

2010

Towards Voluntary Interoperable Open Access Licenses for the Global Earth Observation System of Systems (GEOSS)

Harlan J Onsrud, *University of Maine - Main*

James Campbell

Bastiaan van Loenen

Towards Voluntary Interoperable Open Access Licenses for the Global Earth Observation System of Systems (GEOSS) *

Harlan Onsrud¹, James Campbell¹, Bastiaan van Loenen²

¹University of Maine (onsrud@spatial.maine.edu, campbell@spatial.maine.edu)

²Delft University of Technology (b.vanloenen@tudelft.nl)

Abstract

Access to earth observation data has become critically important for the wellbeing of society. A major impediment to achieving widespread sharing of earth observation data is lack of an operational web-wide system that is transparent and consistent in allowing users to legally access and use the earth observations of others without seeking permission from data contributors or investigating terms of usage on a case-by-case basis. This article explores approaches to supplying a license-based system to overcome this impediment in the context of the Global Earth Observation System of Systems. It discusses the benefits and drawbacks of the explored approaches and suggests an integrated legal and technological approach for supplying an effective web-wide sharing environment for earth observation data.

Keywords: Licensing, Usage Rights, Open Access, GEOSS

* This work is licensed under the Creative Commons Attribution-Non commercial Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

1. INTRODUCTION

Among the stated goals of the Global Earth Observation System of Systems (GEOSS) is to proactively link existing and planned Earth observing systems around the world so that the readings and processed results from thousands of different Earth-orbiting, airborne, and land-based sensors and related data may be more effectively used to support decision making for a variety of societal contexts. The GEOSS vision, similar to other smaller scale visions that have come before, is (1) to enable content providers to link data and services through common technical standards and processes, and (2) to enable decision-makers from all walks of life to gain user-friendly access to an extraordinary range of reliable and up-to-date information on their desktops or through mobile devices.

As of this writing, members of the Group on Earth Observations (GEO) include 80 national governments and the European Commission. GEO also recognizes 58 intergovernmental, international, and regional participating organizations that have mandates related to Earth observations. Most of the material to be shared through GEOSS is envisioned as being contributed by national government agencies from across the globe, although the current architecture and policies allow for contributions from almost any source.

The GEOSS common infrastructure, which is currently at the test and prototype stage, allows the registration and linking of services, data, and products. None of these resources are hosted centrally. Instead, the GEOSS architecture provides an interface for registering such resources (GEO, 2010a) and a service for potential users to find the resources (GEO, 2010b). A catalogue or registry service is required to find these geospatial resources because geographic data and services appropriate for specific uses are currently very difficult to find through traditional Web browser searches. Web search engines specialize in processing words and phrases or in finding embedded code but earth observation data does not typically consist of words or phrases (although metadata does) nor, to date, has a system been deployed generally for embedding coded identifiers in geographic data sets or portions of such datasets. If generally deployed, such identifiers would also facilitate the tracking of datasets and portions of them over time.

A Data Sharing Task Force has been assembled to develop consensus-based implementation guidelines for GEOSS that would encourage practices by government agencies and others consistent with an already adopted set of data sharing principles. The member approved ten-year implementation plan includes the phrase: *“there will be full and open exchange of data, metadata, and products shared within GEOSS.”* (GEO, 2005) This phrase and the other overarching principles for data sharing as stated in the plan are subject to broad interpretation. Most parties at the table appear to agree that governments and others should not register data and services made available through the GEOSS

architecture unless these are made available in conformance with the GEOSS data sharing principles. Participating member governments, however, have very different views on how the agreed upon principles should be interpreted.

The current GEOSS definition of “full and open exchange” is contained in the Implementation Guidelines for GEOSS Data Sharing Principles (GEO, 2009) and states that “full and open exchange means that data, metadata and products made available through the GEOSS are made accessible with minimal time delay and with as few restrictions as possible, on a non-discriminatory basis, at minimum cost for no more than the cost of reproduction and distribution.” This definition is useful primarily in policy contexts and would be very difficult to enforce in licensing contexts.

Currently, the GEOSS architectural design allows registration of datasets, products, and services but the legal status of registered resources is reported inconsistently. In practice, the legal status is often indeterminable. As a result, it is unclear whether one may legally use a registered resource or a combination of them to generate a derivative product. Thus, the ability to use each and every registered resource in a potential pool of tens of thousands must be explored outside the GEOSS infrastructure on a case-by-case basis. This lack of legal clarity is so inefficient that the GEOSS solution today provides very little added-value over general web-based searches. If not corrected, this inefficiency will result in a disincentive for government agencies and other organizations to register resources through the GEOSS system. The result will be a system of much reduced value to intended users.¹

2. GOAL OF THIS ARTICLE

In this article, we argue that GEOSS is unlikely to succeed in its primary goals unless it offers a consistent and transparent set of standard licenses that contributors can agree to abide by in offering to share data through the GEOSS architecture. Developing these licenses will require engaging the legal scholarly community. If the GEOSS community fails to appropriately address the practical legal realities at the outset in its implementation architecture, the resulting pool of linked data and resources will be polluted by conflicting property rights that will inevitably invite legal disputes and endanger the contributions of time, effort, and resources made by scientists and agency administrators.

¹ It should be noted that GEOSS is a work in progress. In furtherance of the GEOSS data sharing principles in the Ten Year Implementation Plan, Implementation Guidelines for the GEOSS Data Sharing Principles, Document 7 (Rev2) were accepted at GEO-VI in November 2009 in Washington, D.C. (GEO, 2009) and an action plan is being pursued as of the date of this article. In parallel, an update of the GEOSS Architecture Implementation Pilot (AIP) that includes a suggested exploratory licensing implementation was issued 29 January 2010 and a third call for participation in the pilot (AIP-3) has been issued. (GEO, 2010c)

Simply put, not cooperatively engaging the legal community up front is almost certain to antagonistically engage legal practitioners later on. The practical legal realities need to be addressed from the outset in the system design if the GEOSS system is to meet its goals of enabling increased use of earth observation data to more effectively support decision making.

A draft version of the Implementation Guidelines for the GEOSS Data Sharing Principles stated that “GEO should consider utilizing machine-readable, common-use licensing approaches for copyrighted data products that place primary responsibility for compliance on the users rather than enforcing compliance through technical controls on data access.” This article outlines how such an approach might be implemented.

We suggest an implementation solution for GEOSS whereby the primary economic, policy, and legal concerns of most potential contributors to GEOSS can be accommodated within the stated GEOSS principle of data sharing. Our suggested implementation avoids a one-size-fits-all solution and advocates legal simplicity. It promotes transparency in the legal status of GEOSS data offerings, allows governments and others to experiment in making data available under various open access conditions, and encourages the movement over time of all contributed data towards an open access status unfettered by legal restrictions.

3. THE PROBLEM²

If a human, or a human-initiated software agent, extracts from a geographic dataset or database without the legal authority to do so, substantial liability exposure is incurred. By analogy, 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.

If one were to steal a CD from a music store, the maximum typical fine in the U.S. might be about \$1000. However, if one were to download the same ten songs contained on the CD from the Internet, liability exposure could be as high as \$1.5 million. (Statutory damages for copyright violations if behaviour is not proven as “wilful” are set at “not less than \$750 or more than \$30,000” per infringement, while statutory damages for “wilful” behaviour are not more than \$150,000 per infringement. The Recording Industry Association of America (RIAA) has been very effective in using these very high damage limits to extract large settlements from individual students and others accused of illegal downloading of music files.) In a similar manner, if a web mapping service or an automated data mining

² This section is a direct quote with minor edits from Onsrud, 2009.

software program is used to draw together data from ten sources distributed across the Web, the same potential \$1.5 million liability exposure could arise. Similar tight intellectual property laws and increasing penalties are prevalent across the globe.

Suppose an agency, organization, or GEO hosts only a Web mapping service, or provides only a Web-based interface for searching and accessing data hosted by others. Now suppose the site's end users apply the site's sharing technology to extract data from other sites to draw their own maps. Is the web mapping site in its role as a peer-to-peer linking or "mashup" host site relieved of liability?

Legal cases in the U.S., such as *A&M Records, Inc v. Napster* (2000), and the recent *Pirate Bay* case in Sweden can be instructive in this regard, even though the circumstances in each are very specific. In both cases, the operators of the web sites were found guilty of promoting copyright infringement even though they claimed they only facilitated the sharing of files by others, and even though many instances of sharing were legal. In the *Pirate Bay* case, decided on 17 April 2009, each of the four defendants was sentenced to a year in jail and fined roughly \$900,000 in damages. (Carrier, 2009) In a legal context, therefore, in providing a map service or peer data sharing service, a question that could arise is not simply whether the operation of the service provides a valuable public benefit or whether legal uses are being made of the site, but also whether this is an operation for facilitating widespread illegal sharing and use of geographic data selections or files without the explicit permission of the owners of those materials.

Ensuring the quality of data requires completeness, and completeness in the global legal environment requires that one (a) knows the legal status of the data that one draws from others, and (b) has strong confidence that one has legal authority to use the data in a specific context. 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.

A core problem the GEOSS community is confronting with regard to understanding usage rights is the acquisition by database and dataset developers of automatic copyright upon creation, whether creators want it or not. In most jurisdictions of the world, if one creates an original doodle on the back of a napkin, copyright occurs instantaneously in that doodle with no need to register the right, regardless of whether most people would find this instantaneous copyright to be reasonable.

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, including the United States. 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.).

In truth, one cannot know for certain whether there exists some legal originality in a posted geographic data set (and neither can any lawyer) until the gavel falls in a court of law on a case-by-case basis. Thus, most lawyers across the globe advise their clients that they should always assume that a party will emerge to sue unless explicit permission to use is acquired when drawing from the materials of others.

For example, assume that a user has just used the GEOSS portal to link to and extract data elements from fourteen geographic datasets in an automated data mining or web mapping use. Many of us assume that the vast majority of those fourteen sites 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 fourteen sites has posted a license provision that some specific use breaches, and the posted contract or license provisions of some of the sites are very likely to be in conflict with each other.

Some of the fourteen 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.

Sending a request to the addresses of those fourteen sites asking them for explicit permission to use the data extracted from their web sites is burdensome and nonsensical to the typical user. Many of the recipients of the message might also view the request as silly since, after all, they would not have posted the data in the first place if they did not want it to be used. Often, there will simply be no response to a request for permission to use material. Practically speaking, however, one should assume that one or more legal claimants are lurking and ready to sue or prosecute if explicit permission has not been obtained from each and every owner.

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.

Common sense argues that we need strong copyright protections in order to prevent digital thieves from stealing our geographic data offerings against our will. In the world of paper, the law has established a nuanced balance between the rights of creators to benefit from their original works and the rights of new creators to build from the past contributions of those who came before. This balance does not exist in the digital realm.

There are no legal exceptions on the Internet for uses that do not involve making a *copy*, for example, loaning a physical book many times, or selling a used CD. That is because on the Internet virtually every use involves creating a new copy.

In addition, uses normally allowed by default in a particular jurisdiction (e.g. the right to extract such things as facts, the right to *fair use* or comparable concepts) are readily negated on the Internet by imposing contracts, e.g., by posting conditions of use notices, or requiring users to click a license before allowing the downloading of data or software. Laws across the globe are skewing the legal landscape that has been disturbed by changes in technology even further toward copyright owners, rather than restoring the balance between copyright owners and users. Changed technological conditions, as well as legislatures endeavouring to protect copyright owners from theft have shifted society to what Lawrence Lessig refers to as a “permissions culture” in the world of the Internet. (Lessig, 2004) One now almost always needs permission to extract or extend from the work of others. As Lessig states, this is a world dominated by lawyers, and it is the world in which GEOSS seeks to operate.

4. A SUGGESTED SOLUTION FOR GEOSS: ONE-RULE-FITS-ALL

To make GEOSS feasible from a legal perspective, some technologists have suggested that the problem may be solved simply by requiring all those submitting or sharing data through the GEOSS architecture to adhere to a single legal license arrived at through a consensus process. Whether a government or organization chooses to share data through this license environment would, of course, be completely voluntary. It would also be nearly impossible.

Arriving at a practical single consensus definition of “full and open exchange” for use in guiding development of such a license is a substantial challenge. The draft White Paper on the GEOSS Data Sharing Principles (GEO, 2008) suggested that “full and open exchange” should mean that “data and information derived from publicly funded research are made available with as few restrictions as possible, on a non-discriminatory basis, for no more than the cost of reproduction and distribution.” (GEO, 2005 with definition drawn from National Research Council, 1997) This definition covered only data and information derived from “publicly funded research,” and the phrase “as few restrictions as possible” was open to wide interpretation.

In further developing the Implementation Guidelines for the GEOSS Data Sharing Principles, some representatives from member nations wanted to expand coverage of the definition such that all rights by all parties in all materials registered through GEOSS would be waived to the greatest extent possible in all jurisdictions of the globe. This would place all data linked through GEOSS in the public domain and thus create a legal status for the entire linked contents of GEOSS similar to that of federal government records in the United States.

Representatives of other nations suggested a much more restrictive interpretation whereby only data and information derived exclusively from “publicly funded research” need be provided on a non-discriminatory basis for no more than the cost of reproduction and distribution; only metadata need be made available through GEOSS without restrictions on its subsequent use; and data developed in whole or in part by government for purposes other than research may impose a wide variety of restrictions, including monetary charges, and still be acceptable for disseminating through the GEOSS portal.

Most parties at the table attempting to reach a one-rule-fits-all compromise solution were suggesting positions somewhere between the two just stated, but even these were often in conflict with each other.

One benefit of the one-rule-fits-all approach is that contributors could be required in the registry process to agree that the single rule applies to the data or resources being registered and made available through the GEOSS portal, and that all other previous licensing terms concerning the resource are superseded

by the GEOSS license terms. Thus, in the previous example of drawing data from fourteen different sources, if all of the resources were accessed through the GEOSS architecture, the user would have a clear and consistent picture concerning the legal status of the data sets and allowed subsequent uses.

Regardless of which “one rule” is reached most of the parties at the table will find a take-it-or-leave-it single rule difficult to support. Some contributors will want to impose fewer restrictions in their interpretation of “full and open access” while others will want to impose greater or conflicting restrictions in their interpretation.

If the “one rule” solution is declared as only a minimum for making contributions to or through the GEOSS architecture, and if additional restrictions can be attached to individual data sets, we arrive back at the situation of an unknown, conflicted pool of data and resources in which the actual rights in the various data sets that one might use in a single application are not known and the conditions of use of each and every source data set must be investigated.

Neither people nor agencies will contribute to a pooled resource if they do not have a clear legal right to draw combinations of data from the pool suitable for their purposes. Thus, in our opinion, the technological expedient of conforming to a single simple rule for current contributions is unlikely to achieve the ultimate goal of widespread sharing of public sector geographic information. A more flexible approach should be pursued.

5. ANOTHER SUGGESTED SOLUTION: MYRIAD RULES

Another approach advocated by others is to allow those organizations and individuals making data or resources available through GEOSS to allow their own user-created independent licenses or contracts to control under the assumption that the license terms meet the definition of providing the “full and open exchange” of data and products “recognizing relevant international instruments and national policies and legislation.” Contributors would be given great deference to follow their own definition for “full and open,” even if their definition departs substantially from that of the official GEOSS definition. This is essentially the *status quo* for entries coming into the existing GEOSS registry. In this approach each contributor has the obligation to explicitly state in the metadata the legal terms of use and specify any limitations. Many contributors are not doing so. Further, problems arise in this approach since it depends upon potentially tens of thousands of contributors of space-based, aerial, and *in situ* earth observation data to summarize accurately and comprehensively the license terms in their metadata. Within most organizations, licensing language is constructed by lawyers yet the metadata is constructed and submitted typically by technical experts. Even if accurate legal summarizations in the metadata can be achieved, a single license or contract term such as *copyright* is likely to have

slightly different, or even vastly different, legal meanings and ramifications from one national jurisdiction to the next.

It has been suggested that controlled vocabularies could help in this process. In our opinion, using a controlled vocabulary in entering metadata for legal terms provides little benefit unless each contributor truly understands the meaning of the legal terms they are selecting in the context of their own national laws and reflects those laws accurately. For example, the *limits of applying copyright to data* is quite challenging even for lawyers to understand, and there are probably few potential geographic data contributors who could choose from among several controlled definitions for the term *copyright* to select the one that best matched the meaning as supported by the laws of their jurisdictions. With most contributors lacking understanding to make informed choices, even if the GEOSS community went to the considerable effort to create legally supportable controlled vocabularies, wrong choices in conflict with local laws would likely be rampant. Those wrong choices would lead to disputes over which terms should be controlling in the event of conflicts, and wrong metadata vocabulary selections would probably lead at least some users into unsettling liability exposure situations. In addition, controlled vocabulary terms could miss or mismatch at least some jurisdictional definitions simply due to the very large number of jurisdictions in the world. Such complexity creates great fodder for lawyers but generally would not enhance or support wide spread sharing.

The Internet currently has hundreds of thousands of openly hosted valuable geographic data sets being offered under myriad terms of use. Some of the terms of use are explicitly stated, while the terms of use for other data sets are controlled through the default laws of the nations where the data is hosted. The resulting complexity substantially impedes the ability of many to use or share data, derivative products, and online services that depend on geographic data. Development of the GEOSS common infrastructure provides a rare opportunity to guide the community out of the legal morass within which all of us are currently ensnared.

6. A BETTER SOLUTION: VOLUNTARY INTEROPERABLE OPEN ACCESS LICENSES

In order to encourage motivated participation by larger numbers of government agencies and other organizations to make their earth observations available through GEOSS, we urge that a limited number of well-defined license options be made available for those contributing data through the registry. The goal is to (1) accommodate the needs of most agencies desiring to make at least some data available through GEOSS while minimizing limitations on use and reuse of data that is immediately accessible through GEOSS, and (2) grow the pool of GEOSS linked data that enters the public domain over time.

Open access license options that we believe should be considered for possible inclusion in the common infrastructure of GEOSS include those set forth in Table 1. However, through the assessment of each option that follows, we have concluded that some of the listed options should NOT be included in the proposed interoperable open access arrangement for GEOSS. Our set of recommended licenses for deployment in the common infrastructure follows in Table 2 at the end of this section.

Table 1: Open Access License Options Assessed for Possible Inclusion in GEOSS

I. Dedication to the Public Domain
II. Creative Commons Licenses
a. Attribution Required
b. Non-commercial Use Only
c. No Derivative Works
d. Share Alike
III. Specialized GEOSS Open Access Licenses
a. Research and Education Uses Only
b. Developing Country Users or Uses Only
c. Environmental Uses Only
d. Humanitarian Assistance Uses Only
e. GEOSS Societal Benefit Areas Only
f. Cost of Dissemination Fee Required for Initial Download
IV. Non-Standard Open Access License

Any of the above license options chosen for inclusion in the common infrastructure would be synchronized. That is, if the license under I was chosen by a user, the system would not allow licenses to be chosen under II, III or IV nor vice versa. Within license options II and III, one or all of the sub-option license provisions could be chosen. In all cases, the license selected would indicate that the licensing language chosen for the linked contribution supersedes any previous licensing language regarding the same dataset. (Note that there is no need to actually deliver a copy of the dataset to the GEOSS portal to have the license control, although archiving each dataset licensed through GEOSS might be advisable for later evidentiary purposes in the event of disputes)

The Dedication to the Public Domain instrument (under I. in Table 1) should be used in those instances in which an agency or organization chooses to impose no limitations on use or reuse of the data that is being linked through GEOSS. This option is highly recommended to promote widespread use, provide the greatest flexibility for users from all sectors of society, provide the greatest legal clarity, and negate almost all legal impediments to sharing. If contributors affirmatively waive all rights to the greatest extent possible, there are no

restrictions to conflict with each other. If attribution for contributors is still desired, other means than a copyright derived right may be utilized, e.g., trademark or other legal device.

Widespread use of the Dedication to the Public Domain option would ensure that geographic data could be used legally across the Web for the broadest range of purposes. In conjunction with this selection option, we recommend use of CC0, the Creative Commons Zero or No Rights Reserved commitment (Creative Commons, 2010a). Creative Commons Zero language waives all copyright and database rights to the extent that one may have these rights in any jurisdiction, and is the best current option to ensure that data can be used legally for general web mapping and feature services, data mining, copying, and extraction.

In the event that a Dedication to the Public Domain is not fully supported legally in a particular jurisdiction, the CCO license “fall-back” position “grants to each affected person a royalty-free, non transferable, non sub-licensable, non exclusive, irrevocable and unconditional license to exercise Affirmer’s Copyright and Related Rights in the Work...” This essentially grants all users all of the same rights under copyright law that the owner enjoys in the jurisdiction. Note that this waiver instrument is also advocated for general use by Science Commons for data sets, and, in the best of worlds, GEOSS should advocate the migration of all geographic datasets towards this option, either immediately or over time.

The Creative Commons Licenses (Table 1, II) should be used when an agency or organization desires to reserve one or more of the rights listed. Among the optional restrictions one may choose to impose include requiring attribution, restricting uses to non-commercial purposes, not allowing derivative works, and allowing others to distribute derivative works but only under the condition that those works use a license identical to your license (Creative Commons, 2010b). If a user wants to use the data without adhering to the provisions listed they have a legal obligation to contact the agency or organization directly to negotiate a specific agreement. Every license helps the contributor to “retain your copyright (and) announce that other people’s fair use, first sale and free expression rights are not affected by the license.” “Every license allows licensees (i.e., users), provided they live up to (the license provisions), to copy the work, to distribute the work, to display or perform it publicly, to make digital public performances (e.g., web casting), and to shift the work into another (medium). Every license applies worldwide, lasts for the duration of the work’s copyright and is not revocable.” (See Creative Commons Baseline Rights at Creative Commons, 2010c.) The Creative Commons No Derivative Works license option (Table 1, II c.) is fine for freely sharing a book or music file that you do not want others to change when they use it but it is highly impractical for geographic data. Thus we do not recommend its use in the GEOSS sharing environment. Unlike a music or video file, almost every use of an earth observation data file or database will involve

some extraction of a portion of the dataset, merging with other data, or processing or revisualization of the data. Within the GEOSS environment a contributor's selection of this choice would largely negate the ability to use that geographic data to create anything new.

Use of the Creative Commons Share Alike provision (Table 1, II d.), we believe, also should not be offered as a choice within the GEOSS sharing environment. This is because it causes "license piling" problems. Under the Share Alike license, all subsequent users of a contributed dataset are required to make the license of their newly derived product, map, or dataset identical to the same Share Alike license. Yet, to use our previous example, some or all of those fourteen other datasets that a user may have used to derive their new product, map, or dataset are likely to have use conditions in direct conflict with this requirement. For example, a user would be banned from combining any data drawn from a commercial source (e.g. Google Earth) in creating their derived product. Thus, the Share Alike provision would largely negate the ability of users to draw extracts from that dataset for many purposes. Wikipedia, to use another example, has employed the GNU General Documentation License for its entries (another license with Share Alike provisions) since its founding. This license conflicts with all other licenses except itself, as does the CC Share Alike option, and it took years of work to make it possible for those who wished to create new works using excerpts from Wikipedia to combine that material with material that used Creative Commons or any other "some rights reserved" license. In our opinion, neither the Creative Commons No Derivatives nor the Share Alike license is workable in the GEOSS context.

The benefit of using Creative Commons (CC) licenses is that the licenses have been meticulously developed and debated by leading legal scholars, they are well known and widely used across the globe, and the legal provisions have been translated into numerous languages. Major web search engines automatically pick up embedded html code indicating that the returned sites contain CC licensed material. Thus, web-wide searches for open access geographic data might often lead to the GEOSS web site if these licenses were used and if the GEOSS web portal site was appropriately designed.

To some, one drawback of using CC licenses is that they may not be altered in any way. To others, this might be viewed as a great benefit. Another consideration in using CC licenses is that they may not apply to some geographic datasets in some jurisdictions. CC licenses are intended for use with "creative works" or those that meet the legal standard for "originality," regardless of the jurisdiction and further are intended to be "ironclad" for such works.

As mentioned previously, most large geographic datasets have sufficient selection, coordination, and arrangement of elements contained within them to be "creative." Validity issues sometimes arise with datasets because small

extractions from large datasets may not copy the creative aspects of the dataset and simple standard datasets may not rise to the legal requirement of being “creative” or “original” in some jurisdictions. However, while the validity of such a license might be in question in a specific instance based on such a challenge, use of such a license by a data contributor might very well express sufficient evidence to a court that a binding contract was indeed intended and should be supported. Thus, while not “ironclad,” such licenses may still be sufficient in most cases. Resolving the potential validity of CC licenses as applied to a specific dataset in a specific instance would, of course, need to be assessed on a case-by-case basis. Thus it would be up to administrators from data contributing organizations to determine whether such licenses would be sufficient for their purposes in their jurisdiction.

Specialized GEOSS Open Access Licenses (Table 1, III) might be used by a contributor to GEOSS when the options under I or II fail to meet the needs of an agency or organization but the organization still desires to make substantial amounts of earth observation data openly available for some purposes. We do not recommend all the options listed due to the increasing complexity caused by each additional license but we raise them here in the interest of fully exploring the possibilities.

The third principle in the GEOSS Data Sharing Principles as specified in the Ten Year Implementation Plan distinguishes data used for “research and education,” and thus contributors might be allowed to contribute data for only this purpose if they so desired (Table 1, III a.). The Ten Year Implementation Plan also stresses the engagement of users in developing countries, facilitation of access to data about developing countries, and assistance for developing countries. As such, some contributors may want to allow use of their data contributions for these purposes but not for other purposes (Table 1, III b.). High-level societal needs to which GEOSS intends to respond as expressed in the Ten Year Implementation Plan and as advocated for special treatment by some members of the GEO Task Force include protecting the environment and reducing disaster losses. As such, further contributors may want to allow use of their data contributions for environmental protection (Table 1, III c.) or humanitarian assistance purposes (Table 1, III d.) but not for others. Others may want to allow open use for any of the societal benefit areas that GEOSS is intended to support (Table 1, III e.). Of course one might also click one, a few, or all of the provisions under III and a license would be generated automatically that would incorporate all of the specified allowable uses. Standard licenses would need to be developed by legal scholars to define the uses allowed under each of the conditions just discussed since they do not yet exist. In all of the above instances, users at a minimum should be able to freely download the data without human intervention upon clicking an agreement to abide by the relevant Specialized GEOSS License.

Option III f. in Table 1 allows a data contributor, at its option, to place users on

notice that they will be required to pay a cost of reproduction and dissemination fee before being able to download and use the dataset. Monetary transfers raise the spectre that the transfer might be treated by the courts as a “sale” of data which then may raise issues of implied warranties of merchantability or fitness for use. Thus, liability exposure may increase substantially under monetary exchange arrangements and the validity of claims of sovereign immunity by government agencies may decrease (Onsrud, 1999). Licensing language could be added in this case to minimize the potential liability exposure although, in our opinion, the preferred option is to make the data available without cost conditions and avoid any of the potential pitfalls noted above.

If a data contributor clicks license option IV in Table 1, it means that the agency or organization contributor believes the specific data use conditions they are offering meet the definition of being “full and open exchange of data” under their interpretation but the standard licensing provisions offered by the GEOSS architecture are insufficient for their purposes. When users confront such offerings, they would need to study the specific license or published conditions being offered by the agency or might need to go into an individualized negotiation process and separate contract to acquire and use the data. This knowledge of further non-standard requirements is valuable information for the potential user since it is now transparent that conditions beyond those imposed by most other GEOSS contributors will be imposed if this specific data is used.

Under options III and IV we highly recommend that the license language be constructed to automatically convert all contributions licensed under these categories to licenses under categories I or II, at the contributors choice, after the contribution has been linked through the GEOSS system for a set period, say, for example, four years. If no post four-year license was “clicked” by the contributor, the default post four-year license would be the option I license. This approach would migrate all GEOSS licensed data sets to the least restrictions possible after waiting a reasonable period to allow agencies and organizations to extract most or all of the commercial value out of their offerings (National Research Council, 2004).





A major drawback of the list shown in Table 1 is that each new license, while offering increased flexibility and options for contributors, also creates increased complexity for users.

7. RECOMMENDED LICENSE OPTIONS FOR USE IN GEOSS

We therefore recommend that no more than six license options be offered to GEOSS contributors and our recommendations are indicated in Table 2. The first five indicated options should meet 80% of the needs of potential GEOSS contributors and the last option should satisfy the other 20%. Under this

arrangement, there is room for everyone in the world to either license open access contributions through the GEOSS infrastructure, or to simply retain their current independent licenses (or develop new ones) and still register their resources through the architecture.

Table 2: Recommended License Options for Use in GEOSS

Type of License	Symbol
I. Dedication to the Public Domain	
II. Creative Commons Attribution Required License	
a. Attribution Required	
b. Non-commercial Use Only	
III. Specialized GEOSS Open Access Licenses	
a. GEOSS Societal Benefits Areas Only	SB
b. Cost of Dissemination Fee Required for Initial Download	\$
IV. Non-Standard License	Other

7.1. BENEFITS OF THE SUGGESTED APPROACH

The overall licensing approach suggested above accomplishes several major objectives.

First, it supports a generalized global solution for conveying and acquiring intellectual property rights in data sets on a web-wide basis. No intellectual property negotiations are required to understand and legally use the data that is made available through the GEOSS architecture.

Second, it provides licensing interoperability for the sharing of data by standardizing the detailed license provisions across government agency and organization offerings made available through GEOSS. This eliminates conflicts among current licenses that require myriad individualized negotiations to resolve. In such a conflicted licensing environment, sharing does not take place because few want to take the time and effort to resolve the conflicts, conflicts that not infrequently are intractable. Non-transparent and inconsistent licenses have been identified as a major barrier to the sharing of data across the geospatial community (MICUS Management Consulting GmbH, 2008; Groot, et al, 2007; National Research Council, 2004; Spatial Technologies Industry Association, 2001; KPMG, 2001; Ravi bedrijvenplatform, 2000; Meixner and Frank, 1997). This approach addresses this problem directly.

Third, this approach enables the use of simple symbols to convey to potential data users any legal conditions imposed on the data efficiently and effectively. What a user can or cannot do with each data set becomes much more obvious and transparent. In returning search results on the GEOSS portal that meet the user's thematic and geographic requirements, each dataset in the listing returned on the user's screen would now have one or more symbols by each entry. Sample suggested icons are provided in Table 2. Thus, the user can quickly identify and reject those data offerings that will not meet their legal use requirements. At least one other geoportal web site, the National GeoRegister of the Netherlands, has begun to reference the legal status of datasets in this manner.

Fourth, primary responsibility for compliance with the licenses is on the shoulders of the users, rather than on data suppliers using technological controls that often greatly impede the ability to effectively share or use data. We fully realize that any derived product will be controlled by the most restrictive license attached to any of the source data sets used but the proposed approach allows users to readily avoid using data with unacceptable restrictions. It provides an incentive for contributors to make their data less restrictive over time in order to allow it to be more widely used.

Fifth, the approach suggested sets up the ability to embed code within each data offering processed through GEOSS, and thereby make licenses machine-readable on a web-wide basis.

Let us assume that a user would like to find an appropriate digital elevation model for all of Europe. Table 3 shows the results of a hypothetical search through GEOSS using our suggested licensing approach and how two organizations might each choose to license their data set offerings.

If the two hypothetical organizations were to register their data sets through GEOSS with the licenses indicated, it is readily apparent that the 30-meter resolution data set may be acquired and used for free for the nine societal benefit areas of GEOSS (disaster, health, energy, climate, weather, ecosystem, agriculture, and biodiversity). Any other planned uses of that data set would require negotiation with the contributor(s).

The second data set of a lower resolution is available at a higher cost and is probably available for any use for those willing to pay the price. This may or may not be a barrier to use. Because the second hypothetical contributor chose the Other option, each potential user would need to thoroughly investigate the terms of the non-standard license. The conditions of use for both data sets have become much more transparent for all potential users.

It should be noted that the voluntary interoperable open access license

environment just described could also be deployed in a commercial context. The geographic and earth observation data sales and services market is extremely fragmented currently. Offering a global and industry consistent suite of standard commercial data license options in a similar web-wide deployment would make the commercial market much more transparent and efficient and allow large numbers of suppliers of earth observation 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 pursuing data sale economic models have not sought to unify or grow the market along these lines but a successful GEOSS implementation for voluntary open access licensing might provide the example for that later expansion.

Table 3: Hypothetical Result of a GEOSS Search

Data Sets Matching Search Criteria	Applicable License
<i>Data Set Title:</i> ASTER Global Digital Elevation Model (G DEM) <i>Abstract:</i> grid of elevation postings: 1 arc second, app. 30 meter resolution <i>Price:</i> zero <i>Contributor:</i> METI and NASA	SB
<i>Data Set Title:</i> EuroDEM <i>Abstract:</i> grid of elevation postings: app. 60 meter resolution, <i>Price:</i> €710,000 for five year period and unlimited number of users <i>Contributor:</i> EuroGeographics and the contributing European National Mapping Agencies	Other

8. CONCLUSION

A major goal of GEOSS is to promote the sharing of earth observation data that would be useful to the broad global community in addressing nine societal benefit areas: disasters, health, energy, climate, water, weather, ecosystems, agriculture, and biodiversity.

To achieve this goal, a simple, flexible, and usable system for indicating the legal usage status of earth observation data is necessary.

A legal commons has already 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. See, for example, the Advanced Search functions of the Google search engine.

An analogous legal commons environment for contributed geodata could result by implementing clear and simple licensing provisions using the recommendations above. The GEOSS architecture would be viewed as a very great success if within five years it could claim that millions of geographic data sets are now accessible through open access licenses and readily findable through web wide searches.

Not all geographic data creators and collectors will contribute their data to a legal commons environment, but many would. Naturally, if specific data contributions would raise concerns regarding national security, proprietary interests, privacy, confidentiality, indigenous rights, and/or conservation of sensitive ecological, natural, archaeological or cultural resources if made openly available, the contributing agency or organization should go through an evaluation process and not publish such data through GEOSS if the evaluation indicates that the drawbacks outweigh the benefits of making the data widely available.

A GEOSS earth observation data commons would constitute a valuable web-wide resource providing assured legal authorization to copy datasets, extract from databases, provide web mapping and web feature services, and engage in data mining. The legal right to carry out such activities under the current status of web and geospatial technology development and in the current global legal environment, however, is very much in question. It is to a data commons architecture, in our opinion, that GEOSS should be aspiring.

We strongly suggest that the geospatial community work closely with the legal scholarly community to develop and embrace some of the legal approaches already found to be successful in facilitating the legal sharing of resources generally across the Web by adapting them in a manner similar to the one suggested here.

ACKNOWLEDGMENTS

The authors wish to thank and acknowledge Robert Chen, Anne Fitzgerald, Doug Nebert, Paul Uhlir and John Wilbanks for their reviews and comments on an earlier version of this article but we did not accept all of their suggestions and take full credit for any errors, inconsistencies or controversial statements that may yet remain.

REFERENCES

- A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001), affirming, 114 F.Supp.2d 896 (N.D. Cal. 2000).
- Carrier, M.A. (2009). "BitTorrent: Legal Nightmare or Future Business Model?" at <http://blog.oup.com/2009/04/bittorrent/>, [accessed 13 April 2010].
- CODATA (2008). *White Paper on the GEOSS Data Sharing Principles [Review Draft]*, Paris: CODATA, at [http://www.earthobservations.org/documents/dsp/Draft White Paper for GEOSS Data Sharing Policies_27Sept08.pdf](http://www.earthobservations.org/documents/dsp/Draft%20White%20Paper%20for%20GEOSS%20Data%20Sharing%20Policies_27Sept08.pdf), [accessed 13 April 2010].
- Creative Commons (2010a). CC0 at <http://creativecommons.org/license/zero/>, [accessed 13 April 2010].
- Creative Commons (2010b). Creative Commons Licenses, at <http://creativecommons.org/about/licenses/meet-the-licenses>, [accessed 13 April 2010].
- Creative Commons (2010c). Baseline Rights, at http://wiki.creativecommons.org/Baseline_Rights, [accessed 13 April 2010].
- DeAcosta v. Brown, 146 F.2d 408 (1944), cert. denied, 325 U.S. 862 (1945).
- Fitzgerald, A. and K. Pappalardo, (2009). "Sharing environmental data: the role of an information policy framework and copyright licensing", *Proceedings of eResearch, November 9-13, 2009*, at <http://www.eresearch.edu.au/fitzgerald2009>, [accessed 13 April 2010].
- Groot N., A. Bregt, J. Crompvoets, and Y. Pluijmers (2007). Toegankelijkheid publieke geodata vooral gehinderd door juridische barrières [Accessibility of public sector geodata mostly hampered by legal barriers], VI Matrix 15 (8): 6-8.
- Group on Earth Observations (2005), *10-Year Implementation Plan*, at [http://www.earthobservations.org/documents/10-Year Implementation Plan.pdf](http://www.earthobservations.org/documents/10-Year%20Implementation%20Plan.pdf), [accessed 14 April 2010].
- Group on Earth Observations (2008), *White Paper on the GEOSS Data Sharing Principles [Review Draft]* CODATA, Paris 27 September, 2008 [http://www.earthobservations.org/documents/dsp/Draft White Paper for GEOSS Data Sharing Policies_27Sept08.pdf](http://www.earthobservations.org/documents/dsp/Draft%20White%20Paper%20for%20GEOSS%20Data%20Sharing%20Policies_27Sept08.pdf)
- Group on Earth Observations (2009). *Implementation Guidelines for GEOSS Data Sharing Principles, Document 7 (Rev 2)*, at [http://www.earthobservations.org/documents/geo_vi/07_Implementation Guidelines for the GEOSS Data Sharing Principles Rev2.pdf](http://www.earthobservations.org/documents/geo_vi/07_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf), [accessed 13 April 2010].

- Group on Earth Observations (2010a). GEOSS Components and Services Registry, at http://www.earthobservations.org/gci_cr.shtml, [accessed 14 April 2010].
- Group on Earth Observations (2010b). The GEO Portal, at http://www.earthobservations.org/gci_gp.shtml, [accessed 14 April 2010].
- Group on Earth Observations (2010c). *GEOSS AIP Architecture*, at http://www.earthobservations.org/documents/cfp/20100129_cfp_aip3_architecture.pdf, [accessed 14 April 2010].
- KPMG Consulting Inc. [G .Sears] (2001). *Geospatial Data Policy Study*, GeoConnections Policy Advisory Node, Ottawa, Canada: GeoConnections.
- Lessig, L. (2004), Chapter 10, *Free Culture: The Nature and Future of Creativity*, New York: Penguin Press.
- McCurry, D. (2007). *Provenance Tracking in a Commons of Geographic Data*, Master's Thesis, Orono: University of Maine.
- Meixner H., and A.U. Frank (1997). *Study on policy issues relating to geographic information in Europe*, IMPACT program, European Commission.
- MICUS Management Consulting GmbH (2008). *Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors*, European Commission, at http://ec.europa.eu/information_society/policy/psi/docs/pdfs/micus_report_december2008.pdf, [accessed 14 April 2010].
- National Research Council (1997). *Bits of Power: Issues in Global Access to Scientific Data*, Washington: National Academy Press.
- National Research Council (2004). *Licensing Geographic Data and Services*, Washington D.C: National Academies Press.
- Onsrud, H. J. (2009). Liability for Spatial Data Quality, at <http://www.spatial.maine.edu/~onsrud/LiabilityForSptlQlty.pdf>, [accessed 14 April 2010].
- Onsrud, H. J. (1999). "Liability in the Use of GIS and Geographical Datasets", in Longley, D., Goodchild, M., Maguire, D. and D. Rhind (eds.). *Geographical Information Systems: Vol.2 Management Issues and Applications*, New York: John Wiley and Sons, Inc., pp. 643-652.
- Pira International Ltd, University of East Anglia, KnowledgeView Ltd (2000). *Commercial exploitation of Europe's public sector information*, European Commission, at

http://www.epsipius.net/psi_library/reports/commercial_exploitation_of_europe_s_public_sector_information_pira_study, [accessed 14 April 2010].

Ravi bedrijvenplatform (2000). *Economische effecten van laagdrempelige beschikkingstelling van overheidsinformatie* [Economic effects of making public sector information readily available], Amersfoort, the Netherlands: Ravi bedrijvenplatform.

Spatial Technologies Industry Association (2001). *Phase 1 report: Increase private sector awareness of, and enthusiastic participation in the National Spatial Data Infrastructure (NSDI)*, Washington DC: FGDC, at http://www.fgdc.gov/library/whitepapers-reports/sponsored-reports/stia/stia_cover.doc, [accessed 14 April 2010].

QSIO (Queensland Spatial Information Office) (2006). *Government Information and Open Content Licensing: An Access and Use Strategy*, QSIC (Queensland Spatial Information Council), Brisbane, Queensland, Australia, at <http://www.qsic.qld.gov.au>, [accessed 14 April 2010].