- Campbell-Kelly, M., Aspray, W., Ensmenger, N. and Yost, J.R., 2018. *Computer: A history of the information machine*. Routledge.
- Caldwell, C.A. and Millen, A.E., 2008. Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior*, *29*(3), pp.165-171.
- Cavalli-Sforza, L.L. and Feldman, M.W., 1981. *Cultural transmission and evolution: A quantitative approach*. Princeton University Press.
- Chamberlain S., Boettiger K., Zhu H., Jahn N., Ram K, 2020. Rcrossref. https://cran.r-project.org/web/packages/rcrossref/rcrossref.pdf
- Chamberlain S. 2021. *Fulltext*. https://cran.r-project.org/web/packages/fulltext/fulltext.pdf
- Chamberlain S., Boettiger K., Ram K., rOpenSci 2021. *Rplos*. https://cran.r-project.org/web/packages/rplos/index.html
- Corballis, M.C., 2009. The evolution of language.
- Barnosky, A.D., 1987. Punctuated equilibrium and phyletic gradualism. In *Current mammalogy* (pp. 109-147). Springer, Boston, MA.
- Birch, J. and Heyes, C., 2021. The cultural evolution of cultural evolution. *Philosophical Transactions of the Royal Society B*, *376*(1828), p.20200051.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process.* University of Chicago Press.
- Boyd, R. and Richerson, P.J., 2005. *The origin and evolution of cultures*. Oxford University Press.
- Boyd, R. and Silk, J.B., 2014. *How humans evolved*. WW Norton & Company.
- Christensen, G., Dafoe, A., Miguel, E., Moore, D.A. and Rose, A.K., 2019. A study of the impact of data sharing on article citations using journal policies as a natural experiment. *PloS one*, *14*(12), p.e0225883.
- Brown, G.R. and Richerson, P.J., 2014. Applying evolutionary theory to human behaviour: Past differences and current debates. *Journal of Bioeconomics*, *16*(2), pp.105-128.

- Darwin, C., 2008 [1871]. *The descent of man, and selection in relation to sex*. Princeton University Press.
- Dawkins, R. 2016. *The extended selfish gene*. Oxford university press.
- de Vrieze, J., 2021. Large survey finds questionable research practices are common.
- Diamond, A.M., 2000. The complementarity of scientometrics and economics. *The web of knowledge: A festschrift in honor of eugene garfield*, pp.321-336.
- Durand, R., Bruyaka, O. and Mangematin, V., 2008. Do science and money go together? The case of the French biotech industry. *Strategic Management Journal*, *29*(12), pp.1281-1299.
- Fantini D. 2019. *Easypubmed*. https://cran.r-project.org/web/packages/easyPubMed/easyPubMed.pdf
- Fortunato, L. and Galassi, M., 2021. The case for free and open source software in research and scholarship. *Philosophical Transactions of the Royal Society A*, *379*(2197), p.20200079.
- Fortunato, L., 2018. Evolution of marriage systems.
- Foster, K.R., 2004. Diminishing returns in social evolution: the not-so-tragic commons. *Journal of evolutionary biology*, *17*(5), pp.1058-1072.
- Funk, C., 2017. Mixed messages about public trust in science. *Issues in Science and Technology*, *34*(1), pp.86-88.
- Github. https://github.com
- Glänzel, W., 2007. Characteristic scores and scales: A bibliometric analysis of subject characteristics based on long-term citation observation. *Journal of Informetrics*, *1*(1), pp.92-102.
- Harmand, S., Lewis, J.E., Feibel, C.S., Lepre, C.J., Prat, S., Lenoble, A., Boës, X., Quinn, R.L., Brenet, M., Arroyo, A. and Taylor, N., 2015. 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature*, *521*(7552), pp.310-315.
- Haddaway, N.R., Grainger, M.J. and Gray, C.T., 2021. citationchaser: An R package and Shiny app for forward and backward citations chasing in academic searching.
- Henrich, J., 2004. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *American antiquity*, *69*(2), pp.197-214.

- Hepburn, Brian and Hanne Andersen, "Scientific Method", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.),
- Hill, K., Barton, M. and Hurtado, A.M., 2009. The emergence of human uniqueness: Characters underlying behavioral modernity. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *18*(5), pp.187-200.
- Hill, T., Lewicki, P. and Lewicki, P., 2006. *Statistics: methods and applications: a comprehensive reference for science, industry, and data mining*. StatSoft, Inc..
- Holden, C. and Mace, R., 2009. Phylogenetic analysis of the evolution of lactose digestion in adults. *Human biology*, *81*(5/6), pp.597-619.
- Hunt, J., Breuker, C.J., Sadowski, J.A. and Moore, A.J., 2009. Male–male competition, female mate choice and their interaction: determining total sexual selection. *Journal of evolutionary biology*, *22*(1), pp.13-26.
- Hutchins, B.I., Yuan, X., Anderson, J.M. and Santangelo, G.M., 2016. Relative Citation Ratio (RCR): A new metric that uses citation rates to measure influence at the article level. *PLoS biology*, *14*(9), p.e1002541.
- Juristo, N. and Vegas, S., 2010. Replication, reproduction and re-analysis: Three ways for verifying experimental findings. In *Proceedings 1st Int. Workshop on Replication in Empirical Software Eng. Research (RESER 2010)*.
- Laland, K.N., Brown, G.R. and Brown, G., 2011. *Sense and nonsense: Evolutionary perspectives on human behaviour*. Oxford University Press.
- Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., Wray, G.A., Hoekstra, H.E. and Futuyma, D.J., 2014. Does evolutionary theory need a rethink?. *Nature News*, *514*(7521), p.161.
- Mace, R., 2014. Human behavioral ecology and its evil twin. *Behavioral Ecology*, *25*(3), pp.443-449.
- Mauss, M., 2002. *The gift: The form and reason for exchange in archaic societies*. Routledge.
- McElreath, R. 2021. Discussion with Prof. Richard McElreath: Science as Amateur Software Development. https://www.youtube.com/watch?v=zwRdO9_GGhY
- Mesoudi, A. and Thornton, A., 2018. What is cumulative cultural evolution?. *Proceedings of the Royal Society B*, *285*(1880), p.20180712.

- Mesoudi, A., Whiten, A. and Laland, K.N., 2004. Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the perspective of The Origin of Species. *Evolution*, *58*(1), pp.1-11.
- Mingers, J. and Leydesdorff, L., 2015. A review of theory and practice in scientometrics. *European journal of operational research*, *246*(1), pp.1-19.
- Minocher, R., Atmaca, S., Bavero, C., McElreath, R. and Beheim, B., 2020. Reproducibility improves exponentially over 63 years of social learning research.
- Morin, A., Urban, J. and Sliz, P., 2012. A quick guide to software licensing for the scientist-programmer.
- Moore, M.W. and Perston, Y., 2016. Experimental insights into the cognitive significance of early stone tools. *PLoS One*, *11*(7), p.e0158803.
- Munafò, M.R., Nosek, B.A., Bishop, D.V., Button, K.S., Chambers, C.D., Du Sert, N.P., Simonsohn, U., Wagenmakers, E.J., Ware, J.J. and Ioannidis, J.P., 2017. A manifesto for reproducible science. *Nature human behaviour*, *1*(1), pp.1-9.
- Nature. https://www.nature.com
- Nowak, M.A., 2006. Five rules for the evolution of cooperation. *science*, *314*(5805), pp.1560-1563.
- O'Connor, C., 2019. The natural selection of conservative science. *Studies in History and Philosophy of Science Part A*, *76*, pp.24-29.
- Public Library of Science. https://plos.org
- Paolicelli, R.C., Bolasco, G., Pagani, F., Maggi, L., Scianni, M., Panzanelli, P., Giustetto, M., Ferreira, T.A., Guiducci, E., Dumas, L. and Ragozzino, D., 2011. Synaptic pruning by microglia is necessary for normal brain development. *science*, *333*(6048), pp.1456-1458.
- Pritchard, A., 1969. Statistical bibliography or bibliometrics. *Journal of documentation*, *25*(4), pp.348-349.
- Ram, Y., Liberman, U. and Feldman, M.W., 2019. Vertical and oblique cultural transmission fluctuating in time and in space. *Theoretical population biology*, *125*, pp.11-19.
- Renn, J. and Sauer, T., 2007. Pathways out of classical physics. In *The genesis of general relativity* (pp. 113-312). Springer, Dordrecht.

- Roberts, J., 2009. An author's guide to publication ethics: a review of emerging standards in biomedical journals. *Headache: The Journal of Head and Face Pain*, *49*(4), pp.578-589.
- Sackstein, S., Spark, L. and Jenkins, A., 2015. Are e-books effective tools for learning? Reading speed and comprehension: iPad® i vs. paper. *South African Journal of Education*, *35*(4).
- Scott-Phillips, T.C., Dickins, T.E. and West, S.A., 2011. Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, *6*(1), pp.38-47.
- Sciencemag. https://www.sciencemag.com
- Semaw, S., Rogers, M.J., Quade, J., Renne, P.R., Butler, R.F., Dominguez-Rodrigo, M., Stout, D., Hart, W.S., Pickering, T. and Simpson, S.W., 2003. 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *Journal of Human Evolution*, *45*(2), pp.169-177.
- Smaldino, P.E. and McElreath, R., 2016. The natural selection of bad science. *Royal Society open science*, *3*(9), p.160384.
- Statista 2021. Social Media Users. https://www.statista.com
- Sommer, V. and Parish, A.R., 2010. Living Differences. In *Homo Novus–A Human Without Illusions* (pp. 19-33). Springer, Berlin, Heidelberg.
- Smaldino, P. and O'Connor, C., 2020. Interdisciplinarity can aid the spread of better methods between scientific communities.
- Stackoverflow. https://stackoverflow.com
- Tinbergen, N., 1963. On aims and methods of ethology. *Zeitschrift für tierpsychologie*, *20*(4), pp.410-433.
- Van Noorden, R. and Chawla, D.S., 2019. Hundreds of extreme self-citing scientists revealed in new database. *Nature*, *572*(7771), pp.578-580.
- West, S.A., Griffin, A.S. and Gardner, A., 2007. Evolutionary explanations for cooperation. *Current biology*, *17*(16), pp.R661-R672.
- Whiten, A., McGuigan, N., Marshall-Pescini, S. and Hopper, L.M., 2009. Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), pp.2417-2428.

- Wikipedia, 2021. *Most common words in English*. https://en.wikipedia.org/wiki/Most_common_words_in_English
- Wikipedia, 2021. *Twitter*. https://en.wikipedia.org/wiki/Twitter
- Wikipedia, 2021. *Facebook*. https://en.wikipedia.org/wiki/Facebook Zahavi, A., & Zahavi, A. (1997). The handicap principle. New York: Oxford University Press.
- Zeifman, D.M., 2001. An ethological analysis of human infant crying: answering Tinbergen's four questions. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *39*(4), pp.265-285.
- Zimmermann, F.K., 1971. *Induction of mitotic gene conversion by mutagens*. Forstbotanisches Institut, Freiburg i. B..
- Zuse, Konrad (2010) [1984], *The Computer – My Life* Translated by McKenna, Patricia and Ross, J. Andrew from: *Der Computer, mein Lebenswerk (1984)*, Berlin/Heidelberg: Springer-Verlag.