

# Legal Interoperability in Support of Spatially Enabling Society

Harlan J Onsrud

onsrud@spatial.maine.edu

University of Maine, USA

## ABSTRACT

Spatial data is critically important for the wellbeing of society. Yet appropriate spatial data is often very difficult to find and, when found, the legal ability to use it is often in question. Lack of an operational web-wide capability allowing users to legally access and use the geospatial data of others without seeking permission on a case-by-case basis remains as an entrenched major impediment to general spatial enablement for all sectors in society. This chapter presents a legal interoperability vision for offering, acquiring, and using spatial data and proposes an operational environment for gaining much greater legal clarity and efficiency in wide scale sharing and licensing of such data.

**KEYWORDS:** Legal Interoperability, Licensing, Usage Rights, Open Access, Commercial Licenses

## 1. INTRODUCTION

Society and science have a problem. Spatial data is critically important for the wellbeing of society. It is used by all sectors of society including by individuals, businesses, non-profit organizations, and local, state, national, and international government agencies and organizations. These entities use it for addressing challenges and solving problems related to transportation, housing, disasters, health, energy, climate, water, weather, ecosystems, agriculture, and biodiversity. It is used daily for immediate pragmatic decision-making and for more thoughtful long-term planning. In short, we use spatial data and services pervasively across all sectors to satisfy immediate needs and to address our most pressing long-term societal challenges.

Yet appropriate spatial, geographic or location data is often very difficult to find. If and when found, the suitability of the data for specific uses is often unknown, whether one can legally use the data is frequently in question, integrity is not typically traceable and liability exposure in using the data is often problematic. These challenges need to be addressed before effective use of spatial data and services can fully develop to spatially enable citizens more broadly in every day living and business.

One of the reasons that spatial data is so difficult to find is that current search engines process words, phrases, and sometimes embedded code. However, spatial data does not consist of words or phrases nor has identifying code yet been deployed generally in spatial data sets to aid in finding them. That is, one cannot ask a search engine to find all image and vector datasets that contain the visible outline or representation of a particular building and expect a germane listing of datasets to result. However, lack of ability to readily find through word searches resources such as geospatial data sets, location-referenced data and scientific and technical datasets in general existed long before the Internet. As a result, the geospatial

community in conjunction with the information and library communities has developed processes for reporting metadata, services and products and further capabilities for finding the reported works. The language of repositories, portals, clearinghouses, registries and archives permeate disciplinary discourse. This imposition of creating and reporting metadata and supporting registries and repositories creates a cost or overhead that many in the community resist, particularly if the primary benefits in enhanced ability to find data and resources accrue to others.

Regardless of these challenges, the technological means for achieving interoperability of many forms of geospatial data, products and services, establishing viable registries, and enhancing findability are within sight and deemed achievable by many within the technical community. What is not yet in sight for many is the means or ability to achieve legal interoperability.

I define legal interoperability for data as a functional environment in which

- differing use conditions imposed on datasets drawn from multiple disparate sources are readily determinable, typically through automated means, with confidence,
- use conditions imposed on datasets do not disallow creation of derivative products that incorporate data carrying different use conditions, and
- users may legally access and use the data of others without seeking permission on a case-by-case basis.

When these characteristics exist, the legal use conditions outcome for a derivative dataset or product may be automatically computed. Lack of an operational web-wide capability allowing users to legally access and use the geospatial data of others without seeking permission on a case-by-case basis remains as an entrenched major impediment to general spatial enablement for all sectors in society. (National Research Council, 2004)

## **2. THE TECHNOLOGICAL VISION**

Assume that in an ideal operational environment, a potential user searches for and finds seven geo-referenced datasets out of thousands accessible over the Internet that each meets her requirements for geographic coverage, subject material, resolution, and technical standards compatibility. By extracting all or portions of these seven data sets, perhaps using a web mapping or feature service, she creates a new product or map that meets her needs and which she desires to further distribute and make available on the Internet and through other publication avenues. As noted, many in the technical community are striving towards achievement of a technological vision that would allow such actions to better take place. While the vision is laudable, it fails to take into account and automatically assess whether the seven found data sets are legally interoperable as well as technically interoperable. This is a substantial world-wide problem that will continue to undermine and impede the ability of individuals, businesses, non-profits, and government agencies to use the geospatial data of others even when technical interoperability is a much more prevalent operational reality. The operational vision of the geospatial community needs to be expanded.

## **3. THE PROBLEM<sup>1</sup>**

Copyright automatically exists in expression upon creation of that expression in tangible form whether the creator wants it or not in most jurisdictions. This means that for most works discovered on the Internet or elsewhere one should assume that some party has an ownership

---

<sup>1</sup> This section summarizes and directly extracts from Onsrud, 2009.

right in the work. Just because one finds a music file readily available on the Web does not mean that legal authority has been granted to download it or to incorporate all or part of the file content into a derivative work. The same general rule applies to geographic data files. By far the greatest legal liability exposure for those pursuing exchange and integration of geographic data in a networked world of interoperable web services and data mining is incurred through the violation, whether intended or unintended, of copyright, database legislation, and similar intellectual property protections. Such laws are in place and are being continually strengthened by international treaties and by national legislatures around the world.

Therefore, if someone creates a story, song, image, or dataset and places it openly on the Web, it is not necessarily free for anyone to copy without permission. *Common practice, not getting caught, or small likelihood of being sued* are not equivalent to having a clear legal right to copy. The lawyerly response to the question of whether one can copy particular data or datasets is that the answer will depend on answers to several additional questions. As a general proposition, however, there will be some legally protected originality in the vast majority of digital works made accessible through the Internet. If legal originality exists, the law assumes one must acquire permission from the copyright holder to copy, distribute, or display the work, or to generate derivative products from it.

One might argue that “data” or “empirical values” drawn from a database are all legally equivalent to “facts”, and therefore are not protected by copyright in many jurisdictions. Even if that were true in a specific jurisdiction in a specific case, the creative selection, coordination, and arrangement of “facts” is protected by copyright in most jurisdictions. Further, the explicit legal tests for qualifying for, or determining what is protected by, copyright vary from jurisdiction to jurisdiction. Some jurisdictions protect *sweat of the brow* and *industriousness*, even absent originality or creativity, and many jurisdictions supply protections for datasets and databases that extend well beyond those granted by copyright (e.g., database protection legislation, unfair competition regulations, moral rights, catalogue rules, etc.).

Our hypothetical user searched for and found seven web accessible geo-referenced data sets she wants to use. Many of us might assume that contributors to the seven sites from which she downloaded the datasets, all or in part, probably placed their datasets on the Web and are adhering to data format and other interoperability standards so that others might freely benefit from their postings. Yet, the laws of most nations generally hold that we must *not* make this assumption. The likelihood is high that one or more of those seven sites has posted terms of use that one or more of her intended uses breaches. Furthermore, the posted contract or license provisions of some of the sites are very likely to be in conflict with each other. Some of the seven sites will have no license language or use restrictions posted. In those instances, the intellectual property laws of some particular nation will apply by default. A user is required to meet the national requirements of all of the involved sites unless the user has explicit permission stating otherwise.




Copyright liability is a *strict liability* concept in many jurisdictions: no intent to break the law, or even having knowledge of breaking the law, is necessary to be found guilty. Even innocent or accidental infringement may produce liability (e.g. *DeAcosta v. Brown*, 1945). For each violating extraction or copying, the potential damages are huge, and the possibility of incurring damages is having a chilling effect on using the geographic data of others.

#### 4. AN EXPANDED VISION INCORPORATING LEGAL INTEROPERABILITY

In an ideal legal interoperability vision, when the user in our hypothetical finds the seven geo-referenced data sets that ideally meet her thematic and technical interoperability needs, she would be informed automatically of the legal status of each data set that she desires to use in her web mapping application. In fact, she could limit her search to report back only those data sets that meet her specific legal requirements in addition to her technological requirements.

For illustrative purposes, however, let us assume she has searched based on only technical and thematic constraints and now discovers the legal status conditions for each of seven technically ideal data sets as set forth in Table 1. An appropriately designed web-wide metadata reporting and search environment provides her with the search result that the seven found data sets each has a different legal status as shown. The legal status of each dataset would be automatically identified either through embedded code in the dataset and/or through the metadata for the data set.

Table 1. Legal Status of Seven Source Geospatial Data Sets

Discovered Dataset	Type of License	Suggested Specific License <sup>2</sup>
1.	Dedication to the Public Domain (cc 0)	
2.	Attribution License (cc by)	
3.	Attribution Non-Commercial License (cc by-nc)	
4.	Standard Industry-wide License: User Selection of Applicable Conditions	
5.	Offeror License: Idiosyncratic to offering firm, agency, group or individual	
6.	Conditions of Use Published	
7.	No Legal Status Information Provided	

*Seventh Data Set:* The seventh data set represents the status quo for most geographic datasets accessible through the Internet today. That is, the user has little to no indication as to the legal status of the data set. As noted in the previous section, this means she has a legal obligation in this and the vast majority of similar circumstances to find the owner and seek the owner’s permission to use the data set as a component or part of the expression in her derivative product. This is true even though the data set may be openly posted on the web with no technological restrictions on gaining access to it. To use the data set without such explicit permission greatly increases liability exposure for the user. She does so at her peril.

*First Data Set:* Assuming an appropriately designed information infrastructure supporting consistent legal status reporting, the potential user now has a clear indication for the first data

<sup>2</sup> Creative Commons, 2010. Additional licenses under categories 1 and 2 respectively that would largely have the same ramifications include Public Domain Dedication & Licence (PDDL) (<http://www.opendatacommons.org/licenses/pddl/>) and Open Data Commons - Attribution License (ODC-By) (<http://www.opendatacommons.org/licenses/by/>)

set that a purported owner has made the data set legally available to the world by dedicating the data to the public domain or, in the event that a public domain dedication is not enforceable under the laws of some nations, waiving all copyright and other rights to the greatest extent possible allowed by law in the contributors jurisdiction.

*Second and Third Data Sets:* In the second and third data sets the potential user now has a clear indication that purported owners have made the data freely available for her and anyone else to use in derivative products as long as she lives up to the provisions in the standard open access license language. That is, if she uses the second data set she must provide attribution to the party that has affirmatively licensed that data to the world. If she uses the third data set she must give attribution to the contributor as well as not use the data for a commercial purpose.

Note that the licenses for the first three data sets are legally interoperable with each other in that if a derivative product is created from the three data sets, the provisions of all three licenses may be supported. Assuming that significant portions of each of the source data sets are expressed in the derivative product, the license with the most stringent conditions would control the use of the derivative work. That is, our hypothetical user would not be able to use her derivative map or other product for a commercial purpose even though only a portion of it was derived from the data source with the commercial restriction. If she wanted to use the derived map for commercial purposes she legally would need to acquire permission of the owner of the third data set to do so or find an alternative to the third data set without the restriction.

In this instance let us assume that the user wishes to use the resultant derivative work for non-commercial purposes. Thus, in the present case, if she uses only these three data sets, she knows she has an affirmative legal right to use all three of them. No case-by-case human communication was required to gain legal clarity in her derivative product and she knew the legal status of her derivative product automatically and almost instantaneously.

*Fourth Data Set:* The fourth data legal status result assumes that the global market for geographic data has become much more unified through a global approach to open intellectual property rights management. The geospatial data sales and services market is extremely fragmented currently. Offering a global and industry consistent suite of standard commercial data license options in a web-wide deployment would make the commercial market much more transparent and efficient and allow large numbers of suppliers of geospatial data to better compete with each other in offering data and services. This would likely grow the demand for such products. To date, commercial companies and those government agencies across the globe pursuing data sales economic models have not sought to unify or grow the market along these lines. One might think of this as an E-Bay model for selling geographic data in that the very smallest to the very largest sellers could make their offerings available in a consistent manner through such an operational environment.

In the proposed approach, any data producer would click through an extensive list of universally consistent standard commercial license provisions and provide a matrix of prices for data offerings based on the volume to be licensed by a user and on various combinations of restrictions imposed on the data. An interested user would click to remove only those license restrictions on the data that would allow them to use the data for their intended purpose. The fewer restrictions the user removes with a click and the larger the quantity/coverage/themes purchased, the lower the price would be typically in acquiring the data set. The pricing matrix would be defined by the owner/offeror when making their georeferenced data available and the pricing matrix could be changed by the owner at any time in

order to support ready flexibility in responding to changing marketplace conditions. A specific data supplier might disallow the removal of some provisions at any price. By example, many data suppliers might not allow removal of standard risk allocation provisions (e.g. liability waivers) except through direct negotiations.

An additional concept here is that after an initial dataset download is made by a user upon payment, technology is used not to lock out use or access by other than the purchaser but technology is instead used to make purchasers accountable. It does so by attaching personally identifiable information to purchased files so that those who blatantly and widely abuse the license by distributing their purchased file to others in violation of the license provisions can be identified and pursued. (Onsrud, 2008)

Again, in this instance, no case-by-case human communication was required to gain legal clarity in her derivative product. Upon clicking the licensing terms she needed and paying the fee by credit card, our user was able to download her desired dataset within minutes and knew the clear legal status of both her downloaded data set and that of her derivative product automatically and almost instantaneously. If designed appropriately, these commercial license variations would again be legally interoperable with the licenses of 1 and 3 in that if a derivative product is created from the first four data sets (assuming that the user is automatically notified that an exception must be acquired for the non-commercial use provision of 3), the provisions of all four licenses of the source data sets will be supported and again the license with the most stringent conditions would control the use of the derivative work.

*Fifth Data Set:* The metadata or embedded code for the fifth data set indicates that the dataset is offered under a publicly published non-standard license that may be unique to the particular firm making the offering. It might be the “standard license” offered by the firm but would not be “standard” in a global commercial sense. An appropriately designed web-wide metadata reporting and search environment should lead any potential user of this geo-referenced data offering to the URL where the user may read the explicit terms of the full license. Notice here that the web mapping exercise or other derivative product generation needs to halt from a legal perspective until either human discourse takes place and an agreement is reached or the user reads the full license and agrees to it, typically through a “click license” process.

At this point, automated determination of legal interoperability is no longer realistically possible. The full license must be read by the human user and an evaluation made by the human reader to determine whether any of the license provisions are incompatible with any of the licenses under which datasets 1 through 4 were offered. By example, the licensing language applied to data set 5 may require something that one of the previous licenses explicitly bans. Thus, the licenses are incompatible and the only way to resolve the inherent conflict is to negotiate an alternative for one or both of the conflicting licenses. This need to stop and humanly read a non-standard license and potentially negotiate new licenses greatly impedes the generation of products by potential users. As such, our hypothetical user may want to avoid dealing with the supplier of data set 5 if she has any alternative data set that might meet her technical needs, even if less optimal technically and even if more expensive, if it was legally interoperable.

When licenses are incompatible it is often very difficult and sometimes next to impossible to identify one or more creators that can authoritatively negotiate an alternative for the

conflicting provisions. Crowd sourced geospatial resources are particularly problematic. Open Street Map (OSM) has recently addressed this problem by requiring recently joining contributors to grant to the OSM Foundation authority to be able to grant exceptions.<sup>3</sup> In this way, those who are unable to live with the no-derivative works provision of the standard OSM license now have a single authority to petition that may grant an exception, at least for more recent contributions.

Although data set 5 may not be interoperable legally, having in the search results a direct link to the full and complete standard license of the company, agency or crowd-sourcing group is at least a significant improvement over the current situation of having little to no idea of the legal status of the dataset after finding it through a search. This eliminates the often arduous step of tracking down the standard license or sales contract offered by a firm. The standard contract and license offerings of many firms may not even be publicly available or openly posted for a wide range of reasons.

*Sixth Data Set:* The sixth data set that our user found in her search has recorded in the metadata a link to the general conditions of use for the data set. Although not expressed in conventional contractual or license language and although the user is not forced technologically to agree to anything before downloading the data set, these posted conditions may be viewed by the courts as implied or actual contractual offers depending upon the jurisdiction. As such, if the user downloads and uses the data without gaining explicit permission of the provider to alternative terms, the user may be held by the courts to have impliedly or actually consented to the provisions. Again the onus is on the user to explore whether the conditions of use of this particular dataset inherently conflict with any of the other licenses. Similar to the conditions of datasets 5 and 7, the additional overhead burden in investigating and resolving the legal status may make the data too onerous to use.

## **5. CONCLUSION**

To spatially enable all sectors of society to better address immediate and long-term challenges related to transportation, housing, disasters, health, energy, climate, water, weather, ecosystems, agriculture, and biodiversity, there is a need to greatly enhance ready ability to find and use relevant geospatial data and services. While substantial progress is being made in achieving these goals in terms of technical interoperability, we have yet a long road ahead in achieving legal interoperability. Lack of an operational web-wide capability allowing users to legally access and use the geospatial data of others from wide-ranging sources without seeking permission on a case-by-case basis is a major impediment to general spatial enablement for all sectors in society.

“A legal commons has been created for creative works on the Web through the use of Creative Commons (CC) licenses. With a few clicks, in less than a minute, one may create ironclad licenses for creative work to make it practically and legally accessible to others. Well over 200 million such open access licenses have already been created, and the advanced functions of most major web search engines allow one to restrict web searches to return hits to only those sites with the standard CC license specified in the search.” (Onsrud et. al., 2010) An analogous global legal commons and marketplace environment for geospatial data and services could result by implementing clear and simple licensing provisions using the

---

<sup>3</sup> For OSM contributors joining on or after May 11, 2010, the OSM Foundation should be able to grant the exception for their specific data contributions. See “What is the licence for OpenStreetMap's geodata?” at [http://wiki.openstreetmap.org/wiki/Legal\\_FAQ](http://wiki.openstreetmap.org/wiki/Legal_FAQ)

recommendations in this chapter. Success or failure would be easy to measure. A successful implementation would be one in which, within five years, millions of geographic data sets become available to all through standardized open access and commercial licenses and readily findable through web wide searches. Widespread spatial enablement of society among all sectors will require that the practitioner, business, government and research communities join together to pursue such an approach.

#### **ACKNOWLEDGEMENTS**

The author wishes to thank and acknowledge Steven F. Browdy, Kevin Pomfret and Mikel Maron for their reviews and comments on an earlier version of this chapter but I did not accept all of their suggestions and take full credit for any errors, inconsistencies or controversial statements that may remain.

#### **REFERENCES**

- Creative Commons (2010), Creative Commons Licenses, at <http://creativecommons.org/about/licenses/meet-the-licenses>, [accessed 15 June 2010].
- DeAcosta v. Brown, 146 F.2d 408 (1944), cert. denied, 325 U.S. 862 (1945).
- National Research Council (2004), Licensing Geographic Data and Services, Washington D.C: National Academies Press.
- Onsrud, Harlan J. (2008), "Implementing Geographic Information Technologies Ethically", ArcNews, Fall 2008, Vol 30 Number 3, pp. 1-8 at <http://www.esri.com/news/arcnews/fall08articles/implementing-gi-technologies.html> [accessed 3 July 2010].
- Onsrud, H. (2009), Liability for Spatial Data Quality. In Spatial Data Quality: From Process to Decisions (Devillers, R. and Goodchild, H.), CRC Press, Boca Raton, FL, pp. 187-195 at <http://www.spatial.maine.edu/~onsrud/LiabilityForSptlQlty.pdf> [accessed 14 June 2010].
- Onsrud, H.J., James Campbell and Bastiaan van Loenen (2010), "Towards Voluntary Interoperable Open Access Licenses for the Global Earth Observation System of Systems (GEOSS)", International Journal for Spatial Data Infrastructure Research, Vol 5 (2010), pp. 194-215 at <http://ijmdir.jrc.ec.europa.eu/index.php/ijmdir/article/view/168/203> [accessed 3 July 2010].