



Briefing paper on: Intellectual Property Considerations for Computable Biomedical Knowledge (CBK)

Briefing Advisor: Ana Enriquez

Briefing Paper Authors: Emily Dibble, Brad Iott, and Allen Flynn

Disclaimer: This brief is to surface ideas. It is not a source of legal guidance. Legal terms appear in a different font.

Different types of CBK are subject to different legal frameworks for Intellectual Property (IP)

SHARING CBK AND THE LAW

This briefing paper focuses on the law as it pertains to people who want to access and use knowledge generated by others. On behalf of those people, it asks, “*What governs whether or not knowledge in the world, including CBK, can be reused?*” This is mainly a question of copyright. This paper does not explore the important related question pertinent to many knowledge generators, “*How can new IP be protected?*”, which is a question of copyright too, but also of patent law and other protections besides.

TWO INSTANCES OF CBK, ONLY ONE OF WHICH IS SUBJECT TO COPYRIGHT PROTECTION

This briefing paper introduces the idea that, from the perspective of copyright law, not all instances of CBK have the same standing. To help make this point, examples of two different instances are given next.

CBK HOLDING KNOWLEDGE IN A MATHEMATICAL-STATISTICAL FORM

Consider the following actual example of knowledge published in the form of a statistical prediction rule (SPR). A paper reporting results of a federally-funded sub-study, made possible by a larger clinical trial sponsored by 8 medical-device manufacturers, includes the fine details of a SPR enabling anyone reading the study to learn precisely how to calculate a patient’s bleeding risk after heart catheterization (Yeh; 2016). *What governs whether or not a third party can make and distribute this bleeding risk SPR by encoding it and publicly sharing its newly encoded form online?*

CBK HOLDING KNOWLEDGE IN A “LOGICAL-LEXICAL-VISUAL” OR “CREATIVE” FORM

Now consider a contrasting real-world example of knowledge published in the form of a written guideline. The National Cancer Care Network (NCCN), a private, non-profit alliance, has published a 2017 version of its Cancer-and-Chemotherapy-Induced Anemia guideline. This NCCN guideline provides both a “decision tree” and very detailed instructions telling what to do in the case of chemo-induced anemia. *What governs whether or not a third party can make and distribute the decision tree and corresponding written instructions in this guideline by encoding them and publicly sharing their encoded form online?*

Thought Question:

What other examples of *knowledge in the world* can you think of that may be *materially different* than these?

Many Different Players are Involved in the CBK Story

Numerous, different actors play unique roles in the creation, sharing, and use of CBK. Sometimes, biomedical knowledge is generated by researchers and shared through written publications. This human-readable knowledge may then be transformed into CBK through knowledge engineering. Such CBK can then be used privately or publicly. Other times, CBK results from analytic processes that generate it directly in a fully computable form. In these cases, no further transformation may be necessary because the CBK already exists. Considering these situations, to help make CBK findable and accessible, it is anticipated that librarians have a vital role to play to steward and manage CBK.

Once CBK exists and is made available, its users may find and access it in a variety of ways, e.g., from websites, libraries, or other repositories. These users will have a wide variety of backgrounds, interests, and needs. The different and unique roles and responsibilities of knowledge generators, knowledge engineers, librarians, and CBK users are all very important to the “CBK Story” and its related IP considerations.

General Intellectual Property Considerations

There are four major elements of intellectual property: copyright, patents, trade secrets, and trademark. Patents, exclusive rights to an invention, and trade secrets, information generally not known or ascertainable, are applicable issues to consider for CBK. Trademarks, used to protect designs and signs, are not a primary concern for CBK. This briefing largely focuses on copyright, specifically on two elements, copyrightability and fair use. Although not in the scope of this briefing, it is important to note that both copyright holders and users of copyrighted materials have certain rights by default. Creators of CBK can also obtain a license (usually for a cost) for the use of copyrighted work in an instance of computable biomedical knowledge. Additionally, some CBK instances may be subject to copyright, which may facilitate their monetization. In this case, the user of a CBK instance may obtain a license to use it. Adding some complexity, sometimes the copyrightability of a statement of fact can be hard to determine. For example, is a regression equation a statement of fact or is it a creative product of research?

Copyrightability

Copyrightability, is a key issue for CBK because it helps determine whether the content and composition of the “source knowledge” in question is protected by copyright. Copyrightability in the US is determined by creativity rather than effort, which is unlike some nations who account for the “sweat of the brow.” Moreover, a fact itself is generally unprotected by copyright. A classic example of this distinction comes from *Feist Publications, Inc. v. Rural Telephone Service Co.*, a case in which the court ruled that Feist’s phonebook, which was made by copying telephone numbers from the Rural Telephone Service directory, was not copyrightable because it was simply a collection of statements of fact. Besides fair use, which is discussed in more detail below, two other copyright considerations are *scènes à faire* and merger:

Scènes à faire is a legal concept describing works or components of works that are so conventional or typical of their genre that they are excluded from copyright protection.

Examples of items that cannot be copyrighted due to *scènes à faire*:

- The use of passive voice in scientific writing.
- The use of common phrases in screening questions, such as “on average, how many days per week do you __,” “Has anyone in your immediate family __,” etc.
- Standard notation of any kind.

Merger applies if there is only one way (or only a few ways) of expressing a fact, making the expression of the fact is uncopyrightable. Merger happens when such an uncopyrightable fact “merges” with a specific expression of that fact (because there are very few ways of expressing it), making it so that the specific expression of the fact is also unprotected by copyright.

Examples of items that cannot be copyrighted due to merger:

- The simple statement of a fact, such as “Elevated IIS induces the targeting of GJ proteins to lysosomes and degradation.”
- There exist numerous formulas to calculate GFR (a marker of human kidney function). Because each equation was created to model specific data, and they can only be written in a few ways, descriptions of these equations are subject to merger. As a result, a written description of a GFR equation found in CBK may not implicate copyright at all, because the GFR equation itself is uncopyrightable.

(Consider the example given above of a statistical prediction rule (SPR) in this light. Is an SPR copyrightable?)

Fair Use

Certain uses of copyrighted materials do not require the permission of the copyright holder; these uses are termed “fair use”. Four factors are used to determine fair use: (1) the purpose and character of the use, (2) the nature of the copyrighted work, (3) the amount and substantiality of the portion used in relation to the work as a whole, and (4) the effect of the use upon the potential market for, or value of, the work. All four factors need not apply for a use of knowledge to be deemed fair use. Two additional legal concepts also pertain, intermediate copying and transformativeness.

Intermediate copying is a legal concept that states that the use of copyrighted material during an intermediary step in a process of creation is often fair use, if the end result of the process is also deemed fair use or is otherwise legal. Examples of intermediate copying are:

- Reverse engineering of software to create a novel software product
- Copying to facilitate text mining (where the texts themselves are not published in final research), copying text into a corpus for training AI, or copying in order to create an index for a search engine.

Transformative use is the use of copyrighted material for a different purpose than for what it was originally intended. If a use is transformative, that improves the fair use analysis on behalf of the user.

Examples of transformativeness include:

- Using a calculus textbook to teach how math lesson plans are designed rather than using it to teach calculus.
- Copying media into servers for indexing in order for search engines to produce results for users’ queries.
- Maintaining a database of student papers for an anti-plagiarism tool which compares submissions to existing papers rather than using the materials to inform on the subject matter.

(Consider the example given above of an NCCN guideline in this light.

Is it transformative to make an NCCN guideline into an instance of CBK? Why or why not?)

CONCLUSION

The creation, distribution, and use of CBK falls subject to many intellectual property considerations. The nuance and complexity of these matters, and the complexity of the legal frameworks and precedents that apply, indicate that collaboration with IP experts and legal advocates is very important to ensure the legal protection of CBK for creators, consumers, and all other CBK stakeholders.

REFERENCES

Copyright Basics Guide. (2017). University of Michigan Library. <http://guides.lib.umich.edu/copyrightbasics>

Yeh, R.W., Secemsky, E.A., Kereiakes, D.J., Normand, S.L.T., Gershlick, A.H., Cohen, D.J., Spertus, J.A., Steg, P.G., Cutlip, D.E., Rinaldi, M.J. and Camenzind, E., 2016. Development and validation of a prediction rule for benefit and harm of dual antiplatelet therapy beyond 1 year after percutaneous coronary intervention. *Jama*, 315(16), pp.1735-1749.