Copyright for Scientists*

https://muskulus.folk.ntnu.no/blog/copyright-for-scientists/

Michael Muskulus

2021-05-30

Contents

Copyright for scientists

TL;DR: Copyright protects creative forms of expressions, not the content. In particular, facts or ideas cannot be copyrighted. On the side, we note that therefore mathematical proofs are not copyrightable. Even if something is copyrighted, it is usually possible to use *quotations* (including images) from such works in your own scientific publications. If in doubt, it is best to get *permission* from the copyright holder. Otherwise, one needs to check national copyright law - usually quotations are allowed *in one's own creative works*, but only if this is *necessary for the discussion* and the copied amount is *not larger than necessary*. Re-use of copyrighted works for illustrative, aesthetic or entertaining purposes is usually *not possible* without

explicit permission. Similar rules apply to public *presentations*, e.g. at conferences or when making lectures available on the internet. Lectures or teaching materials given to students at non-commercial educational institutions, however, typically enjoy less restrictions and can include parts of copyrighted works also for illustrative or educational purposes, at least to some extent - but then cannot be published on the internet or otherwise made publicly available anymore! A problematic consequence of copyright law is that one is usually *not allowed to share one's own journal articles* with the general public, if one has transferred the copyright to a publisher. In particular, uploading the published article to, or sharing via a commercial platform such as ResearchGate is not allowed. What *is* possible, is sharing of a *preprint* or often even an *accepted manuscript* (before final editing by the publisher). And it is usually allowed to *share your own published articles with students or other researchers*! In general, as scientists we should be suspicious of copyright, as it reduces the efficiency of sharing our knowledge with society - which is our primary mission, and what we are paid to do! Instead we should aim to license our works for free use by anyone (but requiring attribution and not allowing changes), e.g. through Open Access publication or *other forms* of publication where we retain more control over the copyright. Finally, copyright in China, although officially similar to the Western world, seems to work quite differently in practice. Maybe there is something interesting that can be learned from this?

Introduction

> Copyright? Copy RIGHT: Steal ideas, steal facts but do not steal words. Dan Poynter

I got inspired to write this post after a recent experience in a Marie Curie Skłodowska Innovative Training Network we are partnering in, where there was considerable uncertainty about copyright issues when publishing a joint technical report and review on the project's webpage. While thinking about the particular issues involved there and trying to give advice, I thought it would be valuable to collect and summarize what I know about the subject - and use this opportunity to actually understand it better myself.

Disclaimer: I have no formal background in law, and my knowledge of the subject is necessarily superficial. It remains your **own responsibility** to check any facts or advice that I give here before acting upon them! In addition, note that I might have personal opinions that I voice here, so although I will try to be as objective as possible, this cannot be considered a fully neutral review.

That out of the way, let us start with the most important questions first: *What is copyright*, and *why does it exist*?

The first thing to realize is that there is no universal, world-wide copyright law, so differences exist among countries. However, most countries have joined international treaties governing intellectual property that regulate certain minimum requirements or at least establish some common framework for national copyright laws.[1] The general principles discussed below should

[1]The first international copyright treaty was the *Berne Convention* of 1886 that has tried to establish a kind of mutual recognition of copyright among its signatories. It was later extended by the *WIPO* (World Intellectual Property Organization) *Copyright Treaty* of 1996 that establishes and clarifies copyright protection also for computer programs (which are considered as "literary works") and for the content of databases. The most important nowadays is the *Agreement on Trade-Related Aspects of Intellectual Property Rights* (TRIPS). In the European Union this agreement has been implemented by the 2001 *Information Society Directive* (also known as the EU *Copyright Directive*). It was updated in 2019 by the highly controversial *Directive on Copyright in the*

therefore apply in most countries. But if in doubt - check your local copyright law!

The U.S. Copyright Office nicely summarizes what copyright is on their webpage:

> Copyright is a type of intellectual property that protects *original works of authorship* as soon as an author *fixes* the work in a *tangible form of expression*.[2]

This says that only creative ("original") works can be protected by copyright.[3] Unfortunately it is not completely clear what this means, and there are important exceptions. Usually very short "works" are excluded, for example. In particular names and slogans cannot be copyrighted. It becomes even more tricky when considering that only a "form of expression" is protected. This is an important point and seems to be often misunderstood. You have to realize: copyright protects a particular "form of expression", such as a specific way *how* a story is told, *how* a musical idea is expressed in a song, or *how* some data is visualized. Copyright does *not* protect the content, such as the underlying ideas, facts or data![4] In science we are mostly talking about copyright in the context of publications. But what is the point of publications actually, if not to communicate ideas and facts to others? So it would be really strange if what we read would be protected so that we could not put it to some use and let it influence us.

On the other hand, copyright does not protect a work only by disallowing identical copies. Works that are "substantially similar" are usually able to *infringe* (the technical term for a violation of copyright) someone's rights as well. There are many well-known examples, e.g. in photography.[5] When it comes to writing, my interpretation is that simply editing a copied text (replacing some words, changing the order of some sentences, adding or removing certain parts) would still constitute a copyright violation, as long as there remains some similarity with the original (and maybe even beyond that?). It seems that all these modern plagiarism checking software solutions available nowadays could (and probably soon will) also be used to detect possible copyright violations![6]

I will take a closer look at some of these issues below. For now, we should realize that copyright exists because of *commercial interests*.[7] Someone, let us call them a *creator*, has developed a new way to express an idea, implement an algorithm, or illustrate some facts. As it is very easy to copy or reuse such expressions nowadays, but potentially takes a lot of time, expertise and effort -

Digital Single Market.

[2] U.S. Copyright Office: *What is copyright?* https://copyright.gov/what-is-copyright/ [Accessed 2021-06-07]

[3] What is a *creative* work? One could get very philosophical here. In fact, courts seem to generally think that some kind of creative force is needed, and that therefore only *humans* can be creative. In other words, a human component is needed for something to be copyrightable. The hot discussion at the moment is whether computer-generated music can be copyrighted or not.[^4]

[4] There are other mechanisms that are meant for protecting ideas and data. The most important one is (surprise!) *secrecy*, e.g. keeping so-called trade secrets. A much more complicated alternative, but one that is often considered, would be *patenting*.

[5] Kaitlyn Ellison: *5 famous copyright infringement cases (and what you can learn)* https://99designs.com.au/blog/tips/5-famous-copyright-infringement-cases/ [Accessed 2021-06-07]

[6] I should mention the difference here: Plagiarism is an *ethical* violation that does not affect anyone else's rights or financial interests. The main damage is to the field itself, the profession and community of scientists. So although plagiarism can (and should) have severe consequences, it is often somewhat unclear how it should be handled. Copyright infringement, on the other hand, is a *legal* violation and more straightforward to handle. There will typically be a copyright owner who can claim economic damages and sue. This case will then be addressed by a court.

[7] There are other rights connected to creative works that can be applicable, apart from copyright. Whereas copyright is mainly concerned about financial aspects, so-called *moral rights* are concerned with the right of authors to be identified (or remain anonymous) and that their works are not distorted (the right to integrity of the work), e.g. by removing parts, if this would affect the authors reputation. Even if the copyright to a work is sold or otherwise transferred, moral rights might still exist. In science, we have additional ethical obligations to consider, such as not plagiarising others (or ourselves) that apply, even if no legal (copyright) issues are at hand.

not to speak about creativity and inspiration - to develop them, it seems natural that creators want protection for their work such that they can benefit financially from it (should they want to). This is an example where we are led to think that legislation should fix an obvious economic asymmetry that is perceived as unfair.

Finally, it should be mentioned that copyright is *automatically* obtained once a work has been created. There is no need to apply for it and not even a need to remind people about it (such as a copyright notice).[8] In particular, I should point out that images and text "found" *on the internet* are usually copyrighted and *cannot* be freely re-used by others without proper permission.

Permitted and fair use

What is one allowed to do with a copyrighted work that does not require explicit permission from the copyright owner, so called *unlicensed* use?

To begin with, most people are probably not aware of all the restrictions imposed by copyright! In fact, copyright is not only about copying a work, there are usually also provisions for guaranteeing so-called *distribution rights* to authors. This gives copyright owners the right to sell, lend or otherwise distribute their works - and restricts others from doing so. In particular, if one buys a copyrighted work (or a copy thereof), but not the copyright itself, this does not automatically give one the right to sell or lend it to others! Astonishingly as this might seem, it is really the case that additional rules are needed for such a right. In the U.S. the most important of these is the so-called *first-sale doctrine*, whereas in Europe a similar role is taken by the so-called *principle of exhaustion*. These rules state that the right of distribution only applies to the first time that a copyrighted work is sold. It is these additional rules that make it possible that libraries and second-hand stores can exist (and that one can lend a book or give a CD to a friend!). Importantly, these rules currently only apply to physical objects, and not to digital works...[9]

Another issue is the question of what are actually permitted *uses* of a copyrighted work and its content. For example, buying a book does not transfer the copyright, one only obtains ownership of a physical copy. Reading the book is obviously allowed, as this is its purpose. Uses that are surely problematic are reproduction and making the contents available to others. But could taking notes or presenting its contents to others also get one into trouble?

In the discussions below I will focus on the German situation. This is mainly because it is easy to understand (after the reform of the German Copyright Act in 2018 that simplified the existing law enormously) and well documented. The situation in other countries will typically differ - but usually not in very large ways. Still keep in mind to check the legislation that applies in your country. Or, when in doubt, consider to simply *ask for permission*, which is often the easiest and best way to cover your back![10]

[8] It is possible to officially register copyright, but this is mostly relevant in the US only. There one needs to have a work officially registered with the U.S. Copyright Office if one wants to sue someone else for infringement. Although this can still be done later, if necessary, it is usually advisable to register works with commercial potential (such as books, songs or certain works of arts) before publishing them. The fee is very reasonable (as of writing, it seems to amount to 45 USD).

[9] Anyone who has ever obtained an *e-book* should know about the rule of exhaustion, although in a negative way - since it is currently considered to apply only to physical copies! In other words, non-physical, digital representations such as e-books or computer code cannot usually be sold again or given to others. This is one of the reasons why nowadays software is sold only digitally. A physical copy of software, such as a CD, *could* be sold again to others. A digital copy *cannot* be given to others, so companies will potentially earn more by selling only these.

[10] Note however that one has to be sure to ask the correct copyright owner for the correct permission, so even in this case one needs a basic understanding of the copyright law. Finally, there are cases where this is not possible,

Making copies

In many countries it is perfectly permissible to take notes and even copies of whole works for *personal* or *private use* (which must not be work-related or with any commercial implications), e.g. for a personal archive or for safety reasons.[11]

Copies for research or academic work fall into a different category, and more restrictions apply. The German Copyright Act, for example, allows researchers to copy **up to 75 percent** of a given work (e.g. a book) for use in **non-commercial research**. The main difference with *personal use* is that one is allowed to *work* with the materials and be *paid* for this activity by an employer. Papers in academic journals and small-scale works (up to 25 pages) can be **fully used** even.[12]

For teaching and **educational purposes** at an established **non-commercial** educational establishment (including high schools and kindergardens) **15 percent** can nowadays be copied out of any work without explicit permission.[13] This rule applies also to students copying from textbooks, newspapers or videos. And as above, articles from scientific journals and small-scale works can be fully used.

The German law even contains special provisions for *data mining*. Such applications may use whole works for academic research, provided that the copied information is not used for other purposes, and either deleted after concluding the research, or given over to a suitable, responsible archive without public access.

The right to quote

The German Copyright Act, as a more or less representative example, allows to use quotations to copyrighted works without the need to obtain permission, for the purpose of **discussing** parts of a (or even a whole) copyrighted work. The requirements are that the quotation is *used in one's own creative work* and that it includes an *indication of the original authors*. Small-scale quotations are permissible under these conditions, and even large-scale quotations (such as a complete image!) are allowed - but these only in an academic work. In both cases the content must be *necessary and of adequate length for the discussion*, and cannot be used in an ornamental, aesthetic or entertaining function only.

The Norwegian intellectural property law, as another example that I am familiar with, is very similar. It allows quotations for *critical* purposes, but also for *scientific* purposes, which could be interpreted slightly broader than the German law to also include, for example, illustrating or reviewing previous works. There also seems to be no restriction on the length, but the principle of using an adequate amount applies as well.

---

for example when the copyright owner is unknown or cannot be easily contacted. And of course there are lots of examples where the copyright owner does not give permission - but one can still be allowed to use the work by the law! (Such situations are common when critizing or paroyding a work, for example.)

[11] Although personal use is usually unproblematic, there are certain restrictions: Personal copies must not be used for work-related or commercial purposes, and must not be made accessible to others. Moreover, it is usually only allowed to copy from original works or other legal copies, and only one copy can be made.

[12] Federal Ministry of Education and Research (2020): *Copyright in academic work - an overview for research, teaching and libraries*. Report, Germany.

[13] This used to be much stricter previously - to the point of being completely impractical - where highschool textbooks were concerned. As of 2009, for example, teachers at highschools were only allowed to copy once per year (and class) from any given textbook, for a maximum of 15 percent and not exceeding 20 pages.[ˆ11]

For illustrative and educational purposes

The use of copyrighted materials for purely illustrative or pedagogical purposes also falls under the above rules. An important case are **presentations** and lecture notes that include copyrighted *images* of others, e.g. from the internet. It is allowed to include such materials without explicit permissions - at a nonprofit educational institution - if the result is not made public. This applies to presentations and materials for both research and teaching purposes, and in Germany up to **15 percent** of a work can be used (as well as whole images and everything from scientific articles). These works can be stored and distributed internally, e.g., given to students, a small group of other researchers, or published inside a protected webside.

A potential problem arises when there is a wish to make learning materials (e.g. recordings of lectures, or the used lecture notes and teaching aids) *publicly available*, e.g. on the internet. In this case the materials lose their special status of an educational activity ("for a specific closed group at a non-commercial educational institution. . . ") and can now only use quotations for critical purposes, but no longer reproductions for educational or illustrative purposes. In other words, it would be impossible to make many entertaining, interesting and meaningful lectures publicly available - unless the used copyrighted works will be licensed, which will quickly become time-consuming and potentially expensive. What is the alternative? In some subjects, e.g. mathematics, high-quality lectures can be given without the need for using copyrighted works. But many other subjects will find it difficult or impossible to adapt (For example: How can one teach anatomy without being able to show pictures? How can one discuss political philosophy without being able to read the relevant essays and articles in their entirety?).

In term papers, bachelor or master theses

A similar issue arises here than in the previous paragraph. Interestingly, the German Copyright Act seems to allow the use of illustrations also in term papers and bachelor or master theses, considering these to also fall into the context of an educational activity. However, a precondition seems to be that these works than cannot be made publicly available.

This seems somewhat misguided as this is in conflict with the idea that a master thesis constitutes a scientific work. My interpretation - and recommendation - therefore is to treat student theses to the same standard as scientific publications, i.e., to only allow the use of quotations.

Fair use

The U.S. have so-called *fair use* provisions that provide a more flexible framework for the unlicensed use of copyrighted materials. A similar legal construction does not exist in Europe, although in practice there is quite a lot of overlap with existing European and national legislation. Importantly, fair use is not a right or in any way guaranteed, but is rather a set of guidelines that can be used to *defend* against an infringement claim. In the end it lies with a court to decide if they accept the fair use defense or not. So what constitutes fair use?

Typically judges consider four items when deciding if the use of a copyrighted work in ones own work constitutes a case of fair use[14]:

1. The *purpose of ones own work* is important. If it is for nonprofit educational purposes it is much more likely to be considered fair use than if it is has a commercial purpose (such as being used in a textbook that can be sold, or being in a scientific article that earns

[14] For an illuminating collection of examples of fair use rulings, see: *Summaries of fair use cases* https://fairuse.stanford.edu/overview/fair-use/cases/ [Accessed: 2021-06-07]

money for its publisher!). Also, *transformative purposes* that create something new out of a copyrighted work, e.g. by critising, commenting, or parodying it, are more likely to be considered fair use than e.g. using a copyrighted work for purely illustrative purposes.
2. The *nature of the copyrighted work* is important. Quotations from factual works or textbooks are more likely to be considered fair use than uses of artistic works and fiction. It is also harder to justify fair use for unpublished works (remember that these automatically enjoy copyright protection after being created).
3. The *amount that is used* is obviously important for judging whether fair use applies. In contrast to the above mentioned German Copyright Act the U.S. fair use provisions do not specify a maximum percentage or amount - it is up to a court to weight each case individually.
4. The *effect upon the market* the use has. If the use deprives the copyright owner of income it is much harder to justify fair use than if there are no financial implications.

Examples relevant for scientists

Here are some thoughts about applications of copyright law to situations common in an university context. This is basically a review of the above considerations, now oriented more toward applications. The redundancy is intended!

Presentations

As discussed above, presentations used for *teaching purposes* at *nonprofit educational institutions* can usually use *parts* of coprighted works without requiring explicit permission.

What about presentations at *scientific events*, such as conferences and workshops, that are open to the public (potentially, upon payment of a registration fee)? If the event is closed and the participants will only use the presented material for their own academic research, then small-case use seems to be permissible (e.g. in Germany up to 15 percent of a copyrighted work can be presented and distributed to others). On a large international conference, however, also participants from industry and other people with commercial interests (such as independent inventors or for profit writers) can be expected to be present. I would therefore conclude that in this case the use of copyrighted material of others is not allowed, without additional permissions - unless it is for the explicit purpose of discussion. In other words: it should be OK to show passages or even images from other publications, *if* the presentation comments, criticises or otherwise discusses these. It would *not* be OK to show funny images or beautiful illustrations that only serve to make the presentation look nicer or be more effective!

The same right to quote usually also applies to presentations given in other contexts, e.g. in a company or for profit. Again, it should be permissible to use parts of copyrighted works if necessary for the discussion. All other uses would usually require explicit permission from the copyright holders.

Teaching materials

Similar to presentations, it is usually possible to copy and make available parts of copyrighted works for educational purposes. Such unlicensed use is often restricted to established, nonprofit, educational institutions and to the context of a specific course, i.e., the materials can only be given to students actively attending the course.

What about recordings or videos of lectures? First of all, let us be clear that - even at university -

it is not allowed to record lectures or presentations without the knowledge and permission of the presenter (and potentially more stakeholders might be involved, e.g. the administration might also need to be informed and asked, and the rest of the audience might also need to give their permission). That said, what if a lecturer provides a recording of a lecture to students? This is obviously a copyrighted work that can only be used by the participants of that course, and solely for educational purposes. That said, the right to quote allows students to use small parts (e.g. up to 5 minutes of a video) in their own works (such as in a thesis, or in a personal blog post on their private webpage), if they are needed for discussion (not only for illustration) of something. (Exercise: Consider... if these teaching materials use other copyrighted works in an educational or illustrative capacity only, in the part that is quoted... would this be still a permissible quotation? Now *this* is the kind of copyright law nightmare that keeps mathematicians awake at night!)

For completeness, let us mention that written exams also fall under the rules of copyright. In other words, students are not allowed to reproduce or make available exam questions to others, unless again as a quotation, i.e., as part of a creative work where this is necessary for the discussion.

Textbooks

When writing a textbook or monograph[15], the obvious case for copyright are the illustrations and the possible use of copyrighted material for discussions. We have basically covered these issues above. But note that books usually have serious financial implications! Even open access, freely given away books can impact the markets of *other* books. It is therefore *highly recommended* (and usually required by any serious publisher) to *obtain explicit permissions* for all copyrighted images and illustrations used in a book. An alternative is the use of *stock images* one pays for - here it is important to buy a suitable license that covers all possible current and potential *future* uses.[16]

With regard to the contents of the book, there can be still more subtle copyright issues. Obviously the presentation of the material in a book is a particular expression (or "story") that enjoys copyright protection. Although the purpose of a textbook is educational, the availability of the book outside of a closed educational institution means that this has to be considered a use on more or less commercial terms. Even then, parts of copyrighted works can usually be included for discussion purposes, as we have discussed over and over again.

What about the *choice of topics* for a textbook, or the *choice of exercises*? For the latter, people seem to generally agree that providing example solutions for all exercises in a textbook is usually not possible, since one thereby has to reproduce all exercises themselves - which is not a small-scale quotation anymore. At first thought, even using a small number of exercises from another book in one's own book should also not be possible without explicit permission - as developing exercises is a creative act, right?

This seems all very subtle. Let us turn to mathematics where we hopefully can get a better understanding of the issues involved. Is a *mathematical proof* a creative expression that can be copyrighted? The well-known mathematician Keith Devlin might actually think so, when he states that

[15] Just to clear up the distinction: A *textbook* is usually a pedagogic exposition of a topic, field or research area. The main audience are students and people trying to enter a new field. A *monograph* is usually a presentation of original research, and typically aimed at experts.

[16] Helen Sedwick: *Self-publisher's legal handbook.* Ten Gallon Press, 2017.

> Proofs are stories that convince suitably qualified others that a certain statement is true.[17]

But here's the thing: It is commonly believed that *mathematical proofs are an idea, not a creative expression*. Although the idea for a proof can be highly creative and difficult to invent, the proof itself is usually somewhat mechanical, with a more or less fixed notation, and does not involve a lot of artistic freedom.[18] In other words: the *idea* behind a mathematical proof can be very special and rely on a unique "story" or approach - but ideas are not copyrightable.

When it comes to *exercises* in a textbook, I would think a similar argument can be made. The exact formulation of an exercise is usually not an original, creative expression of the (possibly quite interesting) idea behind it. What *seems to be* a creative expression, however, is the *choice* of exercises and topics presented in a textbook. There is a lot of freedom and creativity in how a textbook teaches an understanding of a field, what it highlights, which approach it favors, and which exercises promise the best learning experience for readers. The result of these choices could be protected by copyright! On the other hand, in certain fields or when discussing narrow technical details there are only few viable options for presenting and discussing a certain topic, and only a few exercises that seem to make sense. I would argue that in such cases there is no claim to copyright protection of the limited choices made (or the underlying "generic ideas" used) by an author.

Similarly, if a certain principle or fact is illustrated by a highly specific drawing or plot, one cannot use a similar representation - unless it is so obvious or trivial a representation (such as plotting a curve or datapoints in a well known standard way) that no creativity is involved.

To summarize this discussion: The choice of materials and approach taken in a textbook could potentially fall under copyright, as well as nontrivial, interesting ways of illustrating facts or concepts. In practice it is therefore advisable to use a somewhat unique approach to one's own book, in order to be sufficiently different from all other existing textbooks. That said, it should be possible to reuse some exercises, methods of proof, and specific ways of presenting the material from other works without explicit permissions. After all, "knowledge is free" and cannot be copyrighted!

Scientific papers

As said in the beginning, facts and data in general do not fall under copyright protection, whereas artistic expressions do. That seems to imply, in particular, that given a plot published in a scientific article, one is allowed to extract the data that is shown and use it (for anything, really)! One can also plot the data, but if the particular way that data is shown in the original paper is special, one might need to be careful not to do something that is too similar.

Figures from other scientific articles can potentially be used in one's own scientific works without the need for explicit permission, as discussed above. However, in the interest of avoiding doubt and since it has become so routine to do so, it is highly advisable to ask for permission from the copyright owner (usually the publisher).

[17] Keith Devlin (2019): *What is a mathematical proof?* https://www.mathvalues.org/masterblog/what-is-a-mathematical-proof [Accessed: 2021-06-07]

[18] I picked up this argument from Jim Lewis at this StackExchange post: https://math.meta.stackexchange.com/questions/1853/copyright-of-mathematical-formulas. It seems entirely reasonable to me, but it would be nice to elaborate more on this.

Reprints and sharing research

It took me some time to realize that this *could* be a potential problem (although it is not, really, anymore). When publishing research in a traditional non-open access journal (e.g. such as Science or one of the journals run by Elsevier) one usually transfers the copyright to the publisher. These contracts often specify that authors retain the right to use their own papers for teaching and research purposes. But do they also allow you to share your work with others? If someone asks for a copy of your paper, are you allowed to send them a PDF of your published manuscript? And, to talk about the elephant in the room, are you allowed to upload the manuscript to ResearchGate or similar sites and distribute it via their systems?

In earlier times - when it was not possible to send articles to others electronically this was rather straightforward. In those times, after publication of an article, one would usually obtain (read: buy!) a number of reprints from the publisher, and these one could then freely send or give to others (a consequence of the first-sale doctrine!). This way of sharing research is still possible, in fact!

And as discussed above, other researchers are typically allowed to copy and use academic papers from others for non-commercial research purposes. So are you allowed to send them your manuscript directly? An important principle is that permissible copies are only allowed from legal versions of the work. In other words, other researchers can go to a library that has the article in question and obtain a copy therefrom. And they could copy from an author's copy of an article - if they could get their hands on it in a legal way. And as authors are usually allowed to copy and distribute articles from scientific journals (including their own) to colleagues - for scientific purposes - this seems possible. Now this argument is a little bit contrived. It would be much easier if publishers could simply allow such sharing of research with colleagues explicitly.

And this is what seems to have happened, in many cases. Elsevier, for example, grants authors the right to send the final published manuscript to colleagues and other researchers (but not to the general public!)[19]. And the accepted manuscript (before final editing and formatting by the publisher) can even be published publicly on a personal webpage. But note that you *cannot* share your published article via commercial platforms such as ResearchGate! (Exercise: Why would that be? What would stop your colleague from obtaining a legal copy of your article if it would be sent to him using the tools of such a platform?)

A look at China?

China does not seem to have a particular positive image in the mind of many people when it comes to copyright (or other forms of intellectual property), but in fact the concept can be found in Chinese sources as early as the Song dynasty (960-1279), and explicit copyright laws were establishd in 1910. In comparison with the Western world this is indeed somewhat late, as the Psalter of St Columba (560-600) already mentions copyright issues, and explicit copyright laws were established in the beginning of the 18th century. More interestingly, China enjoyed a long period without any copyright legislation, starting with the cultural revolution 1949.[20]

This all changed with the 1990 Chinese Copyright Law, and nowadays China is signatory of most international copyright agreements.

But there seems to still be something special about the way the Chinese deal with copyright

[19] https://www.elsevier.com/about/policies/sharing [Accessed 2021-06-09]

[20] Yiping Yang: The 1990 copyright law of the People's Republic of China. *UCLA Pasific Basin Law Journal* **11** (1993): 260-284.

issues on a practical level. As Andrew "bunnie" Huang writes about the distinct "gongkai" style of Chinese innovation:

> The West has a "broadcast" view of IP and ownership: good ideas and innovation are credited to a clearly specified set of authors or inventors, and society pays them a royalty for their initiative and good works. China has a "network" view of IP and ownership: one attains the far-reaching sight necessary to create good ideas and innovations by standing on the shoulders of others, and people trade these ideas as favors. In a system with such a loose attitude towards IP, sharing with the network is necessary, as tomorrow your friend could be standing on your shoulders, and you'll be looking to them for favors.[21]

It is apparent that if the Chinese way is really inherently different from the Western approach toward intellectural property, maybe there is something important to be learned from it? At the very least we should not assume that the Western concept of copyright is the only possibility in which creative works can be made available and shared in a meaningful way.

What should we do?

It can be argued that our main responsibility as scientists is towards society... After all, we are more or less paid by society to develop knowledge for the benefit of everyone. It is only natural that we share this knowledge as much and with the least restrictions as possible. As scientists we should not have commercial interests, and we **should not actually need copyright** - apart from *protecting access* to our work for others![22]

So does it go wrong somewhere? Apparently mostly with publishers requiring a transfer of copyright and then putting our creative works behind expensive paywalls. Are there no alternatives? Unfortunately the quality of most journal submissions has become so low that *peer review* is more or less a must nowadays (whereas it was apparently not very often done before the 1960s - but this is topic for another blog post!) The bean counting introduced by administrators and the ensuing pressure to publish have contributed their part. Whereas in earlier times technical reports (non-peer-reviewed!) were also important publication avenues, nowadays everything has to be published in journals, it seems. Open Access journal publication could be the solution. Unfortunately the shift in market forces that it introduces, coupled with the mentioned challenges, allows unprofessional and fake publishers without sufficient regard for scientific quality to prosper in its wake.

Nevertheless - it seems our current best practice for scientific publications would be Open Access in a well established and trustworthy journal, using a Creative Commons license that allows others to freely copy and use our work (as long as it is attributed and not changed).

I think the key here is *attribution*, and getting our priorities right. If we are serious about science, then we should share our ideas, our data, our software, without any restrictive licenses based on copyright. This is the way we (as humanity) will make progress in the fastest possible way. This is also economically most efficient - and we will realize the benefits of free trade (if reciprocal).

[21] Andrew "bunnie" Huang: *The hardware hacker - Adventures in making & breaking hardware.* No Starch Press, 2017.

[22] The use of copyright to protect access to software for everyone is taken to its extreme in the "copyleft" philosophy pioneered by Richard M. Stallman and the Free Software Foundation. The GNU General Public License (GPL) for software has a unique mechanism for keeping access free - as any work using code under the GPL is required to also use the GPL. Thereby, not only does free software remain free forever, also any software using free software becomes free. Since this is problematic for commercial entities, more permissible licenses have been invented, such as the well-known Creative Commons licenses.

By this I mean, for example, if you use some open source software for your work in science and engineering (and who does not?) - making your own work and publications freely available is one easy way to give something back! Just like they do in China...

Summary of key insights

1. Copyright protects a **creative expression**. It does *not* protect ideas, facts or data.
2. Copyright protects a *form of expression*. It is not necessary for a copy to be *identical* for copyright to apply, only to be **substantially similar**. In particular, even highly modified and edited text could consitute a copyright violation, if it looks similar enough to an existing source.
3. Copyright protection is obtained **automatically**. One does *not* need to apply for copyright (with the exception of the U.S., where one should consider this carefully).
4. **Images and text from the internet** are usually copyrighted. One needs to obtain proper permissions if one wants to re-use them.
5. The principle of exhaustion allows one to **sell or lend copies** of copyrighted works that one has legally acquired. This only applies to **physical objects** currently, however. In particular, one is usually not allowed to sell (or otherwise to make available) electronic books or papers that one has bought.
6. It is usually possible to make full copies of most copyrighted works for **personal use**.
7. The law typically allows copies for **educational purposes** (such as studying or teaching) in a **non-commercial** setting at an **educational institution**. Small-scale works (such as images) and academic journal papers can often be fully copied and used, but from large-scale works (such as books) only parts can be used (typically in the order of 15 percent or 1-2 chapters).
8. The same applies for **research purposes** at an **educational institution**, only a potentially larger share of the copyrighted work can be copied and used (e.g. typically 75-100 percent of a large-scale work).
9. **Academic research in data mining** in the European Union might be able to fully use copyrighted works.
10. Copyrighted works can usually be **quoted** in ones **own creative works** in an appropriate way (adequate length) if they are **discussed**. This applies to book reviews, articles in magazines, etc. In academic works it might even be possible to reproduce whole images, if necessary for the discussion.
11. **Images** and other content from the literature or the internet can be used in **presentations** without explicit permissions, subject to the above rules about educational purposes. Such presentations or materials can be made available to students (only).
12. When it comes to **presentations** that are performed outside of an institutional educational environment or a closed group of academics, there is usually still a right to *quote* parts of copyrighted works for **discussion** purposes. Purely illustrative or entertaining uses are usually not allowed. Student qualification works (e.g. **bachelor or master theses**) should be considered similarly as *academic works*, not in an educational context.
13. The U.S. has **fair use** provisions that are guidelines for what might be permissible unlicensed uses of copyrighted works. There is a lot of overlap with the European legislation, but whereas European copyright often tries to be explicit (e.g. fixing the percentage that one is allowed to copy), in the U.S. the decision always lies with a court. This creates uncertainty and a certain risk in the U.S., but also a potentially larger freedom in what one might be able to do with copyrighted works.

14. **Mathematical proofs** are usually considered *not* to fall under copyright protection. A similar argument can be made for **exercises in textbooks**. However, the choice of contents and exercises is an artistic expression. A new textbook - or a set of lecture notes - therefore has to be different from all existing books in what it presents about the topic and how it does that.
15. Published facts and data do not enjoy copyright protection. In particular it is allowed to extract **data** from published figures and plots and reuse it.
16. Even if copyright has been transfered to a publisher it is usually possible to share **published articles** with other researchers and students. But one is usually not allowed to upload or share manuscripts via communities such as **ResearchGate**.
17. **Open Access** journal publication might be our current best practice for responsible publication of research. We should also consider publishing materials freely online ourselves (without peer review) as a possible alternative - especially relevant when it comes to supplementary materials, lecture notes, tutorials, data, example cases, etc.
18. Making our scientific work freely available to anyone is one way to **give something back** to the thousands of volunteer **open source software** writers whose code we use when doing research.

Further reading

A short selection of interesting books where readers can find more information and details. In order of probable interest.

- A Sinnreich: *The essential guide to intellectual property.* Yale University Press, 2019.
- B Klemens: *Math you can't use - Patents, copyright, and software.* Brookings Institution Press, 2006.
- American Society of Mechanical Engineers: *Intellectual property - A guide for engineers.* ASME Press, 2001.
- D Bainbridge: *Intellectual property.* Pearson, 2018. (A well-known textbook, with a strong focus on UK copyright law. Some background in law studies might be required.)
- M Boon: *In praise of copying.* Harvard University Press, 2010. (A critical, literate account. More or less the complete opposite of Bainbridge's book.)
- Bundesministerium für Bildung und Forschung: *Copyright in academic work.* Report, 2020. Available from https://www.bmbf.de/de/was-forschende-und-lehrende-wissen-sollten-9523.html (A detailed explanation of the passages from the German copyright law that are relevant for researchers and teachers)
- Forskerforbundet: *Opphavsrett.* Skriftserien **4** (2019). Available from https://www.forskerforbundet.no/om-oss/publikasjoner/skriftserien/ (This is a very informative discussion and commentary of the Norwegian copyright law. Unfortunately only available in Norwegian.)

Acknowledgements

Thanks to *Po Wen Cheng* for pointing me at the German report and for suggesting that I try to clarify what is possible when doing presentations.