Facilitating Collaboration On Geospatial Data Using Social and Legal Norms

Janelle C. Bonanno
Christopher Cotter
Rachel DeLetto
Cynthia Grady

New York Law School’s website can be accessed at www.nyls.edu
Table of Contents

I. INTRODUCTION .....................................................................................................................2

II. COPYRIGHT RIGHTS IN DATABASES ................................................................................2
    A. UNITED STATES .............................................................................................................3
    B. EUROPEAN UNION .........................................................................................................4

III. METHODS OF ASSERTING RIGHTS .....................................................................................5

IV. RECOMMENDATIONS .........................................................................................................7
    A. COMMUNITY AND SOCIAL NORMS ...............................................................................8
    B. INTEROPERABILITY .......................................................................................................9
    C. MARKETING AND MODULARITY ....................................................................................9

V. CONCLUSION .......................................................................................................................11
1. INTRODUCTION

Collaborative production of geospatial data offers tremendous opportunities. Allowing others to build on and contribute to existing datasets enables innovation that would not otherwise be possible. Open standards can facilitate and encourage collaborative innovation, increase the speed of data dissemination, and allow the creation of more complete data sets. Following the January 12, 2010 catastrophic earthquake off the coast of Haiti, for example, it soon became painfully apparent that not only had Haiti’s infrastructure collapsed, but also rescue and aid organizations lacked detailed maps they could use to reach earthquake victims and provide relief to survivors; existing maps of the area were rudimentary at best. In response, volunteers from all over the world began generating and collecting data about navigable roads, physical features such as hospitals and landing strips, and reports of where victims were tapped, and uploading this information to a common database established by the organization OpenStreetMap. The resulting database was capable of producing real-time maps of all of Haiti’s 10,000 square miles. Ideally, this map will continue to be used and improved upon by relief organizations as they collect additional information about the terrain and needs of the country. This amazing effort in Haiti illustrates the value, significance, and potential of “open” geo-databases – databases containing geographic and geospatial data that are made available for others to use and expand upon.

Collaborative production of intellectual works is often in tension, however, with intellectual property norms that grant legal monopolies in those works to “creators.” Those seeking to use or improve on a work must seek permission from the owner of the copyright. The open access and copyleft movements have sought to ease this tension by establishing legal and social norms that facilitate sharing. The GNU Project, which was started in 1984, is aimed at developing free software for operating systems and is now an essential component in the widely used GNU/Linux operating system. Creative Commons was created in 2001 to provide a set of user-friendly licenses that allow creators of works to specify the conditions under which their works will be shared. Geo-data, data with a spatial vector, differs from the software code and creative works covered by such licenses. Intellectual property protections for geo-data and geo-databases are typically weaker than for software and creative works. Because geo-databases often lack copyright protection, Creative Commons or other copyright-based licenses are not legally effective to control or clarify the terms for subsequent users. Even those who want to share their databases may want to ensure that the database remains available to others and that users acknowledge the source of the data. Database owners can compensate for this lack of clarity by imposing contractual conditions on later uses of the database, but such conditions are difficult to enforce against downstream users. Imposing conditions on database use may also render the data incompatible with other data that is shared under conflicting terms and conditions. Some jurisdictions have also responded to the legal


2 For an example of the amount of additional data added to the maps of Haiti available, see The volunteer mappers who helped Haiti, BBC New Magazine, http://news.bbc.co.uk/2/hi/8517057.stm.


5 History, Creative Commons, http://wiki.creativecommons.org/History.
uncertainty surrounding databases by creating new legal regimes with special protections for databases. The result is a tremendous amount of legal uncertainty surrounding the extent to and conditions under which later users can use, build on, and transform a geo-database.

Several organizations have already responded to the need for greater clarity regarding sharing of geo-databases. Among others, Science Commons offers the Open Access Database Protocol (OADP), which provides guidance for individuals seeking to release data consistently with open access principles. Open Data Commons has also created three licenses specifically designed for the release of data: the Public Domain Dedication and License (PDDL), the Open Database License (ODbL), and the Attribution License (ODC-By). These licenses offer database owners several different options for releasing their data and provide significant clarity to downstream users.

The purpose of this paper is to build on these efforts by evaluating the needs of database owners, assessing the extent to which these needs are currently met by existing licenses, and proposing additional recommendations for further facilitating the sharing of geospatial data. (The paper addresses these issues primarily from a U.S. perspective but also examines the challenges to data sharing created by the independent legal protections afforded databases in the European Union.) The paper argues that although existing regimes provide an excellent legal foundation, they should be combined with measures that enhance the use of social norms to facilitate sharing. Fostering community building and promulgating voluntary standards in clear, predicable terms can incentivize the continued openness of geo-databases and minimize the risk of incompatible data sets.

II. COPYRIGHT RIGHTS IN DATABASES

Copyright rights allow the copyright owner to open or limit the use of her property by third parties on terms she chooses. Unlike creative works, however, which are protected by copyright in their entirety, only some portions of a database receive protection. As a result, understanding the extent to which a database is protected by copyright often requires a case-by-case analysis. This is particularly true for geo-databases, which by definition are composed of geo-spatial data that is factual and does not itself enjoy copyright protection. Further, a work’s copyrightability and the terms of its copyright protection can vary widely depending on jurisdiction.

The uncertain status of copyright in geo-databases contributes to a lack of clarity not present with respect to other types of works. These complications increase for downstream users seeking to understand the availability and terms of use associated with a database to which many have contributed. Conditions imposed by different contributors may also be incompatible, thus preventing some or all uses of the data.

The fact that database owners have historically used a variety of licenses, including copyright-based licenses that do not apply to databases, has only exacerbated the confusion. Further, although many are committed to open uses of their databases, “data is rarely made available in a manner that makes it straightforward to ascertain the uses to which it may subsequently be put by a third party.” Without knowledge of the applicable terms third-party users could risk costly litigation for an unauthorized use.

---

A. United States

In the United States, geo-database owners have very limited copyright rights in their databases and generally no rights in their databases’ contents. Any existing rights in the database are also easily lost if the information is transferred to another electronic format.

In the United States, the copyright in each separate contribution contained within a database is distinct from the copyright in the database itself. A database creator only has copyright in the database, not its underlying contents. Each part—the individual pieces of data and the database itself—must qualify for copyright protection separately.

A geo-database’s contents are not likely protected by copyright because geographic data is factual and thus outside the scope of copyright. Copyright law protects only the expression of facts, not facts themselves. For example, a Haitian hospital’s longitudinal and latitudinal position is a fact and therefore not copyrightable under U.S. law. In contrast, a user’s drawing of the hospital embedded into a map at that same longitude and latitude may be copyrightable because the drawing is an expression of the hospital’s appearance.

However, the database itself may receive protection as long as it reflects some minimal level of creativity in the selection or arrangement of its contents. It is not sufficient that the database creator invested considerable time and effort in creating the database. As the Supreme Court explained in *Feist*, the “sine qua non of copyright is originality.” Thus, in the United States, copyright law will protect only databases with an original selection or arrangement. As the Court noted in *Feist*, “[t]his inevitably means that the copyright in a factual compilation is thin.” For example, a database containing the names of all of Haiti’s roads arranged in alphabetical order is simply a compilation of facts. An alphabetical arrangement, though logical, is not original, and thus not protected by copyright. Nor would the database creator have copyright based on the selection of the database’s contents. The database contains the names of all the roads in Haiti. Selecting all of the known Haitian road name data is not original or creative, just comprehensive. In contrast, a database containing the names of all Haitian roads that have a church on them would more likely be copyrightable because selecting just those roads with a church is original.

In most cases, the selection and arrangement of data in a geo-database is unlikely to qualify for copyright protection. Often, the goal of a geo-database creator is to provide contents that are as comprehensive as possible. Including all data that is known or available for collection will not qualify as an original selection. Further, the data’s arrangement likely will not be original. Typically the data will be arranged functionally, and thus unoriginally, such as alphabetically or by location.

In addition, any copyright in a database’s arrangement may be easily lost if the data is transferred to another electronic format. For example, a New York federal appeals court has held that

---

11 Even if the contents of a geo-database were sufficiently creative to qualify for copyright, they might be nonetheless excluded if they were originally created by the U.S. government. Most notably U.S. government works, such as those created by US Geological Survey (USGS) are not copyrightable. 17 U.S.C. § 105. See also USGS.gov, Maps, Imagery, and Publications, http://www.usgs.gov/pubprod/ (last visited Apr. 18, 2010).
12 Courts treat databases as compilations. 17 U.S.C. § 103. § 103 defines compilations as “a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship.”
14 Id.
copyright in a compilation of factual content was lost when the compilation was copied and transferred to a CD-ROM. Because the transfer stripped the compilation of its arrangement, the resulting CD was not protected even if the files were thereafter rearranged in the same order as the original database.\textsuperscript{15}

The scope of copyright in a database and what uses may be made of copyrighted works also varies even within the United States, making it even more difficult for database owners and users to ascertain the extent of the database owner’s copyright rights in the database. For example, in \textit{CCC Information Services}, the Second Circuit adopted a low threshold of originality, holding that logically arranging data does not preclude the arrangement from being original.\textsuperscript{16} Moreover, the fair use defense, which permits certain uses that might otherwise constitute copyright infringement, also requires a highly subjective, fact-specific inquiry and is interpreted differently across jurisdictions. As a result, U.S. copyright law provides database owners with limited and extremely unpredictable protection.

\textbf{B. European Union}

The European Union protects databases by statute and provides them with far greater protection than they are entitled to in the United States. The 2006 EU Database Directive (96/9/EC)\textsuperscript{17} creates a right similar to an intellectual property right (an independent or “sui generis” right) that protects the effort invested by database owners in collecting or compiling data in a database. The Database Directive does not create a new right in the underlying data and does not affect any underlying copyright in the data.\textsuperscript{18} To qualify for protection under the Directive, a database’s creator or owner must be either a national or a habitual resident of a Member State of the European Union.\textsuperscript{19}

The Directive allows database owners to prevent others from extracting or reutilizing the whole or a substantial part of the database if the database is the result of a substantial investment by the creator in obtaining, verifying, or presenting the contents.\textsuperscript{20} The investment can be either quantitative or qualitative, but must be an investment in the \textit{obtaining, verification, or presentation} of the data, not the generation of the data itself.\textsuperscript{21} For example, a database creator who has spent hundreds of hours and thousand of dollars verifying street names has made the type of quantitative investment in data verification that might provide the database with protection under the Directive. In contrast, if a database creator spent the same number of hours and money in simply generating data points, the database may not qualify for protection under the Directive. However, the distinction between obtaining and generating is unclear and has not been the subject of an authoritative interpretation by the European Court of Justice. Thus, the database creator might argue that capturing data points via GPS is not generating but collecting data, in which case it may have made sufficient investment in obtaining data to qualify for protection.

\textsuperscript{15} Matthew Bender & Co. v. West Publishing Co., 158 F.3d 693 (2d Cir. 1998), cert. denied 1999. The selection of the cases in the database was not at issue.

\textsuperscript{16} CCC Information Services, Inc. v. Maclean Hunter Market Reports, Inc., 44 F.3d 61 (2d Cir. 1994).

\textsuperscript{17} See Database Directive, Chapter II: Sui Generis Right.

\textsuperscript{18} Database Directive, Preamble at ¶46 (“The underlying data in a database can be collected again by a third party, but the third party cannot obtain the underlying data by extracting it from the protected database”).

\textsuperscript{19} Database Directive at art. 11(1).

\textsuperscript{20} Database Directive at art. 7(1), Preamble at ¶46.; Database Directive at art. 7(2).

III. METHODS OF ASSERTING RIGHTS

There are several ways in which database owners might clarify and communicate the scope of their claimed rights in the database. Typically, they can do so through one of three methods: a license, a contract, or dedication to the public domain. Because of their limited rights at least under U.S. law, if database owners do want to impose conditions on subsequent uses of their databases, they must use both a license and a contract to achieve this goal. A database owner will choose to dedicate the work to the public domain only where he or she does not desire to impose any conditions on subsequent uses. The purpose of this section is to evaluate the pros and cons of each of these three methods given the legal frameworks currently available to database owners.

The owner of a copyrighted work can use a license to communicate and enforce terms for subsequent uses, including terms that require the work to remain available for new uses. To be legally enforceable, however, the licensee must have an ownership interest in the underlying work. As described above, geo-databases enjoy limited, if any, copyright protection. As a result, the legal enforceability of a license as applied to a geo-spatial database is dubious.

Despite the lack of legal enforceability, databases, including geo-databases, have been released under copyright-only licenses, such as the licenses provided by Creative Commons. Database owners who wanted to preserve the openness of their data found Creative Commons’ CC-BY-SA license particularly attractive for its attribution and share-alike provisions. SA stands for “share-alike,” which requires users who alter, build upon, or transform the work in any way to re-license the resulting work under the same license (i.e., “CC-BY-SA”). BY stands for “attribution,” which requires users to attribute their use of the owner’s work in the manner specified by that owner. Nonetheless, using a copyright-only license for works that are not likely protected by copyright has been critiqued as exacerbating rather than reducing the lack of clarity regarding subsequent uses of databases.

Database owners can also impose contractual conditions on subsequent users. Unlike a license, a contract requires no legal property right; a variety of terms and conditions can be applied to almost anything during a transaction between two parties, as long as the exchange is bargained for and both parties either give up or receive something in return as consideration. Thus, a database owner could seek to impose contractual conditions on database users even if the owner lacks a copyright in the database itself. The enforceability of contractual conditions may be weakened, however, for downstream users. A contract only binds the parties to it and may not be enforced against anyone who is not a party to that contract. In order to bind downstream users, a contract must require the original user to impose similar contractual conditions on any subsequent user. If the original user does not, however, or if the subsequent user ignores the conditions, the database owner’s only remedy is against the original user. Further, like copyright law, contract law differs by jurisdiction, and it may be necessary to evaluate the enforceability of a contract on a jurisdiction-by-jurisdiction basis.

Licenses for creative content are often paired with a supplemental contract. That way, if the license fails for some reason—if the use is found to be fair use or the owner’s copyright rights are limited—the content owner can rely on the contractual terms to achieve the same or similar goals which the license provided. For example, a database owner whose database consists of longitude and latitude vectors can license the use of this database and pair the license with a supplemental contract that would bind the subsequent user even if the selection and arrangement of the database is found not original enough to warrant copyright protection. Indeed, the U.S. Court of Appeal for the Federal Circuit has found such contracts to be enforceable.

Despite the problems inherent in a contractual approach, the limited copyright in databases has led those working on proposed database licenses to combine a license with a contract. The Open Database License (ODbL) is one such example. In 2008, the OpenStreetMap Foundation established a License Working Group to collaborate with OpenDataCommons—a not-for-profit entity working to provide legal tools for the benefit of the open knowledge community—in developing an open and reciprocal database license.\(^\text{24}\) The collaboration launched an extensive international effort that considered many of the open data community’s members’ concerns and opinions and evaluated new license terms against practical examples of how data under the license might be used.\(^\text{25}\) In 2009, OpenDataCommons released the product of this process, the Open Database License (ODbL v1.0).\(^\text{26}\)

The ODbL seeks to clarify the terms under which people can use and contribute to databases subject to its provisions while also preserving the fundamental share-alike and attribution requirements that characterize the Creative Common’s CC-BY-SA license.\(^\text{27}\) The ODbL is both a license and a contract. It provides a license for any rights in the database arising under either copyright law or the EU Database Directive.\(^\text{28}\) To the extent those rights do not exist, it also acts as a contract between the database licensor and the user, under which the user agrees to the ODbL’s conditions in exchange for access to the database.

The ODbL was designed to protect a “database.”\(^\text{29}\) Section 2.4 delineates the relationship between the individual content items that comprise the database and the database as a whole. It provides that the ODbL governs only reuse of the database itself and not the individual contents it contains. The ODbL does not affect any intellectual property, privacy or personality rights in an individual image, description or audiovisual clip added by a contributor. Thus, to the extent that any of this individual content is protected by copyright, reuse of that content would require a license from the original contributor.\(^\text{30}\)

The ODbL imposes both a “share-alike” and an attribution requirement on any subsequent uses of the database itself (a “database use”).\(^\text{31}\) It encourages “non-database uses” of the database by excluding from the share-alike requirement any uses of the database to generate an image (i.e., a map tile), audiovisual material (i.e., television newscast), text, or sound.\(^\text{32}\) Non-database users are also exempt from the ODbL’s attribution requirements unless the work containing the data is “Publicly Used” (although this term is not well defined).\(^\text{33}\) The Attribution License (ODC-By), released by

---


\(^{26}\) Open Data Commons, Open Database License (ODbL) v1.0 Released, http://www.opendatacommons.org/2009/06/29/open-database-license-odbv1.0-released/ (last visited Apr. 18, 2010). In October 2009, the License Working Group proposed that OpenStreetMap adopt the ODbL v1.0, and opened the proposal to the OpenStreetMap community for discussion. See New License Proposal, supra note 24.

\(^{27}\) OpenStreetMap’s license also contains contributor terms which require that individual contributors grant OpenStreetMap a worldwide, royalty-free, non-exclusive, perpetual, irrevocable copyright and database rights license for their contribution that includes the right to sublicense their work. See The OpenStreetMap Foundation, OpenStreetMap Contributor Terms, http://www.osmfoundation.org/wiki/License/Contributor_Terms (last visited Apr. 18, 2010), at § 2; New License Proposal at 3-4. The inclusion of contributor terms makes clear that subsequent users need only attribute use of the database to OpenStreetMap, not to the individuals who contribute data to the OpenStreetMap database.

\(^{28}\) ODbL v1.0, § 2.2.

\(^{29}\) ODbL defines database using analogous terminology to the EU Directive. See supra at n4.

\(^{30}\) ODbL v1.0, § 2.4.


\(^{32}\) These types of exempt uses are referred to in the license as “Produced Works.” ODbL v1.0, § 4.5(b).

\(^{33}\) ODbL v1.0, § 4.3.
Open Data Commons, resembles the ODbL in many ways; it also pairs a license and a contract and limits its protection to the compilation of the database and not its content, but it requires only that subsequent users make proper attribution and does not require subsequent uses to be available for reuse.\textsuperscript{34}

A copyright owner can also make the scope of his or her claimed rights clear by explicitly dedicating the work to the public domain.\textsuperscript{35} Information in the public domain is not subject to any restrictions. There are two primary ways for works to be in the public domain: when a work is ineligible for intellectual property protection, and when the intellectual property protection the work previously enjoyed either lapses or the owner disclaims it.\textsuperscript{36} Disclaiming copyright protection and releasing a work into the public domain is an active decision by the author to forgo any future rights.

Databases that do not qualify for copyright protection are in the public domain. However, even if the database owner does not have any rights in the database, users may still be inhibited from using the database if they do not realize this. A rightsholder (or possible rightsholder, if the copyright status is unknown) can choose to release any rights it might have into the public domain by waiving any potential copyright protection to the fullest extent permitted by law thereby allow potential users to easily understand the conditions under which they can use the database.\textsuperscript{37} The Open Data Commons’s Public Domain Dedication and License (PDDL) allows rightsholders to dedicate their databases to the public domain.\textsuperscript{38} For creative works, the Creative Commons CC0 license disclaims any copyright protection offered by applicable law.\textsuperscript{39}

There are benefits to fully releasing rights over a database into the public domain. Most importantly, doing so allows the widest dissemination of information. This can be extremely important in fields like science, where building upon older work is fundamental to new discoveries. Indeed, many believe dedicating geospatial databases to the public domain is the best way to keep them open.\textsuperscript{40} It will also encourage others to use them because a public domain license assures potential users that they can use the databases without any obligation to the database owner. However, there are also disadvantages to fully releasing rights over a database. One concern is that the rightsholder no longer has any control over what happens to the work, including how it is used, by whom it is used, and if he or she receives any attribution. Releasing all potential rights in a geospatial database, therefore, like any license choice, is a determination made by a rightholder based on her priorities.

\section*{IV. Recommendations}

Although improvements can still be made, existing database licenses such as the ODbL provide an solid legal framework for licensing geo-databases. Such licenses alone, however, are not likely to be effective at increasing the widespread availability of geo-databases. The threat of legal liability, particularly if it is not paired with rigorous enforcement efforts, may not itself be sufficient

---

\textsuperscript{34} Open Data Commons Attribution License (ODC-By) v1.0, http://www.opendatacommons.org/licenses/by/1.0/.

\textsuperscript{35} Dedicating all existing intellectual property rights to the public domain is not a license but rather a release of the rightsholder’s interest. It functions like a perpetual license from the rightsholder to society.

\textsuperscript{36} For example, mathematical equations and scientific formulas are not protectable (e.g., E=mc$^2$ is in the public domain). In contrast, a work of literature, such as Charles Dickens’s \textit{A Tale of Two Cities} was protected by copyright when it was published in late eighteenth century, but that copyright has long since expired. See generally 17 U.S.C. §§ 301-305. Hence, just like E=mc$^2$, \textit{A Tale of Two Cities} is in the public domain and available for use by the public without any restriction.

\textsuperscript{37} In the case of a database, the rightsholder can only release any potential copyright the database. The database rightsholder cannot release the copyright in the underlying data, unless it also owns that copyright.

\textsuperscript{38} ODC Public Domain Dedication and License (PDDL), http://www.opendatacommons.org/licenses/pddl/1.0/.

\textsuperscript{39} See Creative Commons CC0, http://creativecommons.org/choose/zero (last visited Apr. 18, 2010).

\textsuperscript{40} Once something is released into the public domain it remains there indefinitely.
to encourage widespread compliance with conditions that require subsequent uses to remain open. In addition, legal terminology is often a highly ineffective way of communicating the conditions under which a work is released. One of the reasons behind the success of projects such as GNU and Creative Commons has been the decision to pair legal with social norms to encourage compliance with the conditions under which owners released their creative works for reuse. Similar efforts are underway currently for database licensing through the Open Source Initiative. The purpose of this section is to discuss both improvements that might be made to the legal framework established by licenses such as the ODbL and to propose several additional initiatives that could be used to strengthen the social norms that support compliance with these regimes.

A. Community and Social Norms

Licensing and contractual regimes might be paired with structures that encourage the development of community or social norms. Historically, such social norms have been very effective in ensuring such compliance with open source principles, at least in smaller communities. For example, even though Creative Commons licenses are effectively unenforceable for most database licenses, they functioned well to control downstream uses because subsequent users believed they should comply with the attribution and “share alike” requirements they purported to impose. Such licenses might also be respected either because users do not realize their limited legal enforceability or because the community enforces the license outside of the legal system.

Such mechanisms, however, are difficult to sustain on a larger scale. As Miller, Styles and Health explain: “In small, tightly-knit groups where interchange of data may be governed by existing social norms this may rarely present a problem. However, with data interchange and interoperability reaching Web scale, social norms alone cannot be relied upon to enforce fair and appropriate usage of data.” In larger group settings, social norms might be scaled by pairing them with legal norms and institutions. Legal norms establish a baseline for social norms. Despite costs to some users, legal norms, even when not often enforced, create a feeling of stability within a community and a basis for community norms to flourish. Further, the process of agreeing on legal norms can serve a constitutive function, creating community while identifying the values that the community shares.

Much of this has already occurred through the consensus process undertaken in the drafting of the ODbL. This process and the resulting license, however, will be most effective when paired with an institution that can serve as a continuing “home” for discussions about community values and how these values might need to adapt to changing circumstances. The Open Data Commons (ODC) is currently providing this kind of community space. ODC is the home of three different database licenses and has even drafted a set of “Community Norms” that describe “best practices” in data licensing. There are, however, a few additional things that ODC (or an organization like it) might undertake to foster the development of social norms around database licensing. For example, toward its goal of fostering “community-building, education, and public advocacy,” Open Source Initiative (OSI), a non-profit community group focused on open software, adopted an “Open Source Definition,” which it has promoted as an industry standard for “open source” in the software industry. OSI offers to review licenses submitted to it for their compliance with this standard. Like the OSI, community organization focused on database licensing like ODC might become involved in

---

reviewing and approving licenses based on their conformity with such best practices.\textsuperscript{45} Such an organization might also offer to provide technical assistance to organizations seeking to adapt its licenses to their specific needs.

\textbf{B. Interoperability}

Database licensing regimes should also pay particular attention to ensuring interoperability. Imposing conditions on data—even if only to ensure that the data continues to be available for others to use—can make it difficult to combine that data with data from other sources. Databases often combine data from many different sources to create a new, more robust product or service. If each source operates under a different license that requires users to “share-alike” under the specific terms of its own license, the sources cannot be combined into a new, unitary or collective database. Alternatively, even if it were possible to distribute combined data sets under the terms of different licenses simultaneously, the resulting combinations would be subject to an ever-increasing number of conditions. Furthermore, any future change in a license may require purging of any data governed by a different license that did not authorize such a change. Legal interoperability is a significant obstacle to the ultimate goal of encouraging open access to and growth of databases.

There are several technical and institutional techniques that might be employed to enhance interoperability. First, database licenses can exempt certain kinds of uses from “share alike” provisions. The ODbL, for example, exempts “non-database” uses, or uses of data to generate something other than a database. Geo-database owners do not have the same interest in ensuring the continued openness of, for example, maps or other images generated from the data as they do in the data itself, and exempting these uses would encourage the production of these works. Second, the involvement of a community-based standards organization in guiding licensing efforts would also help to address concerns about interoperability. Common licensing provisions from the same source with a central oversight mechanism would reduce the risk of license conflict. Guidelines or “best practices” might also help to standardize language, which would aid users in understanding and comparing licenses. Third, licenses might be made more interoperable through the incorporation of a standard “escape clause” that would change the terms of the share-alike requirement to conform to those of an otherwise incompatible data set. Such an escape clause would provide that if the database is combined with another database under conflicting license terms, both data sets would automatically be subjected to either the most restrictive or the least restrictive conditions imposed by the two licenses.

\textbf{C. Marketing and Modularity}

To be effective, the terms under which a license is released must be clearly understood by subsequent users. The actual legal terminology used in a license is generally ineffective in accomplishing this goal. Further, without broad awareness of their availability, such licenses are unlikely to gain widespread adoption. As a result, some of the most successful licenses are those which have been paired with a marketing campaign to inform users about the availability of the licenses and to communicate the content of their terms. Creative Commons, for example, has gained widespread adherence in part because it was paired with an extremely effective marketing campaign. Many internet users know about the Creative Commons licenses and have used these licenses even if they are not completely sure of the implications of such use.

Creative Commons has also enjoyed widespread adoption in part because of its modular approach. Creative Commons provides content owners with a limited menu of easy-to-understand

\textsuperscript{45} Open Source Initiative, \textit{History}, http://www.opensource.org/history (last visited Apr. 18 2010).
licenses from which they can select to create a personalized level of control over third party uses of their creative works.\textsuperscript{46} This model provides an especially good example of effectively simplifying complicated license terms into manageable components that any layperson can easily understand. A modular framework provides a predictable licensing scheme while preserving flexibility in terms and conditions.

To a large extent, this goal is already accomplished by the three licenses offered by ODC—public domain, attribution only, and attribution plus share-alike (the ODbL). Each of these licenses offers a plain language summary of the license’s legal terms.\textsuperscript{47} A modular approach to database licensing, however, might expand on the Creative Commons approach and provide not only modular licenses but modular provisions. Such a scheme would involve developing a set of predetermined, uniform contract clauses and licensing conditions that database owners could pick and choose from in applying terms and conditions to use of their data, based on their particular needs.

A modular approach may be, however, in some tension with interoperability. Modularity would increase the variety of conditions that might be imposed on different databases. Unless these conditions are interoperable themselves, modularity may render more databases incompatible with one another.\textsuperscript{48} Creative Commons recognizes this problem with respect to creative content,\textsuperscript{49} noting that “[s]ince each of the six CC licenses functions differently, resources placed under different licenses may not necessarily be combined with one another without violating the license terms.”\textsuperscript{50} This disadvantage might be tolerable with respect to creative works, since any limitations in a user’s ability to combine works with one another might be outweighed by the benefits of clearly understanding what uses can be made of the work. However, with respect to databases, where the primary purpose of the database is to combine it with other data to achieve new uses, interoperability limitations likely are not outweighed by other benefits.\textsuperscript{51} As a result, modularity may not be advisable for databases.

To increase understanding by subsequent users, greater attention might also be paid to the drafting of the licenses themselves. For example, while commendable in many respects, the ODbL is not as clearly drafted as it might be. This complexity might discourage users from making subsequent uses of a database released under its terms. In addition, some of the defined terms in the license are ambiguous. For example, there is an ongoing debate regarding what types of uses classify as “Produced Works.”\textsuperscript{52} It is unclear whether a use that superimposes restaurant ratings onto a map image generated from the OSM database would constitute a work covered by the database. Fears that the share-alike license term will attach to proprietary data hinder collaboration with organizations that would add such proprietary data to an open database. To the extent that the licensee wants to encourage such uses, this ambiguity should be resolved.

\textsuperscript{46} Licensors can require attribution (BY) or that any subsequent works created using the original are kept open for reuse by others (SA). They can also prohibit others from using their works for commercial purposes (NC) or creating derivative works (ND).


\textsuperscript{49} Creative Commons, FAQ: Can I combine two different Creative Commons licensed works? Can I combine a Creative Commons licensed work with another non-CC licensed work?, http://wiki.creativecommons.org/FAQ#Can_I_combine_two_different_Creative_Commons_licensed_works.Can_I_combine_a_Creative_Commons_licensed_work_with_another_non-CC_licensed_work.3F_Can_I_combine_a_Creative_Commons_licensed_work_with_another_non-CC_licensed_work.3F (last visited Feb. 16 2010).


\textsuperscript{51} Creative Commons, Case Studies, http://wiki.creativecommons.org/Cases (last visited Mar. 23, 2010).

V. Conclusion

The legal framework established by existing database licenses provides a very good foundation. In addition, as the audience for geo-database expands, social norms will be ineffective in ensuring compliance with conditions of continued openness and should be paired with legal norms. Only through attention to the interaction between legal and social norms will we see the development of the confidence necessary to foster greater database openness. Legal rules help create and establish the boundaries for community norms; community norms in turn provide the basis for the legitimacy of legal rules. Both can be channeled by the creation of institutions that provide guidance on the meaning and application of these norms and rules. This paper attempts to provide modest recommendations concerning the continued development of these legal, social and institutional frameworks in the area of database licensing.